The Intuitive Guide to

Fourier Analysis
& Spectral Estimation

with MATLAB ®

This book will deepen your understanding of Fourier analysis making it easier to advance to more complex topics in digital signal processing and data analysis in mathematics, physics, astronomy, bio-sciences, and financial engineering. With numerous examples, detailed explanations, and plots, we make the difficult concepts clear and easy to grasp.

Fourier transform developed slowly, from the Fourier series 200 years ago to Fourier transform as implemented by the FFT today. We tell you this story, in words and equations and help you understand how each step came about.

- We start with the development of Fourier series using harmonic sinusoids to represent periodic signals in continuous and discrete-time domains.
- From here, we examine the complex exponential to represent the Fourier series basis functions.
- Next, we describe the development of the continuous-time and discrete-time Fourier transforms (CTFT, DTFT) for non-periodic signals.
- We show how the DTFT is modified to develop the Discrete Fourier Transform (DFT), the most practical type of the Fourier transform.
- We look at the properties and limitations of the DFT and its algorithmic cousin, the FFT. We examine the use of Windows to reduce leakage effects due to truncation.
- We examine the application of the DFT/FFT to random signals and the role of auto-correlation function in the development of the power spectrum.
- Lastly, we discuss methods of spectral power estimation. We focus on non-parametric power estimation of stationary random signals using the Periodogram and the Autopower.

If you would like to get updates and notices of error corrections to this book, please email me at charldsp@gmail.com.

And if this book has helped you, please do post a review on Amazon for us. Thank you.
Charan Langton and Victor Levin

The Intuitive Guide to

Fourier Analysis & Spectral Estimation

with MATLAB ®

Charan Langton

Victor Levin

Mountcastle Academic

The Intuitive Guide to Fourier Analysis and Spectral Estimation

The authors gratefully acknowledge the many websites and their authors who freely share their knowledge on the Internet, making this a rich world to both showcase our work and to learn from others.

Book site for supplemental material: www.complextoreal.com/fftguide

The Library of Congress cataloging data has been applied for.

Print ISBN-13: 9780913063293
1. Fourier transform 2. Signal processing 3. Stochastic signal analysis 4. Spectral estimation

First print edition, September 2017
1 2 3 4 5 6
Printed in the United States of America

Book cover design by Josep Blanco
Book interior design and Latex typesetting by Patricio Prada
Editing by Rena Tishman, Grima Sharma

Contents

Foreward

Spectrum analysis is the single most important topic in the field of digital signal processing (DSP) and in many other fields. And spectrum analysis is performed by implementing the various forms of the Fourier transform. Thus a solid knowledge of Fourier transforms is mandatory for practitioners of DSP and any field requiring spectral analysis. That's because, for all but the most simplistic signals, the best way to understand the true nature of a real-world signal is by examining its spectrum. With this notion in mind, I believe no amount of time spent studying the Fourier transform is a waste of time.

Having said that, the question is: "What's the best way for you to learn Fourier transforms as they are related to DSP?" Years ago learning Fourier transforms was a daunting task. Your only choice was to plow through academic books written by university professors, and those books often relied on page after page of mathematical equations to enlighten their readers. (For example, I have a spectrum analysis book on my bookshelf that has a grand total of two figures in its first 165 pages!) Back then, unfortunately, many readers suffered a kind of Death by Algebra because engineering textbooks too rich in mathematics are hard to digest.

Times have changed. Today we have available to us signal processing books that are much more palatable and easier to read, while still being technically informative. With its careful and gentle plain-language text and beautiful illustrations, you're holding such a book in your hands right now! Let's be clear now. This book does contain all the mathematics you need to learn the Fourier transform, but the important distinction is that the equations are explained in terms that you will understand.

I first met Charan years ago in Silicon Valley when she was an working at Loral Space Systems. It was enjoyable meeting a lively smart lady who was also interested in signal processing. After a few years I realized she wasn't just interested, but rather she was captivated by DSP. It was no surprise when she began a part-time career instructing, writing about, and spreading the gospel of DSP through her website.

It’s with pleasure that I recommend this book to beginners in the field of DSP. It’s clear that its authors are dedicated specialists of whom the English poet Chaucer would say, "Gladly would they learn and gladly teach." I wish this book had been available decades ago when I first struggled to learn the Fourier transform.

Richard G. Lyons
Besser Associates
Mt. View, California
Author of: “Understanding Digital Signal Processing”

Preface

The Fourier transform shows up in so many different fields, that there isn't an another concept that spans across so many disciplines. Even the famous Heisenberg uncertainty principle is just a restating of the Fourier transform. There is no other mathematical concept that gives one as much bang for the buck as does the Fourier analysis. If you have a strong grasp of this concept, it can help you solve problems in interference, communications, probability theory, cryptography, acoustics, optics, control systems, and the list goes on. Like *e* and the *golden ratio*, it is one of those concepts, which defies boundaries and show up everywhere.

In college, not enough time is devoted to this subject. All electrical engineering students either take a Transform theory class or are exposed to it in bits and pieces in a DSP course. In DSP classes, the application of Fourier transform is thrown in with other material such as linear systems analysis. However, DSP only makes sense when you understand Fourier analysis *well*. All fundamental concepts of DSP are based on Fourier analysis. In mathematics and physics, this subject is covered only cursorily and left as an optional class in other sciences. In other fields such as biology, astronomy and geo-sciences, practitioners often lack an understanding of the basis of spectral estimation.

This book is intended to give both students and practicing engineers a deeper understanding of Fourier analysis, as a stand-alone topic from its emergence as Fourier series, its application to both analog and discrete signals and finally to spectral estimation using the Fourier transform. While this subject is at the pinnacle of human achievement, engineering education, due to the limited time available, often fails to impart its beauty and scope to students. It often takes years to appreciate this and happens mostly on the job when you have to apply these concepts to something "real."

Our main goal in this book is to help you fully understand what is happening when you compute an FFT of a discrete signal. It is intended to deepen your understanding of the transform theory on which much of numerical analysis is built. There is a whole long story behind this apparently easy-to-compute but hard-to-understand concept. We start the story with the Fourier series in its original trigonometric form as imagined by Baron Fourier, and then progress through all its developments with contributions from other notables along the way to the end point, the spectral estimation of random signals using the discrete Fourier transform. In the last two chapters of this book, we cover application of the Fourier analysis to the non-parametric spectral analysis of random signals.

The Fourier transform, a special case of the Laplace transform, is a fundamental tool for the analysis of stationary signals. In this book, we only cover Fourier analysis and although it leads to all sorts of other important transforms, we feel it is best not to confuse the issue by introducing other transforms. They all deserve a book of their own.

One of the hardest parts of writing this book was deciding where to start. To explain any concept in signal processing requires a good understanding of what comes before it. And of course, to build that understanding means you have to be comfortable with the fundamentals of that concept. To understand the ins and outs of modern-day digital spectral estimation, we start with the basics of sine waves. It is like starting a history lesson about WWII at the Middle ages! But this background is needed to really get into the topic.

Another difficult part is deciding how to describe these concepts. Do we go with the mantra that an equation speaks a thousand words or should a thousand words accompany the equation? Should we repeat ourselves? Take for example, the following sentence which is completely true about the discrete time Fourier transform (DTFT):

"The DTFT is a transformation that maps a discrete-time signal onto a continuous function of ω that spans from 0 to 2π. Alias spectrum appear outside this range."

It is completely true and yet, transfers no quick, shiny nugget of knowledge. Translated into English, it may read as:

"The DTFT is a mathematical machine that takes an input signal as a function of time and produces an output signal as a function of frequency. The input signal is discrete, meaning it only has amplitude values (called samples) at equally spaced time intervals. The output signal is a function of frequency ω and is continuous, and has a value, although it might be zero, at all frequencies. The output signal is called the *spectrum*. This output signal ranges from 0 to ∞ radians, but because the spectrum repeats outside of 0 to 2π, we only need consider this one section. The repetition occurs because if there is a sine wave of frequency ω_x that fits some given samples, then other sine waves of frequencies $(\omega_x + 2\pi k)$ will also fit the same samples."

That is a lot work to write out, and even this is not sufficient to convey full understanding. However, that is what is needed to develop an intuitive understanding. Unfortunately short dense sentences and pages of equations is what students get in the textbooks. Most of the focus in school is also on solving hard, tricky problems or remembering transform pairs. This part of learning is important, but without developing an intuitive understanding, the work is only half-done.

Fourier analysis and transform is a wonderfully simple but deceptively deep subject. We want this book to be a supplement for your education so that you will come to appreciate its beauty and depth. We explain what is going on behind the equations. We try to make these *complex* ideas come *alive* with the use of plots. A picture may be worth a thousand words, but we have decided to *also* give you the thousand words. Hence, you may say we examine the signals in a brand new domain, called the *Word Domain* in which we explore our signals in plain language. True understanding and an intuitive feel comes only when you can describe a mathematical concept in *words*.

Our goal is to help you master Fourier analysis from its beginning with the Fourier series, through CTFT and DTFT, all the way to the discrete Fourier transform (DFT), and spectral estimation, in a painless manner.

The first five chapters set the stage for the DFT. We start with the easy to understand trigonometric form of the Fourier series in Chapter 1, and then its more *complex* form in Chapter 2. From there, we go to discrete time signals in Chapter 3 which introduce new complexity to the topic. The development of the Fourier transform from the Fourier series, specifically the continuous time Fourier transform (CTFT) is discussed next. We combine the last two chapters to get to the discrete-time Fourier transform (DTFT) in Chapter 5. From here, it is a manageable leap to the DFT, our main quarry in Chapter 6. From there we spend the last three chapters on how the Fourier transform is used in "real life". Chapter 7 explains how windows can improve the spectrum by mitigating leakage, an inevitable consequence of DFT. Chapters 8 and 9 explain spectral estimation of stationary signals, specifically the non-parametric spectral estimation of random signals.

Altogether this book should help fill in the details and big concepts in Fourier analysis and, importantly, show you how to use them with comfort and ease. At the end of each chapter, we include some questions to test your conceptual understanding. These questions do not require a lot of calculation, and should be answered verbally as much as possible.

We want to thank some important people who offered comments, and encouragement. At Loral, we would like to thank Tom Watson and John Walker for helpful technical discussions and advice in many fields over the years. I would like to thank Rick Lyons, the author of my favorite DSP book, who read much of the material and helped me think through a few hard topics. His book, "Understanding Digital Signal Processing" is my model of how all engineering books should be written. I credit him for being my inspiration. Rena Tishman, my good friend and a lapsed engineer, read, edited, and corrected the chapters over several iterations of the book. I can't thank her enough for her life-long friendship and support. Ryan Bahneman, Stephen Paine and Paul Johnson read the chapters and provided valuable comments. I am also grateful to Dr. Jerry Gibson at UCSB and Dr. Yannis Tsividis at Columbia University for providing valuable comments. We are also grateful for the detailed(microscopic!) review of the book by Björn Neubauer of the Brunssweig University. Björn caught some subtle errors in indices in Chapter 2 as well as some big ones in the calculations. An unexpected source of help came from Patricio Parada. Patricio not only designed the book in Latex, being an engineer himself, also corrected errors. We are very grateful for his help. For editing, we would like to thank Grima Sharma, who added linguistic polish to the book. Nina Levin, the publisher and our taskmaster, made sure we got the book done in this century.

We also want to thank The Mathworks, Inc. for their support with MATLAB. The MATLAB code for many of the plots in this book as well as other related materials will be available at the book site.

This is the first edition of this book. Errors are bound to be there! May we ask your consideration in dropping us an email if you find any errors in this version. Thank you.

Charan Langton
Victor Levin
www.complextoreal.com
charldsp@gmail.com

Chapter 1

Trigonometric Representation of Continuous-Time Periodic Signals

Jean-Baptiste Joseph Fourier
1768 – 1830

Jean-Baptiste Joseph Fourier was a French mathematician and physicist. He was appointed to the École Normale Supérieure, and subsequently succeeded Joseph-Louis Lagrange at the École Polytechnique. He is best known for developing the Fourier series and its applications to problems of heat transfer and vibrations. The Fourier transform and Fourier's Law are named in his honor. Fourier also did very important work in the field of astronautics, as well as discovering the greenhouse effect for which he is not so well known. – From Wikipedia

What is Fourier Analysis

When sunlight hits rain-soaked air, an interesting phenomenon happens. Water drops take the ostensibly pure white light and split it into multiple colors. The mathematical description of this process, the subject of this book, was first tackled by Isac Newton approximately 400 years ago. However, even though Newton was able to show that white light is, in fact, composed of other colors, he was unable to make the jump to the idea that light can be described as waves. He called the component colors of white light "specter" or ghosts, from which we get the word spectrum. It took the development of trigonometric series and the recognition of the fact that light can be thought of as a composite wave before Fourier could apply these ideas to the problem of heat transfer. Although the concept of harmonic trigonometric series already existed by the time he worked on the heat transfer problem, Fourier's contribution is considered so important that the whole field of trigonometric waveform analysis and synthesis now bears his name, *Fourier analysis*.

Fourier developed the following partial differential equation called the *Diffusion* equation to describe heat transfer through solids and other media. Here, v is a function representing the measure of heat and K is the heat diffusion constant of the material.

$$\frac{\partial v}{\partial t} = K \frac{\partial^2 v}{\partial x^2} \tag{1.1}$$

Fourier observed that the most general solution to this equation was given as a linear summation of sinusoids, i.e., sine and cosine waves of the form:

$$v(k,x) = \sum_{k=0}^{\infty} \left(a_k \sin kx + b_k \cos kx \right) \tag{1.2}$$

This led Fourier to conclude that an arbitrary wave can be represented as a sum of an infinite number of weighted sinusoids, i.e., sine and cosine waves. This sinusoid summation concept is now known as the **Fourier series**. This book is all about this simple but important idea.

Fourier analysis is applicable to a wide variety of disciplines and not just signal processing, where it is now an essential tool. In addition, Fourier analysis is used in image processing, geothermal and seismic studies, stochastic biological processes, quantum mechanics, acoustics, and even finance.

The Fourier analysis of waves or signals is similar to the concept of compound analysis in chemistry. Instead of atoms coming together to form a myriad of compounds, in signal processing sinusoids can be thought of as doing the same thing. A particular set of these sinusoids is called the **basis set**. Just as a compound may consist of two units of one element and four units of another, an arbitrary wave can consist of two units of one base wave and four units of another. Hence, we can create a particular wave by putting together some basis waves from the set. This process is called **Synthesis**. Conversely, the process of decomposing an arbitrary wave into a set of basis waves is termed **Analysis**. These two complementary and linear processes fall under the name of **Fourier analysis** and its analog, the **Fourier transform**. In Fourier analysis, the basis set of waves is periodic sinusoids.

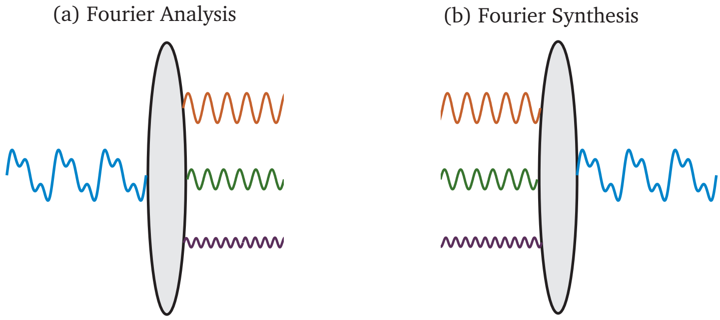

Figure 1.1: Fourier analysis is used to understand composite waves. (a) Analysis: breaking a given signal into sine and cosine components and (b) Synthesis: adding certain sine and cosines to create a desired signal.

Possibly, it was the solution of Eq. (1.1) that led Fourier to notice that summation of harmonic sine and cosine waves leads to some interesting looking periodic waves. From this, he posited that conversely it is also possible to create any periodic signal by the summation of a particular set of harmonic sinusoids. This may not seem like a big thing now but it was a revolutionary insight at the time.

Fourier's discovery was met with incredulity at first. Rightfully so, many of his contemporaries did not accept that his idea was truly general and applied to *all* signals. After some years of work by Fourier as well as other famous mathematicians of the age, his theorem was upheld, albeit not under all conditions and not for all types of signals. Subsequent development led to the Fourier transform, the extension of Fourier's original idea to *nonperiodic* signals. However, this computationally demanding concept languished for over 100 years, until the development of the Fast Fourier Transform (FFT), by J.W. Cooley and John Tukey in 1965. The FFT, an algorithmic technique, made the computation of Fourier series simpler and quicker and finally allowed Fourier analysis to be recognized and used widely. It is now the premier tool of analysis in many fields.

Frequency and Time Domain Views of a Signal

Consider the wave in Fig. 1.2. One would be hard pressed to guess its equation. Yet, it is just a sum of three waves as shown in Fig. 1.3(a), 1.3(b) and 1.3(c) of differing frequencies and amplitudes.

Therefore, although we know that the wave of Fig. 1.2 is created from the sum of three regular-looking sinusoids of Fig. 1.3, how could we figure this out, if we did not already know the answer?

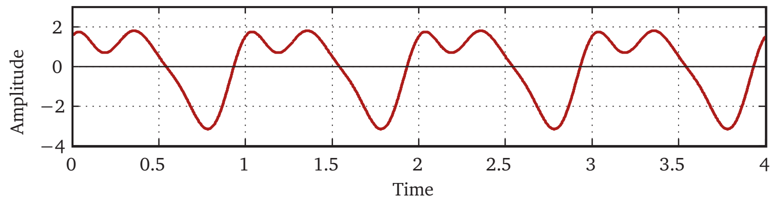

Figure 1.2: An arbitrary periodic wave for which we would like to know its equation.

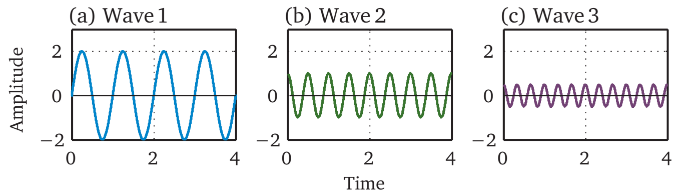

Figure 1.3: The components of the arbitrary wave of figure 1.2.

Spectrum of a signal

Let us look at Fig. 1.4 which shows a general signal in three dimensions. When we see a signal in time, we are looking at it in what is called, the **Time Domain**. What we are actually observing is a *composite* signal. It is a sum of components which we are unable to see. In this three-dimensional view, when we look at a signal from the *side view*, each component appears as a single vertical line at its own discrete frequency. This *side view* of the signal is called the **Frequency Domain**. Another name for this view is the **Signal Spectrum**.

The Fourier series and its further refinements, such as the Fourier transform, are a form of math that takes a function of one variable (such as time) to another (such as frequency.) Hence this type of process is classified as a **transform** and studied under **transform theory**. A mathematical operation, such as a typical function we see in calculus, does not change the variable after processing and is called an **operator**. In common language, we say that a transform allows us to see a signal from a different dimension or a point of view, as we see in Fig. 1.4. In signal processing, this alternate view of a signal, made possible by the *Fourier transformation* is often called a spectrum.

The spectrum is a way to quantify the component frequencies. By *quantify* we mean identifying the amplitude of each of the components. The spectrum of the signal shown in Fig. 1.2 is composed of just three frequencies and can be drawn as illustrated in Fig. 1.5(a). This is called a **one-sided amplitude spectrum**. The x-axis in Fig. 1.5(a) represents the component frequencies of the signal, whereas the y-axis is the amplitude of those frequencies. Fig. 1.5(a) shows the amplitude spectrum with the amplitude of each frequency, whereas Fig. 1.5(b) shows an alternate form of the spectrum called the power spectrum. The **power spectrum**, a more typical representation of the spectrum shows instead of amplitude, the quantity amplitude-squared (which is equal to the instantaneous power), in units of Decibels (dB).

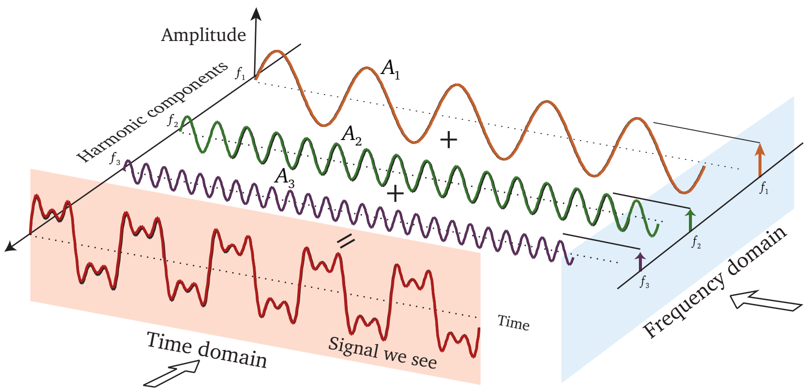

Figure 1.4: A time domain wave seen from the frequency domain provides useful information. The view from the frequency domain is called the spectrum of the wave.

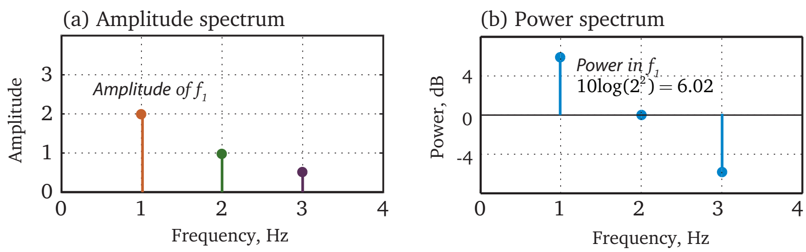

Figure 1.5: The frequency view of the arbitrary wave of Fig. 1.2. (a) The amplitude; and (b) the power in dBs.

We note that Fig. 1.5(a) is in fact the side view from Fig. 1.4, showing the amplitude of each of the components, f_1, f_2 and f_3. The graphical representation, such as the one shown in Fig. 1.5 is a unique *signature* of the signal at a *particular time*. This spectrum is a distinctive and variable quality of a signal. It is *not* a static thing but changes as the signal and its components change. Signals have distinctive spectrum and we can tell a lot about a signal by looking only at its spectrum. The signals for which Fourier analysis is considered valid, must have a non-changing spectrum. This property is generally called **stationarity**.

Fundamental waves and their harmonics

The basic building blocks of Fourier analysis come from a set of **harmonic sinusoids**, called the **basis set**. The basis set is our tinker-toy from which we can construct a variety of waves. The set contains an infinite number of sinusoids of differing frequencies related in a special way known as *harmonic*.

The top row of Fig. 1.6 shows a sinusoid of an arbitrary frequency, f_0. Let us call this arbitrary frequency the *fundamental frequency*. We specify some more waves based on this wave called the **harmonics**. Each harmonic frequency is an integer multiple of the frequency of the fundamental. Fig. 1.6 shows that the second wave has half the wavelength and twice the frequency of the first one and so on. Each kth wave has a wavelength of T_0/k and a frequency of kf_0 with k being consecutive integers greater than or equal to 1. All such waves for $k > 1$ are called harmonics of the fundamental. The frequency of the fundamental is of course *arbitrary*, it can be any number whatsoever, but its harmonics are strictly integer multiples of the fundamental frequency, such as: f_0, $2f_0$, $3f_0$, ... kf_0.

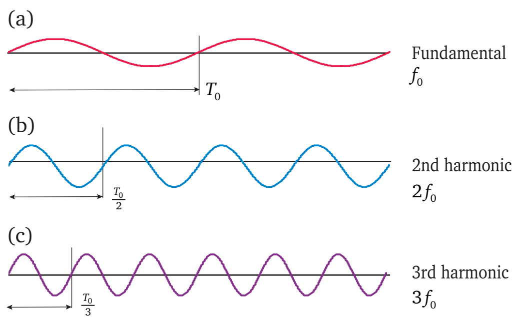

Figure 1.6: An arbitrary fundamental wave and its harmonics.

Let us start with the expression of a set of complex sinusoids; a pair of cosine and a sine of a certain frequency:

$$\begin{aligned} c(t) &= \cos(2\pi f_0 t) \\ s(t) &= \sin(2\pi f_0 t) \end{aligned} \tag{1.3}$$

Here, f_0 is an arbitrary frequency (measured in cycles/second or Hz). We will call this frequency the **fundamental**. The *basic period* T, of this sinusoid is the inverse of the frequency. Note that if a sinusoid is periodic with period T, then it is also periodic with period $2T$, $3T$, etc. for integer multiples of the basic period. From this we obtain the definition of harmonic waves.

A set of waves is harmonic if its frequency is an integer multiple of the fundamental wave's frequency. We can write this as a set.

$$\begin{aligned} c(t) &= \cos(2\pi f_k t) \\ s(t) &= \sin(2\pi f_k t) \end{aligned} \tag{1.4}$$

We have introduced an index, k in Eq. (1.4) such that each harmonic frequency is equal to k times the fundamental frequency, $f_k = kf_0$, with k any arbitrary integer, $0, 1, 2, \ldots, \infty$.

In Fourier series formulation, the index k spans all positive integers to infinity, including zero. Note that the fundamental wave or the first harmonic is often defined as the one for $k = 1$. Hence, the frequency of the first harmonic and the fundamental frequency are the same. The wave obtained for $k = 0$ is of course nothing but a flat line. The wave for $k = 2$ is called the second harmonic and so on for higher values of the index k.

Let us rewrite the definition of the harmonics allowing the **phase** and the **amplitude** to vary. We give each harmonic a unique amplitude and phase and rewrite the harmonic signals as:

$$\begin{aligned} c_k(t) &= a_k \cos(2\pi k f_0 t + \phi_k) \\ s_k(t) &= b_k \sin(2\pi k f_0 t + \phi_k) \end{aligned} \tag{1.5}$$

The wave $c_k(t)$ is a cosine wave of kth harmonic frequency or kf_0, its amplitude being a_k and the phase in radians being ϕ_k. The signal $s_k(t)$ is a sine wave with similarly unique amplitude and phase for the same harmonic frequency, kf_0. Now although the frequencies are still related by multiples of integer k, we are allowing the amplitude and the phase of each harmonic to be different. Such waves are still considered harmonic. The amplitude coefficient a_k for a cosine and b_k for a sine are now arbitrary.

What is **amplitude**? Amplitude can be thought of as the value of a wave's height above its mean at a particular time. Amplitude can be a positive or a negative quantity, depending on where it is being measured. It is the instantaneous value of the wave's height. There is, however, an another definition of the amplitude. If a wave is given by the expression: $f(t) = a\cos(\omega t)$, then the instantaneous value, $f(t)$ is scaled by the coefficient, a. The peak value of the instantaneous amplitude for this wave is never greater than a, hence this coefficient is called the **peak amplitude** of the wave. In common language, the word *peak* is often left off. Hence, there is confusion whether we are talking about the instantaneous value of the amplitude, $f(t)$ or the peak value of the amplitude, a. We need to be aware of which *amplitude* we are talking about. Most often we are interested in the *coefficient* of the wave which is its peak amplitude. In Eq. (1.5), the terms a_k and b_k are the amplitudes (peak) of the waves.

What is **phase**? The argument of a sinusoid $\sin(\theta)$, is in fact an angle, or the term θ. However, in signal processing, we often need to represent a sinusoid as a function of time. We do this by writing the sinusoid as $\sin(\omega t)$, where ω is the radial frequency defined in terms of angles per time. Both forms of the sinusoid argument, the θ and its equivalent form ωt, are called the *instantaneous phase* of the sinusoid. We can shift a sinusoid, or in fact any wave in time, by adding an independent term, ϕ, to the instantaneous phase, θ, writing the sinusoid as $\sin(\theta + \phi)$. This second term, ϕ, is often assumed to be fixed with time (for linear systems) and is commonly called *the phase*. This can be confusing, because we have two quantities here, both called phase. The total phase at any time is made up of these two parts, one a fixed quantity and the other changing with time. However, generally when we say phase, we are referring not to the instantaneous quantity ωt, but to the fixed quantity, ϕ. This term is more properly called the *phase shift* but the qualifier word *shift* is often left off. In Eq. (1.5), the term ϕ_k is the phase (shift) for the kth wave.

Figure 1.7(a) shows that the amplitude of the sinusoid is 1.0, even though the full excursion is 2.0. Figure 1.7(b) shows that a phase shift of $\phi = \pi/2$ has turned the sine wave into a cosine wave. This phase shift has no effect on the amplitude or the frequency of the wave, hence phase, amplitude and frequency are independent terms.

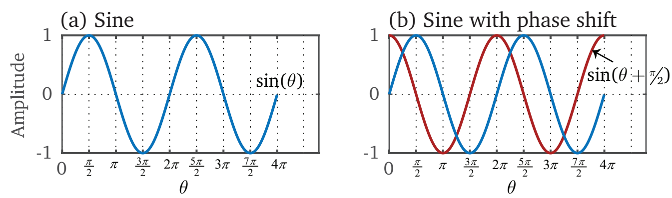

Figure 1.7: The phase of a sinusoid consists of its instantaneous and changing phase, ($\theta = \omega t$), plus a constant phase term, ϕ, called the phase shift as shown in (b).

The sine and cosine waves differ in phase by $\pi/2$ radians, or 90°. When two signals differ in phase by −90° or +90°, they are said to be in *quadrature*, hence sine and cosine waves are in quadrature. For linear Fourier analysis, it is assumed that phase (the ϕ term) and frequency are not a function of time, i.e., both the phase and the frequency of a signal do not change over time. This is one of the fundamental assumptions of Fourier analysis. All signals subjected to Fourier analysis are assumed to have fixed and unchanging frequencies as well as the phase term, ϕ. This assumption implies that the signal is *stationary*, a concept discussed in Chapter 8.

In Fig. 1.8, two harmonics of a sinusoid of frequency 1 Hz are seen, with sines on the left side and cosines on the right. Both the sines and cosines have a peak amplitude of 1. Depending on your field of interest, the units of amplitude can be pretty much anything. In this book, the signal amplitude units will be in volts and phase in radians.

The cosine always reaches its peak amplitude at time $t = 0$, and has a phase of $\pi/2$ at that time. The sine, no matter what the frequency, always has an amplitude of zero at time $t = 0$, which is equivalent to a phase of 0 radians. Hence, no matter how many sine waves of different frequencies and amplitudes are added together, they will never achieve any amplitude other than zero at time = 0, and can represent only those waves that are zero-valued at time $t = 0$. Similarly, the addition of various cosines will not achieve an amplitude of 0 at time $t = 0$.

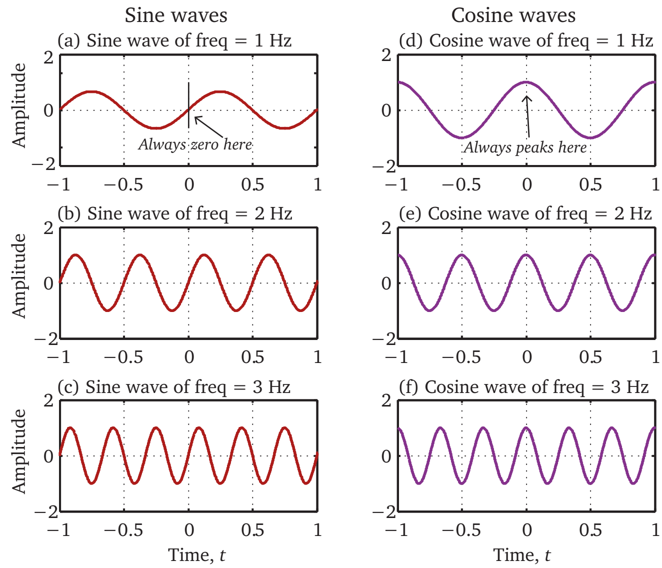

Figure 1.8: The fundamental of $f_0 = 1$ Hz and its two cosine and sine harmonics. All sines start at time $t = 0$, at amplitude $= 0$ and all cosines at the peak amplitude.

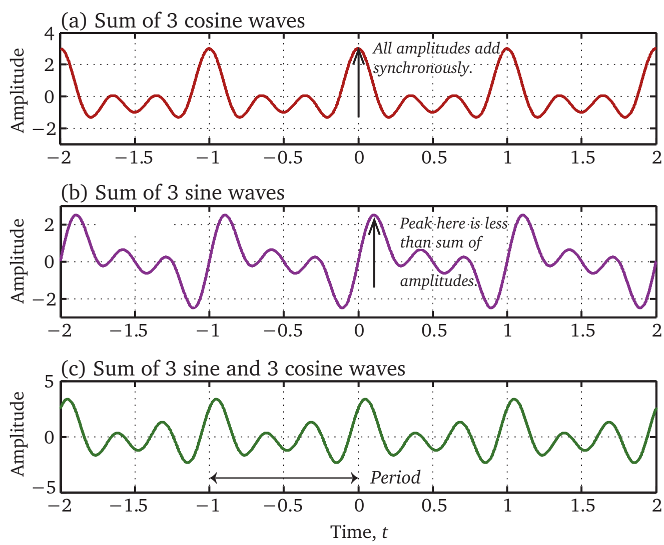

Figure 1.9: Sums of constant amplitude harmonics. (a) Sum of three cosines starts at an amplitude of 3; (b) sum of sines starts at 0; and (c) summation of both sines and cosines also starts at 3 since sines add nothing at $t = 0$.

Harmonics as Basis Functions

In Fig. 1.9 three cosine and three sine waves with amplitude of 1 are added together, respectively. After the addition of these three waves, it is seen that the cosines in

Fig. 1.9(a) add such that the peak is equal to 3. This happens because the cosine is an *even* wave. The cosines add *constructively* at time $t = 0$ and at other times that are integer-periods away. The sum of the three sine waves added together in Fig. 1.9(b) however, looks strange. This asymmetry is a consequence of the sine wave being an *odd* wave. In Fig. 1.9(c), the three sine and three cosine waves are all added together. The behavior of this summation defies an easy explanation.

To examine this effect further, we add together an even larger number, 20 cosine and 20 sine harmonics; each of the same amplitude and 0 phase. Fig. 1.10(a) shows clearly that the cosine sum is symmetric about the time $t = 0$ point. Similarly, Fig. 1.10(b), shows that the sine sum is non-symmetric. Now, we add both of these waves per the following equation:

$$s(t) = \sum_{k=0}^{\infty} \left| \cos(k\omega_0 t) + \sin(k\omega_0 t) \right|$$

The summation of harmonic sine and cosine waves of equal amplitudes gives a function that is approaching an impulse train, i.e., a signal consisting of very narrow pulses. This result in Fig. 1.10(c) gives a clue to an important property of sinusoids that will be used in subsequent chapters. This property says that a summation of an infinite number of harmonic sinusoids (sine and cosine) results in an impulse train.

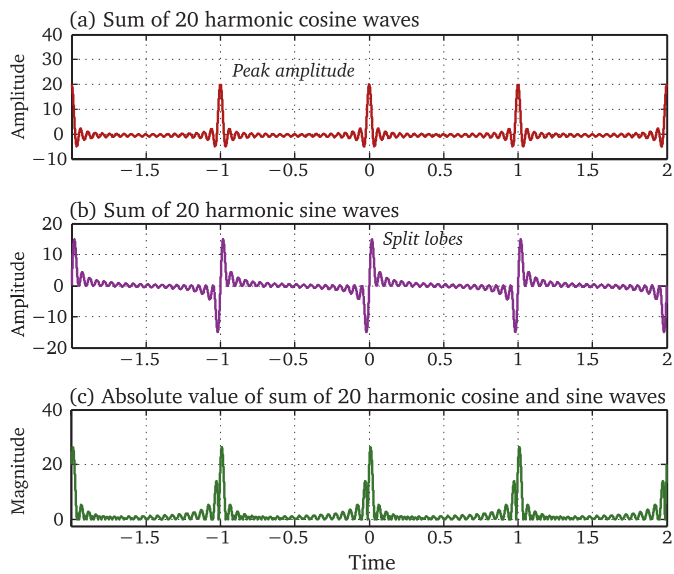

Figure 1.10: Sums of large numbers of both sine and cosine harmonics approach an impulse train. Note that the sum of cosines creates an even blip (a), whereas for sines it is an odd blip (b). In (c), we see the absolute value of the sum (a) and (b).

What if we allow the amplitudes and phases of these sinusoid to vary? An example of a wave created using this idea is shown in Fig. 1.11, where each sine and cosine wave has a different amplitude and a different phase. Note that the frequency of the composite wave is equal to the frequency of the fundamental, which is 1 Hz.

$$s(t) = 0.1\cos(2\pi t - 0.5) + 0.3\cos(4\pi t) - 0.4\cos(6\pi t - 0.1) - 0.5\sin(2\pi t + 0.1) - 0.8\sin(4\pi t - 0.3) + 0.67\sin(6\pi t + 0.19) \quad (1.6)$$

The most important thing to note is that by adding any number of harmonics, and allowing the amplitude and phase of each to vary, we can create or mimic many other waves. Figure 1.11 shows an example of just one such "interesting" looking wave created by using only three different sinusoids of distinctly different amplitudes and phases. This is the main idea behind Fourier synthesis and analysis.

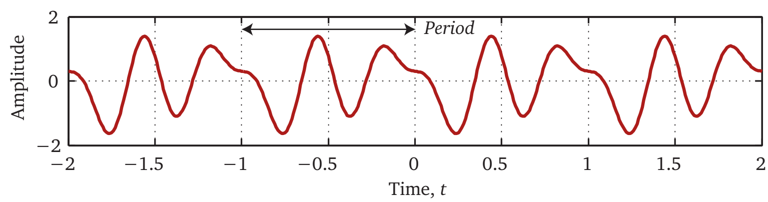

Figure 1.11: A wave comprising arbitrary amplitude harmonics of Eq. (1.6) *begins to look like a real signal.*

Evenness and oddness of sinusoids

Sinusoids have many interesting properties, a few of which are very useful in Fourier analysis. One important property of harmonic sine and cosine waves is asymmetry or *oddness* of the wave. All *sine waves* are considered odd functions because they obey the following definition of an odd function.

$$f(x) = -f(-x) \quad (1.7)$$

If you look at a sine wave, you see that it starts with an amplitude of 0 at time $t = 0$. If we were to flip it about the y-axis, the images would not overlap. However, if one of the sides was first flipped about the x-axis, then they do overlap. This describes the *oddness* of signals. It requires two flips for values to coincide, as we can see from the two negative signs in Eq. (1.7).

The cosine waves on the other hand are called *even* functions by a similar definition. The two sides of a cosine wave, if flipped about the y-axis, would overlap, hence there is only one negative in the equation below for even symmetry.

$$f(x) = f(-x) \quad (1.8)$$

By the superposition principle, if multiple odd waves are summed together, the resulting wave will remain *odd*. In contrast, if multiple *even* waves are summed, then the resulting wave will remain even, and a mixture will have no symmetry. This becomes important when synthesizing, which is the process of putting some waves together to make a desired wave. If a wave to be synthesized is purely an odd, or an even function, then it will only contain sines, or cosines, respectively, depending on its symmetry.

Making waves

Using the idea of harmonic summation, we can create a variety of waveforms. All we need to do is to change the amplitudes and the phases of the harmonics as we see fit. Certain combinations of these parameters lead to great-looking and useful waves that are periodic with the frequency of the fundamental.

Square waves

Now we examine the construction of a square wave. A square wave is not actually square in any particular sense. It is a wave where each period looks somewhat like two rectangles of opposite signs, as seen in Fig. 1.12(d). Yes, it has wiggles in it, and, it is not actually square. The square waves are very useful in signal processing and are used for data transmission. It is amazing that we can create them by just adding a bunch of sinusoids. The more sinusoids we add to the summation, the better the wave looks, with the wiggles getting smaller.

In Fig. 1.12, we have created a square wave by adding together only a few harmonic sinusoids. As the square wave shown in Fig. 1.12(a) is an odd function, we know that only sine waves are needed in its construction, a process also called synthesis. In Fig. 1.12(b), we have added to the fundamental of frequency ω_0, only one sine wave of frequency $3\omega_0$ and amplitude of $\frac{1}{3}$. In Fig. 1.12(c), we add one more sine wave of frequency $5\omega_0$ and of amplitude $\frac{1}{5}$, and in the last case another sine wave of frequency $7\omega_0$ and of amplitude $\frac{1}{7}$ is added. This last result looks good.

Figure 1.13(a) shows the same type of wave but shifted so that it is an even wave, per Eq. (1.8). This even wave can be created using just the cosines. For both cases, we start with a sinusoid of the fundamental frequency. This seems like a good start. But why? Because a periodic wave created from the addition of harmonics will always be periodic with the period of the fundamental, the lowest frequency of the signal. We then add more harmonic sine waves for the odd version of the square wave in Fig. 1.12(a) and cosine waves for the even version in Fig. 1.13(a) and watch the evolution of the square wave. With only three terms in the addition, the results are pretty decent looking square waves.

Each addition of a sine wave (or a cosine) with a specific frequency and amplitude makes the synthesized wave appear closer to a square wave. We are, in fact, cooking up

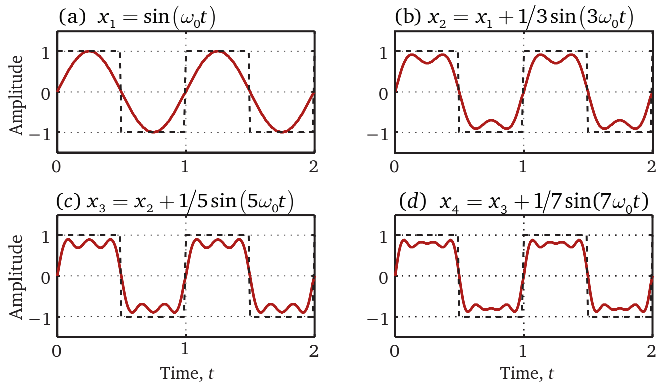

Figure 1.12: Synthesizing an odd square wave by adding odd harmonics of differing amplitudes. (a) Start with fundamental of 1 Hz, (b) add harmonic of 3 Hz, (c) add harmonic of 5 Hz, (d) add harmonic of 7 Hz.

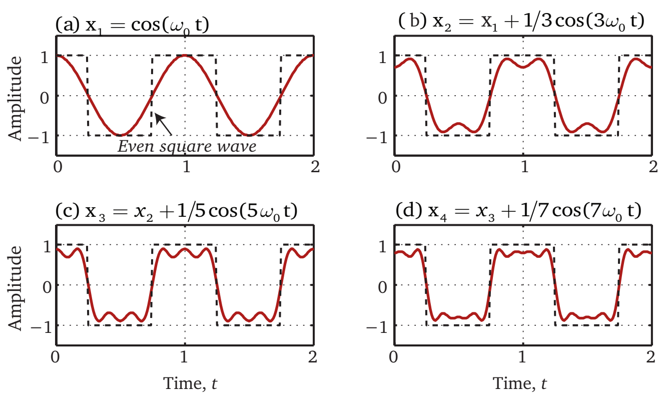

Figure 1.13: Synthesizing an even square wave by adding even harmonics of differing amplitudes.

interesting recipes for making all kinds of waves using specific "quantities" of sinusoids. The quantities we vary are the amplitude and the phase of each harmonic. Collectively, the *amplitude* and the *phase* of a particular harmonic is called its **coefficient**. So to create a particular wave, we are controlling or changing the coefficients of the harmonics.

Here are the *recipes* for the two types of square waves, the odd and the even. The ingredient list is limited, as we used only three terms beyond the fundamental, which is underlined.

$$\begin{aligned} x_1 &= \underline{\sin(\omega_0 t)} + \frac{1}{3}\sin(3\omega_0 t) + \frac{1}{5}\sin(5\omega_0 t) + \frac{1}{7}\sin(7\omega_0 t) \\ x_2 &= \underline{\cos(\omega_0 t)} - \frac{1}{3}\cos(3\omega_0 t) + \frac{1}{5}\cos(5\omega_0 t) - \frac{1}{7}\cos(7\omega_0 t) \end{aligned} \tag{1.9}$$

Of note, in both of these "recipes", only odd harmonics (i.e. the harmonic index k is odd.) are used. This is true for both the odd and even versions of the square wave. The reason why only odd harmonics are used is that sinusoids of even harmonics cancel the odd harmonics as we see in Fig. 1.14 and hence mixing of odd and even harmonics does not allow us to create useful waves. We see in Fig. 1.14 that the two even harmonics of frequencies 2 and 4 Hz are poles apart from the odd harmonics of frequencies 1 and 3 Hz (at t = 0.5 sec.). We find that summation of consecutive harmonics begins to approach an impulse-like signal, as we see in Fig. 1.10 due to this destructive behavior.

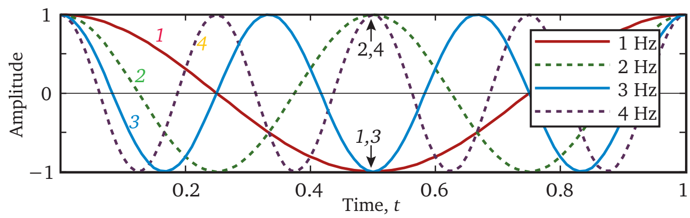

Figure 1.14: Even order harmonics are destructive and hence are not used in combination with odds to create most communications signals such as square waves. Note here that frequencies 1 and 3 Hz are antipodal to frequencies of 2 and 4 Hz.

The square wave is probably the single most important wave in signal processing and we will keep coming back to it in each chapter. The summation with k as the odd integer can go from $-\infty$ to $+\infty$ to form a really great looking square wave. The more terms we add, the closer we get to what we are trying to achieve. But in fact for square waves, we are never able to create a perfect square wave, as Gibbs phenomena takes hold near the corners. The corners never do become the true right angles as we would wish. This tells us that this form of harmonic representation, despite our best efforts, may not result in a perfect reconstruction for every signal, for example square waves.

Gibbs phenomenon

The square waves created in Fig. 1.12 and Fig. 1.13 have a fair amount of **ripple** (formal name for the wiggle) in the center part of the square wave. We can increase the number of terms to see if it goes away. Figure 1.15 shows that even as K, the number of terms in the square wave summation is increased, the overshoot at the corners never goes away. There remains, approximately an 18 % overshoot (in height) even with a very large number of terms in the summation, although its width does decrease. This behavior, called the **Gibbs phenomenon**, discovered by Henry Wilbraham (1848) and by J. Willard Gibbs (1899), is a clear demonstration that Fourier representation can not accurately reconstruct *all* waves, particularly those that are piece-wise continuous but with discontinuities. Hence, waveforms that have discontinuities in amplitude, always cause this effect and hence are avoided in signal processing. Instead of sharp-edged square waves, we use shaped waves with gentle corners and curves, ones that can be represented with better fidelity using just a finite number of sinusoids.

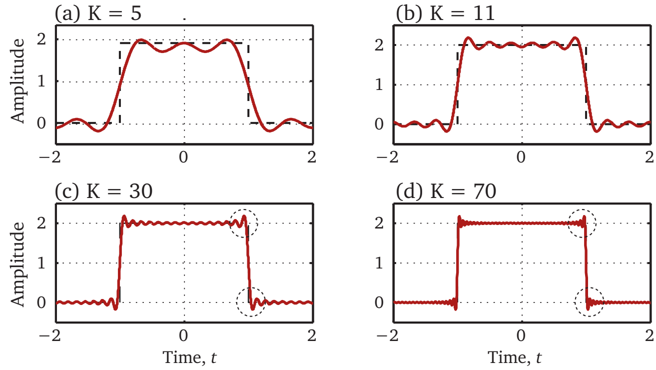

Figure 1.15: Gibbs phenomenon does not allow for a perfect Fourier representation of a square wave.

Creating a sawtooth wave

Let us look at one more special signal, a sawtooth wave. The sawtooth wave is an odd function, hence, composed only of sine waves. Its equation is given as

$$\text{sawtooth}(t) = \frac{2}{\pi}\sum_{k=1}^{\infty}(-1)^{k+1}\frac{\sin(2\pi k f_0 t)}{k} \tag{1.10}$$

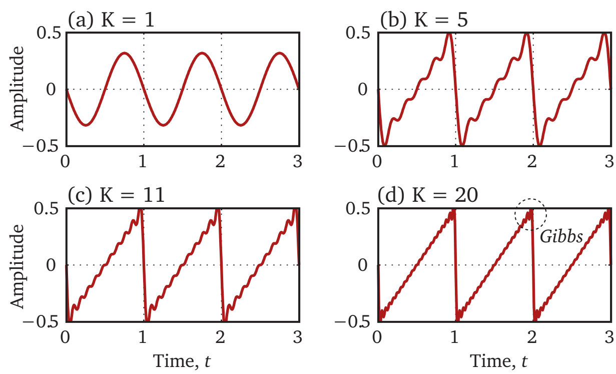

Figure 1.16: Evolution of a sawtooth wave, (a) fundamental sinusoid of desired frequency, (b) addition of five harmonics, (c) addition of eleven harmonics, and (d) with addition of 20 harmonics, a pretty decent looking saw-tooth wave presents itself.

Note that in Eq. (1.10) the coefficient, $\frac{(-1)^{k+1}}{k}$ is inversely proportional to the harmonic index, k and hence the contribution of higher frequencies is decreasing. This is also true for the square waves. We write this equation allowing the index k to go to ∞. In real life, we can often get by with just a few terms. Notice that Gibbs phenomena is present at the corners for this representation also. Hence, smooth waves lend to best Fourier synthesis.

Generalizing the Fourier Series Equation

A Fourier series is a general equation consisting of the summation of weighted harmonics, whereby manipulating the weighting (hence coefficients) we can represent any periodic wave. We called it a recipe but its mathematical name is **Fourier representation**.

$$f(t) = \sum_{k=1}^{K} a_k \cos(2\pi k f_0 t) + \sum_{k=1}^{K} b_k \sin(2\pi k f_0 t) \tag{1.11}$$

The coefficient a_k represents the coefficient of the kth cosine wave, and b_k of the kth sine wave. What is K? This is the largest harmonic index we use in any particular summation. If we use only 10 terms (or harmonics), then $K = 10$. We can control this parameter depending on the accuracy desired. For a general representation, we set K to ∞.

The sum of sines and cosines is always symmetrical about the x-axis so there is no possibility of representing a wave with a DC offset using the form in Eq. (1.11). (*The term DC comes from direct current but it is used in signal processing to mean a constant as well as the mean power of a signal.*) To create a wave of nonzero mean, a new term must be added to Eq. 1.11. The constant, a_0 is included in Eq. (1.12) so we can create waves that can move up (or down) from the x-axis.

$$\boxed{f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(2\pi f_k t) + \sum_{k=1}^{\infty} b_k \sin(2\pi f_k t)} \tag{1.12}$$

Eq. (1.12) is called the **Fourier series equation**. The coefficients a_0, a_k, b_k are called the **Fourier Series Coefficients (FSC)**. The process of Fourier analysis consists of computing these three types of coefficients, given an arbitrary periodic wave, $f(t)$.

Multiple ways of writing the Fourier series equation

There are several different forms of the Fourier series equation in literature, and that can make understanding this equation harder.

The following representation uses the radial frequency, $\omega_k = 2\pi f_k$ to make the equation simpler to type. We can write this form of Fourier series as:

$$\boxed{f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(\omega_k t) + \sum_{k=1}^{\infty} b_k \sin(\omega_k t)} \tag{1.13}$$

Now, we define T_0 as the period of the fundamental frequency as: $T_0 = \frac{1}{f_0}$. Then the period of the kth harmonic becomes T_0/k and its frequency, $f_k = k/T_0$. We can alternately write the Fourier series equation by adopting this form of the frequency in Eq. (1.14).

$$\boxed{f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi k}{T_0} t\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{2\pi k}{T_0} t\right)} \tag{1.14}$$

We can shorten the Fourier series equation by starting at zero frequency, hence index k starts at 0 instead of 1. Now the DC term disappears as it is included as the zero frequency coefficient obtained by setting the index $k = 0$. Here is this form with the DC term gone:

$$\boxed{f(t) = \sum_{k=0}^{\infty} a_k \cos(2\pi f_k t) + b_k \sin(2\pi f_k t)} \tag{1.15}$$

This can be simplified further by getting rid of the sine term altogether. Sines and cosines are really the same thing, one is just the shifted version of the other. The representation in Eq. (1.13) can be written solely with cosines with a shift. This way sine becomes a cosine with a $\pi/2$ shift and we get this form of the equation.

$$\boxed{f(t) = a_0 + \sum_{k=1}^{\infty} c_k \cos(\omega_k t + \phi_k)} \tag{1.16}$$

In this form, each harmonic, whether a sine or a cosine, can be thought of as a cosine of some phase. Hence each harmonic is represented by two cosines, one with a zero phase shift and the other with a shift of $\pi/2$. Now the whole expression uses only cosine wave and the index, k spans not from 0 to 1, but from $-\infty$ to $+\infty$.

$$f(t) = c_1 \cos(\omega_1 t \pm \phi_1) + c_2 \cos(\omega_2 t \pm \phi_2) + c_3 \cos(\omega_3 t \pm \phi_3) + \ldots$$

Fourier series in complex exponential form

In its *most* important representation, the *complex representation*, the Fourier series is written as:

$$\boxed{f(t) = \sum_{k=-\infty}^{\infty} C_k \underline{e^{j2\pi \frac{k}{T_0} t}}} \tag{1.17}$$

Here, we introduce a new term, the **complex exponential**, as underlined in Eq. (1.17). The complex exponential (CE) represents both a sine and a cosine in one concise form. In the next chapter, we will discuss this function in detail. The expanded form of the Fourier series in terms of the complex exponential looks like this:

$$f(t) = C_0 + C_1 e^{j2\pi \frac{1}{T_0} t} + C_2 e^{j2\pi \frac{2}{T_0} t} + C_3 e^{j2\pi \frac{3}{T_0} t} + \ldots$$

This form shows that we can create a periodic function by summing together complex exponentials. Although the complex form of the Fourier series is scary looking, it is the most commonly used form. In the next chapter, we will look at how it is derived and why we use it in Fourier analysis. All these different representations of the Fourier series are *identical* and mean exactly the same thing. They are all different ways in which you see the Fourier series equation written in books.

The Fourier Analysis

The process of adding together a bunch of sinusoids to create useful waves is called the **Synthesis** process. Synthesis of waveforms, is of course, interesting but what is really useful is the reverse process, that is, to take an arbitrary periodic signal and figure out its components. It's like trying to figure out the ingredients of a particular dish. By ingredients, we mean frequencies in the signal that contain significant power or amplitudes. This is the main use of Fourier analysis. It is called, not surprisingly, the **Analysis** part. What this involves is to make a guess of the fundamental frequency, f_0, and then computing the amplitudes (coefficients) of a certain number of harmonics. There is no guarantee that the fundamental chosen will result in finding all the main signal components exactly. Nevertheless, in most cases, we have a pretty good idea of signal components a-priori. So the process works well enough.

The usefulness of the process can be seen in the equation of the sawtooth wave in Eq. (1.10). The Fourier series allows us to create an estimate of the wave using a few or many terms. Hence, Fourier series represents an *estimate* of the true representation, the accuracy of which depends on the number of terms used.

The Fourier analysis process consists of finding the **series coefficients**. When we talk about Fourier series coefficients (FSC), we are talking about the amplitudes of the sine and cosine harmonics, and nothing else. Once we decide on a fundamental frequency -a starting point for the analysis- we already know all the harmonic frequencies since they are integer multiples of the fundamental frequency. All we have to do now is to compute the coefficients. The following are the three types of coefficients we need to compute.

1. The DC offset or the coefficient of the 0th frequency, $k = 0$.
2. Coefficients of the cosine $a_k \cos(2\pi k f_0 t)$ with $k = 1, 2, 3, \ldots, \infty$.
3. Coefficients of the sine $b_k \sin(2\pi k f_0 t)$ with $k = 1, 2, 3, \ldots, \infty$.

We will discuss each of these three types of coefficients separately and see how to compute them.

Computing a_0, the DC coefficient

We are given an arbitrary periodic signal, $f(t)$, of period T. The Fourier series says that the signal $f(t)$ is equivalent to a summation of K sinusoidal harmonics. Our task is to find the coefficients of each of these harmonics starting with $k = 0$ to $k = K$. This task is called Fourier analysis.

$$f(t) = \boxed{a_0} + \sum_{k=1}^{\infty} a_k \cos(2\pi f_k t) + b_k \sin(2\pi f_k t) \tag{1.18}$$

The constant a_0 in the Fourier series equation represents the DC offset. If our target wave has a nonzero DC component (if its average amplitude value is not zero), then we know that $a_0 \neq 0$. But before we compute it, let's take a look at a useful property of sine and cosine waves. Both sine and cosine waves are symmetrical about the x-axis.

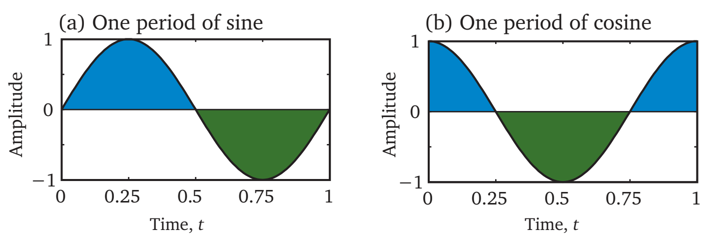

Figure 1.17: The area under both sine and cosine over one period is zero, no matter what their frequency.

When you integrate a sine or a cosine wave over one period, you always get zero. The area above the x-axis cancels out the area below it. This is always true over one period as we can see in Fig. 1.17. The same is also true for the sum of sine and cosine waves. Any wave made by summing sine and cosine waves also has zero area over one period. If we were to integrate the given signal $f(t)$ over one period, as in Eq. (1.18), the area obtained will have to come from the coefficient a_0 only. None of the sinusoids makes any contribution to the integral and they will all fall out. Hence, the calculation of the DC term becomes easy simply because the integral of the harmonics is zero.

$$\int_0^{T_0} f(t)\, dt = \int_0^{T_0} a_0\, dt + \int_0^{T_0} \cancelto{0}{\left(\sum_{k=1}^{\infty} a_k \cos(2\pi f_k t) + b_k \sin(2\pi f_k t) \right)} dt$$

We compute a_0 by computing the integral of the wave over one period. The area under one period of this wave is equal to

$$\int_0^{T_0} f(t)\, dt = \int_0^{T_0} a_0\, dt$$

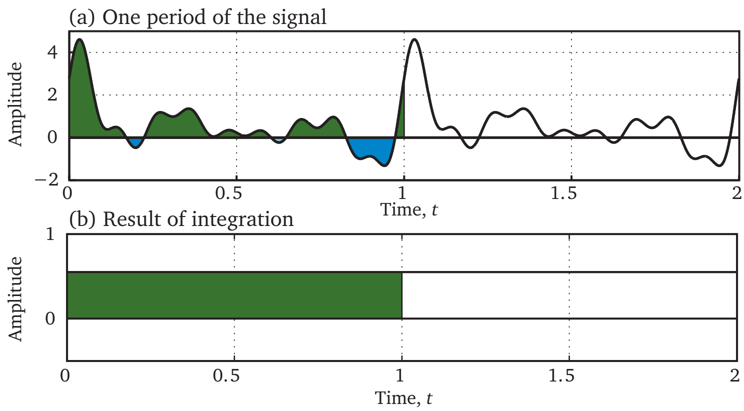

Figure 1.18: (a) The periodic signal before integration. (b) After integration of exactly one period, only the DC component is left.

Integrating this simple equation, we get,

$$\int_0^{T_0} f(t)\,dt = a_0 T_0$$

We can now write a very easy equation for the first coefficient, a_0

$$\boxed{a_0 = \frac{1}{T_0}\int_0^{T_0} f(t)\,dt} \tag{1.19}$$

Summary: To compute the DC coefficient, integrate the target signal $f(t)$ over one period. The result of the integration, normalized by the period, is equal to the 0th coefficient. Hence the area under the given signal comes only from the 0th coefficient.

Computing b_k, the coefficients of sine harmonics

We are assuming that our target signal is composed of sines and cosine harmonics. Now we multiply the target signal by just one harmonic of kth frequency. We will get various different types of combinations. We will get sine harmonics multiplied by sine and cosine harmonics of both the same and different frequencies, or in other words a lot of terms to solve.

Here is a target signal we wish to represent with just two harmonics (1 and 3).

$$f(t) = a_1\cos(\omega_0 t) + b_1\sin(\omega_0 t) + a_3\cos(3\omega_0 t) + b_3\sin(3\omega_0)$$

We want to compute the coefficient of the first sine harmonic or the term b_1. To do that we multiply the signal $f(t)$ by this same harmonic. We get in the product, various combinations of sines and cosines:

$$\int_t^{t+T_0} \cos(\omega_0 t) \times \sin(\omega_0 t)\, dt$$
$$\int_t^{t+T_0} \sin(\omega_0 t) \times \sin(\omega_0 t)\, dt$$
$$\int_t^{t+T_0} \cos(3\omega_0 t) \times \sin(\omega_0 t)\, dt$$
$$\int_t^{t+T_0} \sin(3\omega_0 t) \times \sin(\omega_0 t)\, dt$$

Now compute the coefficient, a_1, we multiply the representation by $\cos(\omega_0 t)$. We get these various types of products:

$$\int_t^{t+T_0} \cos(\omega_0 t) \times \cos(\omega_0 t)\, dt$$
$$\int_t^{t+T_0} \sin(\omega_0 t) \times \cos(\omega_0 t)\, dt$$
$$\int_t^{t+T_0} \cos(3\omega_0 t) \times \cos(\omega_0 t)\, dt$$
$$\int_t^{t+T_0} \sin(3\omega_0 t) \times \cos(\omega_0 t)\, dt$$

From these two examples, we see that in total there exist, just six different types of products, no matter how many harmonics we want to use to represent the target signal. These types, in terms of sine and cosine products are:

1. Cosine times a sine of the same frequency, $\cos(k\omega_0 t) \times \sin(k\omega_0 t)$
2. Sine times a sine of same frequency, $\sin(k\omega_0 t) \times \sin(k\omega_0 t)$
3. Cosine times a sine of a different frequency, $\cos(k\omega_0 t) \times \sin(m\omega_0 t)$
4. Sine times a sine of a different frequency, $\sin(k\omega_0 t) \times \sin(m\omega_0 t)$
5. Cosine times a cosine of same frequency, $\cos(k\omega_0 t) \times \cos(k\omega_0 t)$
6. Cosine times a cosine of a different frequency, $\cos(k\omega_0 t) \times \cos(m\omega_0 t)$

We need to compute the integral of each of these types of products over one period for the Fourier equation. Conveniently, it turns out, that the integral of most of these sinusoidal products, over one period is zero, except for Type 2 and Type 5, when the frequencies of the sinusoids coincide. This makes the problem so much more tractable!

It is easy to show that these terms are zero. However, let us examine first what happens when we integrate a type 2 product, a sine times a sine of the same frequency, over one period. We notice that the product of the two sine waves of the same frequency, as shown in Fig. 1.19, lies entirely above the x-axis and has a net positive area which is proportional to the coefficient b_k. From integral tables we can compute this area as equal to

$$\int_0^{T_0} b_k \sin(k\omega_0 t) \times \sin(k\omega_0 t)\, dt = b_k \frac{T_0}{2}$$

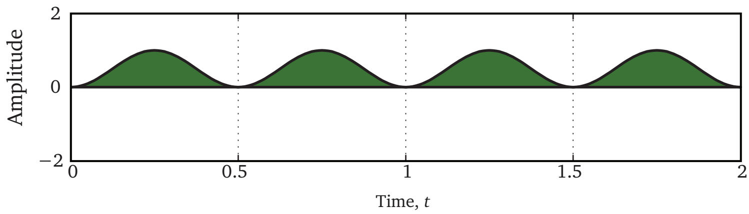

Figure 1.19: A sine wave multiplied by a sine wave of the same frequency has nonzero area under one period.

Now let us take a look at the integral of a Type 4 product, i.e. a sine wave multiplied by a sine wave of a different frequency.

$$\sin(k\omega_0 t) \times \sin(\underline{m}\omega_0 t)$$

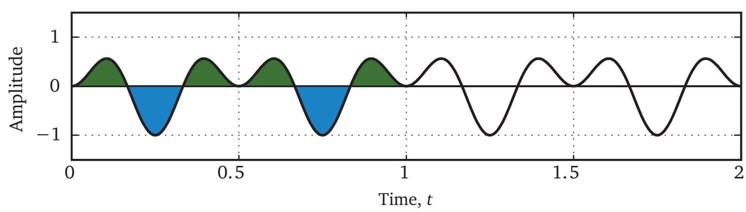

Figure 1.20: A sine wave multiplied by a sine wave of a different frequency has zero area under one period.

Figure 1.20 shows in the time domain the product of two sine waves of different frequencies. It may not be obvious from the figure, but the total area in one period of this product is zero. We make an important observation that the area in one period of a sine wave multiplied by *any* of its harmonics is zero. We conclude that when we multiply a signal by any of its harmonics, and integrate the product over one period, then the result can be used to find the contribution of just that harmonic and none other. The integral tells us how much of that frequency is present in the target signal. All other sinusoids in the signal contribute nothing. The individual contribution to the signal $f(t)$ by the

kth harmonic can be written as:

$$\int_0^{T_0} b_k \sin(k\omega_0 t) \times \sin(m\omega_0 t)\, dt = 0 \quad \text{for } k \neq m.$$
$$\int_0^{T_0} b_k \sin(k\omega_0 t) \times \sin(m\omega_0 t)\, dt = \frac{T_0}{2} b_k \quad \text{for } k = m.$$

The same is true for cosine waves or types 5 and 6.

$$\int_0^{T_0} a_k \cos(k\omega_0 t) \times \cos(m\omega_0 t)\, dt = 0 \quad \text{for } k \neq m.$$
$$\int_0^{T_0} a_k \cos(k\omega_0 t) \times \cos(m\omega_0 t)\, dt = \frac{T_0}{2} a_k \quad \text{for } k = m.$$

We see that the result of the integration of the product of two harmonics when their frequencies are unequal is **zero**. It is nonzero only when the waves have the *same* frequency. Hence, if we multiply our signal by its K harmonics, and integrate K times, the result for each case is the coefficient of the harmonic being used for multiplication.

Let us now look at the product of a sine wave and a cosine wave of the same frequency, or Type 1. For this case, as shown in Fig. 1.21, the net area under the product is also zero.

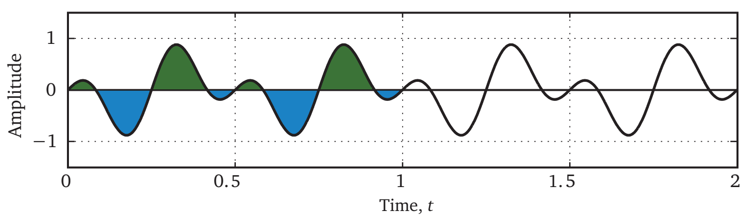

Figure 1.21: A sine wave multiplied by a cosine has total area of zero under one period.

The integral of Type 3 products is also zero. Hence, we note this important result; *the area under the product of a sine and a cosine over one period is zero whether the frequencies are the same or not.* Sines and cosines just don't agree. This is also the concept of orthogonality. We say that these waves are orthogonal to each other as they contribute nothing to the integral. Summarizing the results:

$$\int_0^{T} b_k \cos(k\omega_0 t) \times \sin(m\omega_0 t)\, dt = 0 \tag{1.20}$$

A practical interpretation of these properties is that sine and cosine waves act as filtering signals. When a signal, is multiplied by a sinusoid of a particular frequency, the integral is proportional to the content of that multiplying sinusoid, (by which we mean its amplitude). Hence, sinusoids behave as narrow-band filters. This is the fundamental concept of a *filter*.

Here are the key results we will be using in calculating the coefficients of the Fourier equation:

1. If you multiply a sine or a cosine wave by any of its harmonics, the area under the product is zero.
2. If you multiply a sine or a cosine of a particular frequency by itself, the area under the product is proportional to the Fourier coefficient of that frequency.
3. The area under a sine wave multiplied by a harmonic cosine is always zero. (Because sine and cosine are orthogonal!)

We use these observations to compute the b_k coefficients. We successively multiply the target signal, $f(t)$ by a sine wave of a specific harmonic frequency and then integrate over one period as in the equation below. All terms are zero except one. This term is given by:

$$\int_0^{T_0} b_k \sin(k\omega_0 t)\sin(k\omega_0 t)\,dt = \frac{b_k T_0}{2}$$

From this we obtain the coefficient of the sine, b_k as follows

$$\boxed{b_k = \frac{2}{T_0}\int_0^{T_0} f(t)\sin(k\omega_0 t)\,dt} \tag{1.21}$$

The coefficient b_k is computed by taking the target signal over one period, successively multiplying it with a sine wave of kth harmonic frequency and then integrating. The result of the integration is then multiplied (or normalized) by $2/T_0$ to obtain the coefficient for that particular harmonic. If we do this K times, we get K independent b_k coefficients.

Computing a_k, the coefficients of cosine harmonics

Now we need to do something similar for cosine coefficients. The process of computing the coefficients of the cosine harmonics is exactly the same as the one used for the sine waves. Instead of multiplying the target signal, $f(t)$ by a sine wave, we multiply the target signal sequentially by a cosine wave of frequency, $k\omega_0$. We get exactly the same result as when we compute the sine coefficients. Only one term will remain in the big long multiplication for each value of k. That term is

$$\int_0^{T_0} a_k \cos(k\omega_0 t)\cos(k\omega_0 t)\,dt = \frac{a_k T_0}{2}$$

The coefficient a_k hence can be calculated by multiplying $f(t)$ by the kth harmonic and integrating the expression. The result is proportional to the coefficient. This expression is nearly identical to Eq. (1.21) for the b_k coefficients.

$$\boxed{a_k = \frac{2}{T_0}\int_0^{T_0} f(t)\cos(k\omega_0 t)\,dt} \tag{1.22}$$

In all, we do this $2K$ times, with K computations each for the sine and, K for the cosine.

Conceptually, the process of computing the coefficients consists of filtering the target signal, one frequency at a time. Hence, a spectrum can be seen as tiny little filter fingers that pull out the quantity (amplitude) of that harmonic in the signal.

In some simple cases, when given a signal that is composed of sinusoids, we do not need to do any calculations, as the coefficients are there in plain sight for us to see as in the following example.

Example 1.1. What are the FSC of this signal:

$$f(t) = 0.8\sin(6\pi t) - 0.3\cos(6\pi t) + 0.75\cos(12\pi t)$$

The signal is already in the form of a Fourier series. Just by looking we can tell that the coefficient a_0 is 0. We set the fundamental frequency at 1 Hz, because then both frequencies in the signal, 3 Hz and 6 Hz fall on a harmonic. The coefficients by observation are: $a_0 = 0$, $a_3 = -0.3$, $b_3 = 0.8$ and $a_6 = 0.75$. Of course we could have set the fundamental as 3 Hz, in which case the coefficients would be given as: $a_0 = 0$, $a_1 = -0.3$, $b_1 = 0.8$ and $a_2 = 0.75$. Nothing has changed except the index.

We just did our first Fourier analysis without lifting a pencil! Now something a little more complicated.

Example 1.2. What are the FSC of this signal:

$$f(t) = \cos^3(2\pi t)$$

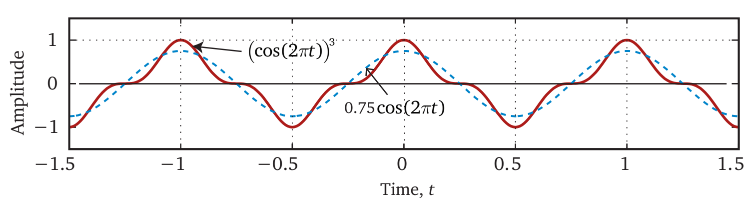

Figure 1.22: The signal $\cos^3(2\pi t)$*, in solid curve, and its first order representation in dashed line.*

From Fig. 1.22, the plot of this signal in the time-domain, we can spot two pieces of information. The first is that the period of the wave is 1 second (its frequency is 1 Hz.), and second, the symmetry of the signal is even. Because it is an even signal, it can be represented solely with cosine components. (But, of course, that makes sense as the function is cosine-cubed.) But because this a cubed function, at first we might not know what to do next. Fortunately, we can expand this function and put it in the Fourier series form, in which case the coefficients would become obvious. But instead, we will find the coefficients the hard way, and that means integration. First we calculate the a_0 coefficient using Eq. (1.19) to determine the DC offset.

$$a_0 = \int_0^1 \cos^3(2\pi t)\, dt = 0$$

This integration gives us a zero and the result agrees with the graph in Fig. 1.22. The function has a symmetry about the x-axis meaning there is no DC offset. Now using Eq. (1.22) we can determine the a_k coefficients (of cosine) starting with the fundamental harmonic. We set fundamental frequency f_0 equal to 1 Hz.

$$a_1 = 2\int_0^1 \cos^3(2\pi t)\cos(1 \times 2\pi t)\, dt = 0.75$$

The Fourier representation of this signal is now given by $0.75\cos(2\pi t)$. Figure 1.22 plots in the dashed line the representation of the signal using only the a_1 coefficient. It is clear that one coefficient is not enough to properly represent the signal. We should calculate a few more coefficients. However, it is amazing, how close we are with just one term!

$$a_2 = 2\int_0^1 \cos^3(2\pi t)\cos(2 \times 2\pi t)\, dt = 0$$
$$a_3 = 2\int_0^1 \cos^3(2\pi t)\cos(3 \times 2\pi t)\, dt = 0.25$$

With $a_0 = 0$, $a_1 = 0.75$, $a_2 = 0$, and $a_3 = 0.25$, the representation is given as:

$$f(t) = 0.75\cos(2\pi t) + 0.25\cos(6\pi t)$$

This is the Fourier series representation of the signal given to us as $\cos^3(2\pi t)$. If we continue to calculate more coefficients (for $k > 3$), we will see that they are all zero after a_3. From this, we can say that function cosine cubed is, in fact, made up of only two cosines of frequency 1 and 3 Hz. In this case, Fourier series representation is an exact representation of the signal, but this is not always the case.

Example 1.3. What are the FSC of this signal?

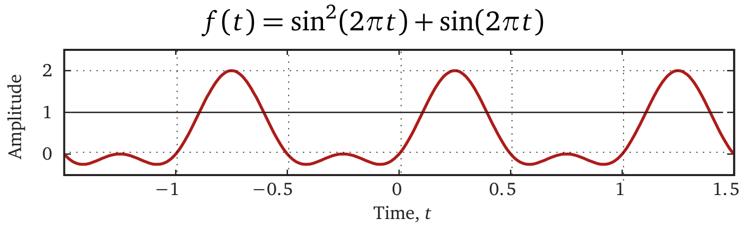

Figure 1.23: The signal, $f(t) = \sin^2(2\pi t) + \sin(2\pi t)$.

The analysis of this signal uses the ideas from the first two examples. The signal has a linear part, $\sin(2\pi t)$, and a non-linear part, $\sin^2(2\pi t)$. The best way to compute its Fourier series representation is to convert the non-linear term to a linear term. Using trigonometric identities, we do this conversion and rewrite the signal with all linear terms as:

$$f(t) = \frac{1}{2} + \sin(2\pi t) - \frac{1}{2}\cos(4\pi t)$$

From here, now that the signal is very nearly in Fourier form, we can expect to have nonzero a_0, b_1 and a_2 coefficients. The values of these will match the coefficients of the above equation. Solving for these the hard way, using integration, we show that this is indeed the case. The Fourier representation of this signal is exact. This only happens if the original signal is composed only of harmonic sinusoids. In real life, this situation is unlikely. Practical signals are almost never composed solely of harmonic sinusoids. That is our challenge. When we compute the Fourier representation, we are indeed *estimating* the form of the signal. And for this estimation, we use Fourier analysis.

$$a_0 = \int_0^1 \big(\sin^2(2\pi t) + \sin(2\pi t)\big)\, dt = \frac{1}{2}$$
$$a_1 = 2\int_0^1 \big(\sin^2(2\pi t) + \sin(2\pi t))\cos(1\cdot 2\pi t)\big)\, dt = 0$$
$$a_2 = 2\int_0^1 \big(\sin^2(2\pi t) + \sin(2\pi t))\cos(2\cdot 2\pi t)\big)\, dt = -\frac{1}{2}$$
$$b_1 = 2\int_0^1 \big(\sin^2(2\pi t) + \sin(2\pi t))\sin(1\cdot 2\pi t)\big)\, dt = 1$$
$$b_2 = 2\int_0^1 \big(\sin^2(2\pi t) + \sin(2\pi t))\sin(2\cdot 2\pi t)\big)\, dt = 0$$

Note that there is a $\frac{1}{2}$ in the equation of the signal, and this is the same thing as a DC offset, hence this is the a_0 coefficient. There is one cosine at a frequency of 2 Hz with amplitude of $-\frac{1}{2}$, which means that the coefficient a_2 would be $-\frac{1}{2}$. There is a sine at frequency of 1 Hz of amplitude 1, hence b_1 should be 1 and, it is. We could have solved this problem by observation only, without doing any calculations.

Coefficients become the spectrum

How do we go from coefficients to a spectrum? Assume that we have a signal which we have analyzed and have found that it has nine harmonics with $f_0 = 2.5$ Hz. The coefficients of the nine harmonics starting with $k = 1$ are given by:

$$a_k = [0.25, 0.2, 0.7, 0.5, -0.2, 0.2, 0.1, -0.05, 0.02]$$
$$b_k = [0.4, -0.3, 0.7, 0.7, 0.3, 0.275, 0.25, 0.2, 0.1]$$

Here, a_k are the coefficients of the cosine harmonics and b_k are the coefficients of the sine harmonics. We plot these in a bar graph in Fig. 1.24 as a function of the index, k. Both sine and cosine of the same frequency are plotted next to each other. This is a spectrum that in essence displays the recipe of the signal. It tells us how much of each harmonic, i.e., its amplitude, we need to recreate the signal. Given this spectrum, you should be able to write the Fourier series equation for this signal.

Commonly a spectrum is given in terms of "power" but the coefficients we compute via Fourier analysis are not power. They are *amplitudes*. The term spectrum is often used synonymously with power spectrum but is not the same as the power spectrum.

In signal processing, the coefficients computed for cosine are called *Real* and those for sine, are called *Imaginary*. Of course, there is nothing imaginary about the coefficients of sine, they are just as real as the cosine coefficients. It is just one of the many confusing terms we come across in signal processing.

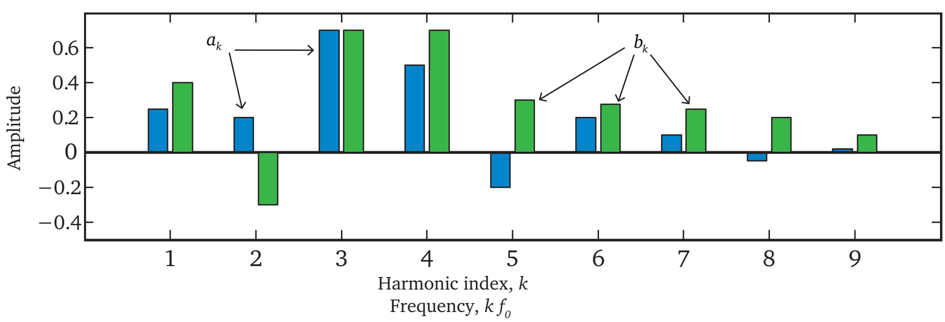

Figure 1.24: The coefficients of the harmonics.
For each harmonic index k representing a frequency of kf_0, there are two coefficients, one for sine and the other for cosine. These coefficients can be positive or negative and represent a formula for creating a Fourier representation of the test wave.

While the spectrum in Fig. 1.24 is in terms of sine and cosine coefficients, you may say that this is not the way we usually see a spectrum in books. A spectrum ought to have just one number for each frequency. So, we *combine* the sine and cosine coefficients into two real-world terms, called **magnitude and phase**.

As harmonic sine and cosine are orthogonal to each other, and have the same frequency, we can compute their *sum* by thinking of them as vectors. We can compute this by doing the root-sum-square (RSS) of the two coefficients. This RSS term is called the ***magnitude***. Now we plot the modified spectrum using magnitude on the y-axis instead of amplitudes of the sine and cosine. The magnitude is *one* number for each frequency and is computed by this expression.

$$\text{Magnitude: } c_k = \sqrt{a_k^2 + b_k^2} \tag{1.23}$$

While the amplitude can indeed be negative, the magnitude is always positive. The effect of the sign of the amplitude is now seen in a term called phase, which we calculate by

$$\text{Phase: } \phi_k = \tan^{-1} \frac{b_k}{a_k} \tag{1.24}$$

This phase term is different from the previous two types of phases we mentioned, as this is a combination of the phase of the sine and the cosine harmonic of a particular frequency. The range of arctan is from $-\pi/2$ to $+\pi/2$, so it must be unwrapped to compute the full phase. What it represents is the relative phase shift between the cosine and the sine harmonic. The phase spectrum is usually quite noisy due to computational effects.

We show these two quantities on a pair of plots, one for the magnitude and the other for this new phase term, both plotted as a function of the harmonic frequencies or the index, as shown in Fig. 1.25. The first one is called the **magnitude spectrum** and the second is called the **phase spectrum**.

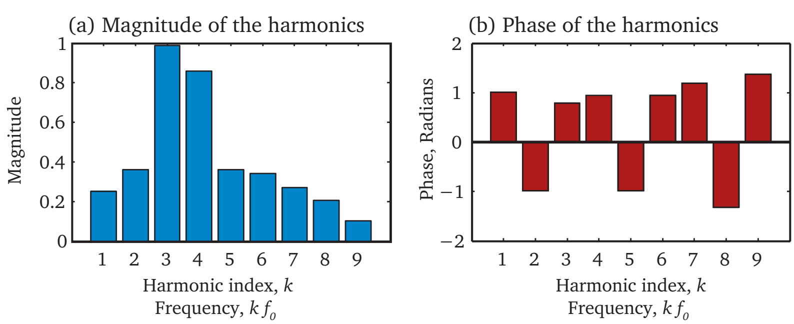

Figure 1.25: The magnitude and phase spectrum.
In (a), we see the RSS of both the sine and cosine coefficients, representing the magnitude of that harmonic frequency. In (b), we plot the combined phase for each frequency. The magnitude on the right is usually quite instructive but phase is hard to comprehend as it changes quickly from one frequency to the next.

Hence, a spectrum can be plotted in two ways: as a pair of plots of the coefficients of the sine and the cosine harmonics, or, as the magnitude and phase. (In Fig. 1.24, we have plotted both the real and imaginary coefficients on just on one figure.) The spectrum based on the real(cosine) and imaginary(sine) coefficients, changes depending on the starting point of the analysis. Whereas in the second form, the plot of magnitude and the phase are not a function of the starting point. Hence practicing engineers prefer the second form.

The magnitude spectrum, computed from Eq. (1.23) is the preferred form in industry. Its companion, the phase spectrum usually does not offer information we can use quickly, and often goes ignored. Confusingly, most people make no distinction when talking about amplitude or magnitude, and often these two terms are used interchangeably.

The power spectrum

The process of doing Fourier analysis consists of computing the amplitude of each harmonic and then converting it to the magnitude and phase spectrum. Fig. 1.26 shows a power spectrum with the y-axis given in dBs. To convert a magnitude spectrum to a power spectrum, we use Parseval's theorem. This tells us that the total power in a signal is the sum of the powers in each harmonic. The power of each harmonic is defined as the square of its magnitude. Hence to represent power, you square each magnitude value and then compute its dB value by $10 \cdot \log10(c_k)^2$. Alternatively, you can compute the base 10 log of the coefficient and then multiply it by 20.

$$\text{Power in the } k\text{th harmonic} = 20\log(c_k) \tag{1.25}$$

The power spectrum is often normalized to maximum power, such as **bin 3** in Fig. 1.26. (The bin is a form of identifier of the harmonic and is equivalent to the harmonic index k.) The level of each component is the dB equivalent of its ratio to the maximum power. All component levels relative to the maximum power, are said to be a certain number of dBs below the maximum. The maximum level is zero dB, with all other values shown as negative. These values are also called **power spectral density** (PSD) because they are a form of density of the power across a bandwidth.

Looking at the FSC, note that there are just K (the largest k) number of such coefficients, one for each of the harmonics. The spectrum from the FSC is thus considered *discrete*. In the time domain the signal is continuous, however, in the frequency domain, the one-sided spectrum developed from the FSC is *discrete*. It has only K terms. Since the index k is positive, the plot starts at 0 and hence, is called, a **one-sided spectrum**.

This is an important property of the Fourier series. A one-sided spectrum consists only of a finite number of components, each an independent value. The x-axis is given either

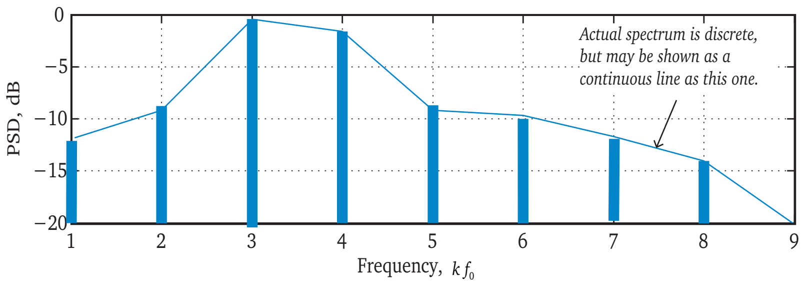

Figure 1.26: A traditional Power spectrum created from the Fourier coefficients.

by the index k, but more commonly, by the actual frequency ranging from 0, f_0, $2f_0$, $3f_0$,..., etc.

Fourier analysis applies only to periodic waves. And if that is all we could do with Fourier analysis, then it would not have a lot of use. For real signals we can never tell what the period is nor where it starts. Usually no periodicity can be seen. In fact, a real signal may not be periodic at all. In this case, further developments of the theory allow us to extend the "period" to infinity so we just pick any section of a signal or even the whole signal and call it "The Period" representing the whole signal. This idea is the basis of the Fourier transform that we will discuss in Chapter 4.

In this chapter we discussed the Fourier series representation of an arbitrary periodic wave in terms of weighted harmonic sine and cosine waves. In the next chapter, we will look at how we can do the same using the confusing and scary-looking complex exponentials.

Summary of Chapter 1

In this chapter, we introduced the concept of using sinusoids to represent an arbitrary periodic wave. We also introduced the concept of the fundamental frequency and its harmonics. The Fourier series representation consists of finding unique weightings of these harmonics to represent a particular periodic wave. These weights are called the Fourier series coefficients (FSC). When plotted as a function of frequency, these coefficients represent the spectrum of the signal. The spectrum calculated using FSC is discrete although the signal is continuous in time.

Terms used in this chapter:

- **Fundamental frequency** - The smallest frequency of the signal to be represented by Fourier series.
- **Harmonic frequencies** - All integer multiples of the fundamental frequency.
- **Sinusoids** - Sine or cosine wave.
- **Harmonic coefficients** - The amplitude of a harmonic.

- **Real and imaginary** - Cosine is said to exist in the real plane and sine in the imaginary plane.
- **Magnitude** - The RSS value of the amplitudes of the sine and cosine of a particular harmonic. It is always positive.
- **Phase** - The fixed phase shift value of a wave at $t = 0$, often specified in radians.
- **One-sided spectrum** - A spectrum for positive harmonic index, $k = 0, 1, 2$, etc.
- **Power spectrum** - Showing distribution of power rather than amplitude.

1. The most common trigonometric form of the Fourier series is given by

$$f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(2\pi f_k t) + \sum_{k=1}^{\infty} b_k \sin(2\pi f_k t)$$

2. The coefficients of the Fourier series are easily computed by

$$a_0 = \frac{1}{T_0} \int_0^{T_0} f(t)\, dt$$
$$b_k = \frac{2}{T_0} \int_0^{T_0} f(t) \sin(k\omega_0 t)\, dt$$
$$a_k = \frac{2}{T_0} \int_0^{T_0} f(t) \cos(k\omega_0 t)\, dt$$

3. Many periodic signals can be represented by the weighted sum of harmonics sinusoids. The representation is an estimation and may not be an exact replication.
4. Harmonic sinusoids are orthogonal to each other; hence, the integral of their products (or cross product) is zero.
5. The linearity property of the Fourier series implies that a change of the coefficients of one harmonic does not affect the coefficients of the other harmonics.
6. A time or phase shift of the signal does not affect the magnitude of the coefficients.
7. To synthesize a signal based on the Fourier series, we pick a fundamental frequency first. All harmonics are integer multiples of the fundamental.
8. We designate harmonics by the letter k. Hence, for all integer k, all kf_0 frequencies are harmonics of the fundamental frequency f_0.
9. Fourier analysis means to find the Fourier coefficients of the Fourier series.
10. The Fourier coefficients are discrete because the harmonics are discrete. Hence the spectrum of a periodic continuous-time signal is discrete.
11. A spectrum can be a spectrum of amplitude, magnitude or power. These are all different and used for different purposes.

Questions

1. Can you state in words the principle behind the Fourier series.
2. What is the first and the third harmonic of a sinusoid of frequency 3 Hz?
3. Can any signal have harmonics?
4. When is one harmonic orthogonal to another? What trigonometric property tells us that this is true?
5. Three non-harmonic sinusoids are added together. What is the period of the summed wave?
6. If we add three non-harmonic sinusoids together, is the resulting signal periodic?
7. Is the sum of N harmonics also harmonic to the fundamental?
8. Are harmonics within a harmonic set *also* harmonic to each other?
9. If we make the amplitude of a signal, a function of time, what effect will that have on its frequency?
10. A change in phase of a cosine wave means it is still orthogonal to a sine wave of the same frequency; true or false?
11. If the phase of a harmonic is changed, does it remain harmonic to an another wave?
12. If the amplitude of a harmonic is changed, does it remain a harmonic to an another wave?
13. What is the maximum amplitude of N harmonic cosine waves added together. What is it for sine waves?
14. We need to represent a wave that starts at time t = 0. What type of harmonics will be in its representation?
15. The summation of odd and even waves can be used to create any waveform we want, true or not?
16. Is Fourier series representation an accurate representation of a wave? Why not?
17. Can we create a Fourier series representation of any wave?
18. Why do we consider the set of harmonics a *basis set*? What constitutes a basis set?
19. Sine and cosine waves are a basis set for Fourier analysis. Can you give an example of another set of basis functions.
20. What quality of sine and cosine makes them suitable as a basis set?
21. Fourier series analysis is considered a linear process. Why?
22. What do the coefficients of a Fourier series represent? What does the a_0 coefficient represent?
23. What are the Fourier series coefficients of this signal? $s = A + B\sin(2\pi f t)$.
24. We want to compute the FSC of this signal. $s = \sin(6.5t) - \cos(4.75t)$. What should we pick as the fundamental frequency so as to accurately represent this signal.

25. If the target wave is shifted by a certain phase, what happens to its coefficients?
26. How many coefficients would you need to describe this wave? $x(t) = 2 + B\sin(4\pi t + \pi/2) - \cos(12\pi t)$. Find the coefficients of this signal.
27. What is the fundamental period of this Fourier series representation? $x(t) = (1/2)(\sin(4\pi t) + (1/3)\sin(12\pi t) + (1/5)\sin(20\pi t) + (1/7)\sin(28\pi t)$. What are the coefficients of the first four cosine and sine harmonics?
28. Given these equations, what are the Fourier series coefficients, a_0, a_1, a_2 for each case.
    (a) $y = \frac{1}{2} + \frac{3}{4}\sin(\pi x) - \frac{3}{5}\cos(2\pi x)$.
    (b) $y = \frac{3}{4}\cos(2\pi x) - \frac{3}{5}\cos(3\pi x)$.
29. What is the difference between amplitude and magnitude?
30. The amplitude of a harmonic varies from -1 v to +1 v. What is its peak amplitude and its peak-to-peak amplitude? What is its magnitude? What is the power of the harmonic? What is the value of the power in dB?
31. Examine Fig. 1.13 and give the first four coefficients of the even square wave.