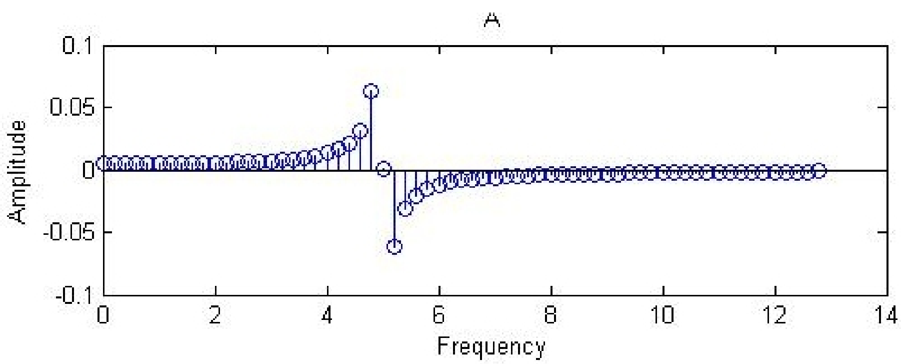
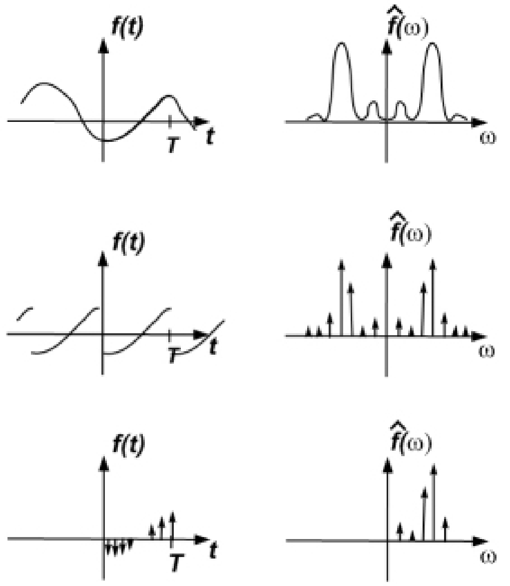
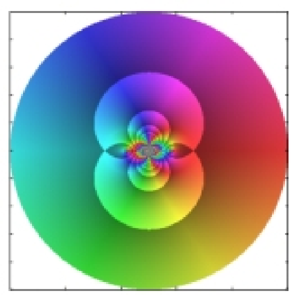
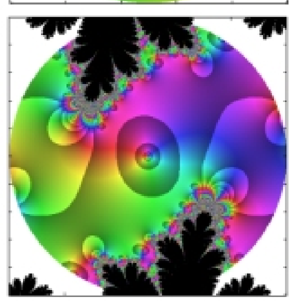
RUSSELL L. HERMAN

AN INTRODUCTION TO FOURIER AND COMPLEX ANALYSIS WITH APPLICATIONS TO THE SPECTRAL ANALYSIS OF SIGNALS









R. L. HERMAN - VERSION DATE: JANUARY 13, 2016

PUBLISHED BY R. L. HERMAN

This text has been reformatted from the original using a modification of the Tufte-book documentclass in LaTeX. See TUFTE-LATEX.GOOGLECODE.COM.

Sixth printing, January 2016

Contents

Dedicated to those students who have endured the various editions of AN INTRODUCTION TO FOURIER AND COMPLEX ANALYSIS WITH APPLICATIONS TO THE SPECTRAL ANALYSIS OF SIGNALS.

Introduction

"A mathematical theory is not to be considered complete until you have made it so clear that you can explain it to the first man whom you meet on the street." David Hilbert (1862-1943)

This is an introduction to topics in Fourier analysis and complex analysis. These notes have been class tested several times since 2005.

This book is based on a course in applied mathematics originally taught at the University of North Carolina Wilmington in 2004 and set to book form in 2005. The notes were used and modified in several times since 2005. The course is an introduction to topics in Fourier analysis and complex analysis. Students are introduced to Fourier series, Fourier transforms, and a basic complex analysis. As motivation for these topics, we aim for an elementary understanding of how analog and digital signals are related through the spectral analysis of time series. There are many applications using spectral analysis. These course is aimed at students majoring in mathematics and science who are at least at their junior level of mathematical maturity.

At the root of these studies is the belief that continuous waveforms are composed of a number of harmonics. Such ideas stretch back to the Pythagoreans study of the vibrations of strings, which led to their program of a world of harmony. This idea was carried further by Johannes Kepler (1571-1630) in his harmony of the spheres approach to planetary orbits. In the 1700's others worked on the superposition theory for vibrating waves on a stretched spring, starting with the wave equation and leading to the superposition of right and left traveling waves. This work was carried out by people such as John Wallis (1616-1703), Brook Taylor (1685-1731) and Jean le Rond d'Alembert (1717-1783).

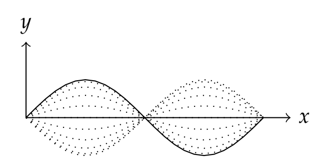


Figure 1: Plot of the second harmonic of a vibrating string at different times.

In 1742 d'Alembert solved the wave equation

$$c^2 \frac{\partial^2 y}{\partial x^2} - \frac{\partial^2 y}{\partial t^2} = 0,$$

where y is the string height and c is the wave speed. However, this solution led himself and others, like Leonhard Euler (1707-1783) and Daniel Bernoulli (1700-1782), to investigate what "functions" could be the solutions of this equation. In fact, this led to a more rigorous approach to the study of analysis by first coming to grips with the concept of a function. For example, in 1749 Euler sought the solution for a plucked string in which case the initial condition $y(x,0) = h(x)$ has a discontinuous derivative! (We will see how this led to important questions in analysis.)

Solutions of the wave equation, such as the one shown, are solved using the Method of Separation of Variables. Such solutions are studies in courses in partial differential equations and mathematical physics.

In 1753 Daniel Bernoulli viewed the solutions as a superposition of simple vibrations, or harmonics. Such superpositions amounted to looking at solutions of the form

$$y(x,t) = \sum_k a_k \sin\frac{k\pi x}{L}\cos\frac{k\pi ct}{L},$$

where the string extend over the interval $[0, L]$ with fixed ends at $x = 0$ and $x = L$.

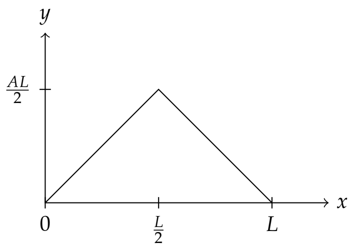


Figure 2: Plot of an initial condition for a plucked string.

However, the initial conditions for such superpositions are

$$y(x,0) = \sum_k a_k \sin\frac{k\pi x}{L}.$$

It was determined that many functions could not be represented by a finite number of harmonics, even for the simply plucked string given by an initial condition of the form

$$y(x,0) = \begin{cases} Ax, & 0 \le x \le L/2, \\ A(L-x), & L/2 \le x \le L. \end{cases}$$

Thus, the solution consists generally of an infinite series of trigonometric functions.

Such series expansions were also of importance in Joseph Fourier's (1768-1830) solution of the heat equation. The use of such Fourier expansions has become an important tool in the solution of linear partial differential equations, such as the wave equation and the heat equation. More generally, using a technique called the Method of Separation of Variables, allowed higher dimensional problems to be reduced to one-dimensional boundary value problems. However, these studies led to very important questions, which in turn opened the doors to whole fields of analysis. Some of the problems raised were

The one dimensional version of the heat equation is a partial differential equation for $u(x,t)$ of the form

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}.$$

Solutions satisfying boundary conditions $u(0,t) = 0$ and $u(L,t) = 0$, are of the form

$$u(x,t) = \sum_{n=0}^{\infty} b_n \sin\frac{n\pi x}{L} e^{-n^2\pi^2 kt/L^2}.$$

In this case, setting $u(x,0) = f(x)$, one has to satisfy the condition

$$f(x) = \sum_{n=0}^{\infty} b_n \sin\frac{n\pi x}{L}.$$

This is another example leading to an infinite series of trigonometric functions.

1. What functions can be represented as the sum of trigonometric functions?
2. How can a function with discontinuous derivatives be represented by a sum of smooth functions, such as the above sums of trigonometric functions?
3. Do such infinite sums of trigonometric functions actually converge to the functions they represent?

There are many other systems in which it makes sense to interpret the solutions as sums of sinusoids of particular frequencies. One example comes from the study of ocean waves. Ocean waves are affected by the gravitational pull of the moon and the sun (and many other forces). These periodic forces lead to the tides, which in turn have their own periods of motion. In an analysis of ocean wave heights, one can separate out the tidal components by making use of Fourier analysis. Typically, we views the tide height $y(t)$ as a continuous function. One sits at a specific location and measures the movement of the ocean surface as a function of time. Such a function, or time series, is called an analog function. Another common analog signal is an audio signal, giving the amplitude of a sound (musical note, noise, speech, etc.) as a function of time (or space). However, in both of these cases, we actually observe a part of the signal. This is because we can only sample a finite amount of data over a finite time interval. Thus, we have only the values $y_n = y(t_n)$. However, we are still interested in the spectral (frequency) content of our signals even if the signal is not continuous in time.

For example, for the case of ocean waves we would like to use the discrete signal (the sampled heights) to determine the tidal components. For the case of audio signals, we may want to save a finite amount of discretized information as an audio file to play back later on our computer.

So, how are the analog and discrete signals related? We sample an analog signal, obtaining a discrete version of the signal. By sampling an analog signal, we might wonder how the sampling affects the spectral content of the original analog signal. What mathematics do we need to understand these processes? That is what we will study in this course. We will look at Fourier trigonometric series, integral transforms, and discrete transforms. However, we will actually begin with a review of infinite series. We will recall what infinite series are and when they do, or do not, converge. Then we will be ready to talk about the convergence of series of sinusoidal functions, which occur in Fourier series.

We will see how Fourier series are related to analog signals. A true representation of an analog signal comes from an infinite interval and not a finite interval, such as that the vibrating string lives on. This will lead to Fourier Transforms. In order to work with continuous transforms, we will need a little complex analysis. So, we will spend a few sections on an introduction to complex analysis. This consists of the introduction of complex function, their derivatives, series representations, and integration in the complex plane.

Having represented continuous signals and their spectral content by Fourier transforms, we will then see what needs to be done to represent discrete signals. We end the course by investigating the connection between these two types of signals and some of the consequences of processing analog data through real measurement and/or storage devices.

However, the theory of Fourier analysis is much deeper than just looking at sampling time series. The idea of representing functions as an expansion of oscillatory functions extends far intro both physics and mathematics. In physics, oscillatory and wave motion are crucial in electromagnetism, optics and even quantum mechanics. In mathematics, the concepts of expansion of functions in sinusoidal functions is the basis of expanding functions over an infinite dimensional basis. These ideas can be expanded beyond the sinusoidal basis, as we will see later in the book. Thus, the background to much of what we are doing involves delving into infinite dimensional vector spaces. Hopefully, the basics presented here will be useful in your future studies.

The topics to be studied in the book are laid out as follows:

1. Sequences and Infinite Series
2. Fourier Trigonometric Series
3. Vector Spaces
4. Generalized Fourier Series
5. Complex Analysis
6. Integral Transforms
7. Analog vs Discrete Signals
8. Signal Analysis

At this point I should note that most of the examples and ideas in this book are not original. These notes are based upon mostly standard examples from assorted calculus texts, like Thomas and Finney[1], advanced calculus texts like Kaplan's *Advanced Calculus*, [2], texts in mathematical physics[3], and other areas[4]. A collection of some of these well known sources are given in the bibliography and on occasion specific references will be given for somewhat hard to find ideas.

[1] G. B. Thomas and R. L. Finney. *Calculus and Analytic Geometry*. Addison-Wesley Press, Cambridge, MA, ninth edition, 1995

[2] K. Kaplan. *Advanced Calculus*. Addison Wesley Publishing Company, fourth edition, 1991

[3] G. Arfken. *Mathematical Methods for Physicists*. Academic Press, second edition, 1970

[4] A. J. Jerri. *Integral and Discrete Transforms with Applications and Error Analysis*. Marcal Dekker, Inc, 1992

1

Review of Sequences and Infinite Series

"Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth." Sherlock Holmes (by Sir Arthur Conan Doyle, 1859-1930)

In this chapter we will review and extend some of the concepts and definitions related to infinite series that you might have seen previously in your calculus class [1] [2] [3]. Working with infinite series can be a little tricky and we need to understand some of the basics before moving on to the study of series of trigonometric functions.

The material in this chapter is a review of material covered in a standard course in calculus with some additional notions from advanced calculus. It is provided as a review before encountering the notion of Fourier series and their convergence as seen in the next chapter.

For example, one can show that the infinite series

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

converges to $\ln 2$. However, the terms can be rearranged to give

$$1 + \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{5}\right) + \left(\frac{1}{7} - \frac{1}{4} + \frac{1}{9}\right) + \left(\frac{1}{11} - \frac{1}{6} + \frac{1}{13}\right) + \cdots = \frac{3}{2}\ln 2.$$

In fact, other rearrangements can be made to give any desired sum!

[1] G. B. Thomas and R. L. Finney. *Calculus and Analytic Geometry*. Addison-Wesley Press, Cambridge, MA, ninth edition, 1995

[2] J. Stewart. *Calculus: Early Transcendentals*. Brooks Cole, sixth edition, 2007

[3] K. Kaplan. *Advanced Calculus*. Addison Wesley Publishing Company, fourth edition, 1991

Other problems with infinite series can occur. Try to sum the following infinite series to find that

$$\sum_{k=2}^{\infty} \frac{\ln k}{k^2} \sim 0.937548\ldots.$$

A sum of even as many as a million terms only gives convergence to four or five decimal places.

As we will see,

$$\ln(1+x) = x - \frac{x}{2} + \frac{x}{3} - \ldots.$$

So, inserting $x = 1$ yields the first result - at least formally! It was shown in Cowen, Davidson and Kaufman (in *The American Mathematical Monthly*, Vol. 87, No. 10. (Dec., 1980), pp. 817-819) that expressions like

$$\begin{aligned} f(x) &= \frac{1}{2}\left[\ln\frac{1+x}{1-x} + \ln(1-x^4)\right] \\ &= \frac{1}{2}\ln\left[(1+x)^2(1+x^2)\right] \end{aligned}$$

lead to alternate sums of the rearrangement of the alternating harmonic series. See Problem 6

The series

$$\frac{1}{x} - \frac{1}{x^2} + \frac{2!}{x^3} - \frac{3!}{x^4} + \frac{4!}{x^5} - \cdots, \quad x > 0$$

diverges for all x. So, you might think this divergent series is useless. However, truncation of this divergent series leads to an approximation of the integral

$$\int_0^{\infty} \frac{e^{-t}}{x+t}\, dt, \quad x > 0.$$

So, can we make sense out of any of these, or other manipulations, of infinite series? We will not answer all of these questions, but we will go back and review what you have seen in your calculus classes.

1.1 Sequences of Real Numbers

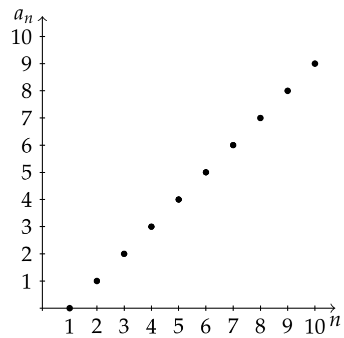


Figure 1.1: Plot of the terms of the sequence $a_n = n - 1$, $n = 1, 2, \ldots, 10$.

We begin with the definitions for sequences and series of numbers. A sequence is a function whose domain is the set of positive integers, $a(n)$, $n \in N$ $[N = \{1, 2, \ldots.\}]$.

Examples are

1. $a(n) = n$ yields the sequence $\{1, 2, 3, 4, 5, \ldots\}$,
2. $a(n) = 3n$ yields the sequence $\{3, 6, 9, 12, \ldots\}$.

However, one typically uses subscript notation and not functional notation: $a_n = a(n)$. We then call a_n the nth term of the sequence. Furthermore, we will denote sequences by $\{a_n\}_{n=1}^{\infty}$. Sometimes we will only give the nth term of the sequence and will assume that $n \in N$ unless otherwise noted.

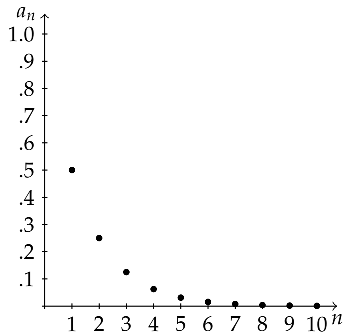


Figure 1.2: Plot of the terms of the sequence $a_n = \frac{1}{2^n}$, for $n = 1, 2, \ldots, 10$.

Another way to define a particular sequence is recursively. A recursive sequence is defined in two steps:

1. The value of first term (or first few terms) is given.
2. A rule, or recursion formula, to determine later terms from earlier ones is given.

Example 1.1. *A typical example is given by the Fibonacci*[4] *sequence. It can be defined by the recursion formula $a_{n+1} = a_n + a_{n-1}$, $n \geq 2$ and the starting values of $a_1 = 0$ and $a_1 = 1$. The resulting sequence is $\{a_n\}_{n=1}^{\infty} = \{0, 1, 1, 2, 3, 5, 8, \ldots\}$. Writing the general expression for the nth term is possible, but it is not as simply stated. Recursive definitions are often useful in doing computations for large values of n.*

[4] Leonardo Pisano Fibonacci (c.1170-c.1250) is best known for this sequence of numbers. This sequence is the solution of a problem in one of his books: *A certain man put a pair of rabbits in a place surrounded on all sides by a wall. How many pairs of rabbits can be produced from that pair in a year if it is supposed that every month each pair begets a new pair which from the second month on becomes productive* http://www-history.mcs.st-and.ac.uk

1.2 Convergence of Sequences

Next we are interested in the behavior of sequences as n gets large. For the sequence defined by $a_n = n - 1$, we find the behavior as shown in Figure 1.1. Notice that as n gets large, a_n also gets large. This sequence is said to be divergent.

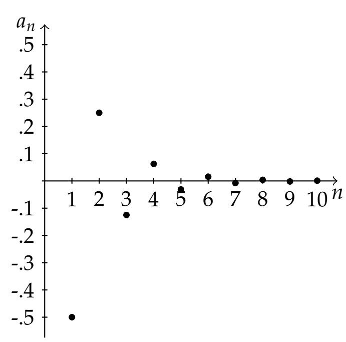


Figure 1.3: Plot of the terms of the sequence $a_n = \frac{(-1)^n}{2^n}$, for $n = 1, 2, \ldots, 10$.

On the other hand, the sequence defined by $a_n = \frac{1}{2^n}$ approaches a limit as n gets large. This is depicted in Figure 1.2. Another related series, $a_n = \frac{(-1)^n}{2^n}$, is shown in Figure 1.3 and it is also seen to approach 0.. The latter sequence is called an alternating sequence since the signs alternate from term to term. The terms in the sequence are $\{-\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \ldots\}$

The last two sequences are said to converge. In general, a sequence a_n converges to the number L if to every positive number ϵ there corresponds an integer N such that for all n,

$$n > N \Rightarrow |a_n - L| < \epsilon.$$

If no such number exists, then the sequence is said to diverge.

In Figures 1.4-1.5 we see what this means. For the sequence given by $a_n = \frac{(-1)^n}{2^n}$, we see that the terms approach $L = 0$. Given an $\epsilon > 0$, we ask for what value of N the nth terms $(n > N)$ lie in the interval $[L - \epsilon, L + \epsilon]$. In these figures this interval is depicted by a horizontal band. We see that for convergence, sooner, or later, the tail of the sequence ends up entirely within this band.

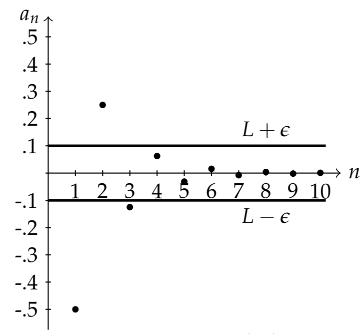


Figure 1.4: Plot of $a_n = \frac{(-1)^n}{2^n}$ for $n = 1 \dots 10$. Picking $\epsilon = 0.1$, one sees that the tail of the sequence lies between $L + \epsilon$ and $L - \epsilon$ for $n > 3$.

If a sequence $\{a_n\}_{n=1}^{\infty}$ converges to a limit L, then we write either $a_n \to L$ as $n \to \infty$ or $\lim_{n\to\infty} a_n = L$. For example, we have already seen in Figure 1.3 that $\lim_{n\to\infty} \frac{(-1)^n}{2^n} = 0$.

1.3 *Limit Theorems*

Once we have defined the notion of convergence of a sequence to some limit, then we can investigate the properties of the limits of sequences. Here we list a few general limit theorems and some special limits, which arise often.

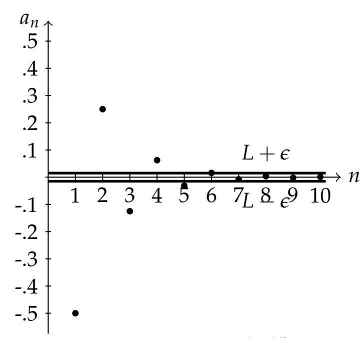


Figure 1.5: Plot of $a_n = \frac{(-1)^n}{2^n}$ for $n = 1 \dots 10$. Picking $\epsilon = 0.015$, one sees that the tail of the sequence lies between $L + \epsilon$ and $L - \epsilon$ for $n > 4$.

Limit Theorem

Theorem 1.1. *Consider two convergent sequences $\{a_n\}$ and $\{b_n\}$ and a number k. Assume that $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$. Then we have*

1. $\lim_{n\to\infty}(a_n \pm b_n) = A \pm B.$
2. $\lim_{n\to\infty}(kb_n) = kB.$
3. $\lim_{n\to\infty}(a_n b_n) = AB.$
4. $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B}, \quad B \neq 0.$

Some special limits are given next. These are generally first encountered in a second course in calculus.

Special Limits

Theorem 1.2. *The following are special cases:*

1. $\lim_{n\to\infty} \frac{\ln n}{n} = 0.$
2. $\lim_{n\to\infty} n^{\frac{1}{n}} = 1.$
3. $\lim_{n\to\infty} x^{\frac{1}{n}} = 1, \quad x > 0.$
4. $\lim_{n\to\infty} x^n = 0, \quad |x| < 1.$
5. $\lim_{n\to\infty}(1 + \frac{x}{n})^n = e^x.$
6. $\lim_{n\to\infty} \frac{x^n}{n!} = 0.$

The proofs generally are straightforward. For example, one can prove the first limit by first realizing that $\lim_{n\to\infty} \frac{\ln n}{n} = \lim_{x\to\infty} \frac{\ln x}{x}$. This limit in its current form is indeterminate as x gets large $(x \to \infty)$ since the numerator and the denominator get large for large x. In such cases one employs

L'Hopital's Rule is used often in computing limits. We recall this powerful rule here as a reference for the reader.

Theorem 1.3. *Let c be a finite number or $c = \infty$. If $\lim_{x\to c} f(x) = 0$ and $\lim_{x\to c} g(x) = 0$, then*

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}.$$

If $\lim_{x\to c} f(x) = \infty$ and $\lim_{x\to c} g(x) = \infty$, then

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}.$$

[5] We should note that we are assuming something about limits of composite functions. Let a and b be real numbers. Suppose f and g are continuous functions, $\lim_{x\to a} f(x) = f(a)$ and $\lim_{x\to b} g(x) = b$, and $g(b) = a$. Then,

$$\begin{aligned} \lim_{x\to b} f(g(x)) &= f\left(\lim_{x\to b} g(x)\right) \\ &= f(g(b)) = f(a). \end{aligned}$$

L'Hopital's Rule. We find that

$$\lim_{x\to\infty} \frac{\ln x}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0.$$

The second limit in Theorem 1.2 can be proven by first looking at

$$\lim_{n\to\infty} \ln n^{1/n} = \lim_{n\to\infty} \frac{1}{n} \ln n = 0$$

from the previous limit case. Now, if $\lim_{n\to\infty} \ln f(n) = 0$, then $\lim_{n\to\infty} f(n) = e^0 = 1$. Thus proving the second limit.[5]

The third limit can be done similarly. The reader is left to confirm the other limits. We finish this section with a few selected examples.

Example 1.2. *Evaluate* $\lim_{n\to\infty} \frac{n^2+2n+3}{n^3+n}$.

Divide the numerator and denominator by n^2. Then,

$$\lim_{n\to\infty} \frac{n^2+2n+3}{n^3+n} = \lim_{n\to\infty} \frac{1+\frac{2}{n}+\frac{3}{n^2}}{n+\frac{1}{n}} = \lim_{n\to\infty} \frac{1}{n} = 0.$$

Another approach to this type of problem is to consider the behavior of the numerator and denominator as $n \to \infty$. As n gets large, the numerator behaves like n^2, since $2n+3$ becomes negligible for large enough n. Similarly, the denominator behaves like n^3 for large n. Thus,

$$\lim_{n\to\infty} \frac{n^2+2n+3}{n^3+n} = \lim_{n\to\infty} \frac{n^2}{n^3} = 0.$$

Example 1.3. *Evaluate* $\lim_{n\to\infty} \frac{\ln n^2}{n}$.

Rewriting $\frac{\ln n^2}{n} = \frac{2\ln n}{n}$, we find from identity 1 of Theorem 1.2 that

$$\lim_{n\to\infty} \frac{\ln n^2}{n} = 2 \lim_{n\to\infty} \frac{\ln n}{n} = 0.$$

Example 1.4. *Evaluate* $\lim_{n\to\infty} (n^2)^{\frac{1}{n}}$.

To compute this limit, we rewrite

$$\lim_{n\to\infty} (n^2)^{\frac{1}{n}} = \lim_{n\to\infty} (n)^{\frac{1}{n}} (n)^{\frac{1}{n}} = 1,$$

using identity 2 of Theorem 1.2.

Example 1.5. *Evaluate* $\lim_{n\to\infty} (\frac{n-2}{n})^n$.

This limit can be written as

$$\lim_{n\to\infty} \left(\frac{n-2}{n}\right)^n = \lim_{n\to\infty} \left(1+\frac{(-2)}{n}\right)^n = e^{-2}.$$

Here we used identity 5 of Theorem 1.2.

1.4 *Infinite Series*

In this section we investigate the meaning of infinite series, which are infinite sums of the form

$$a_1 + a_2 + a_2 + \dots. \tag{1.1}$$

A typical example is the infinite series

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots. \tag{1.2}$$

There is story described in E.T. Bell's "Men of Mathematics" about Carl Friedrich Gauß (1777-1855). Gauß' third grade teacher needed to occupy the students, so she asked the class to sum the first 100 integers thinking that this would occupy the students for a while. However, Gauß was able to do so in practically no time. He recognized the sum could be written as $(1+100)+(2+99)+\dots(50+51)=50(101)$. This sum is a special case of

$$\sum_{k=1}^{n} k=\frac{n(n+1)}{2}.$$

This is an example of an arithmetic progression that is a finite sum of terms. .

E. T. Bell. *Men of Mathematics*. Fireside Books, 1965

How would one evaluate this sum? We begin by just adding the terms. For example,

$$\begin{aligned}
1+\frac{1}{2}&=\frac{3}{2},\\
1+\frac{1}{2}+\frac{1}{4}&=\frac{7}{4},\\
1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}&=\frac{15}{8},\\
1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}&=\frac{31}{16},\dots.
\end{aligned} \tag{1.3}$$

The values tend to a limit. We can see this graphically in Figure 1.6.

In general, we want to make sense out of Equation (1.1). As with the example, we look at a sequence of partial sums . Thus, we consider the sums

$$\begin{aligned}
s_1 &= a_1,\\
s_2 &= a_1+a_2,\\
s_3 &= a_1+a_2+a_3,\\
s_4 &= a_1+a_2+a_3+a_4,\dots.
\end{aligned} \tag{1.4}$$

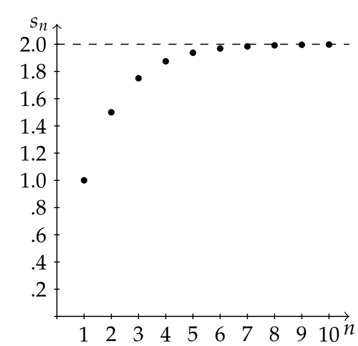


Figure 1.6: Plot of $s_n=\sum_{k=1}^{n}\frac{1}{2^{k-1}}$ for $n=1\dots10$.

In general, we define the nth partial sum as

$$s_n=a_1+a_2+\dots+a_n.$$

If the infinite series (1.1) is to make any sense, then the sequence of partial sums should converge to some limit. We define this limit to be the sum of the infinite series, $S=\lim_{n\to\infty}s_n$. If the sequence of partial sums converges to the limit L as n gets large, then the infinite series is said to have the sum L.

We will use the compact summation notation

$$\sum_{n=1}^{\infty}a_n=a_1+a_2+\dots+a_n+\dots.$$

Here, n will be referred to as the index and it may start at values other than $n=1$.

1.5 Geometric Series

INFINITE SERIES OCCUR often in mathematics and physics. In this section we look at the special case of a geometric series. A geometric series is of the form

Geometric series are fairly common and will be used throughout the book. You should learn to recognize them and work with them.

$$\sum_{n=0}^{\infty}ar^n=a+ar+ar^2+\dots+ar^n+\dots. \tag{1.5}$$

Here a is the first term and r is called the ratio. It is called the ratio because the ratio of two consecutive terms in the sum is r.

Example 1.6. *For example,*

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots$$

is an example of a geometric series. We can write this using summation notation,

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots=\sum_{n=0}^{\infty}1\left(\frac{1}{2}\right)^{n}.$$

Thus, $a=1$ is the first term and $r=\frac{1}{2}$ is the common ratio of successive terms. Next, we seek the sum of this infinite series, if it exists.

The sum of a geometric series, when it exists, can easily be determined. We consider the nth partial sum:

$$s_n = a + ar + \ldots + ar^{n-2} + ar^{n-1}. \tag{1.6}$$

Now, multiply this equation by r.

$$rs_n = ar + ar^2 + \ldots + ar^{n-1} + ar^n. \tag{1.7}$$

Subtracting these two equations, while noting the many cancelations, we have

$$\begin{aligned}(1-r)s_n &= (a + ar + \ldots + ar^{n-2} + ar^{n-1}) \\ &\quad -(ar + ar^2 + \ldots + ar^{n-1} + ar^n) \\ &= a - ar^n \\ &= a(1-r^n). \end{aligned} \tag{1.8}$$

Thus, the nth partial sums can be written in the compact form

$$s_n = \frac{a(1-r^n)}{1-r}. \tag{1.9}$$

The sum, if it exists, is given by $S=\lim_{n\to\infty} s_n$. Letting n get large in the partial sum (1.9), we need only evaluate $\lim_{n\to\infty} r^n$. From the special limits in the Appendix we know that this limit is zero for $|r|<1$. Thus, we have

Geometric Series

The sum of the geometric series exists for $|r|<1$ and is given by

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}, \qquad |r|<1. \tag{1.10}$$

The reader should verify that the geometric series diverges for all other values of r. Namely, consider what happens for the separate cases $|r|>1$, $r=1$ and $r=-1$.

Next, we present a few typical examples of geometric series.

Example 1.7. $\sum_{n=0}^{\infty} \frac{1}{2^n}$

In this case we have that $a = 1$ *and* $r = \frac{1}{2}$. *Therefore, this infinite series converges and the sum is*

$$S = \frac{1}{1 - \frac{1}{2}} = 2.$$

Example 1.8. $\sum_{k=2}^{\infty} \frac{4}{3^k}$

In this example we first note that the first term occurs for $k = 2$. *It sometimes helps to write out the terms of the series,*

$$\sum_{k=2}^{\infty} \frac{4}{3^k} = \frac{4}{3^2} + \frac{4}{3^3} + \frac{4}{3^4} + \frac{4}{3^5} + \ldots.$$

Looking at the series, we see that $a = \frac{4}{9}$ *and* $r = \frac{1}{3}$. *Since* $|r|<1$, *the geometric series converges. So, the sum of the series is given by*

$$S = \frac{\frac{4}{9}}{1 - \frac{1}{3}} = \frac{2}{3}.$$

Example 1.9. $\sum_{n=1}^{\infty} \left(\frac{3}{2^n} - \frac{2}{5^n}\right)$

Finally, in this case we do not have a geometric series, but we do have the difference of two geometric series. Of course, we need to be careful whenever rearranging infinite series. In this case it is allowed [6]. *Thus, we have*

$$\sum_{n=1}^{\infty} \left(\frac{3}{2^n} - \frac{2}{5^n}\right) = \sum_{n=1}^{\infty} \frac{3}{2^n} - \sum_{n=1}^{\infty} \frac{2}{5^n}.$$

Now we can add both geometric series to obtain

$$\sum_{n=1}^{\infty} \left(\frac{3}{2^n} - \frac{2}{5^n}\right) = \frac{\frac{3}{2}}{1 - \frac{1}{2}} - \frac{\frac{2}{5}}{1 - \frac{1}{5}} = 3 - \frac{1}{2} = \frac{5}{2}.$$

[6] A rearrangement of terms in an infinite series is allowed when the series is absolutely convergent. (See the Appendix.)

Geometric series are important because they are easily recognized and summed. Other series which can be summed include special cases of Taylor series and *telescoping series.* Next, we show an example of a telescoping series.

Example 1.10. $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ *The first few terms of this series are*

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \ldots.$$

It does not appear that we can sum this infinite series. However, if we used the partial fraction expansion

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1},$$

then we find the kth partial sum can be written as

$$\begin{aligned} s_k &= \sum_{n=1}^{k} \frac{1}{n(n+1)} \\ &= \sum_{n=1}^{k} \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{k} - \frac{1}{k+1}\right). \end{aligned} \tag{1.11}$$

We see that there are many cancelations of neighboring terms, leading to the series collapsing (like a retractable telescope) to something manageable:

$$s_k = 1 - \frac{1}{k+1}.$$

Taking the limit as $k \to \infty$*, we find* $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$

1.6 Convergence Tests

GIVEN A GENERAL INFINITE SERIES, it would be nice to know if it converges, or not. Often, we are only interested in the convergence and not the actual sum as it is often difficult to determine the sum even if the series converges. In this section we will review some of the standard tests for convergence, which you should have seen in Calculus II.

First, we have the nth Term Divergence Test. This is motivated by two examples:

1. $\sum_{n=0}^{\infty} 2^n = 1 + 2 + 4 + 8 + \ldots.$
2. $\sum_{n=1}^{\infty} \frac{n+1}{n} = \frac{2}{1} + \frac{3}{2} + \frac{4}{3} + \ldots.$

In the first example, it is easy to see that each term is getting larger and larger, and thus the partial sums will grow without bound. In the second case, each term is bigger than one. Thus, the series will be bigger than adding the same number of ones as there are terms in the sum. Obviously, this series will also diverge.

The nth Term Divergence Test.

This leads to the nth Term Divergence Test:

Theorem 1.4. *If* $\lim a_n \neq 0$ *or if this limit does not exist, then* $\sum_n a_n$ *diverges.*

This theorem does not imply that just because the terms are getting smaller, the series will converge. Otherwise, we would not need any other convergence theorems.

For the next theorems, we will assume that the series has nonnegative terms.

The Comparison Test.

Comparison Test

The series $\sum a_n$ converges if there is a convergent series $\sum c_n$ such that $a_n \leq c_n$ for all $n > N$ for some N. The series $\sum a_n$ diverges if there is a divergent series $\sum d_n$ such that $d_n \leq a_n$ for all $n > N$ for some N.

This is easily seen. In the first case, we have

$$a_n \leq c_n, \forall n > N.$$

Summing both sides of the inequality, we have

$$\sum_n a_n \leq \sum_n c_n.$$

If $\sum c_n$ converges, $\sum c_n < \infty$, the $\sum a_n$ converges as well. A similar argument applies for the divergent series case.

For this test, one has to dream up a second series for comparison. Typically, this requires some experience with convergent series. Often it is better to use other tests first, if possible.

Example 1.11. *Determine if $\sum_{n=0}^{\infty} \frac{1}{3^n}$ converges using the Comparison Test.*

We already know that $\sum_{n=0}^{\infty} \frac{1}{2^n}$ converges. So, we compare these two series. In the above notation, we have $a_n = \frac{1}{3^n}$ and $c_n = \frac{1}{2^n}$. Because $\frac{1}{3^n} \leq \frac{1}{2^n}$ for $n \geq 0$ and $\sum_{n=0}^{\infty} \frac{1}{2^n}$ converges, then $\sum_{n=0}^{\infty} \frac{1}{3^n}$ converges by the Comparison Test.

Limit Comparison Test

The Limit Comparison Test.

If $\lim_{n\to\infty} \frac{a_n}{b_n}$ is finite, then $\sum a_n$ and $\sum b_n$ converge together or diverge together.

Example 1.12. *Determine if $\sum_{n=1}^{\infty} \frac{2n+1}{(n+1)^2}$ converges using the Limit Comparison Test.*

In order to establish the convergence, or divergence, of this series, we look to see how the terms, $a_n = \frac{2n+1}{(n+1)^2}$, behave for large n. As n gets large, the numerator behaves like $2n$ and the denominator behaves like n^2. Therefore, a_n behaves like $\frac{2n}{n^2} = \frac{2}{n}$. The factor of 2 does not really matter.

This leads us to compare the infinite series $\sum_{n=1}^{\infty} \frac{2n+1}{(n+1)^2}$ with the series $\sum_{n=1}^{\infty} \frac{1}{n}$. Then,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{2n^2+n}{(n+1)^2} = 2.$$

We can conclude that these two series both converge or both diverge.

If we knew the behavior of the second series, then we could finish the problem. Using the next test, we will prove that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. Therefore, $\sum_{n=1}^{\infty} \frac{2n+1}{(n+1)^2}$ also diverges by the Limit Comparison Test. Another example of this test is given in Example 1.14.

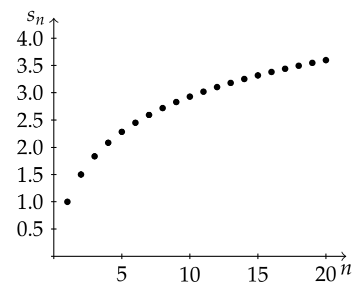


Figure 1.7: Plot of the partial sums of the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$.

Integral Test

The Integral Test.

Consider the infinite series $\sum_{n=1}^{\infty} a_n$, where $a_n = f(n)$. Then, $\sum_{n=1}^{\infty} a_n$ and $\int_1^{\infty} f(x)\,dx$ both converge or both diverge. Here we mean that the integral converges or diverges as an improper integral.

Example 1.13. *Does the harmonic series, $\sum_{n=1}^{\infty} \frac{1}{n}$, converge?*

We are interested in the convergence or divergence of the infinite series $\sum_{n=1}^{\infty} \frac{1}{n}$ that we saw in the Limit Comparison Test example. This infinite series is famous and is called the harmonic series. The plot of the partial sums is given in Figure 1.7. It appears that the series could possibly converge or diverge. It is hard to tell graphically.

In this case we can use the Integral Test. In Figure 1.8 we plot $f(x) = \frac{1}{x}$ and at each integer n we plot a box from n to $n+1$ of height $\frac{1}{n}$. We can see from the figure that the total area of the boxes is greater than the area under the curve. Because the area of each box is $\frac{1}{n}$, then we have that

$$\int_1^{\infty} \frac{dx}{x} < \sum_{n=1}^{\infty} \frac{1}{n}.$$

But, we can compute the integral

$$\int_1^{\infty} \frac{dx}{x} = \lim_{x\to\infty} (\ln x) = \infty.$$

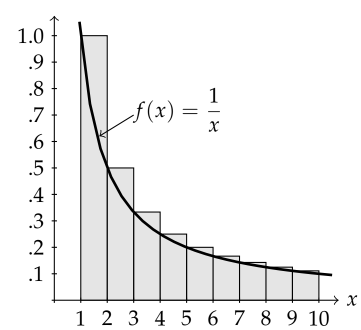


Figure 1.8: Plot of $f(x) = \frac{1}{x}$ and boxes of height $\frac{1}{n}$ and width 1.

Thus, the integral diverges. However, the infinite series is larger than this! So, the harmonic series diverges by the Integral Test.

The Integral Test provides us with the convergence behavior for a class of infinite series called a p-series . These series are of the form $\sum_{n=1}^{\infty} \frac{1}{n^p}$. Recalling that the improper integrals $\int_1^\infty \frac{dx}{x^p}$ converge for $p > 1$ and diverge otherwise, we have the p-test :

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converges for } p > 1$$

p-series and p-test.

and diverges otherwise.

Example 1.14. *Does the series* $\sum_{n=1}^{\infty} \frac{n+1}{n^3-2}$ *converge?*

We first note that as n gets large, the general term behaves like $\frac{1}{n^2}$ *since the numerator behaves like n and the denominator behaves like* n^3*. So, we expect that this series behaves like the series* $\sum_{n=1}^{\infty} \frac{1}{n^2}$*. Thus, by the Limit Comparison Test,*

$$\lim_{n\to\infty} \frac{n+1}{n^3-2}(n^2) = 1.$$

These series both converge, or both diverge. However, we know that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *converges by the p-test since* $p = 2$*. Therefore, the original series converges.*

The Ratio Test.

Ratio Test

Consider the series $\sum_{n=1}^{\infty} a_n$ for $a_n > 0$. Let

$$\rho = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

Then, the behavior of the infinite series can be determined from the conditions

$$\begin{array}{ll} \rho < 1, & \text{converges} \\ \rho > 1, & \text{diverges.} \end{array}$$

Example 1.15. *Use the Ratio Test to determine the convergence of* $\sum_{n=1}^{\infty} \frac{n^{10}}{10^n}$*.*

We compute

$$\begin{aligned} \rho &= \lim_{n\to\infty} \frac{a_{n+1}}{a_n} \\ &= \lim_{n\to\infty} \frac{(n+1)^{10}}{n^{10}} \frac{10^n}{10^{n+1}} \\ &= \lim_{n\to\infty} \left(1+\frac{1}{n}\right)^{10} \frac{1}{10} \\ &= \frac{1}{10} < 1. \end{aligned} \tag{1.12}$$

Therefore, the series is said to converge by the Ratio Test.

Example 1.16. *Use the Ratio Test to determine the convergence of* $\sum_{n=1}^{\infty} \frac{3^n}{n!}$*.*

In this case we make use of the fact that[7] $(n+1)! = (n+1)n!$*. We compute*

[7] Note that the Ratio Test works when factorials are involved because using $(n+1)! = (n+1)n!$ helps to reduce the needed ratios into something manageable.

$$\begin{aligned}\rho &= \lim_{n\to\infty}\frac{a_{n+1}}{a_n}\\ &= \lim_{n\to\infty}\frac{3^{n+1}}{3^n}\frac{n!}{(n+1)!}\\ &= \lim_{n\to\infty}\frac{3}{n+1}=0<1.\end{aligned} \tag{1.13}$$

This series also converges by the Ratio Test.

nth Root Test

The nth Root Test.

Consider the series $\sum_{n=1}^{\infty} a_n$ for $a_n > 0$. Let

$$\rho = \lim_{n\to\infty} a_n{}^{1/n}.$$

Then the behavior of the infinite series can be determined using

$$\begin{aligned}\rho < 1, &\quad \text{converges}\\ \rho > 1, &\quad \text{diverges.}\end{aligned}$$

Example 1.17. *Use the nth Root Test to determine the convergence of* $\sum_{n=0}^{\infty} e^{-n}$.

We use the nth Root Test: $\lim_{n\to\infty}\sqrt[n]{a_n} = \lim_{n\to\infty} e^{-1} = e^{-1} < 1$. *Thus, this series converges by the nth Root Test.*[8]

[8] Note that the Root Test works when there are no factorials and simple powers are involved. In such cases special limit rules help in the evaluation.

Example 1.18. *Use the nth Root Test to determine the convergence of* $\sum_{n=1}^{\infty}\frac{n^n}{2^{n^2}}$.

This series also converges by the nth Root Test.

$$\lim_{n\to\infty}\sqrt[n]{a_n} = \lim_{n\to\infty}\left(\frac{n^n}{2^{n^2}}\right)^{1/n} = \lim_{n\to\infty}\frac{n}{2^n} = 0 < 1.$$

Absolute and Conditional Convergence

We next turn to series that have both positive and negative terms. We can toss out the signs by taking absolute values of each of the terms. We note that since $a_n \le |a_n|$, we have

$$-\sum_{n=1}^{\infty}|a_n| \le \sum_{n=1}^{\infty} a_n \le \sum_{n=1}^{\infty}|a_n|.$$

If the sum $\sum_{n=1}^{\infty}|a_n|$ converges, then the original series converges. For example, if $\sum_{n=1}^{\infty}|a_n| = S$, then by this inequality, $-S \le \sum_{n=1}^{\infty} a_n \le S$

Convergence of the series $\sum|a_n|$ is useful, because we can use the previous tests to establish convergence of such series. Thus, we say that a series converges absolutely if $\sum_{n=1}^{\infty}|a_n|$ converges. If a series converges, but does not converge absolutely, then it is said to converge conditionally.

Conditional and absolute convergence.

Example 1.19. *Show that the series* $\sum_{n=1}^{\infty}\frac{\cos\pi n}{n^2}$ *converges absolutely.*

This series converges absolutely because $\sum_{n=1}^{\infty}|a_n| = \sum_{n=1}^{\infty}\frac{1}{n^2}$ *is a p-series with* $p = 2$.

Finally, there is one last test that we recall from your introductory calculus class. We consider the alternating series, given by $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$. The convergence of an alternating series is determined from Leibniz's Theorem.[9]
.

[9] Gottfried Wilhelm Leibniz (1646-1716) developed calculus independently of Sir Isaac Newton (1643-1727).

Theorem 1.5. *The series* $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ *converges if*

1. a_n*'s are positive.*
2. $a_n \geq a_{n+1}$ *for all* n*.*
3. $a_n \to 0$.

Convergence of alternating series.

The first condition guarantees that we have alternating signs in the series. The next condition says that the magnitude of the terms gets smaller and the last condition imposes further that the terms approach zero.

Example 1.20. *Establish the type of convergence of the alternating harmonic series,* $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$.

First of all, this series is an alternating series. The a_n*'s in Leibniz's Theorem are given by* $a_n = \frac{1}{n}$*. Condition 2 for this case is*

$$\frac{1}{n} \geq \frac{1}{n+1}.$$

This is certainly true, as condition 2 just means that the terms are not getting bigger as n increases. Finally, condition 3 says that the terms are in fact going to zero as n increases. This is true in this example. Therefore, the alternating harmonic series converges by Leibniz's Theorem.

Note: The alternating harmonic series converges conditionally, since the series of absolute values $\sum_{n=1}^{\infty}\left|\frac{(-1)^{n+1}}{n}\right| = \sum_{n=1}^{\infty}\frac{1}{n}$ *gives the (divergent) harmonic series. So, the alternating harmonic series does not converge absolutely.*

Example 1.21. *Determine the type of convergence of the series* $\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n}$.

$\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n}$ *also passes the conditions of Leibniz's Theorem. It should be clear that the terms of this alternating series are getting smaller and approach zero. Furthermore, this series converges absolutely!*

1.7 *Sequences of Functions*

Our immediate goal is to prepare for studying Fourier series, which are series whose terms are functions. So, in this section we begin to discuss series of functions and the convergence of such series. Once more we will need to resort to the convergence of the sequence of partial sums. This means we really need to start with sequences of functions. A sequence of functions is simply a set of functions $f_n(x)$, $n = 1, 2, \ldots$ defined on a common domain D. A frequently used example will be the sequence of functions $\{1, x, x^2, \ldots\}$, $x \in [-1, 1]$.

Evaluating each sequence of functions at a given value of x, we obtain a sequence of real numbers. As before, we can ask if this sequence converges. Doing this for each point in the domain D, we then ask if the resulting collection of limits defines a function on D. More formally, this leads us to the idea of pointwise convergence.

Pointwise convergence of sequences of functions.

A sequence of functions f_n converges pointwise on D to a limit g if

$$\lim_{n\to\infty} f_n(x) = g(x)$$

for each $x \in D$. More formally, we write that

$$\lim_{n\to\infty} f_n = g \text{ (pointwise on } D)$$

if given $x \in D$ and $\epsilon > 0$, there exists an integer N such that

The symbol $\forall$ means "for all."

$$|f_n(x) - g(x)| < \epsilon, \quad \forall n \geq N.$$

Example 1.22. *Consider the sequence of functions*

$$f_n(x) = \frac{1}{1+nx}, \quad |x| < \infty, \quad n = 1, 2, 3, \ldots.$$

The limits depends on the value of x. We consider two cases, $x = 0$ and $x \neq 0$.

1. $x = 0$. *Here* $\lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} 1 = 1$.
2. $x \neq 0$. *Here* $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{1}{1+nx} = 0$.

Therefore, we can say that $f_n \to g$ pointwise for $|x| < \infty$, where

$$g(x) = \begin{cases} 0, & x \neq 0, \\ 1, & x = 0. \end{cases} \tag{1.14}$$

We also note that for a sequence that converges pointwise, N generally depends on both x and ϵ, $N(x, \epsilon)$. We will show this by example.

Example 1.23. *Consider the functions $f_n(x) = x^n, x \in [0,1], n = 1, 2, \ldots.$*

We recall that the definition for pointwise convergence suggests that for each x we seek an N such that $|f_n(x) - g(x)| < \epsilon, \forall n \geq N$. This is not at first easy to see. So, we will provide some simple examples showing how N can depend on both x and ϵ.

1. $x = 0$. *Here we have $f_n(0) = 0$ for all n. So, given $\epsilon > 0$ we seek an N such that $|f_n(0) - 0| < \epsilon, \forall n \geq N$. Inserting $f_n(0) = 0$, we have $0 < \epsilon$. Since this is true for all n, we can pick $N = 1$.*
2. $x = \frac{1}{2}$. *In this case we have $f_n(\frac{1}{2}) = \frac{1}{2^n}$, for $n = 1, 2, \ldots.$*
   As n gets large, $f_n \to 0$. So, given $\epsilon > 0$, we seek N such that

   $$\left|\frac{1}{2^n} - 0\right| < \epsilon, \quad \forall n \geq N.$$

   This result means that $\frac{1}{2^n} < \epsilon$.
   Solving the inequality for n, we have

   $$n > N \geq -\frac{\ln \epsilon}{\ln 2}.$$

   We choose $N \geq -\frac{\ln \epsilon}{\ln 2}$. Thus, our choice of N depends on ϵ.
   For, $\epsilon = 0.1$, this gives

   $$N \geq -\frac{\ln 0.1}{\ln 2} = \frac{\ln 10}{\ln 2} \approx 3.32.$$

   So, we pick $N = 4$ and we have $n > N = 4$.

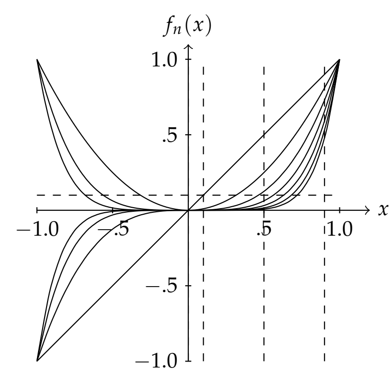


Figure 1.9: Plot of $f_n(x) = x^n$ showing how N depends on $x = 0, 0.1, 0.5, 0.9$ (the vertical lines) and $\epsilon = 0.1$ (the horizontal line). Look at the intersection of a given vertical line with the horizontal line and determine N from the number of curves not under the intersection point.

3. $x = \frac{1}{10}$. *This case can be examined like the last example.*
   We have $f_n(\frac{1}{10}) = \frac{1}{10^n}$, *for* $n = 1, 2, \ldots$. *This leads to* $N \geq -\frac{\ln \epsilon}{\ln 10}$ *For* $\epsilon = 0.1$, *this gives* $N \geq 1$, *or* $n > 1$.
4. $x = \frac{9}{10}$. *This case can be examined like the last two examples. We have* $f_n(\frac{9}{10}) = \left(\frac{9}{10}\right)^n$, *for* $n = 1, 2, \ldots$. *So given an* $\epsilon > 0$, *we seek an* N *such that* $\left(\frac{9}{10}\right)^n < \epsilon$ *for all* $n > N$. *Therefore,*

$$n > N \geq \frac{\ln \epsilon}{\ln\left(\frac{9}{10}\right)}.$$

   For $\epsilon = 0.1$, *we have* $N \geq 21.85$, *or* $n > N = 22$.

So, for these cases, we have shown that N can depend on both x and ϵ. *These cases are shown in Figure 1.9.*

There are other questions that can be asked about sequences of functions. Let the sequence of functions f_n be continuous on D. If the sequence of functions converges pointwise to g on D, then we can ask the following.

1. Is g continuous on D?
2. If each f_n is integrable on $[a, b]$, then does

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b g(x)\, dx?$$

3. If each f_n is differentiable at c, then does

$$\lim_{n\to\infty} f_n'(c) = g'(c)?$$

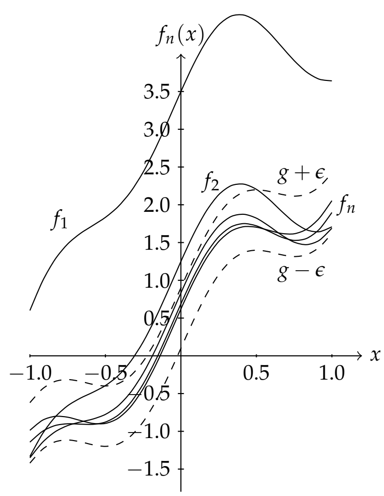


Figure 1.10: For uniform convergence, as n gets large, $f_n(x)$ lies in the band $g(x) - \epsilon, g(x) - \epsilon$.

Uniform convergence of sequences of functions.

It turns out that pointwise convergence is not enough to provide an affirmative answer to any of these questions. Though we will not prove it here, what we will need is uniform convergence.

Consider a sequence of functions $\{f_n(x)\}_{n=1}^\infty$ on D. Let $g(x)$ be defined for $x \in D$. Then the sequence converges uniformly on D, or

$$\lim_{n\to\infty} f_n = g \text{ uniformly on } D,$$

if given $\epsilon > 0$, there exists an N such that

$$|f_n(x) - g(x)| < \epsilon, \quad \forall n \geq N \text{ and } \forall x \in D.$$

This definition almost looks like the definition for pointwise convergence. However, the seemingly subtle difference lies in the fact that N does not depend upon x. The sought N works for all x in the domain. As seen in Figure 1.10 as n gets large, $f_n(x)$ lies in the band $(g(x) - \epsilon, g(x) + \epsilon)$.

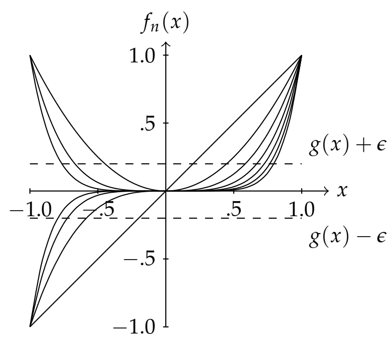


Figure 1.11: Plot of $f_n(x) = x^n$ on [-1,1] for $n = 1, \ldots, 10$ and $g(x) \pm \epsilon$ for $\epsilon = 0.2$.

Example 1.24. *Does the sequence of functions* $f_n(x) = x^n$, *converge uniformly on* $[0, 1]$?

Note that in this case as n gets large, $f_n(x)$ *does not lie in the band* $(g(x) - \epsilon, g(x) + \epsilon)$ *as seen in Figure 1.11. Therefore, this sequence of functions does not converge uniformly on* $[-1, 1]$.

Example 1.25. *Does the sequence of functions* $f_n(x) = \cos(nx)/n^2$ *converge uniformly on* $[-1, 1]$?

For this example, we plot the first several members of the sequence in Figure 1.12. We can see that eventually ($n \geq N$) *members of this sequence do lie inside a band of width* ϵ *about the limit* $g(x) \equiv 0$ *for all values of* x. *Thus, this sequence of functions will converge uniformly to the limit.*

Finally, we should note that if a sequence of functions is uniformly convergent then it converges pointwise. However, the examples should bear out that the converse is not true.

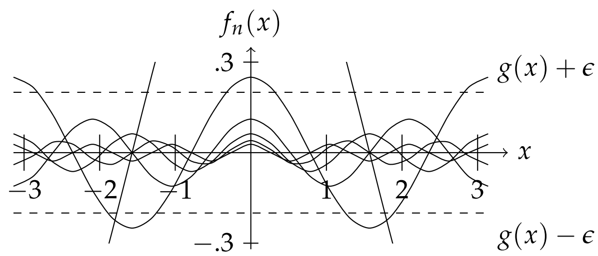


Figure 1.12: Plot of $f_n(x) = \cos(nx)/n^2$ on $[-\pi, \pi]$ for $n = 1 \ldots 10$ and $g(x) \pm \epsilon$ for $\epsilon = 0.2$.

1.8 Infinite Series of Functions

We now turn our attention to infinite series of functions, which will form the basis of our study of Fourier series. An infinite series of functions is given by $\sum_{n=1}^{\infty} f_n(x)$, $x \in D$. Using powers of x, an example of an infinite series of functions might be $\sum_{n=1}^{\infty} x^n$, $x \in [-1, 1]$. In order to investigate the convergence of this series, we could substitute values for x and determine if the resulting series of numbers converges. In general, to investigate the convergence of infinite series of functions, we would consider the Nth partial sums

$$s_N(x) = \sum_{n=1}^{N} f_n(x)$$

and ask if this sequence converges? We will begin to answer this question by defining pointwise and uniform convergence of infinite series of functions.

Pointwise convergence of an infinite series.

The infinite series $\sum f_j(x)$ converges pointwise to $f(x)$ on D if given $x \in D$, and $\epsilon > 0$, there exists an N such that

$$|f(x) - s_n(x)| < \epsilon$$

for all $n > N$.

Uniform convergence of an infinite series.

The infinite series $\sum f_j(x)$ converges uniformly to $f(x)$ on D given $\epsilon > 0$, there exists and N such that

$$|f(x) - s_n(x)| < \epsilon$$

for all $n > N$ and all $x \in D$.

Again, we state without proof the following important properties of uniform convergence for infinite series of functions:

1. Uniform convergence implies pointwise convergence.
2. If f_n is continuous on D, and $\sum_n^\infty f_n$ converges uniformly to f on D, then f is continuous on D.
3. If f_n is continuous on $[a,b] \subset D$, $\sum_n^\infty f_n$ converges uniformly on D to g, and $\int_a^b f_n(x)\,dx$ exists, then
$$\sum_n^\infty \int_a^b f_n(x)\,dx = \int_a^b \sum_n^\infty f_n(x)\,dx = \int_a^b g(x)\,dx.$$
4. If f_n' is continuous on $[a,b] \subset D$, $\sum_n^\infty f_n$ converges pointwise to g on D, and $\sum_n^\infty f_n'$ converges uniformly on D, then
$$\sum_n^\infty f_n'(x) = \frac{d}{dx}\left(\sum_n^\infty f_n(x)\right) = g'(x)$$
for $x \in (a,b)$.

Uniform convergence gives nice properties under some additional conditions, such as being able to integrate, or differentiate, term by term.

Since uniform convergence of series gives so much, like term by term integration and differentiation, we would like to be able to recognize when we have a uniformly convergent series. One test for such convergence is the Weierstraß M-Test[10].

[10] Karl Theodor Wilhelm Weierstraß (1815-1897) was a German mathematician who may be thought of as the father of analysis.

Theorem 1.6. Weierstraß M-Test *Let $\{f_n\}_{n=1}^\infty$ be a sequence of functions on D. If $|f_n(x)| \le M_n$, for $x \in D$ and $\sum_{n=1}^\infty M_n$ converges, then $\sum_{n=1}^\infty f_n$ converges uniformly of D.*

Proof. First, we note that for $x \in D$,
$$\sum_{n=1}^\infty |f_n(x)| \le \sum_{n=1}^\infty M_n.$$
Since by the assumption that $\sum_{n=1}^\infty M_n$ converges, we have that $\sum_{n=1}^\infty f_n$ converges absolutely on D. Therefore, $\sum_{n=1}^\infty f_n$ converges pointwise on D. So, we let $\sum_{n=1}^\infty f_n = g$.

We now want to prove that this convergence is, in fact, uniform. So, given an $\epsilon > 0$, we need to find an N such that
$$\left|g(x) - \sum_{j=1}^n f_j(x)\right| < \epsilon$$
if $n \ge N$ for all $x \in D$. Therefore, for any $x \in D$, we find a bound on $\left|g(x) - \sum_{j=1}^n f_j(x)\right|$:
$$\begin{aligned}
\left|g(x) - \sum_{j=1}^n f_j(x)\right| &= \left|\sum_{j=1}^\infty f_j(x) - \sum_{j=1}^n f_j(x)\right| \\
&= \left|\sum_{j=n+1}^\infty f_j(x)\right| \\
&\le \sum_{j=n+1}^\infty |f_j(x)|, \quad \text{by the triangle inequality} \\
&\le \sum_{j=n+1}^\infty M_j. \qquad (1.15)
\end{aligned}$$

Now, the sum over the M_j's is convergent, so we can choose N such that

$$\sum_{j=n+1}^{\infty} M_j < \epsilon, \quad n \geq N.$$

Combining these results, we have

$$\left| g(x) - \sum_{j=1}^{n} f_j(x) \right| \leq \sum_{j=n+1}^{\infty} M_j < \epsilon$$

for all $n \geq N$ and $x \in D$. Thus, we conclude that the series $\sum f_j$ converges uniformly to g, which we write $\sum f_j \to g$ uniformly on D. □

We now give an example of how to use the Weierstraß M-Test.

Example 1.26. *Show that the series $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$ converges uniformly on $[-\pi, \pi]$.*

Each term of the infinite series is bounded by $\left|\frac{\cos nx}{n^2}\right| = \frac{1}{n^2} \equiv M_n$. We also know that $\sum_{n=1}^{\infty} M_n = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$. Thus, we can conclude that the original series converges uniformly, as it satisfies the conditions of the Weierstraß M-Test.

1.9 *Special Series Expansions*

Examples of infinite series are geometric series, power series, and binomial series. These are discussed more fully in Sections 1.5, 1.10, and 1.11, respectively. These series are defined as follows:

1. The sum of the geometric series exists for $|r| < 1$ and is given by

   The geometric series.

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}, \qquad |r| < 1. \tag{1.16}$$

2. A power series expansion about $x = a$ with coefficient sequence c_n is given by $\sum_{n=0}^{\infty} c_n (x-a)^n$.
3. A Taylor series expansion of $f(x)$ about $x = a$ is the series

   Taylor series expansion.

$$f(x) \sim \sum_{n=0}^{\infty} c_n (x-a)^n, \tag{1.17}$$

   where

$$c_n = \frac{f^{(n)}(a)}{n!}. \tag{1.18}$$

4. A Maclaurin series expansion of $f(x)$ is a Taylor series expansion of $f(x)$ about $x = 0$, or

   Maclaurin series expansion.

$$f(x) \sim \sum_{n=0}^{\infty} c_n x^n, \tag{1.19}$$

   where

$$c_n = \frac{f^{(n)}(0)}{n!}. \tag{1.20}$$

   Some common expansions are provided in Table 1.1.

Table 1.1: Common Mclaurin Series Expansions

Series Expansions You Should Know				
e^x	=	$1+x+\frac{x^2}{2}+\frac{x^3}{3!}+\frac{x^4}{4!}+\dots$	=	$\sum_{n=0}^{\infty}\frac{x^n}{n!}$
$\cos x$	=	$1-\frac{x^2}{2}+\frac{x^4}{4!}-\dots$	=	$\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}$
$\sin x$	=	$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\dots$	=	$\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$
$\cosh x$	=	$1+\frac{x^2}{2}+\frac{x^4}{4!}+\dots$	=	$\sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}$
$\sinh x$	=	$x+\frac{x^3}{3!}+\frac{x^5}{5!}+\dots$	=	$\sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}$
$\frac{1}{1-x}$	=	$1+x+x^2+x^3+\dots$	=	$\sum_{n=0}^{\infty}x^n$
$\frac{1}{1+x}$	=	$1-x+x^2-x^3+\dots$	=	$\sum_{n=0}^{\infty}(-x)^n$
$\tan^{-1}x$	=	$x-\frac{x^3}{3}+\frac{x^5}{5}-\frac{x^7}{7}+\dots$	=	$\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{2n+1}$
$\ln(1+x)$	=	$x-\frac{x^2}{2}+\frac{x^3}{3}-\dots$	=	$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}$

5. The binomial series indexbinomial series is a special Maclaurin series. Namely, it is the series expansion of $(1+x)^p$ for p a nonnegative integer.

We also considered the convergence of power series, $\sum_{n=0}^{\infty} c_n(x-a)^n$. For $x=a$, the series obviously converges. In general, if $\sum_{n=0}^{\infty} c_n(b-a)^n$ converges for $b\neq a$, then $\sum_{n=0}^{\infty} c_n(x-a)^n$ converges absolutely for all x satisfying $|x-a|<|b-a|$.

This leads to the three possibilities:

1. $\sum_{n=0}^{\infty} c_n(x-a)^n$ may only converge at $x=a$.
2. $\sum_{n=0}^{\infty} c_n(x-a)^n$ may converge for all real numbers.
3. $\sum_{n=0}^{\infty} c_n(x-a)^n$ converges for $|x-a|<R$ and diverges for $|x-a|>R$.

Interval and radius of convergence.

The number R is called the radius of convergence of the power series and $(a-R,a+R)$ is called the interval of convergence. Convergence at the endpoints of this interval has to be tested for each power series.

The binomial expansion.

Finally, we have the special case of the general binomial expansion for $(1+x)^p$ for p real. The series is given by the

$$(1+x)^p=\sum_{r=0}^{\infty}\frac{p(p-1)\cdots(p-r+1)}{r!}x^r,\quad |x|<1. \tag{1.21}$$

The binomial approximation.

In practice, one only needs the first few terms for $|x|\ll 1$. Then,

$$(1+x)^p\approx 1+px+\frac{p(p-1)}{2}x^2|x|\ll 1. \tag{1.22}$$

1.10 Power Series

ANOTHER EXAMPLE OF AN INFINITE SERIES that the student has encountered in previous courses is the power series. Examples of such series are provided by Taylor and Maclaurin series.

Actually, what are now known as Taylor and Maclaurin series were known long before they were named. James Gregory (1638-1675) has been recognized for discovering Taylor series, which were later named after Brook Taylor (1685-1731) . Similarly, Colin Maclaurin (1698-1746) did not actually discover Maclaurin series, but the name was adopted because of his particular use of series.

A power series expansion about $x = a$ with coefficient sequence c_n is given by $\sum_{n=0}^{\infty} c_n(x-a)^n$. For now we will consider all constants to be real numbers with x in some subset of the set of real numbers.

Consider the following expansion about $x = 0$:

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \dots. \tag{1.23}$$

We would like to make sense of such expansions. For what values of x will this infinite series converge? Until now we did not pay much attention to which infinite series might converge. However, this particular series is already familiar to us. It is a geometric series. Note that each term is gotten from the previous one through multiplication by $r = x$. The first term is $a = 1$. So, from Equation (1.10), we have that the sum of the series is given by

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad |x| < 1.$$

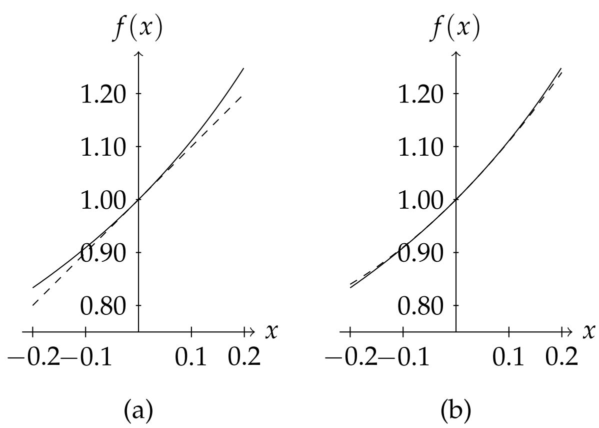


Figure 1.13: (a) Comparison of $\frac{1}{1-x}$ (solid) to $1+x$ (dashed) for $x \in [-0.2, 0.2]$. (b) Comparison of $\frac{1}{1-x}$ (solid) to $1+x+x^2$ (dashed) for $x \in [-0.2, 0.2]$.

In this case we see that the sum, when it exists, is a simple function. In fact, when x is small, we can use this infinite series to provide approximations to the function $(1-x)^{-1}$. If x is small enough, we can write

$$(1-x)^{-1} \approx 1 + x.$$

In Figure 1.13a we see that for small values of x these functions do agree.

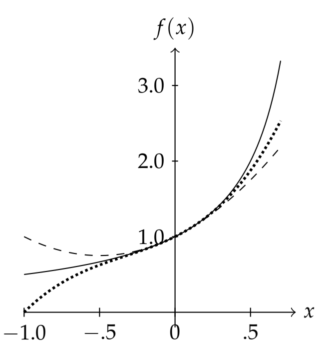


Figure 1.14: Comparison of $\frac{1}{1-x}$ (solid) to $1+x+x^2$ (dashed) and $1+x+x^2+x^3$ (dotted) for $x \in [-1.0, 0.7]$.

Of course, if we want better agreement, we select more terms. In Figure 1.13b we see what happens when we do so. The agreement is much better. But extending the interval, we see in Figure 1.14 that keeping only quadratic terms may not be good enough. Keeping the cubic terms gives better agreement over the interval.

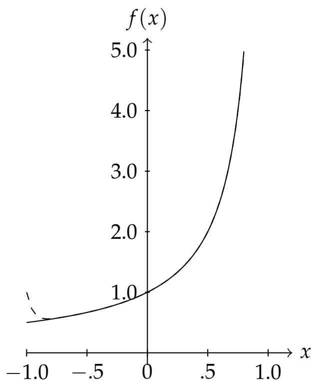


Figure 1.15: Comparison of $\frac{1}{1-x}$ (solid) to $\sum_{n=0}^{20} x^n$ for $x \in [-1,1]$.

Finally, in Figure 1.15 we show the sum of the first 21 terms over the entire interval $[-1,1]$. Note that there are problems with approximations near the endpoints of the interval, $x = \pm 1$.

Such polynomial approximations are called Taylor polynomials. Thus, $T_3(x) = 1 + x + x^2 + x^3$ is the third order Taylor polynomial approximation of $f(x) = \frac{1}{1-x}$.

With this example we have seen how useful a series representation might be for a given function. However, the series representation was a simple geometric series, which we already knew how to sum. Is there a way to begin with a function and then find its series representation? Once we have such a representation, will the series converge to the function with which we started? For what values of x will it converge? These questions can be answered by recalling the definitions of Taylor and Maclaurin series.

Taylor series expansion.

A Taylor series expansion of $f(x)$ about $x = a$ is the series

$$f(x) \sim \sum_{n=0}^{\infty} c_n (x-a)^n, \tag{1.24}$$

where

$$c_n = \frac{f^{(n)}(a)}{n!}. \tag{1.25}$$

Note that we use $\sim$ to indicate that we have yet to determine when the series may converge to the given function. A special class of series are those Taylor series for which the expansion is about $x = 0$. These are called Maclaurin series.

Maclaurin series expansion.

A Maclaurin series expansion of $f(x)$ is a Taylor series expansion of $f(x)$ about $x = 0$, or

$$f(x) \sim \sum_{n=0}^{\infty} c_n x^n, \tag{1.26}$$

where

$$c_n = \frac{f^{(n)}(0)}{n!}. \tag{1.27}$$

Example 1.27. *Expand $f(x) = e^x$ about $x = 0$.*

We begin by creating a table. In order to compute the expansion coefficients, c_n, we will need to perform repeated differentiations of $f(x)$. So, we provide a table for these derivatives. Then, we only need to evaluate the second column at $x = 0$ and divide by $n!$.

n	$f^{(n)}(x)$	$f^{(n)}(0)$	c_n
0	e^x	$e^0 = 1$	$\frac{1}{0!} = 1$
1	e^x	$e^0 = 1$	$\frac{1}{1!} = 1$
2	e^x	$e^0 = 1$	$\frac{1}{2!}$
3	e^x	$e^0 = 1$	$\frac{1}{3!}$

Next, we look at the last column and try to determine a pattern so that we can write down the general term of the series. If there is only a need to get a polynomial approximation, then the first few terms may be sufficient. In this case, the pattern is obvious: $c_n = \frac{1}{n!}$*. So,*

$$e^x \sim \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Example 1.28. *Expand* $f(x) = e^x$ *about* $x = 1$.

Here we seek an expansion of the form $e^x \sim \sum_{n=0}^{\infty} c_n(x-1)^n$*. We could create a table like the last example. In fact, the last column would have values of the form* $\frac{e}{n!}$*. (You should confirm this.) However, we will make use of the Maclaurin series expansion for* e^x *and get the result quicker. Note that* $e^x = e^{x-1+1} = ee^{x-1}$*. Now, apply the known expansion for* e^x *:*

$$e^x \sim e\left(1 + (x-1) + \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3!} + \ldots\right) = \sum_{n=0}^{\infty} \frac{e(x-1)^n}{n!}.$$

Example 1.29. *Expand* $f(x) = \frac{1}{1-x}$ *about* $x = 0$.

This is the example with which we started our discussion. We can set up a table in order to find the Maclaurin series coefficients. We see from the last column of the table that we get back the geometric series (1.23).

n	$f^{(n)}(x)$	$f^{(n)}(0)$	c_n
0	$\frac{1}{1-x}$	1	$\frac{1}{0!} = 1$
1	$\frac{1}{(1-x)^2}$	1	$\frac{1}{1!} = 1$
2	$\frac{2(1)}{(1-x)^3}$	$2(1)$	$\frac{2!}{2!} = 1$
3	$\frac{3(2)(1)}{(1-x)^4}$	$3(2)(1)$	$\frac{3!}{3!} = 1$

So, we have found

$$\frac{1}{1-x} \sim \sum_{n=0}^{\infty} x^n.$$

We can replace $\sim$ by equality if we can determine the range of x-values for which the resulting infinite series converges. We will investigate such convergence shortly.

Series expansions for many elementary functions arise in a variety of applications. Some common expansions are provided in Table 1.2.

We still need to determine the values of x for which a given power series converges. The first five of the above expansions converge for all reals, but the others only converge for $|x| < 1$.

We consider the convergence of $\sum_{n=0}^{\infty} c_n(x-a)^n$. For $x = a$ the series obviously converges. Will it converge for other points? One can prove

Theorem 1.7. *If* $\sum_{n=0}^{\infty} c_n(b-a)^n$ *converges for* $b \neq a$*, then* $\sum_{n=0}^{\infty} c_n(x-a)^n$ *converges absolutely for all* x *satisfying* $|x-a| < |b-a|$.

Table 1.2: Common Mclaurin Series Expansions

Series Expansions You Should Know				
e^x	=	$1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$	=	$\sum_{n=0}^{\infty} \frac{x^n}{n!}$
$\cos x$	=	$1 - \frac{x^2}{2} + \frac{x^4}{4!} - \ldots$	=	$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$
$\sin x$	=	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$	=	$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$
$\cosh x$	=	$1 + \frac{x^2}{2} + \frac{x^4}{4!} + \ldots$	=	$\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$
$\sinh x$	=	$x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$	=	$\sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}$
$\frac{1}{1-x}$	=	$1 + x + x^2 + x^3 + \ldots$	=	$\sum_{n=0}^{\infty} x^n$
$\frac{1}{1+x}$	=	$1 - x + x^2 - x^3 + \ldots$	=	$\sum_{n=0}^{\infty} (-x)^n$
$\tan^{-1} x$	=	$x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots$	=	$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$
$\ln(1+x)$	=	$x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots$	=	$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$

This leads to three possibilities

1. $\sum_{n=0}^{\infty} c_n(x-a)^n$ may only converge at $x = a$.
2. $\sum_{n=0}^{\infty} c_n(x-a)^n$ may converge for all real numbers.
3. $\sum_{n=0}^{\infty} c_n(x-a)^n$ converges for $|x-a| < R$ and diverges for $|x-a| > R$.

Interval and radius of convergence.

The number R is called the radius of convergence of the power series and $(a-R, a+R)$ is called the interval of convergence. Convergence at the endpoints of this interval has to be tested for each power series.

In order to determine the interval of convergence, one needs only note that when a power series converges, it does so absolutely. So, we need only test the convergence of $\sum_{n=0}^{\infty} |c_n(x-a)^n| = \sum_{n=0}^{\infty} |c_n||x-a|^n$. This is easily done using either the ratio test or the nth root test. We first identify the nonnegative terms $a_n = |c_n||x-a|^n$, using the notation from Section 1.4. Then, we apply one of the convergence tests.

For example, the nth Root Test gives the convergence condition for $a_n = |c_n||x-a|^n$,

$$\rho = \lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[n]{|c_n|}|x-a| < 1.$$

Since $|x-a|$ is independent of n,, we can factor it out of the limit and divide the value of the limit to obtain

$$|x-a| < \left(\lim_{n\to\infty} \sqrt[n]{|c_n|}\right)^{-1} \equiv R.$$

Thus, we have found the radius of convergence, R.

Similarly, we can apply the Ratio Test.

$$\rho = \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|}|x-a| < 1.$$

Again, we rewrite this result to determine the radius of convergence:

$$|x-a| < \left(\lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|}\right)^{-1} \equiv R.$$

Example 1.30. *Find the radius of convergence of the series $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.*

Since there is a factorial, we will use the Ratio Test.

$$\rho = \lim_{n\to\infty} \frac{|n!|}{|(n+1)!|}|x| = \lim_{n\to\infty} \frac{1}{n+1}|x| = 0.$$

Since $\rho = 0$, it is independent of $|x|$ and thus the series converges for all x. We also can say that the radius of convergence is infinite.

Example 1.31. *Find the radius of convergence of the series $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$.*

In this example we will use the nth Root Test.

$$\rho = \lim_{n\to\infty} \sqrt[n]{1}|x| = |x| < 1.$$

Thus, we find that we have absolute convergence for $|x| < 1$. Setting $x = 1$ or $x = -1$, we find that the resulting series do not converge. So, the endpoints are not included in the complete interval of convergence.

In this example we could have also used the Ratio Test. Thus,

$$\rho = \lim_{n\to\infty} \frac{1}{1}|x| = |x| < 1.$$

We have obtained the same result as when we used the nth Root Test.

Example 1.32. *Find the radius of convergence of the series $\sum_{n=1}^{\infty} \frac{3^n(x-2)^n}{n}$.*

In this example, we have an expansion about $x = 2$. Using the nth Root Test we find that

$$\rho = \lim_{n\to\infty} \sqrt[n]{\frac{3^n}{n}}|x-2| = 3|x-2| < 1.$$

Solving for $|x-2|$ in this inequality, we find $|x-2| < \frac{1}{3}$. Thus, the radius of convergence is $R = \frac{1}{3}$ and the interval of convergence is $\left(2-\frac{1}{3}, 2+\frac{1}{3}\right) = \left(\frac{5}{3}, \frac{7}{3}\right)$.

As for the endpoints, we first test the point $x = \frac{7}{3}$. The resulting series is $\sum_{n=1}^{\infty} \frac{3^n(\frac{1}{3})^n}{n} = \sum_{n=1}^{\infty} \frac{1}{n}$. This is the harmonic series, and thus it does not converge. Inserting $x = \frac{5}{3}$, we get the alternating harmonic series. This series does converge. So, we have convergence on $[\frac{5}{3}, \frac{7}{3})$. However, it is only conditionally convergent at the left endpoint, $x = \frac{5}{3}$.

Example 1.33. *Find an expansion of $f(x) = \frac{1}{x+2}$ about $x = 1$.*

Instead of explicitly computing the Taylor series expansion for this function, we can make use of an already known function. We first write $f(x)$ as a function of $x-1$, since we are expanding about $x = 1$; i.e., we are seeking a series whose terms are powers of $x-1$.

This expansion is easily done by noting that $\frac{1}{x+2} = \frac{1}{(x-1)+3}$*. Factoring out a* 3*, we can rewrite this expression as a sum of a geometric series. Namely, we use the expansion for*

$$\begin{aligned} g(z) &= \frac{1}{1+z} \\ &= 1 - z + z^2 - z^3 + \ldots . \end{aligned} \tag{1.28}$$

and then we rewrite $f(x)$ *as*

$$\begin{aligned} f(x) &= \frac{1}{x+2} \\ &= \frac{1}{(x-1)+3} \\ &= \frac{1}{3[1+\frac{1}{3}(x-1)]} \\ &= \frac{1}{3}\frac{1}{1+\frac{1}{3}(x-1)}. \end{aligned} \tag{1.29}$$

Note that $f(x) = \frac{1}{3}g(\frac{1}{3}(x-1))$ *for* $g(z) = \frac{1}{1+z}$*. So, the expansion becomes*

$$f(x) = \frac{1}{3}\left[1 - \frac{1}{3}(x-1) + \left(\frac{1}{3}(x-1)\right)^2 - \left(\frac{1}{3}(x-1)\right)^3 + \ldots\right].$$

This can further be simplified as

$$f(x) = \frac{1}{3} - \frac{1}{9}(x-1) + \frac{1}{27}(x-1)^2 - \ldots .$$

Convergence is easily established. The expansion for $g(z)$ *converges for* $|z| < 1$*. So, the expansion for* $f(x)$ *converges for* $|-\frac{1}{3}(x-1)| < 1$*. This implies that* $|x-1| < 3$*. Putting this inequality in interval notation, we have that the power series converges absolutely for* $x \in (-2, 4)$*. Inserting the endpoints, one can show that the series diverges for both* $x = -2$ *and* $x = 4$*. You should verify this!*

Example 1.34. *Prove Euler's Formula:* $e^{i\theta} = \cos\theta + i\sin\theta$*.*

Euler's Formula, $e^{i\theta} = \cos\theta + i\sin\theta$, is an important formula and is used throughout the text.

As a final application, we can derive Euler's Formula ,

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

where $i = \sqrt{-1}$*. We naively use the expansion for* e^x *with* $x = i\theta$*. This leads us to*

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \ldots .$$

Next we note that each term has a power of i*. The sequence of powers of* i *is given as* $\{1, i, -1, -i, 1, i, -1, -i, 1, i, -1, -i, \ldots\}$*. See the pattern? We conclude that*

$$i^n = i^r, \text{ where } r = \text{ remainder after dividing } n \text{ by } 4.$$

This gives

$$e^{i\theta} = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots\right).$$

We recognize the expansions in the parentheses as those for the cosine and sine functions. Thus, we end with Euler's Formula.

We further derive relations from this result, which will be important for our next studies. From Euler's formula we have that for integer n:

$$e^{in\theta} = \cos(n\theta) + i\sin(n\theta).$$

We also have

$$e^{in\theta} = \left(e^{i\theta}\right)^n = (\cos\theta + i\sin\theta)^n.$$

Equating these two expressions, we are led to de Moivre's Formula, named after Abraham de Moivre (1667-1754),

de Moivre's Formula.

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta). \tag{1.30}$$

This formula is useful for deriving identities relating powers of sines or cosines to simple functions. For example, if we take $n = 2$ in Equation (1.30), we find

$$\cos 2\theta + i\sin 2\theta = (\cos\theta + i\sin\theta)^2 = \cos^2\theta - \sin^2\theta + 2i\sin\theta\cos\theta.$$

Looking at the real and imaginary parts of this result leads to the well known double angle identities

$$\cos 2\theta = \cos^2\theta - \sin^2\theta, \quad \sin 2\theta = 2\sin\theta\cos\theta.$$

Here we see elegant proofs of well known trigonometric identities.

$$\begin{aligned}\cos 2\theta &= \cos^2\theta - \sin^2\theta, \quad (1.31)\\ \sin 2\theta &= 2\sin\theta\cos\theta,\\ \cos^2\theta &= \frac{1}{2}(1+\cos 2\theta),\\ \sin^2\theta &= \frac{1}{2}(1-\cos 2\theta).\end{aligned}$$

Replacing $\cos^2\theta = 1 - \sin^2\theta$ or $\sin^2\theta = 1 - \cos^2\theta$ leads to the half angle formulae:

$$\cos^2\theta = \frac{1}{2}(1+\cos 2\theta), \quad \sin^2\theta = \frac{1}{2}(1-\cos 2\theta).$$

We can also use Euler's Formula to write sines and cosines in terms of complex exponentials. We first note that due to the fact that the cosine is an even function and the sine is an odd function, we have

Trigonometric functions can be written in terms of complex exponentials:

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2},$$

$$\sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

$$e^{-i\theta} = \cos\theta - i\sin\theta.$$

Combining this with Euler's Formula, we have that

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}, \quad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

We finally note that there is a simple relationship between hyperbolic functions and trigonometric functions. Recall that

Hyperbolic functions and trigonometric functions are intimately related.

$$\cos(ix) = \cosh x,$$

$$\sin(ix) = -i\sinh x.$$

$$\cosh x = \frac{e^x + e^{-x}}{2}.$$

If we let $x = i\theta$, then we have that $\cosh(i\theta) = \cos\theta$ and $\cos(ix) = \cosh x$. Similarly, we can show that $\sinh(i\theta) = i\sin\theta$ and $\sin(ix) = -i\sinh x$.

1.11 Binomial Series

Another series expansion which occurs often in examples and applications is the binomial expansion. This is simply the expansion of the expression $(a+b)^p$ in powers of a and b. We will investigate this expansion first for nonnegative integer powers p and then derive the expansion for other values of p. While the binomial expansion can be obtained using Taylor series, we will provide a more intuitive derivation to show that

The binomial expansion is a special series expansion used to approximate expressions of the form $(a+b)^p$ for $b \ll a$, or $(1+x)^p$ for $|x| \ll 1$.

$$(a+b)^n = \sum_{r=0}^{n} C_r^n a^{n-r} b^r, \tag{1.32}$$

where the C_r^n are called the *binomial coefficients*.

Lets list some of the common expansions for nonnegative integer powers.

$$\begin{aligned} (a+b)^0 &= 1 \\ (a+b)^1 &= a+b \\ (a+b)^2 &= a^2+2ab+b^2 \\ (a+b)^3 &= a^3+3a^2b+3ab^2+b^3 \\ (a+b)^4 &= a^4+4a^3b+6a^2b^2+4ab^3+b^4 \\ &\cdots \end{aligned} \tag{1.33}$$

We now look at the patterns of the terms in the expansions. First, we note that each term consists of a product of a power of a and a power of b. The powers of a are decreasing from n to 0 in the expansion of $(a+b)^n$. Similarly, the powers of b increase from 0 to n. The sums of the exponents in each term is n. So, we can write the $(k+1)$st term in the expansion as $a^{n-k}b^k$. For example, in the expansion of $(a+b)^{51}$ the 6th term is $a^{51-5}b^5 = a^{46}b^5$. However, we do not yet know the numerical coefficients in the expansion.

Let's list the coefficients for the above expansions.

$$\begin{array}{lccccccccc} n=0: & & & & & 1 & & & & \\ n=1: & & & & 1 & & 1 & & & \\ n=2: & & & 1 & & 2 & & 1 & & \\ n=3: & & 1 & & 3 & & 3 & & 1 & \\ n=4: & 1 & & 4 & & 6 & & 4 & & 1 \end{array} \tag{1.34}$$

This pattern is the famous Pascal's triangle.[11] There are many interesting features of this triangle. But we will first ask how each row can be generated.

[11] Pascal's triangle is named after Blaise Pascal (1623-1662). While such configurations of numbers were known earlier in history, Pascal published them and applied them to probability theory.

Pascal's triangle has many unusual properties and a variety of uses:

- Horizontal rows add to powers of 2.
- The horizontal rows are powers of 11 (1, 11, 121, 1331, etc.).
- Adding any two successive numbers in the diagonal 1-3-6-10-15-21-28... results in a perfect square.
- When the first number to the right of the 1 in any row is a prime number, all numbers in that row are divisible by that prime number. The reader can readily check this for the $n=5$ and $n=7$ rows.
- Sums along certain diagonals leads to the Fibonacci sequence. These diagonals are parallel to the line connecting the first 1 for $n=3$ row and the 2 in the $n=2$ row.

We see that each row begins and ends with a one. The second term and next to last term have a coefficient of n. Next we note that consecutive pairs in each row can be added to obtain entries in the next row. For example, we have for rows $n=2$ and $n=3$ that $1+2=3$ and $2+1=3$:

$$\begin{array}{lccccccc} n=2: & & 1 & & 2 & & 1 & \\ & & & \searrow\ \swarrow & & \searrow\ \swarrow & & \\ n=3: & 1 & & 3 & & 3 & & 1 \end{array} \tag{1.35}$$

With this in mind, we can generate the next several rows of our triangle.

$$\begin{array}{lccccccccccccc} n=3: & & & & 1 & & 3 & & 3 & & 1 & & & \\ n=4: & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\ n=5: & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\ n=6: & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \end{array} \tag{1.36}$$

So, we use the numbers in row $n = 4$ to generate entries in row $n = 5$: $1 + 4 = 5$, $4 + 6 = 10$. We then use row $n = 5$ to get row $n = 6$, etc.

Of course, it would take a while to compute each row up to the desired n. Fortunately, there is a simple expression for computing a specific coefficient. Consider the kth term in the expansion of $(a + b)^n$. Let $r = k - 1$, for $k = 1, \ldots, n + 1$. Then this term is of the form $C^n_r a^{n-r} b^r$. We have seen that the coefficients satisfy

$$C^n_r = C^{n-1}_r + C^{n-1}_{r-1}.$$

Actually, the binomial coefficients, C^n_r, have been found to take a simple form,

$$C^n_r = \frac{n!}{(n-r)!r!} \equiv \begin{pmatrix} n \\ r \end{pmatrix}.$$

This is nothing other than the combinatoric symbol for determining how to choose n objects r at a time. In the binomial expansions this makes sense. We have to count the number of ways that we can arrange r products of b with $n - r$ products of a. There are n slots to place the b's. For example, the $r = 2$ case for $n = 4$ involves the six products: *aabb*, *abab*, *abba*, *baab*, *baba*, and *bbaa*. Thus, it is natural to use this notation.

Andreas Freiherr von Ettingshausen (1796-1878) was a German mathematician and physicist who in 1826 introduced the notation $\begin{pmatrix} n \\ r \end{pmatrix}$. However, the binomial coefficients were known by the Hindus centuries beforehand.

So, we have found that

$$(a + b)^n = \sum_{r=0}^{n} \begin{pmatrix} n \\ r \end{pmatrix} a^{n-r} b^r. \tag{1.37}$$

Now consider the geometric series $1 + x + x^2 + \ldots$. We have seen that such this geometric series converges for $|x| < 1$, giving

$$1 + x + x^2 + \ldots = \frac{1}{1-x}.$$

But, $\frac{1}{1-x} = (1 - x)^{-1}$. This is a binomial to a power, but the power is not an integer.

It turns out that the coefficients of such a binomial expansion can be written similar to the form in Equation(1.37). This example suggests that our sum may no longer be finite. So, for p a real number, $a = 1$ and $b = x$, we generalize Equation(1.37) as

$$(1 + x)^p = \sum_{r=0}^{\infty} \begin{pmatrix} p \\ r \end{pmatrix} x^r \tag{1.38}$$

and see if the resulting series makes sense. However, we quickly run into problems with the coefficients in the series.

Consider the coefficient for $r = 1$ in an expansion of $(1+x)^{-1}$. This is given by

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \frac{(-1)!}{(-1-1)!1!} = \frac{(-1)!}{(-2)!1!}.$$

But what is $(-1)!$? By definition, it is

$$(-1)! = (-1)(-2)(-3)\cdots.$$

This product does not seem to exist! But with a little care, we note that

$$\frac{(-1)!}{(-2)!} = \frac{(-1)(-2)!}{(-2)!} = -1.$$

So, we need to be careful not to interpret the combinatorial coefficient literally. There are better ways to write the general binomial expansion. We can write the general coefficient as

$$\begin{aligned} \begin{pmatrix} p \\ r \end{pmatrix} &= \frac{p!}{(p-r)!r!} \\ &= \frac{p(p-1)\cdots(p-r+1)(p-r)!}{(p-r)!r!} \\ &= \frac{p(p-1)\cdots(p-r+1)}{r!}. \end{aligned} \tag{1.39}$$

With this in mind we now state the theorem:

General Binomial Expansion

The general binomial expansion for $(1+x)^p$ is a simple generalization of Equation (1.37). For p real, we have the following binomial series:

$$(1+x)^p = \sum_{r=0}^{\infty} \frac{p(p-1)\cdots(p-r+1)}{r!} x^r, \quad |x| < 1. \tag{1.40}$$

Often in physics we only need the first few terms for the case that $x \ll 1$:

$$(1+x)^p = 1 + px + \frac{p(p-1)}{2}x^2 + O(x^3). \tag{1.41}$$

The factor $\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}$ is important in special relativity. Namely, this is the factor relating differences in time and length measurements by observers moving relative inertial frames. For terrestrial speeds, this gives an appropriate approximation.

Example 1.35. *Approximate* $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ *for* $v \ll c$.

For $v \ll c$ the first approximation is found inserting $v/c = 0$. Thus, one obtains $\gamma = 1$. This is the Newtonian approximation and does not provide enough of an approximation for terrestrial speeds. Thus, we need to expand γ in powers of v/c.

First, we rewrite γ as

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \left[1 - \left(\frac{v}{c}\right)^2\right]^{-1/2}.$$

Using the binomial expansion for $p = -1/2$, we have

$$\gamma \approx 1 + \left(-\frac{1}{2}\right)\left(-\frac{v^2}{c^2}\right) = 1 + \frac{v^2}{2c^2}.$$

Example 1.36. *Time Dilation Example*

The average speed of a large commercial jet airliner is about 500 mph. If you flew for an hour (measured from the ground), then how much younger would you be than if you had not taken the flight, assuming these reference frames obeyed the postulates of special relativity?

This is the problem of time dilation. Let Δt be the elapsed time in a stationary reference frame on the ground and $\Delta\tau$ be that in the frame of the moving plane. Then from the Theory of Special Relativity these are related by

$$\Delta t = \gamma \Delta \tau.$$

The time differences would then be

$$\begin{aligned} \Delta t - \Delta\tau &= (1-\gamma^{-1})\Delta t \\ &= \left(1 - \sqrt{1-\frac{v^2}{c^2}}\right)\Delta t. \end{aligned} \tag{1.42}$$

The plane speed, 500 mph, is roughly 225 m/s and $c = 3.00 \times 10^8$ m/s. Since $V \ll c$, we would need to use the binomial approximation to get a nonzero result.

$$\begin{aligned} \Delta t - \Delta\tau &= \left(1 - \sqrt{1-\frac{v^2}{c^2}}\right)\Delta t \\ &= \left(1 - \left(1 - \frac{v^2}{2c^2} + \dots\right)\right)\Delta t \\ &\approx \frac{v^2}{2c^2}\Delta t \\ &= \frac{(225)^2}{2(3.00\times 10^8)^2}(1\ h) = 1.01\ ns. \end{aligned} \tag{1.43}$$

Thus, you have aged one nanosecond less than if you did not take the flight.

Example 1.37. Small differences in large numbers: *Compute $f(R,h) = \sqrt{R^2+h^2} - R$ for $R = 6378.164$ km and $h = 1.0$ m.*

Inserting these values into a scientific calculator, one finds that

$$f(6378164, 1) = \sqrt{6378164^2 + 1} - 6378164 = 1 \times 10^{-7}\ m.$$

In some calculators one might obtain 0, in other calculators, or computer algebra systems like Maple, one might obtain other answers. What answer do you get and how accurate is your answer?

The problem with this computation is that $R \gg h$. Therefore, the computation of $f(R,h)$ depends on how many digits the computing device can handle. The best way to get an answer is to use the binomial approximation. Writing $h = Rx$, or $x = \frac{h}{R}$, we have

$$\begin{aligned} f(R,h) &= \sqrt{R^2+h^2} - R \\ &= R\sqrt{1+x^2} - R \\ &\simeq R\left[1 + \frac{1}{2}x^2\right] - R \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2}Rx^2 \\ &= \frac{1}{2}\frac{h}{R^2} = 7.83926 \times 10^{-8}\ m. \end{aligned} \tag{1.44}$$

Of course, you should verify how many digits should be kept in reporting the result.

In the next examples, we generalize this example. Such general computations appear in proofs involving general expansions without specific numerical values given.

Example 1.38. *Obtain an approximation to $(a+b)^p$ when a is much larger than b, denoted by $a \gg b$.*

If we neglect b then $(a+b)^p \simeq a^p$. How good of an approximation is this? This is where it would be nice to know the order of the next term in the expansion. Namely, what is the power of b/a of the first neglected term in this expansion?

In order to do this we first divide out a as

$$(a+b)^p = a^p\left(1+\frac{b}{a}\right)^p.$$

Now we have a small parameter, $\frac{b}{a}$. According to what we have seen earlier, we can use the binomial expansion to write

$$\left(1+\frac{b}{a}\right)^n = \sum_{r=0}^{\infty}\left(\begin{array}{c} p \\ r \end{array}\right)\left(\frac{b}{a}\right)^r. \tag{1.45}$$

Thus, we have a sum of terms involving powers of $\frac{b}{a}$. Since $a \gg b$, most of these terms can be neglected. So, we can write

$$\left(1+\frac{b}{a}\right)^p = 1 + p\frac{b}{a} + O\left(\left(\frac{b}{a}\right)^2\right).$$

Here we used $O()$, big-Oh notation, to indicate the size of the first neglected term.

Summarizing, we have

$$\begin{aligned} (a+b)^p &= a^p\left(1+\frac{b}{a}\right)^p \\ &= a^p\left(1+p\frac{b}{a}+O\left(\left(\frac{b}{a}\right)^2\right)\right) \\ &= a^p + pa^p\frac{b}{a} + a^pO\left(\left(\frac{b}{a}\right)^2\right). \end{aligned} \tag{1.46}$$

Therefore, we can approximate $(a+b)^p \simeq a^p + pba^{p-1}$, with an error on the order of b^2a^{p-2}. Note that the order of the error does not include the constant factor from the expansion. We could also use the approximation that $(a+b)^p \simeq a^p$, but it is not typically good enough in applications because the error in this case is of the order ba^{p-1}.

Example 1.39. *Approximate $f(x) = (a+x)^p - a^p$ for $x \ll a$.*

In an earlier example we computed $f(R,h) = \sqrt{R^2+h^2} - R$ for $R = 6378.164$ km and $h = 1.0$ m. We can make use of the binomial expansion to determine the behavior of similar functions in the form $f(x) = (a+x)^p - a^p$. Inserting the binomial expression into $f(x)$, we have as $\frac{x}{a} \to 0$ that

$$\begin{aligned} f(x) &= (a+x)^p - a^p \\ &= a^p\left[\left(1+\frac{x}{a}\right)^p - 1\right] \\ &= a^p\left[\frac{px}{a} + O\left(\left(\frac{x}{a}\right)^2\right)\right] \\ &= O\left(\frac{x}{a}\right) \quad \text{as } \frac{x}{a} \to 0. \end{aligned} \tag{1.47}$$

This result might not be the approximation that we desire. So, we could back up one step in the derivation to write a better approximation as

$$(a+x)^p - a^p = a^{p-1}px + O\left(\left(\frac{x}{a}\right)^2\right) \quad \text{as } \frac{x}{a} \to 0.$$

We now use this approximation to compute $f(R,h) = \sqrt{R^2+h^2} - R$ for $R = 6378.164$ km and $h = 1.0$ m in the earlier example. We let $a = R^2$, $x = 1$ and $p = \frac{1}{2}$. Then, the leading order approximation would be of order

$$O\left(\left(\frac{x}{a}\right)^2\right) = O\left(\left(\frac{1}{6378164^2}\right)^2\right) \sim 2.4 \times 10^{-14}.$$

Thus, we have

$$\sqrt{6378164^2+1} - 6378164 \approx a^{p-1}px$$

where

$$a^{p-1}px = (6378164^2)^{-1/2}(0.5)1 = 7.83926 \times 10^{-8}.$$

This is the same result we had obtained before. However, we have also an estimate of the size of the error and this might be useful in indicating how many digits we should trust in the answer.

1.12 *The Order of Sequences and Functions*

Often we are interested in comparing the rates of convergence of sequences or asymptotic behavior of functions. This is also useful in approximation theory. We begin with the comparison of sequences and introduce *big-Oh* notation. We will then extend this to functions of continuous variables.

Big-Oh notation.

Let $\{a_n\}$ and $\{b_n\}$ be two sequences. Then, if there are numbers N and K (independent of N) such that

$$\left|\frac{a_n}{b_n}\right| < K \quad \text{whenever} \quad n > N,$$

we say that a_n is *of the order of* b_n. We write this as

$$a_n = O(b_n) \quad \text{as} \quad n \to \infty$$

and say a_n is "big O" of b_n.

Example 1.40. *Consider the sequences given by* $a_n = \frac{2n+1}{3n^2+2}$ *and* $b_n = \frac{1}{n}$.

In this case, we consider the ratio

$$\left|\frac{a_n}{b_n}\right| = \left|\frac{\frac{2n+1}{3n^2+2}}{\frac{1}{n}}\right| = \left|\frac{2n^2+n}{3n^2+2}\right|.$$

We want to find a bound on the last expression as n *gets large. We divide the numerator and denominator by* n^2 *and find that*

$$\left|\frac{a_n}{b_n}\right| = \left|\frac{2+1/n}{3+2/n^2}\right|.$$

Further dividing out a $2/3$*, we find*

$$\left|\frac{a_n}{b_n}\right| = \frac{2}{3}\left|\frac{1+1/2n}{1+2/3n^2}\right|.$$

The last expression is largest for $n = 3$*. This gives*

$$\left|\frac{a_n}{b_n}\right| = \frac{2}{3}\left|\frac{1+1/2n}{1+2/3n^2}\right| \le \frac{2}{3}\left|\frac{1+1/6}{1+2/27}\right| = \frac{21}{29}.$$

Considering the function $f(x) = \frac{2x^2+x}{3x^2+2}$, setting $f'(x) = 0$, we find the maximum is actually at $x = \frac{1}{3}(4+\sqrt{22}) \approx 2.897$. Also, inserting the first few integers yields the sequence $\{0.6000, 0.7143, 0.7241, 0.7200, 0.7143, \ldots\}$. In both cases this supports choosing $n = 3$ in the example.

Thus, for $n > 3$*, we have that*

$$\left|\frac{a_n}{b_n}\right| \le \frac{21}{29} < 1 \equiv K.$$

We then conclude from the definition of big-oh that

$$a_n = O(b_n) = O\left(\frac{1}{n}\right).$$

In practice, one is often given a sequence like a_n, but the second simpler sequence needs to be found by looking at the large n behavior of a_n.

Referring to the last example, we are given $a_n = \frac{2n+1}{3n^2+2}$. We look at the large n behavior. The numerator behaves like $2n$ and the denominator behaves like $3n^2$. Thus, $a_n = \frac{2n+1}{3n^2+2} \sim \frac{2n}{3n^2} = \frac{2}{3n}$ for large n. Therefore, we say that $a_n = O(\frac{1}{n})$ for large n. Note that we are only interested in the n-dependence and not the multiplicative constant since $\frac{1}{n}$ and $\frac{2}{3n}$ have the same growth rate.

In a similar way, we can compare functions. We modify our definition of big-Oh for functions of a continuous variable: $f(x)$ is *of the order of* $g(x)$, or $f(x) = O(g(x))$, as $x \to x_0$ if

$$\lim_{x\to x_0}\left|\frac{f(x)}{g(x)}\right| < K$$

for some finite nonzero constant K independent of x_0.

Example 1.41. *Show that*

$$\cos x - 1 + \frac{x^2}{2} = O(x^4) \quad \textit{as } x \to 0.$$

This should be apparent from the Taylor series expansion for $\cos x$,

$$\cos x = 1 - \frac{x^2}{2} + O(x^4) \text{ as } x \to 0.$$

However, we will show that $\cos x - 1 + \frac{x^2}{2}$ *is of the order of* $O(x^4)$ *using the above definition.*

We need to compute

$$\lim_{x\to 0}\left|\frac{\cos x - 1 + \frac{x^2}{2}}{x^4}\right|.$$

The numerator and denominator separately go to zero, so we have an indeterminate form. This suggests that we need to apply L'Hopital's Rule. (See Theorem 1.3.) In fact, we apply it several times to find that

$$\begin{aligned}
\lim_{x\to 0}\left|\frac{\cos x - 1 + \frac{x^2}{2}}{x^4}\right| &= \lim_{x\to 0}\left|\frac{-\sin x + x}{4x^3}\right| \\
&= \lim_{x\to 0}\left|\frac{-\cos x + 1}{12x^2}\right| \\
&= \lim_{x\to 0}\left|\frac{\sin x}{24x}\right| = \frac{1}{24}.
\end{aligned}$$

Thus, for any number $K > \frac{1}{24}$, *we have that*

$$\lim_{x\to 0}\left|\frac{\cos x - 1 + \frac{x^2}{2}}{x^4}\right| < K.$$

We conclude that

$$\cos x - 1 + \frac{x^2}{2} = O(x^4) \quad \text{as } x \to 0.$$

Example 1.42. *Determine the order of* $f(x) = (x^3 - x)^{1/3} - x$ *as* $x \to \infty$. *We can use a binomial expansion to write the first term in powers of* x. *However, since* $x \to \infty$, *we want to write* $f(x)$ *in powers of* $\frac{1}{x}$, *so that we can neglect higher order powers. We can do this by first factoring out the* x^3 :

$$\begin{aligned}
(x^3 - x)^{1/3} - x &= x\left(1 - \frac{1}{x^2}\right)^{1/3} - x \\
&= x\left(1 - \frac{1}{3x^2} + O\left(\frac{1}{x^4}\right)\right) - x \\
&= -\frac{1}{3x} + O\left(\frac{1}{x^3}\right). \qquad (1.48)
\end{aligned}$$

Now we can see from the first term on the right that $(x^3 - x)^{1/3} - x = O\left(\frac{1}{x}\right)$ *as* $x \to \infty$.

Problems

1. For those sequences that converge, find the limit $\lim_{n\to\infty} a_n$.

a. $a_n = \frac{n^2+1}{n^3+1}$.

b. $a_n = \frac{3n+1}{n+2}$.

c. $a_n = \left(\frac{3}{n}\right)^{1/n}$.

d. $a_n = \frac{2n^2+4n^3}{n^3+5\sqrt{2+n^6}}$.

e. $a_n = n\ln\left(1+\frac{1}{n}\right)$.

f. $a_n = n\sin\left(\frac{1}{n}\right)$.

g. $a_n = \frac{(2n+3)!}{(n+1)!}$.

2. Find the sum for each of the series:

a. $\sum_{n=0}^{\infty} \frac{(-1)^n 3}{4^n}$.

b. $\sum_{n=2}^{\infty} \frac{2}{5^n}$.

c. $\sum_{n=0}^{\infty} \left(\frac{5}{2^n} + \frac{1}{3^n}\right)$.

d. $\sum_{n=1}^{\infty} \frac{3}{n(n+3)}$.

3. Determine if the following converge, or diverge, using one of the convergence tests. If the series converges, is it absolute or conditional?

a. $\sum_{n=1}^{\infty} \frac{n+4}{2n^3+1}$.

b. $\sum_{n=1}^{\infty} \frac{\sin n}{n^2}$.

c. $\sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{n^2}$.

d. $\sum_{n=1}^{\infty} (-1)^n \frac{n-1}{2n^2-3}$.

e. $\sum_{n=1}^{\infty} \frac{\ln n}{n}$.

f. $\sum_{n=1}^{\infty} \frac{100^n}{n^{200}}$.

g. $\sum_{n=1}^{\infty} (-1)^n \frac{n}{n+3}$.

h. $\sum_{n=1}^{\infty} (-1)^n \frac{\sqrt{5n}}{n+1}$.

4. Do the following:

a. Compute: $\lim_{n\to\infty} n\ln\left(1-\frac{3}{n}\right)$.

b. Use L'Hopital's Rule to evaluate $L = \lim_{x\to\infty}\left(1-\frac{4}{x}\right)^x$.
[Hint: Consider $\ln L$.]

c. Determine the convergence of $\sum_{n=1}^{\infty} \left(\frac{n}{3n+2}\right)^{n^2}$.

d. Sum the series $\sum_{n=1}^{\infty} \left[\tan^{-1} n - \tan^{-1}(n+1)\right]$ by first writing the Nth partial sum and then computing $\lim_{N\to\infty} s_N$.

5. Consider the sum $\sum_{n=1}^{\infty} \frac{1}{(n+2)(n+1)}$.

a. Use an appropriate convergence test to show that this series converges.

b. Verify that

$$\sum_{n=1}^{\infty} \frac{1}{(n+2)(n+1)} = \sum_{n=1}^{\infty} \left(\frac{n+1}{n+2} - \frac{n}{n+1} \right).$$

c. Find the nth partial sum of the series $\sum_{n=1}^{\infty} \left(\frac{n+1}{n+2} - \frac{n}{n+1} \right)$ and use it to determine the sum of the resulting *telescoping* series.

6. Recall that the alternating harmonic series converges conditionally.

a. From the Taylor series expansion for $f(x) = \ln(1+x)$, inserting $x = 1$ gives the alternating harmonic series. What is the sum of the alternating harmonic series?

b Since the alternating harmonic series does not converge absolutely, then a rearrangement of the terms in the series will result in series whose sums vary. One such rearrangement in alternating p positive terms and n negative terms leads to the following sum[12]:

$$\begin{aligned} \frac{1}{2} \ln \frac{4p}{n} &= \underbrace{\left(1 + \frac{1}{3} + \cdots + \frac{1}{2p-1}\right)}_{p \text{ terms}} - \underbrace{\left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right)}_{n \text{ terms}} \\ &\quad + \underbrace{\left(\frac{1}{2p+1} + \cdots + \frac{1}{4p-1}\right)}_{p \text{ terms}} - \underbrace{\left(\frac{1}{2n+2} + \cdots + \frac{1}{4n}\right)}_{n \text{ terms}} + \cdots. \end{aligned}$$

[12] This is discussed by Lawrence H. Riddle in the *Kenyon Math. Quarterly*, 1(2), 6-21.

Find rearrangements of the alternating harmonic series to give the following sums; i.e., determine p and n for the given expression and write down the above series explicitly; i.e, determine p and n leading to the following sums.

i. $\frac{5}{2} \ln 2$.

ii. $\ln 8$.

iii. 0.

iv. A sum that is close to π.

7. Determine the radius and interval of convergence of the following infinite series:

a. $\sum_{n=1}^{\infty} (-1)^n \frac{(x-1)^n}{n}$.

b. $\sum_{n=1}^{\infty} \frac{x^n}{2^n n!}$.

c. $\sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{x}{5}\right)^n$

d. $\sum_{n=1}^{\infty} (-1)^n \frac{x^n}{\sqrt{n}}$.

8. Find the Taylor series centered at $x = a$ and its corresponding radius of convergence for the given function. In most cases, you need not employ the direct method of computation of the Taylor coefficients.

a. $f(x) = \sinh x$, $a = 0$.
b. $f(x) = \sqrt{1+x}$, $a = 0$.
c. $f(x) = xe^x$, $a = 1$.
d. $f(x) = \frac{x-1}{2+x}$, $a = 1$.

9. Test for pointwise and uniform convergence on the given set. [The Weierstraß M-Test might be helpful.]

a. $f(x) = \sum_{n=1}^{\infty} \frac{\ln nx}{n^2}$, $x \in [1,2]$.
b. $f(x) = \sum_{n=1}^{\infty} \frac{1}{3^n} \cos \frac{x}{2^n}$ on R.

10. Consider Gregory's expansion

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}.$$

a. Derive Gregory's expansion using the definition

$$\tan^{-1} x = \int_0^x \frac{dt}{1+t^2},$$

expanding the integrand in a Maclaurin series, and integrating the resulting series term by term.

b. From this result, derive Gregory's series for π by inserting an appropriate value for x in the series expansion for $\tan^{-1} x$.

11. Use deMoivre's Theorem to write $\sin^3\theta$ in terms of $\sin\theta$ and $\sin 3\theta$. [Hint: Focus on the imaginary part of $e^{3i\theta}$.]

12. Evaluate the following expressions at the given point. Use your calculator and your computer (such as Maple). Then use series expansions to find an approximation to the value of the expression to as many places as you trust.

a. $\frac{1}{\sqrt{1+x^3}} - \cos x^2$ at $x = 0.015$.
b. $\ln \sqrt{\frac{1+x}{1-x}} - \tan x$ at $x = 0.0015$.
c. $f(x) = \frac{1}{\sqrt{1+2x^2}} - 1 + x^2$ at $x = 5.00 \times 10^{-3}$.
d. $f(R,h) = R - \sqrt{R^2+h^2}$ for $R = 1.374 \times 10^3$ km and $h = 1.00$ m.
e. $f(x) = 1 - \frac{1}{\sqrt{1-x}}$ for $x = 2.5 \times 10^{-13}$.

13. Determine the order, $O(x^p)$, of the following functions. You may need to use series expansions in powers of x when $x \to 0$, or series expansions in powers of $1/x$ when $x \to \infty$.

a. $\sqrt{x(1-x)}$ as $x \to 0$.
b. $\frac{x^{5/4}}{1-\cos x}$ as $x \to 0$.
c. $\frac{x}{x^2-1}$ as $x \to \infty$.
d. $\sqrt{x^2+x} - x$ as $x \to \infty$.

2

Fourier Trigonometric Series

"Profound study of nature is the most fertile source of mathematical discoveries."
Joseph Fourier (1768-1830)

2.1 Introduction to Fourier Series

We will now turn to the study of trigonometric series. You have seen that functions have series representations as expansions in powers of x, or $x - a$, in the form of Maclaurin and Taylor series. Recall that the Taylor series expansion is given by

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n,$$

where the expansion coefficients are determined as

$$c_n = \frac{f^{(n)}(a)}{n!}.$$

From the study of the heat equation and wave equation, Fourier showed that there are infinite series expansions over other functions, such as sine functions. We now turn to such expansions and in the next chapter we will find out that expansions over special sets of functions are not uncommon in physics. But, first we turn to Fourier trigonometric series.

We will begin with the study of the Fourier trigonometric series expansion

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

We will find expressions useful for determining the Fourier coefficients $\{a_n, b_n\}$ given a function $f(x)$ defined on $[-L, L]$. We will also see if the resulting infinite series reproduces $f(x)$. However, we first begin with some basic ideas involving simple sums of sinusoidal functions.

There is a natural appearance of such sums over sinusoidal functions in music. A pure note can be represented as

$$y(t) = A\sin(2\pi f t),$$

As noted in the Introduction, Joseph Fourier (1768-1830) and others studied trigonometric series solutions of the heat and wave equations.

The temperature, $u(x,t)$, of a one dimensional rod of length L satisfies the heat equation,

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}.$$

The general solution, which satisfies the conditions $u(0,t) = 0$ and $u(L,t) = 0$, is given by

$$u(x,t) = \sum_{n=0}^{\infty} b_n \sin\frac{n\pi x}{L} e^{-n^2\pi^2 kt/L^2}.$$

If the initial temperature is given by $u(x,0) = f(x)$, one has to satisfy the condition

$$f(x) = \sum_{n=0}^{\infty} b_n \sin\frac{n\pi x}{L}.$$

The height, $u(x,t)$, of a one dimensional vibrating string of length L satisfies the wave equation,

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}.$$

The general solution, which satisfies the fixed ends $u(0,t) = 0$ and $u(L,t) = 0$, is given by

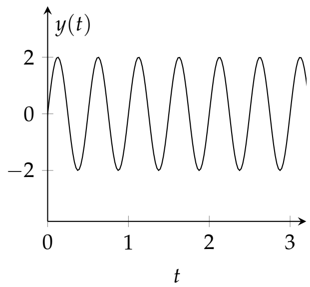
$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} A_n \cos\frac{n\pi ct}{L}\sin\frac{n\pi x}{L} \\ &\quad + B_n \sin\frac{n\pi ct}{L}\sin\frac{n\pi x}{L}. \end{aligned}$$

If the initial profile and velocity are given by $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$, respectively, then one has to satisfy the conditions

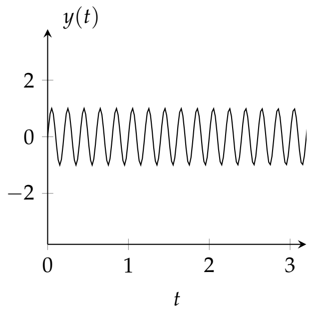
$$f(x) = u(x,0) = \sum_{n=1}^{\infty} A_n \sin\frac{n\pi x}{L}$$

and

$$g(x) = u_t(x,0) = \sum_{n=1}^{\infty} \frac{n\pi c}{L} B_n \sin\frac{n\pi x}{L}.$$



(a) $y(t) = 2\sin(4\pi f t)$



(b) $y(t) = \sin(10\pi f t)$

Figure 2.1: Plots of $y(t) = A\sin(2\pi f t)$ on $[0, 5]$ for $f = 2$ Hz and $f = 5$ Hz.

where A is the amplitude, f is the frequency in Hertz (Hz), and t is time in seconds. The amplitude is related to the volume of the sound. The larger the amplitude, the louder the sound. In Figure 2.1 we show plots of two such tones with $f = 2$ Hz in the top plot and $f = 5$ Hz in the bottom one.

In these plots you should notice the difference due to the amplitudes and the frequencies. You can easily reproduce these plots and others in your favorite plotting utility.

As an aside, you should be cautious when plotting functions, or sampling data. The plots you get might not be what you expect, even for a simple sine function. In Figure 2.2 we show four plots of the function $y(t) = 2\sin(4\pi t)$. In the top left, you see a proper rendering of this function. However, if you use a different number of points to plot this function, the results may be surprising. In this example we show what happens if you use $N = 200, 100, 101$ points instead of the 201 points used in the first plot. Such disparities are not only possible when plotting functions, but are also present when collecting data. Typically, when you sample a set of data, you only gather a finite amount of information at a fixed rate. This could happen when getting data on ocean wave heights, digitizing music, and other audio to put on your computer, or any other process when you attempt to analyze a continuous signal.

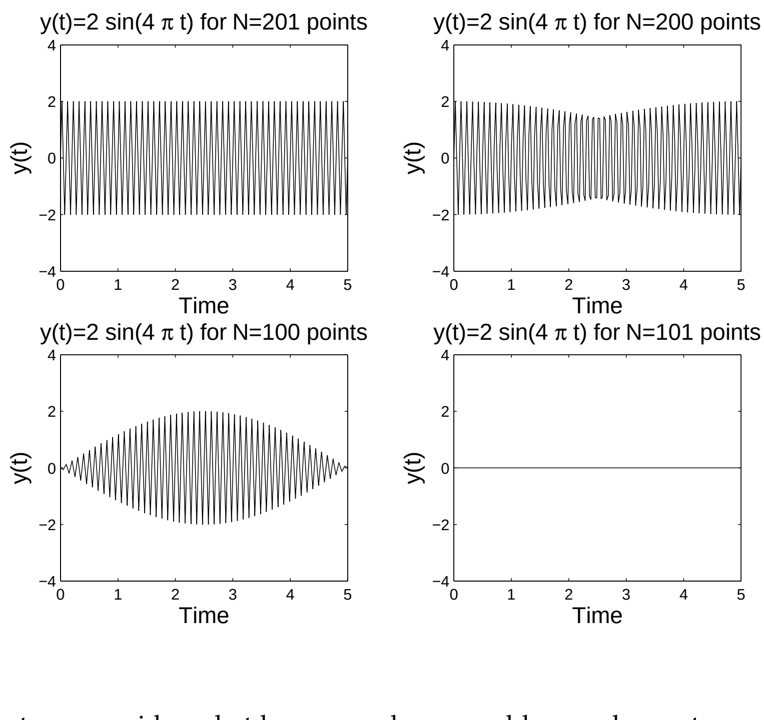


Figure 2.2: Problems can occur while plotting. Here we plot the function $y(t) = 2\sin 4\pi t$ using $N = 201, 200, 100, 101$ points.

Next, we consider what happens when we add several pure tones. After all, most of the sounds that we hear are, in fact, a combination of pure tones with different amplitudes and frequencies. In Figure 2.3 we see what happens when we add several sinusoids. Note that as one adds more and more tones with different characteristics, the resulting signal gets more complicated. However, we still have a function of time. In this chapter we will ask, "Given a function $f(t)$, can we find a set of sinusoidal functions whose sum converges to $f(t)$?"

Looking at the superpositions in Figure 2.3, we see that the sums yield functions that appear to be periodic. This is not unexpected. We recall that a periodic function is one in which the function values repeat over the domain of the function. The length of the smallest part of the domain which repeats is called the period. We can define this more precisely: A function is said to be periodic with period T if $f(t+T) = f(t)$ for all t and the smallest such positive number T is called the period.

For example, we consider the functions used in Figure 2.3. We began with $y(t) = 2\sin(4\pi t)$. Recall from your first studies of trigonometric functions that one can determine the period by dividing the coefficient of t into 2π to get the period. In this case we have

$$T = \frac{2\pi}{4\pi} = \frac{1}{2}.$$

Looking at the top plot in Figure 2.1 we can verify this result. (You can count the full number of cycles in the graph and divide this into the total time to get a more accurate value of the period.)

In general, if $y(t) = A\sin(2\pi f t)$, the period is found as

$$T = \frac{2\pi}{2\pi f} = \frac{1}{f}.$$

Of course, this result makes sense, as the unit of frequency, the hertz, is also defined as s^{-1}, or cycles per second.

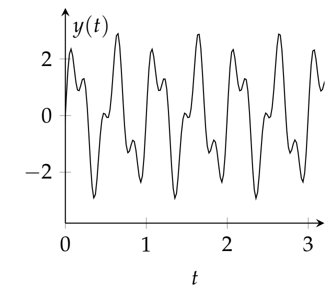
Returning to Figure 2.3, the functions $y(t) = 2\sin(4\pi t)$, $y(t) = \sin(10\pi t)$, and $y(t) = 0.5\sin(16\pi t)$ have periods of 0.5s, 0.2s, and 0.125s, respectively. Each superposition in Figure 2.3 retains a period that is the least common multiple of the periods of the signals added. For both plots, this is 1.0 s $= 2(0.5)$ s $= 5(.2)$ s $= 8(.125)$ s.

Our goal will be to start with a function and then determine the amplitudes of the simple sinusoids needed to sum to that function. We will see that this might involve an infinite number of such terms. Thus, we will be studying an infinite series of sinusoidal functions.

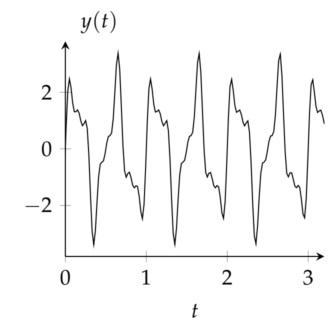
Secondly, we will find that using just sine functions will not be enough either. This is because we can add sinusoidal functions that do not necessarily peak at the same time. We will consider two signals that originate at different times. This is similar to when your music teacher would make sections of the class sing a song like "Row, Row, Row Your Boat" starting at slightly different times.

We can easily add shifted sine functions. In Figure 2.4 we show the functions $y(t) = 2\sin(4\pi t)$ and $y(t) = 2\sin(4\pi t + 7\pi/8)$ and their sum. Note that this shifted sine function can be written as $y(t) = 2\sin(4\pi(t + 7/32))$. Thus, this corresponds to a time shift of $-7/32$.

So, we should account for shifted sine functions in the general sum. Of course, we would then need to determine the unknown time shift as well as the amplitudes of the sinusoidal functions that make up the signal, $f(t)$.

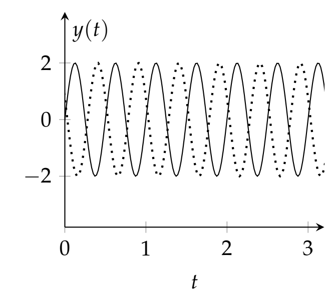


(a) Sum of signals with frequencies $f = 2$ Hz and $f = 5$ Hz.

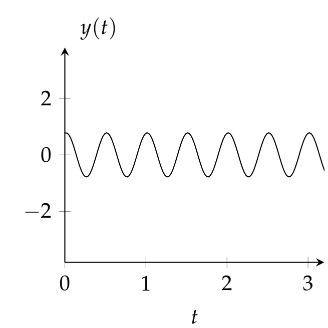


(b) Sum of signals with frequencies $f = 2$ Hz, $f = 5$ Hz, and $f = 8$ Hz.

Figure 2.3: Superposition of several sinusoids.



(a) Plot of each function.



(b) Plot of the sum of the functions.

Figure 2.4: Plot of the functions $y(t) = 2\sin(4\pi t)$ and $y(t) = 2\sin(4\pi t + 7\pi/8)$ and their sum.

While this is one approach that some researchers use to analyze signals, there is a more common approach. This results from another reworking of the shifted function.

We should note that the form in the lower plot of Figure 2.4 looks like a simple sinusoidal function for a reason. Let

$$y_1(t) = 2\sin(4\pi t),$$

$$y_2(t) = 2\sin(4\pi t + 7\pi/8).$$

Then,

$$\begin{aligned} y_1 + y_2 &= 2\sin(4\pi t + 7\pi/8) + 2\sin(4\pi t) \\ &= 2[\sin(4\pi t + 7\pi/8) + \sin(4\pi t)] \\ &= 4\cos\frac{7\pi}{16}\sin\left(4\pi t + \frac{7\pi}{16}\right). \end{aligned}$$

Consider the general shifted function

$$y(t) = A\sin(2\pi f t + \phi). \tag{2.1}$$

Note that $2\pi f t + \phi$ is called the phase of the sine function and ϕ is called the phase shift. We can use the trigonometric identity (2.9) for the sine of the sum of two angles[1] to obtain

$$\begin{aligned} y(t) &= A\sin(2\pi f t + \phi) \\ &= A\sin(\phi)\cos(2\pi f t) + A\cos(\phi)\sin(2\pi f t). \end{aligned} \tag{2.2}$$

[1] Recall the identities (2.9) and (2.10)

$$\begin{aligned} \sin(x+y) &= \sin x\cos y + \sin y\cos x, \\ \cos(x+y) &= \cos x\cos y - \sin x\sin y. \end{aligned}$$

Defining $a = A\sin(\phi)$ and $b = A\cos(\phi)$, we can rewrite this as

$$y(t) = a\cos(2\pi f t) + b\sin(2\pi f t).$$

Thus, we see that the signal in Equation (2.1) is a sum of sine and cosine functions with the same frequency and different amplitudes. If we can find a and b, then we can easily determine A and ϕ:

$$A = \sqrt{a^2 + b^2}, \quad \tan\phi = \frac{b}{a}.$$

We are now in a position to state our goal:

Goal - Fourier Analysis

Given a signal $f(t)$, we would like to determine its frequency content by finding out what combinations of sines and cosines of varying frequencies and amplitudes will sum to the given function. This is called Fourier Analysis.

2.2 *Fourier Trigonometric Series*

As we have seen in the last section, we are interested in finding representations of functions in terms of sines and cosines. Given a function $f(x)$ we seek a representation in the form

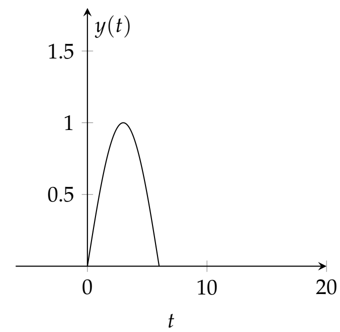
$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\cos nx + b_n\sin nx\right]. \tag{2.3}$$

Notice that we have opted to drop the references to the time-frequency form of the phase. This will lead to a simpler discussion for now and one can always make the transformation $nx = 2\pi f_n t$ when applying these ideas to applications.

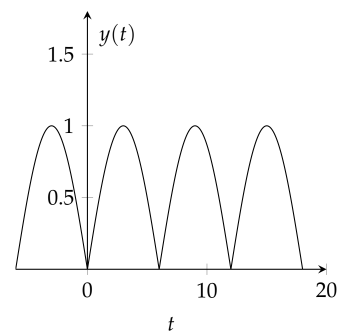
The series representation in Equation (2.3) is called a Fourier trigonometric series. We will simply refer to this as a Fourier series for now. The set of constants a_0, a_n, b_n, $n = 1, 2, \ldots$ are called the Fourier coefficients. The constant term is chosen in this form to make later computations simpler, though some other authors choose to write the constant term as a_0. Our goal is to find the Fourier series representation given $f(x)$. Having found the Fourier series representation, we will be interested in determining when the Fourier series converges and to what function it converges.

From our discussion in the last section, we see that The Fourier series is periodic. The periods of $\cos nx$ and $\sin nx$ are $\frac{2\pi}{n}$. Thus, the largest period, $T = 2\pi$, comes from the $n = 1$ terms and the Fourier series has period 2π. This means that the series should be able to represent functions that are periodic of period 2π.

While this appears restrictive, we could also consider functions that are defined over one period. In Figure 2.5 we show a function defined on $[0, 2\pi]$. In the same figure, we show its periodic extension. These are just copies of the original function shifted by the period and glued together. The extension can now be represented by a Fourier series and restricting the Fourier series to $[0, 2\pi]$ will give a representation of the original function. Therefore, we will first consider Fourier series representations of functions defined on this interval. Note that we could just as easily considered functions defined on $[-\pi, \pi]$ or any interval of length 2π. We will consider more general intervals later in the chapter.



(a) Plot of function $f(t)$.



(b) Periodic extension of $f(t)$.

Figure 2.5: Plot of the function $f(t)$ defined on $[0, 2\pi]$ and its periodic extension.

Fourier Coefficients

Theorem 2.1. *The Fourier series representation of $f(x)$ defined on $[0, 2\pi]$, when it exists, is given by Equation (2.3) with Fourier coefficients*

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx, \quad n = 0, 1, 2, \ldots, \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx \, dx, \quad n = 1, 2, \ldots. \end{aligned} \tag{2.4}$$

These expressions for the Fourier coefficients are obtained by considering special integrations of the Fourier series. We will now derive the a_n integrals in Equation (2.4).

We begin with the computation of a_0. Integrating the Fourier series term by term in Equation (2.3), we have

$$\int_0^{2\pi} f(x)\, dx = \int_0^{2\pi} \frac{a_0}{2}\, dx + \int_0^{2\pi} \sum_{n=1}^{\infty} \left[a_n \cos nx + b_n \sin nx\right]\, dx. \tag{2.5}$$

We will assume that we can integrate the infinite sum term by term. Then we will need to compute

Evaluating the integral of an infinite series by integrating term by term depends on the convergence properties of the series.

$$\begin{aligned} \int_0^{2\pi} \frac{a_0}{2}\, dx &= \frac{a_0}{2}(2\pi) = \pi a_0, \\ \int_0^{2\pi} \cos nx \, dx &= \left[\frac{\sin nx}{n}\right]_0^{2\pi} = 0, \\ \int_0^{2\pi} \sin nx \, dx &= \left[\frac{-\cos nx}{n}\right]_0^{2\pi} = 0. \end{aligned} \tag{2.6}$$

From these results we see that only one term in the integrated sum does not vanish, leaving

$$\int_0^{2\pi} f(x)\, dx = \pi a_0.$$

This confirms the value for a_0.[2]

[2] Note that $\frac{a_0}{2}$ is the average of $f(x)$ over the interval $[0, 2\pi]$. Recall from the first semester of calculus, that the average of a function defined on $[a, b]$ is given by

$$f_{\text{ave}} = \frac{1}{b-a}\int_a^b f(x)\, dx.$$

For $f(x)$ defined on $[0, 2\pi]$, we have

$$f_{\text{ave}} = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, dx = \frac{a_0}{2}.$$

Next, we will find the expression for a_n. We multiply the Fourier series in Equation (2.3) by $\cos mx$ for some positive integer m. This is like multiplying by $\cos 2x$, $\cos 5x$, etc. We are multiplying by all possible $\cos mx$ functions for different integers m all at the same time. We will see that this will allow us to solve for the a_n's.

We find the integrated sum of the series times $\cos mx$ is given by

$$\begin{aligned} \int_0^{2\pi} f(x)\cos mx\, dx &= \int_0^{2\pi} \frac{a_0}{2}\cos mx\, dx \\ &+ \int_0^{2\pi} \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx] \cos mx\, dx. \end{aligned} \tag{2.7}$$

Integrating term by term, the right side becomes

$$\int_0^{2\pi} f(x)\cos mx\, dx = \frac{a_0}{2}\int_0^{2\pi} \cos mx\, dx + \sum_{n=1}^{\infty}\left[a_n \int_0^{2\pi} \cos nx \cos mx\, dx + b_n \int_0^{2\pi} \sin nx \cos mx\, dx\right]. \tag{2.8}$$

We have already established that $\int_0^{2\pi} \cos mx\, dx = 0$, which implies that the first term vanishes.

Next we need to compute integrals of products of sines and cosines. This requires that we make use of some of the well known trigonometric. For quick reference, we list these here.

Useful Trigonometric Identities

$$\sin(x \pm y) = \sin x \cos y \pm \sin y \cos x \tag{2.9}$$

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y \tag{2.10}$$

$$\sin^2 x = \frac{1}{2}(1 - \cos 2x) \tag{2.11}$$

$$\cos^2 x = \frac{1}{2}(1 + \cos 2x) \tag{2.12}$$

$$\sin x \sin y = \frac{1}{2}(\cos(x-y) - \cos(x+y)) \tag{2.13}$$

$$\cos x \cos y = \frac{1}{2}(\cos(x+y) + \cos(x-y)) \tag{2.14}$$

$$\sin x \cos y = \frac{1}{2}(\sin(x+y) + \sin(x-y)) \tag{2.15}$$

We first want to evaluate $\int_0^{2\pi} \cos nx \cos mx\, dx$. We do this using the product identity (2.14). In case you forgot how to derive this identity, we will quickly review the derivation. Using the identities (2.10), we have

$$\begin{aligned} \cos(A+B) &= \cos A \cos B - \sin A \sin B, \\ \cos(A-B) &= \cos A \cos B + \sin A \sin B. \end{aligned}$$

Adding these equations,

$$2\cos A \cos B = \cos(A+B) + \cos(A-B).$$

We can use this result with $A = mx$ and $B = nx$ to complete the integration. We have

$$\begin{aligned} \int_0^{2\pi} \cos nx \cos mx \, dx &= \frac{1}{2}\int_0^{2\pi} [\cos(m+n)x + \cos(m-n)x] \, dx \\ &= \frac{1}{2}\left[\frac{\sin(m+n)x}{m+n} + \frac{\sin(m-n)x}{m-n}\right]_0^{2\pi} \\ &= 0. \end{aligned} \tag{2.16}$$

There is one caveat when doing such integrals. What if one of the denominators $m \pm n$ vanishes? For this problem, $m + n \neq 0$, as both m and n are positive integers. However, it is possible for $m = n$. This means that the vanishing of the integral can only happen when $m \neq n$. So, what can we do about the $m = n$ case? One way is to start from scratch with our integration. (Another way is to compute the limit as n approaches m in our result and use L'Hopital's Rule. Try it!)

For $n = m$ we have to compute $\int_0^{2\pi} \cos^2 mx \, dx$. This can also be handled using a trigonometric identity. Using the half angle formula, Equation (2.12), with $\theta = mx$, we find

$$\begin{aligned} \int_0^{2\pi} \cos^2 mx \, dx &= \frac{1}{2}\int_0^{2\pi} (1 + \cos 2mx) \, dx \\ &= \frac{1}{2}\left[x + \frac{1}{2m}\sin 2mx\right]_0^{2\pi} \\ &= \frac{1}{2}(2\pi) = \pi. \end{aligned} \tag{2.17}$$

To summarize, we have shown that

$$\int_0^{2\pi} \cos nx \cos mx \, dx = \begin{cases} 0, & m \neq n, \\ \pi, & m = n. \end{cases} \tag{2.18}$$

This holds true for $m, n = 0, 1, \ldots$. [Why did we include $m, n = 0$?] When we have such a set of functions, they are said to be an orthogonal set over the integration interval. A set of (real) functions $\{\phi_n(x)\}$ is said to be orthogonal on $[a, b]$ if $\int_a^b \phi_n(x)\phi_m(x) \, dx = 0$ when $n \neq m$. Furthermore, if we also have that $\int_a^b \phi_n^2(x) \, dx = 1$, these functions are called orthonormal.

Definition of an orthogonal set of functions and orthonormal functions.

The set of functions $\{\cos nx\}_{n=0}^{\infty}$ is orthogonal on $[0, 2\pi]$. Actually, the set is orthogonal on any interval of length 2π. We can make them orthonormal by dividing each function by $\sqrt{\pi}$, as indicated by Equation (2.17). This is sometimes referred to as normalization of the set of functions.

The notion of orthogonality is actually a generalization of the orthogonality of vectors in finite dimensional vector spaces. The integral $\int_a^b f(x) f(x)\, dx$ is the generalization of the dot product, and is called the scalar product of $f(x)$ and $g(x)$, which are thought of as vectors in an infinite dimensional vector space spanned by a set of orthogonal functions. We will return to these ideas in the next chapter.

Returning to the integrals in equation (2.8), we still have to evaluate $\int_0^{2\pi} \sin nx \cos mx\, dx$. We can use the trigonometric identity involving products of sines and cosines, Equation (2.15). Setting $A = nx$ and $B = mx$, we find that

Identity (2.15) is found from adding the identities

$$\begin{aligned}\sin(A+B) &= \sin A \cos B + \sin B \cos A,\\ \sin(A-B) &= \sin A \cos B - \sin B \cos A.\end{aligned}$$

$$\begin{aligned}\int_0^{2\pi} \sin nx \cos mx\, dx &= \frac{1}{2}\int_0^{2\pi} [\sin(n+m)x + \sin(n-m)x]\, dx\\ &= \frac{1}{2}\left[\frac{-\cos(n+m)x}{n+m} + \frac{-\cos(n-m)x}{n-m}\right]_0^{2\pi}\\ &= (-1+1) + (-1+1) = 0. \qquad (2.19)\end{aligned}$$

So,

$$\int_0^{2\pi} \sin nx \cos mx\, dx = 0. \qquad (2.20)$$

For these integrals we should also be careful about setting $n = m$. In this special case, we have the integrals

$$\int_0^{2\pi} \sin mx \cos mx\, dx = \frac{1}{2}\int_0^{2\pi} \sin 2mx\, dx = \frac{1}{2}\left[\frac{-\cos 2mx}{2m}\right]_0^{2\pi} = 0.$$

Finally, we can finish evaluating the expression in Equation (2.8). We have determined that all but one integral vanishes. In that case, $n = m$. This leaves us with

$$\int_0^{2\pi} f(x) \cos mx\, dx = a_m \pi.$$

Solving for a_m gives

$$a_m = \frac{1}{\pi}\int_0^{2\pi} f(x) \cos mx\, dx.$$

Since this is true for all $m = 1, 2, \ldots$, we have proven this part of the theorem. The only part left is finding the b_n's This will be left as an exercise for the reader.

We now consider examples of finding Fourier coefficients for given functions. In all of these cases, we define $f(x)$ on $[0, 2\pi]$.

Example 2.1. $f(x) = 3\cos 2x$, $x \in [0, 2\pi]$.

We first compute the integrals for the Fourier coefficients:

$$\begin{aligned}a_0 &= \frac{1}{\pi}\int_0^{2\pi} 3\cos 2x\, dx = 0,\\ a_n &= \frac{1}{\pi}\int_0^{2\pi} 3\cos 2x \cos nx\, dx = 0, \quad n \neq 2,\\ a_2 &= \frac{1}{\pi}\int_0^{2\pi} 3\cos^2 2x\, dx = 3,\\ b_n &= \frac{1}{\pi}\int_0^{2\pi} 3\cos 2x \sin nx\, dx = 0, \forall n.\end{aligned}$$

The integrals for a_0, $a_n, n \neq 2$, and b_n are the result of orthogonality. For a_2, the integral can be computed as follows:

$$\begin{aligned} a_2 &= \frac{1}{\pi}\int_0^{2\pi} 3\cos^2 2x\,dx \\ &= \frac{3}{2\pi}\int_0^{2\pi}\left[1+\cos 4x\right]\,dx \\ &= \frac{3}{2\pi}\left[x+\underbrace{\frac{1}{4}\sin 4x}_{\text{This term vanishes!}}\right]_0^{2\pi} = 3. \end{aligned} \tag{2.21}$$

Therefore, we have that the only nonvanishing coefficient is $a_2 = 3$. So there is one term and $f(x) = 3\cos 2x$.

Well, we should have known the answer to the last example before doing all of those integrals. If we have a function expressed simply in terms of sums of simple sines and cosines, then it should be easy to write the Fourier coefficients without much work. This is seen by writing out the Fourier series,

$$\begin{aligned} f(x) &\sim \frac{a_0}{2}+\sum_{n=1}^{\infty}\left[a_n\cos nx+b_n\sin nx\right]. \\ &= \frac{a_0}{2}+a_1\cos x+b_1\sin x++a_2\cos 2x+b_2\sin 2x+\ldots. \end{aligned} \tag{2.22}$$

For the last problem, $f(x) = 3\cos 2x$. Comparing this to the expanded Fourier series, one can immediately read off the Fourier coefficients without doing any integration. In the next example, we emphasize this point.

Example 2.2. $f(x) = \sin^2 x$, $x \in [0, 2\pi]$.

We could determine the Fourier coefficients by integrating as in the last example. However, it is easier to use trigonometric identities. We know that

$$\sin^2 x = \frac{1}{2}(1-\cos 2x) = \frac{1}{2}-\frac{1}{2}\cos 2x.$$

There are no sine terms, so $b_n = 0$, $n = 1, 2, \ldots$. There is a constant term, implying $a_0/2 = 1/2$. So, $a_0 = 1$. There is a $\cos 2x$ term, corresponding to $n = 2$, so $a_2 = -\frac{1}{2}$. That leaves $a_n = 0$ for $n \neq 0, 2$. So, $a_0 = 1$, $a_2 = -\frac{1}{2}$, and all other Fourier coefficients vanish.

Example 2.3. $f(x) = \begin{cases} 1, & 0 < x < \pi, \\ -1, & \pi < x < 2\pi, \end{cases}$.

This example will take a little more work. We cannot bypass evaluating any integrals this time. As seen in Figure 2.6, this function is discontinuous. So, we will break up any integration into two integrals, one over $[0, \pi]$ and the other over $[\pi, 2\pi]$.

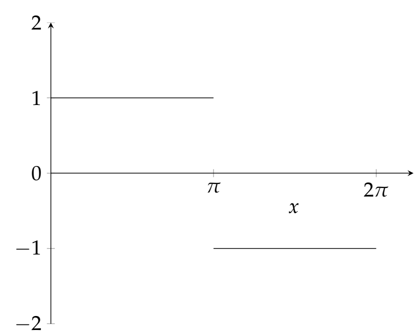


Figure 2.6: Plot of discontinuous function in Example 2.3.

$$a_0 = \frac{1}{\pi}\int_0^{2\pi} f(x)\,dx$$

$$\begin{aligned} &= \frac{1}{\pi}\int_0^{\pi} dx + \frac{1}{\pi}\int_{\pi}^{2\pi}(-1)\,dx \\ &= \frac{1}{\pi}(\pi) + \frac{1}{\pi}(-2\pi+\pi) = 0. \end{aligned} \tag{2.23}$$

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx \\ &= \frac{1}{\pi}\left[\int_0^{\pi}\cos nx\,dx - \int_{\pi}^{2\pi}\cos nx\,dx\right] \\ &= \frac{1}{\pi}\left[\left(\frac{1}{n}\sin nx\right)_0^{\pi} - \left(\frac{1}{n}\sin nx\right)_{\pi}^{2\pi}\right] \\ &= 0. \end{aligned} \tag{2.24}$$

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,dx \\ &= \frac{1}{\pi}\left[\int_0^{\pi}\sin nx\,dx - \int_{\pi}^{2\pi}\sin nx\,dx\right] \\ &= \frac{1}{\pi}\left[\left(-\frac{1}{n}\cos nx\right)_0^{\pi} + \left(\frac{1}{n}\cos nx\right)_{\pi}^{2\pi}\right] \\ &= \frac{1}{\pi}\left[-\frac{1}{n}\cos n\pi + \frac{1}{n} + \frac{1}{n} - \frac{1}{n}\cos n\pi\right] \\ &= \frac{2}{n\pi}(1-\cos n\pi). \end{aligned} \tag{2.25}$$

Often we see expressions involving $\cos n\pi = (-1)^n$ and $1 \pm \cos n\pi = 1 \pm (-1)^n$. This is an example showing how to re-index series containing $\cos n\pi$.

We have found the Fourier coefficients for this function. Before inserting them into the Fourier series (2.3), we note that $\cos n\pi = (-1)^n$. *Therefore,*

$$1 - \cos n\pi = \begin{cases} 0, & n \text{ even}, \\ 2, & n \text{ odd}. \end{cases} \tag{2.26}$$

So, half of the b_n's are zero. While we could write the Fourier series representation as

$$f(x) \sim \frac{4}{\pi}\sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{1}{n}\sin nx,$$

we could let $n = 2k-1$ in order to capture the odd numbers only. The answer can be written as

$$f(x) = \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\sin(2k-1)x}{2k-1},$$

Having determined the Fourier representation of a given function, we would like to know if the infinite series can be summed; i.e., does the series converge? Does it converge to $f(x)$? We will discuss this question later in the chapter after we generalize the Fourier series to intervals other than for $x \in [0, 2\pi]$.

2.3 *Fourier Series over Other Intervals*

In many applications we are interested in determining Fourier series representations of functions defined on intervals other than $[0, 2\pi]$. In this section, we will determine the form of the series expansion and the Fourier coefficients in these cases.

The most general type of interval is given as $[a, b]$. However, this often is too general. More common intervals are of the form $[-\pi, \pi]$, $[0, L]$, or $[-L/2, L/2]$. The simplest generalization is to the interval $[0, L]$. Such intervals arise often in applications. For example, for the problem of a one-dimensional string of length L, we set up the axes with the left end at $x = 0$ and the right end at $x = L$. Similarly for the temperature distribution along a one dimensional rod of length L we set the interval to $x \in [0, 2\pi]$. Such problems naturally lead to the study of Fourier series on intervals of length L. We will see later that symmetric intervals, $[-a, a]$, are also useful.

Given an interval $[0, L]$, we could apply a transformation to an interval of length 2π by simply rescaling the interval. Then we could apply this transformation to the Fourier series representation to obtain an equivalent one useful for functions defined on $[0, L]$.

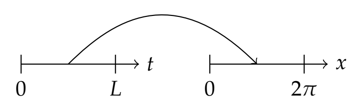


Figure 2.7: A sketch of the transformation between intervals $x \in [0, 2\pi]$ and $t \in [0, L]$.

We define $x \in [0, 2\pi]$ and $t \in [0, L]$. A linear transformation relating these intervals is simply $x = \frac{2\pi t}{L}$ as shown in Figure 2.7. So, $t = 0$ maps to $x = 0$ and $t = L$ maps to $x = 2\pi$. Furthermore, this transformation maps $f(x)$ to a new function $g(t) = f(x(t))$, which is defined on $[0, L]$. We will determine the Fourier series representation of this function using the representation for $f(x)$ from the last section.

Recall the form of the Fourier representation for $f(x)$ in Equation (2.3):

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos nx + b_n \sin nx \right]. \tag{2.27}$$

Inserting the transformation relating x and t, we have

$$g(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos \frac{2n\pi t}{L} + b_n \sin \frac{2n\pi t}{L} \right]. \tag{2.28}$$

This gives the form of the series expansion for $g(t)$ with $t \in [0, L]$. But, we still need to determine the Fourier coefficients.

Recall that

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx.$$

We need to make a substitution in the integral of $x = \frac{2\pi t}{L}$. We also will need to transform the differential, $dx = \frac{2\pi}{L} dt$. Thus, the resulting form for the Fourier coefficients is

$$a_n = \frac{2}{L} \int_0^L g(t) \cos \frac{2n\pi t}{L} \, dt. \tag{2.29}$$

Similarly, we find that

$$b_n = \frac{2}{L} \int_0^L g(t) \sin \frac{2n\pi t}{L} \, dt. \tag{2.30}$$

We note first that when $L = 2\pi$, we get back the series representation that we first studied. Also, the period of $\cos\frac{2n\pi t}{L}$ is L/n, which means that the representation for $g(t)$ has a period of L corresponding to $n = 1$.

At the end of this section, we present the derivation of the Fourier series representation for a general interval for the interested reader. In Table 2.1 we summarize some commonly used Fourier series representations.

Table 2.1: Special Fourier Series Representations on Different Intervals

Fourier Series on $[0, L]$

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\frac{2n\pi x}{L} + b_n \sin\frac{2n\pi x}{L}\right]. \tag{2.31}$$

$$\begin{aligned} a_n &= \frac{2}{L}\int_0^L f(x)\cos\frac{2n\pi x}{L}\,dx. \quad n = 0, 1, 2, \ldots, \\ b_n &= \frac{2}{L}\int_0^L f(x)\sin\frac{2n\pi x}{L}\,dx. \quad n = 1, 2, \ldots. \end{aligned} \tag{2.32}$$

Fourier Series on $[-\frac{L}{2}, \frac{L}{2}]$

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\frac{2n\pi x}{L} + b_n \sin\frac{2n\pi x}{L}\right]. \tag{2.33}$$

$$\begin{aligned} a_n &= \frac{2}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} f(x)\cos\frac{2n\pi x}{L}\,dx. \quad n = 0, 1, 2, \ldots, \\ b_n &= \frac{2}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} f(x)\sin\frac{2n\pi x}{L}\,dx. \quad n = 1, 2, \ldots. \end{aligned} \tag{2.34}$$

Fourier Series on $[-\pi, \pi]$

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos nx + b_n \sin nx\right]. \tag{2.35}$$

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx. \quad n = 0, 1, 2, \ldots, \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx. \quad n = 1, 2, \ldots. \end{aligned} \tag{2.36}$$

Integration of even and odd functions over symmetric intervals, $[-a, a]$.

Even Functions.

At this point we need to remind the reader about the integration of even and odd functions on symmetric intervals.

We first recall that $f(x)$ is an even function if $f(-x) = f(x)$ for all x. One can recognize even functions as they are symmetric with respect to the y-axis as shown in Figure 2.8.

If one integrates an even function over a symmetric interval, then one has that

$$\int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx. \tag{2.37}$$

One can prove this by splitting off the integration over negative values of x, using the substitution $x = -y$, and employing the evenness of $f(x)$. Thus,

$$\begin{aligned}
\int_{-a}^{a} f(x)\,dx &= \int_{-a}^{0} f(x)\,dx + \int_{0}^{a} f(x)\,dx \\
&= -\int_{a}^{0} f(-y)\,dy + \int_{0}^{a} f(x)\,dx \\
&= \int_{0}^{a} f(y)\,dy + \int_{0}^{a} f(x)\,dx \\
&= 2\int_{0}^{a} f(x)\,dx. \qquad (2.38)
\end{aligned}$$

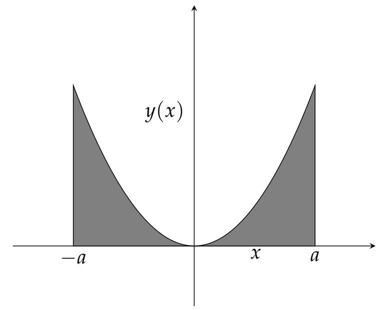


Figure 2.8: Area under an even function on a symmetric interval, $[-a, a]$.

This can be visually verified by looking at Figure 2.8.

A similar computation could be done for odd functions. $f(x)$ is an odd function if $f(-x) = -f(x)$ for all x. The graphs of such functions are symmetric with respect to the origin, as shown in Figure 2.9. If one integrates an odd function over a symmetric interval, then one has that

Odd Functions.

$$\int_{-a}^{a} f(x)\,dx = 0. \qquad (2.39)$$

Example 2.4. *Let $f(x) = |x|$ on $[-\pi, \pi]$ We compute the coefficients, beginning as usual with a_0. We have, using the fact that $|x|$ is an even function,*

$$\begin{aligned}
a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\,dx \\
&= \frac{2}{\pi}\int_{0}^{\pi} x\,dx = \pi \qquad (2.40)
\end{aligned}$$

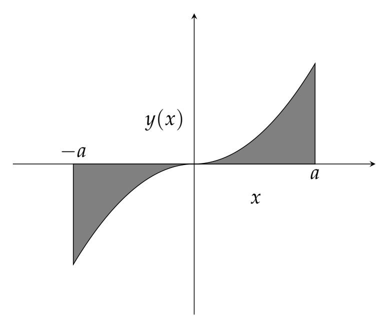


Figure 2.9: Area under an odd function on a symmetric interval, $[-a, a]$.

We continue with the computation of the general Fourier coefficients for $f(x) = |x|$ on $[-\pi, \pi]$. We have

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} |x| \cos nx\,dx = \frac{2}{\pi}\int_{0}^{\pi} x\cos nx\,dx. \qquad (2.41)$$

Here we have made use of the fact that $|x|\cos nx$ is an even function.

In order to compute the resulting integral, we need to use integration by parts ,

$$\int_{a}^{b} u\,dv = uv\Big|_{a}^{b} - \int_{a}^{b} v\,du,$$

by letting $u = x$ and $dv = \cos nx\,dx$. Thus, $du = dx$ and $v = \int dv = \frac{1}{n}\sin nx$.

Continuing with the computation, we have

$$\begin{aligned}
a_n &= \frac{2}{\pi}\int_{0}^{\pi} x\cos nx\,dx. \\
&= \frac{2}{\pi}\left[\frac{1}{n}x\sin nx\Big|_{0}^{\pi} - \frac{1}{n}\int_{0}^{\pi} \sin nx\,dx\right] \\
&= -\frac{2}{n\pi}\left[-\frac{1}{n}\cos nx\right]_{0}^{\pi} \\
&= -\frac{2}{\pi n^2}(1 - (-1)^n). \qquad (2.42)
\end{aligned}$$

Here we have used the fact that $\cos n\pi = (-1)^n$ *for any integer* n*. This leads to a factor* $(1-(-1)^n)$*. This factor can be simplified as*

$$1-(-1)^n = \begin{cases} 2, & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases} \tag{2.43}$$

So, $a_n = 0$ *for* n *even and* $a_n = -\frac{4}{\pi n^2}$ *for* n *odd.*

Computing the b_n*'s is simpler. We note that we have to integrate* $|x|\sin nx$ *from* $x=-\pi$ *to* π*. The integrand is an odd function and this is a symmetric interval. So, the result is that* $b_n = 0$ *for all* n*.*

Putting this all together, the Fourier series representation of $f(x) = |x|$ *on* $[-\pi,\pi]$ *is given as*

$$f(x) \sim \frac{\pi}{2} - \frac{4}{\pi}\sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{\cos nx}{n^2}. \tag{2.44}$$

While this is correct, we can rewrite the sum over only odd n *by re-indexing. We let* $n = 2k-1$ *for* $k = 1,2,3,\ldots$*. Then we only get the odd integers. The series can then be written as*

$$f(x) \sim \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\cos(2k-1)x}{(2k-1)^2}. \tag{2.45}$$

Throughout our discussion we have referred to such results as Fourier representations. We have not looked at the convergence of these series. Here is an example of an infinite series of functions. What does this series sum to? We show in Figure 2.10 the first few partial sums. They appear to be converging to $f(x) = |x|$ fairly quickly.

Even though $f(x)$ was defined on $[-\pi,\pi]$, we can still evaluate the Fourier series at values of x outside this interval. In Figure 2.11, we see that the representation agrees with $f(x)$ on the interval $[-\pi,\pi]$. Outside this interval, we have a periodic extension of $f(x)$ with period 2π.

Another example is the Fourier series representation of $f(x) = x$ on $[-\pi,\pi]$ as left for Problem 7. This is determined to be

$$f(x) \sim 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\sin nx. \tag{2.46}$$

As seen in Figure 2.12, we again obtain the periodic extension of the function. In this case, we needed many more terms. Also, the vertical parts of the first plot are nonexistent. In the second plot, we only plot the points and not the typical connected points that most software packages plot as the default style.

Example 2.5. *It is interesting to note that one can use Fourier series to obtain sums of some infinite series. For example, in the last example, we found that*

$$x \sim 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\sin nx.$$

Now, what if we chose $x = \frac{\pi}{2}$*? Then, we have*

$$\frac{\pi}{2} = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\sin\frac{n\pi}{2} = 2\left[1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\ldots\right].$$

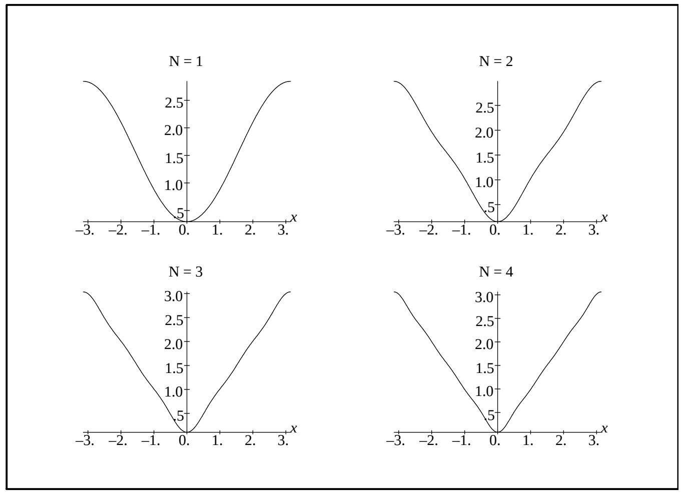


Figure 2.10: Plot of the first partial sums of the Fourier series representation for $f(x) = |x|$.

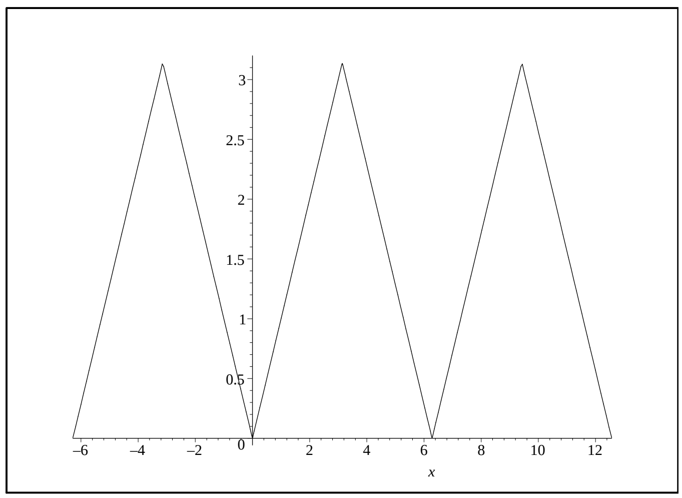


Figure 2.11: Plot of the first 10 terms of the Fourier series representation for $f(x) = |x|$ on the interval $[-2\pi, 4\pi]$.

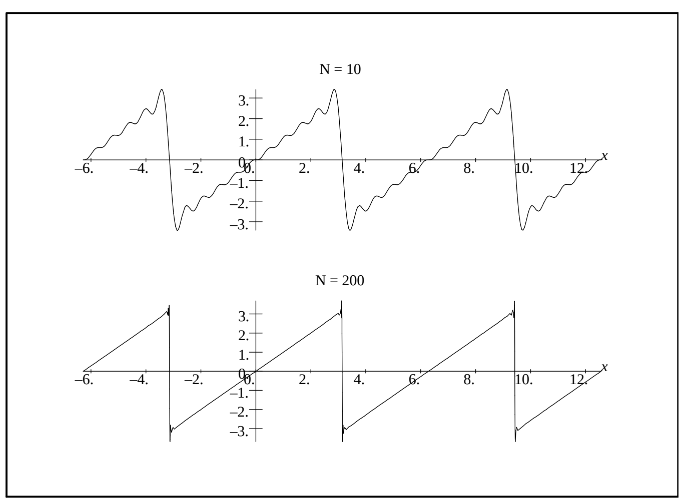


Figure 2.12: Plot of the first 10 terms and 200 terms of the Fourier series representation for $f(x) = x$ on the interval $[-2\pi, 4\pi]$.

This gives a well known expression for π:

$$\pi = 4\left[1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots\right].$$

2.3.1 *Fourier Series on $[a, b]$*

This section can be skipped on first reading. It is here for completeness and the end result, Theorem 2.2 provides the result of the section.

A FOURIER SERIES REPRESENTATION is also possible for a general interval, $t \in [a, b]$. As before, we just need to transform this interval to $[0, 2\pi]$. Let

$$x = 2\pi \frac{t-a}{b-a}.$$

Inserting this into the Fourier series (2.3) representation for $f(x)$, we obtain

$$g(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\frac{2n\pi(t-a)}{b-a} + b_n \sin\frac{2n\pi(t-a)}{b-a}\right]. \tag{2.47}$$

Well, this expansion is ugly. It is not like the last example, where the transformation was straightforward. If one were to apply the theory to applications, it might seem to make sense to just shift the data so that $a = 0$ and be done with any complicated expressions. However, some students enjoy the challenge of developing such generalized expressions. So, let's see what is involved.

First, we apply the addition identities for trigonometric functions and rearrange the terms.

$$\begin{aligned}
g(t) &\sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\frac{2n\pi(t-a)}{b-a} + b_n \sin\frac{2n\pi(t-a)}{b-a}\right] \\
&= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\left(\cos\frac{2n\pi t}{b-a}\cos\frac{2n\pi a}{b-a} + \sin\frac{2n\pi t}{b-a}\sin\frac{2n\pi a}{b-a}\right)\right. \\
&\quad \left. + b_n\left(\sin\frac{2n\pi t}{b-a}\cos\frac{2n\pi a}{b-a} - \cos\frac{2n\pi t}{b-a}\sin\frac{2n\pi a}{b-a}\right)\right] \\
&= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[\cos\frac{2n\pi t}{b-a}\left(a_n\cos\frac{2n\pi a}{b-a} - b_n\sin\frac{2n\pi a}{b-a}\right)\right. \\
&\quad \left. + \sin\frac{2n\pi t}{b-a}\left(a_n\sin\frac{2n\pi a}{b-a} + b_n\cos\frac{2n\pi a}{b-a}\right)\right].
\end{aligned} \tag{2.48}$$

Defining $A_0 = a_0$ and

$$\begin{aligned}
A_n &\equiv a_n\cos\frac{2n\pi a}{b-a} - b_n\sin\frac{2n\pi a}{b-a} \\
B_n &\equiv a_n\sin\frac{2n\pi a}{b-a} + b_n\cos\frac{2n\pi a}{b-a},
\end{aligned} \tag{2.49}$$

we arrive at the more desirable form for the Fourier series representation of a function defined on the interval $[a, b]$.

$$g(t) \sim \frac{A_0}{2} + \sum_{n=1}^{\infty}\left[A_n\cos\frac{2n\pi t}{b-a} + B_n\sin\frac{2n\pi t}{b-a}\right]. \tag{2.50}$$

We next need to find expressions for the Fourier coefficients. We insert the known expressions for a_n and b_n and rearrange. First, we note that under the transformation $x = 2\pi\frac{t-a}{b-a}$, we have

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx \\ &= \frac{2}{b-a}\int_a^b g(t)\cos\frac{2n\pi(t-a)}{b-a}\,dt, \end{aligned} \tag{2.51}$$

and

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx \\ &= \frac{2}{b-a}\int_a^b g(t)\sin\frac{2n\pi(t-a)}{b-a}\,dt. \end{aligned} \tag{2.52}$$

Then, inserting these integrals in A_n, combining integrals, and making use of the addition formula for the cosine of the sum of two angles, we obtain

$$\begin{aligned} A_n &\equiv a_n\cos\frac{2n\pi a}{b-a} - b_n\sin\frac{2n\pi a}{b-a} \\ &= \frac{2}{b-a}\int_a^b g(t)\left[\cos\frac{2n\pi(t-a)}{b-a}\cos\frac{2n\pi a}{b-a} - \sin\frac{2n\pi(t-a)}{b-a}\sin\frac{2n\pi a}{b-a}\right]dt \\ &= \frac{2}{b-a}\int_a^b g(t)\cos\frac{2n\pi t}{b-a}\,dt. \end{aligned} \tag{2.53}$$

A similar computation gives

$$B_n = \frac{2}{b-a}\int_a^b g(t)\sin\frac{2n\pi t}{b-a}\,dt. \tag{2.54}$$

Summarizing, we have shown that:

Theorem 2.2. *The Fourier series representation of $f(x)$ defined on $[a,b]$ when it exists, is given by*

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\cos\frac{2n\pi x}{b-a} + b_n\sin\frac{2n\pi x}{b-a}\right]. \tag{2.55}$$

with Fourier coefficients

$$\begin{aligned} a_n &= \frac{2}{b-a}\int_a^b f(x)\cos\frac{2n\pi x}{b-a}\,dx. \quad n = 0,1,2,\ldots, \\ b_n &= \frac{2}{b-a}\int_a^b f(x)\sin\frac{2n\pi x}{b-a}\,dx. \quad n = 1,2,\ldots. \end{aligned} \tag{2.56}$$

2.4 *Sine and Cosine Series*

IN THE LAST TWO EXAMPLES ($f(x) = |x|$ and $f(x) = x$ on $[-\pi,\pi]$), we have seen Fourier series representations that contain only sine or cosine terms. As we know, the sine functions are odd functions and thus sum to odd functions. Similarly, cosine functions sum to even functions. Such occurrences happen often in practice. Fourier representations involving just sines are called sine series and those involving just cosines (and the constant term) are called cosine series.

Another interesting result, based upon these examples, is that the original functions, $|x|$ and x, agree on the interval $[0, \pi]$. Note from Figures 2.10 through 2.12 that their Fourier series representations do as well. Thus, more than one series can be used to represent functions defined on finite intervals. All they need to do is agree with the function over that particular interval. Sometimes one of these series is more useful because it has additional properties needed in the given application.

We have made the following observations from the previous examples:

1. There are several trigonometric series representations for a function defined on a finite interval.
2. Odd functions on a symmetric interval are represented by sine series and even functions on a symmetric interval are represented by cosine series.

These two observations are related and are the subject of this section. We begin by defining a function $f(x)$ on interval $[0, L]$. We have seen that the Fourier series representation of this function appears to converge to a periodic extension of the function.

In Figure 2.13, we show a function defined on $[0, 1]$. To the right is its periodic extension to the whole real axis. This representation has a period of $L = 1$. The bottom left plot is obtained by first reflecting f about the y-axis to make it an even function and then graphing the periodic extension of this new function. Its period will be $2L = 2$. Finally, in the last plot, we flip the function about each axis and graph the periodic extension of the new odd function. It will also have a period of $2L = 2$.

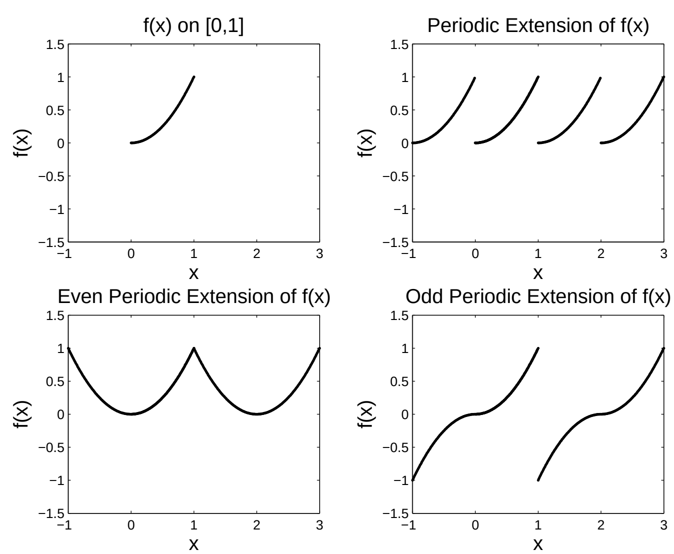


Figure 2.13: This is a sketch of a function and its various extensions. The original function $f(x)$ is defined on $[0, 1]$ and graphed in the upper left corner. To its right is the periodic extension, obtained by adding replicas. The two lower plots are obtained by first making the original function even or odd and then creating the periodic extensions of the new function.

In general, we obtain three different periodic representations. In order to distinguish these, we will refer to them simply as the periodic, even, and odd extensions. Now, starting with $f(x)$ defined on $[0, L]$, we would like to determine the Fourier series representations leading to these extensions. [For easy reference, the results are summarized in Table 2.2]

Table 2.2: Fourier Cosine and Sine Series Representations on $[0, L]$

Fourier Series on $[0, L]$

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos \frac{2n\pi x}{L} + b_n \sin \frac{2n\pi x}{L} \right]. \tag{2.57}$$

$$\begin{aligned} a_n &= \frac{2}{L} \int_0^L f(x) \cos \frac{2n\pi x}{L}\, dx. \quad n = 0, 1, 2, \ldots, \\ b_n &= \frac{2}{L} \int_0^L f(x) \sin \frac{2n\pi x}{L}\, dx. \quad n = 1, 2, \ldots. \end{aligned} \tag{2.58}$$

Fourier Cosine Series on $[0, L]$

$$f(x) \sim a_0/2 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L}. \tag{2.59}$$

where

$$a_n = \frac{2}{L} \int_0^L f(x) \cos \frac{n\pi x}{L}\, dx. \quad n = 0, 1, 2, \ldots. \tag{2.60}$$

Fourier Sine Series on $[0, L]$

$$f(x) \sim \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L}. \tag{2.61}$$

where

$$b_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L}\, dx. \quad n = 1, 2, \ldots. \tag{2.62}$$

We have already seen from Table 2.1 that the periodic extension of $f(x)$, defined on $[0, L]$, is obtained through the Fourier series representation

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos \frac{2n\pi x}{L} + b_n \sin \frac{2n\pi x}{L} \right], \tag{2.63}$$

where

$$\begin{aligned} a_n &= \frac{2}{L} \int_0^L f(x) \cos \frac{2n\pi x}{L}\, dx. \quad n = 0, 1, 2, \ldots, \\ b_n &= \frac{2}{L} \int_0^L f(x) \sin \frac{2n\pi x}{L}\, dx. \quad n = 1, 2, \ldots. \end{aligned} \tag{2.64}$$

Given $f(x)$ defined on $[0, L]$, the even periodic extension is obtained by simply computing the Fourier series representation for the even function

Even periodic extension.

$$f_e(x) \equiv \begin{cases} f(x), & 0 < x < L, \\ f(-x) & -L < x < 0. \end{cases} \tag{2.65}$$

Since $f_e(x)$ is an even function on a symmetric interval $[-L, L]$, we expect that the resulting Fourier series will not contain sine terms. Therefore, the series expansion will be given by [Use the general case in Equation (2.55) with $a = -L$ and $b = L$.]:

$$f_e(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L}. \tag{2.66}$$

with Fourier coefficients

$$a_n = \frac{1}{L} \int_{-L}^{L} f_e(x) \cos \frac{n\pi x}{L}\, dx. \quad n = 0, 1, 2, \ldots. \tag{2.67}$$

However, we can simplify this by noting that the integrand is even and the interval of integration can be replaced by $[0, L]$. On this interval $f_e(x) = f(x)$. So, we have the Cosine Series Representation of $f(x)$ for $x \in [0, L]$ is given as

Fourier Cosine Series.

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L}. \tag{2.68}$$

where

$$a_n = \frac{2}{L} \int_{0}^{L} f(x) \cos \frac{n\pi x}{L}\, dx. \quad n = 0, 1, 2, \ldots. \tag{2.69}$$

Odd periodic extension.

Similarly, given $f(x)$ defined on $[0, L]$, the odd periodic extension is obtained by simply computing the Fourier series representation for the odd function

$$f_o(x) \equiv \begin{cases} f(x), & 0 < x < L, \\ -f(-x) & -L < x < 0. \end{cases} \tag{2.70}$$

The resulting series expansion leads to defining the Sine Series Representation of $f(x)$ for $x \in [0, L]$ as

Fourier Sine Series Representation.

$$f(x) \sim \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L}, \tag{2.71}$$

where

$$b_n = \frac{2}{L} \int_{0}^{L} f(x) \sin \frac{n\pi x}{L}\, dx. \quad n = 1, 2, \ldots. \tag{2.72}$$

Example 2.6. *In Figure 2.13, we actually provided plots of the various extensions of the function $f(x) = x^2$ for $x \in [0, 1]$. Let's determine the representations of the periodic, even, and odd extensions of this function.*

For a change, we will use a CAS (Computer Algebra System) package to do the integrals. In this case, we can use Maple. A general code for doing this for the periodic extension is shown in Table 2.3.

Example 2.7. Periodic Extension - Trigonometric Fourier Series *Using the code in Table 2.3, we have that $a_0 = \frac{2}{3}$, $a_n = \frac{1}{n^2\pi^2}$, and $b_n = -\frac{1}{n\pi}$. Thus, the resulting series is given as*

$$f(x) \sim \frac{1}{3} + \sum_{n=1}^{\infty} \left[\frac{1}{n^2\pi^2} \cos 2n\pi x - \frac{1}{n\pi} \sin 2n\pi x \right].$$

In Figure 2.14, we see the sum of the first 50 terms of this series. Generally, we see that the series seems to be converging to the periodic extension of f. There appear to be some problems with the convergence around integer values of x. We will later see that this is because of the discontinuities in the periodic extension and the resulting overshoot is referred to as the Gibbs phenomenon, which is discussed in the last section of this chapter.

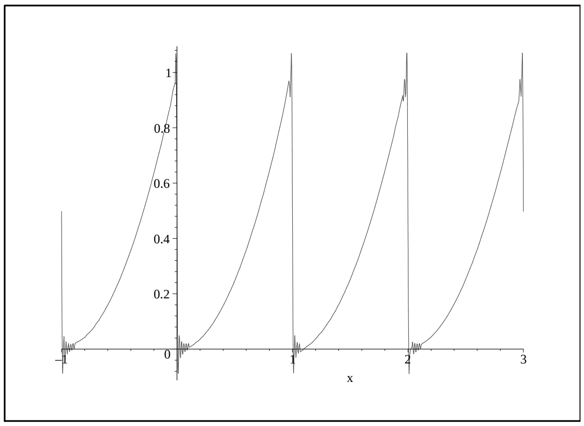


Figure 2.14: The periodic extension of $f(x) = x^2$ on $[0, 1]$.

Table 2.3: Maple code for computing Fourier coefficients and plotting partial sums of the Fourier series.

```
> restart:
> L:=1:
> f:=x^2:
> assume(n,integer):
> a0:=2/L*int(f,x=0..L);
                              a0 := 2/3
> an:=2/L*int(f*cos(2*n*Pi*x/L),x=0..L);
                                     1
                            an := -------
                                    2   2
                                  n~  Pi
> bn:=2/L*int(f*sin(2*n*Pi*x/L),x=0..L);
                                     1
                            bn := - -----
                                    n~ Pi
> F:=a0/2+sum((1/(k*Pi)^2)*cos(2*k*Pi*x/L)
     -1/(k*Pi)*sin(2*k*Pi*x/L),k=1..50):
> plot(F,x=-1..3,title=‘Periodic Extension‘,
      titlefont=[TIMES,ROMAN,14],font=[TIMES,ROMAN,14]);
```

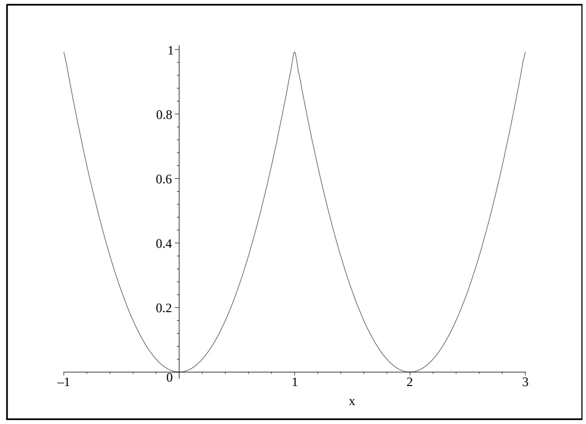
Example 2.8. Even Periodic Extension - Cosine Series

In this case we compute $a_0 = \frac{2}{3}$ and $a_n = \frac{4(-1)^n}{n^2\pi^2}$. Therefore, we have

$$f(x) \sim \frac{1}{3} + \frac{4}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos n\pi x.$$

In Figure 2.15, we see the sum of the first 50 terms of this series. In this case the convergence seems to be much better than in the periodic extension case. We also see that it is converging to the even extension.

Figure 2.15: The even periodic extension of $f(x) = x^2$ on $[0, 1]$.



Example 2.9. Odd Periodic Extension - Sine Series

Finally, we look at the sine series for this function. We find that

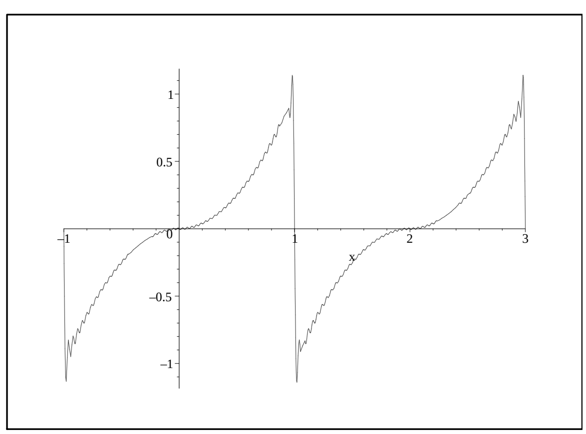
$$b_n = -\frac{2}{n^3\pi^3}(n^2\pi^2(-1)^n - 2(-1)^n + 2).$$

Therefore,

$$f(x) \sim -\frac{2}{\pi^3}\sum_{n=1}^{\infty}\frac{1}{n^3}(n^2\pi^2(-1)^n - 2(-1)^n + 2)\sin n\pi x.$$

Once again we see discontinuities in the extension as seen in Figure 2.16. However, we have verified that our sine series appears to be converging to the odd extension as we first sketched in Figure 2.13.

Figure 2.16: The odd periodic extension of $f(x) = x^2$ on $[0, 1]$.



2.5 *The Gibbs Phenomenon*

We have seen the Gibbs phenomenon when there is a jump discontinuity in the periodic extension of a function, whether the function originally had a discontinuity or developed one due to a mismatch in the values of the endpoints. This can be seen in Figures 2.12, 2.14, and 2.16. The Fourier series has a difficult time converging at the point of discontinuity and these graphs of the Fourier series show a distinct overshoot which does not go away. This is called the Gibbs phenomenon[3] and the amount of overshoot can be computed.

[3] The Gibbs phenomenon was named after Josiah Willard Gibbs (1839-1903) even though it was discovered earlier by the Englishman Henry Wilbraham (1825-1883). Wilbraham published a soon forgotten paper about the effect in 1848. In 1889 Albert Abraham Michelson (1852-1931), an American physicist,observed an overshoot in his mechanical graphing machine. Shortly afterwards J. Willard Gibbs published papers describing this phenomenon, which was later to be called the Gibbs phenomena. Gibbs was a mathematical physicist and chemist and is considered the father of physical chemistry.

In one of our first examples, Example 2.3, we found the Fourier series representation of the piecewise defined function

$$f(x) = \begin{cases} 1, & 0 < x < \pi, \\ -1, & \pi < x < 2\pi, \end{cases}$$

to be

$$f(x) \sim \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\sin(2k-1)x}{2k-1}.$$

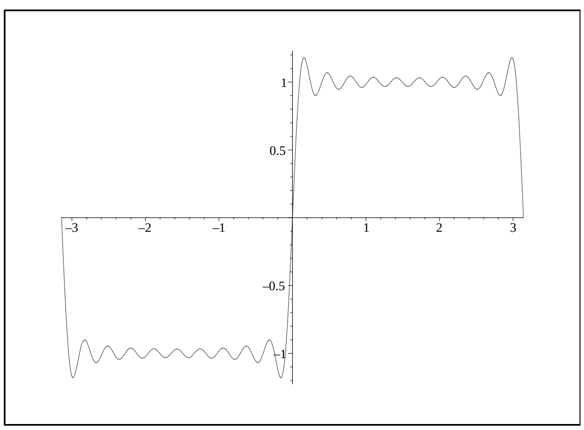


Figure 2.17: The Fourier series representation of a step function on $[-\pi, \pi]$ for $N = 10$.

In Figure 2.17, we display the sum of the first ten terms. Note the wiggles, overshoots and undershoots. These are seen more when we plot the representation for $x \in [-3\pi, 3\pi]$, as shown in Figure 2.18.

We note that the overshoots and undershoots occur at discontinuities in the periodic extension of $f(x)$. These occur whenever $f(x)$ has a discontinuity or if the values of $f(x)$ at the endpoints of the domain do not agree.

One might expect that we only need to add more terms. In Figure 2.19 we show the sum for twenty terms. Note the sum appears to converge better for points far from the discontinuities. But, the overshoots and undershoots are still present. Figures 2.20 and 2.21 show magnified plots of the overshoot at $x = 0$ for $N = 100$ and $N = 500$, respectively. We see that the overshoot

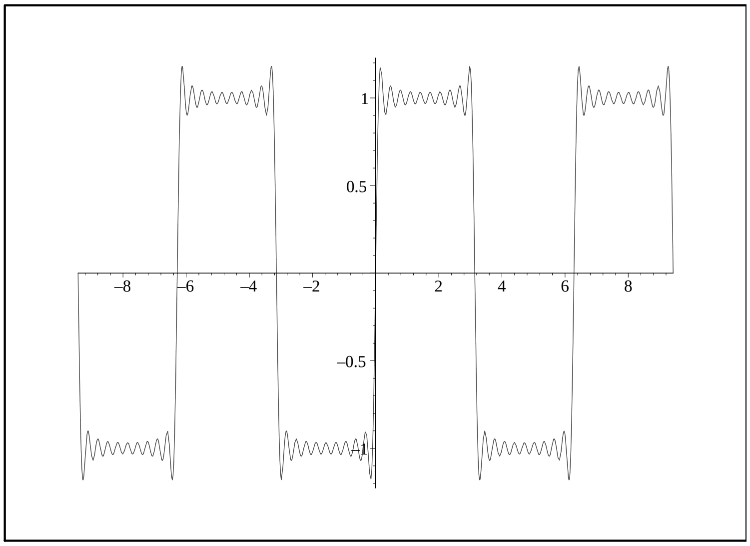


Figure 2.18: The Fourier series representation of a step function on $[-\pi, \pi]$ for $N = 10$ plotted on $[-3\pi, 3\pi]$ displaying the periodicity.

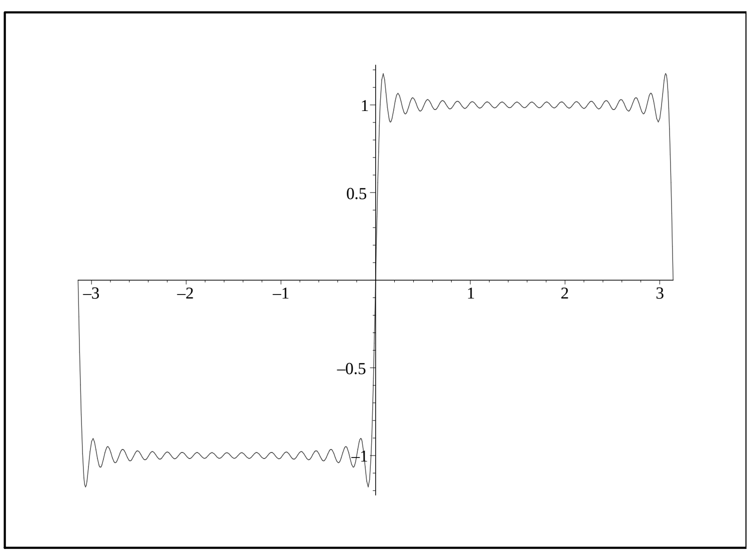


Figure 2.19: The Fourier series representation of a step function on $[-\pi, \pi]$ for $N = 20$.

persists. The peak is at about the same height, but its location seems to be getting closer to the origin. We will show how one can estimate the size of the overshoot.

We can study the Gibbs phenomenon by looking at the partial sums of general Fourier trigonometric series for functions $f(x)$ defined on the interval $[-L, L]$. Writing out the partial sums, inserting the Fourier coefficients, and rearranging, we have

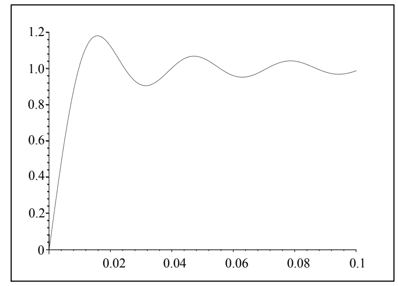


Figure 2.20: The Fourier series representation of a step function on $[-\pi, \pi]$ for $N = 100$.

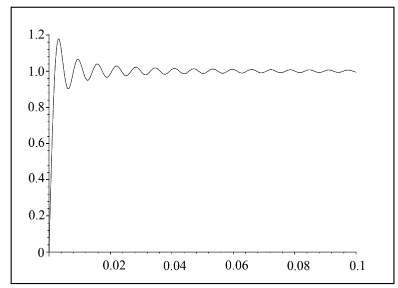


Figure 2.21: The Fourier series representation of a step function on $[-\pi, \pi]$ for $N = 500$.

$$\begin{aligned}
S_N(x) &= a_0 + \sum_{n=1}^{N} \left[a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right] \\
&= \frac{1}{2L} \int_{-L}^{L} f(y)\, dy + \sum_{n=1}^{N} \left[\left(\frac{1}{L} \int_{-L}^{L} f(y) \cos \frac{n\pi y}{L}\, dy \right) \cos \frac{n\pi x}{L} \right. \\
&\quad \left. + \left(\frac{1}{L} \int_{-L}^{L} f(y) \sin \frac{n\pi y}{L}\, dy. \right) \sin \frac{n\pi x}{L} \right] \\
&= \frac{1}{L} \int\limits_{-L}^{L} \left\{ \frac{1}{2} \right. \\
&\quad \left. + \sum_{n=1}^{N} \left(\cos \frac{n\pi y}{L} \cos \frac{n\pi x}{L} + \sin \frac{n\pi y}{L} \sin \frac{n\pi x}{L} \right) \right\} f(y)\, dy \\
&= \frac{1}{L} \int\limits_{-L}^{L} \left\{ \frac{1}{2} + \sum_{n=1}^{N} \cos \frac{n\pi(y-x)}{L} \right\} f(y)\, dy \\
&\equiv \frac{1}{L} \int\limits_{-L}^{L} D_N(y-x) f(y)\, dy
\end{aligned}$$

We have defined

$$D_N(x) = \frac{1}{2} + \sum_{n=1}^{N} \cos \frac{n\pi x}{L},$$

which is called the N-th Dirichlet kernel .

We now prove

Lemma 2.1. *The N-th Dirichlet kernel is given by*

$$D_N(x) = \begin{cases} \dfrac{\sin((N+\frac{1}{2})\frac{\pi x}{L})}{2 \sin \frac{\pi x}{2L}}, & \sin \frac{\pi x}{2L} \neq 0, \\ N + \frac{1}{2}, & \sin \frac{\pi x}{2L} = 0. \end{cases}$$

Proof. Let $\theta = \frac{\pi x}{L}$ and multiply $D_N(x)$ by $2 \sin \frac{\theta}{2}$ to obtain

$$\begin{aligned}
2 \sin \frac{\theta}{2} D_N(x) &= 2 \sin \frac{\theta}{2} \left[\frac{1}{2} + \cos\theta + \cdots + \cos N\theta \right] \\
&= \sin \frac{\theta}{2} + 2 \cos\theta \sin \frac{\theta}{2} + 2 \cos 2\theta \sin \frac{\theta}{2} + \cdots + 2 \cos N\theta \sin \frac{\theta}{2} \\
&= \sin \frac{\theta}{2} + \left(\sin \frac{3\theta}{2} - \sin \frac{\theta}{2} \right) + \left(\sin \frac{5\theta}{2} - \sin \frac{3\theta}{2} \right) + \cdots \\
&\quad + \left[\sin \left(N + \frac{1}{2} \right) \theta - \sin \left(N - \frac{1}{2} \right) \theta \right]
\end{aligned}$$

$$= \sin\left(N+\frac{1}{2}\right)\theta. \tag{2.73}$$

Thus,

$$2\sin\frac{\theta}{2}D_N(x) = \sin\left(N+\frac{1}{2}\right)\theta.$$

If $\sin\frac{\theta}{2} \neq 0$, then

$$D_N(x) = \frac{\sin\left(N+\frac{1}{2}\right)\theta}{2\sin\frac{\theta}{2}}, \qquad \theta = \frac{\pi x}{L}.$$

If $\sin\frac{\theta}{2} = 0$, then one needs to apply L'Hospital's Rule as $\theta \to 2m\pi$:

$$\begin{aligned}
\lim_{\theta\to 2m\pi} \frac{\sin\left(N+\frac{1}{2}\right)\theta}{2\sin\frac{\theta}{2}} &= \lim_{\theta\to 2m\pi} \frac{(N+\frac{1}{2})\cos\left(N+\frac{1}{2}\right)\theta}{\cos\frac{\theta}{2}} \\
&= \frac{(N+\frac{1}{2})\cos\left(2m\pi N + m\pi\right)}{\cos m\pi} \\
&= \frac{(N+\frac{1}{2})(\cos 2m\pi N\cos m\pi - \sin 2m\pi N\sin m\pi)}{\cos m\pi} \\
&= N+\frac{1}{2}. \qquad (2.74)
\end{aligned}$$

□

We further note that $D_N(x)$ is periodic with period $2L$ and is an even function.

So far, we have found that the Nth partial sum is given by

$$S_N(x) = \frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\,dy. \tag{2.75}$$

Making the substitution $\xi = y - x$, we have

$$\begin{aligned}
S_N(x) &= \frac{1}{L}\int_{-L-x}^{L-x} D_N(\xi)f(\xi+x)\,d\xi \\
&= \frac{1}{L}\int_{-L}^{L} D_N(\xi)f(\xi+x)\,d\xi. \qquad (2.76)
\end{aligned}$$

In the second integral, we have made use of the fact that $f(x)$ and $D_N(x)$ are periodic with period $2L$ and shifted the interval back to $[-L, L]$.

We now write the integral as the sum of two integrals over positive and negative values of ξ and use the fact that $D_N(x)$ is an even function. Then,

$$\begin{aligned}
S_N(x) &= \frac{1}{L}\int_{-L}^{0} D_N(\xi)f(\xi+x)\,d\xi + \frac{1}{L}\int_{0}^{L} D_N(\xi)f(\xi+x)\,d\xi \\
&= \frac{1}{L}\int_{0}^{L} \left[f(x-\xi)+f(\xi+x)\right] D_N(\xi)\,d\xi. \qquad (2.77)
\end{aligned}$$

We can use this result to study the Gibbs phenomenon whenever it occurs. In particular, we will only concentrate on the earlier example. For this case, we have

$$S_N(x) = \frac{1}{\pi}\int_{0}^{\pi} \left[f(x-\xi)+f(\xi+x)\right] D_N(\xi)\,d\xi \tag{2.78}$$

for

$$D_N(x) = \frac{1}{2} + \sum_{n=1}^{N} \cos nx.$$

Also, one can show that

$$f(x-\xi) + f(\xi + x) = \begin{cases} 2, & 0 \le \xi < x, \\ 0, & x \le \xi < \pi - x, \\ -2, & \pi - x \le \xi < \pi. \end{cases}$$

Thus, we have

$$\begin{aligned} S_N(x) &= \frac{2}{\pi}\int_0^x D_N(\xi)\,d\xi - \frac{2}{\pi}\int_{\pi-x}^{\pi} D_N(\xi)\,d\xi \\ &= \frac{2}{\pi}\int_0^x D_N(z)\,dz + \frac{2}{\pi}\int_0^x D_N(\pi - z)\,dz. \end{aligned} \tag{2.79}$$

Here we made the substitution $z = \pi - \xi$ in the second integral.

The Dirichlet kernel for $L = \pi$ is given by

$$D_N(x) = \frac{\sin(N+\frac{1}{2})x}{2\sin\frac{x}{2}}.$$

For N large, we have $N + \frac{1}{2} \approx N$; and for small x, we have $\sin\frac{x}{2} \approx \frac{x}{2}$. So, under these assumptions,

$$D_N(x) \approx \frac{\sin Nx}{x}.$$

Therefore,

$$S_N(x) \to \frac{2}{\pi}\int_0^x \frac{\sin N\xi}{\xi}\,d\xi \qquad \text{for large } N, \text{ and small } x.$$

If we want to determine the locations of the minima and maxima, where the undershoot and overshoot occur, then we apply the first derivative test for extrema to $S_N(x)$. Thus,

$$\frac{d}{dx}S_N(x) = \frac{2}{\pi}\frac{\sin Nx}{x} = 0.$$

The extrema occur for $Nx = m\pi$, $m = \pm 1, \pm 2, \ldots$. One can show that there is a maximum at $x = \pi/N$ and a minimum for $x = 2\pi/N$. The value for the overshoot can be computed as

$$\begin{aligned} S_N(\pi/N) &= \frac{2}{\pi}\int_0^{\pi/N} \frac{\sin N\xi}{\xi}\,d\xi \\ &= \frac{2}{\pi}\int_0^{\pi} \frac{\sin t}{t}\,dt \\ &= \frac{2}{\pi}\mathrm{Si}(\pi) \\ &= 1.178979744\ldots. \end{aligned} \tag{2.80}$$

Note that this value is independent of N and is given in terms of the sine integral,

$$\mathrm{Si}(x) \equiv \int_0^x \frac{\sin t}{t}\,dt.$$

2.6 Multiple Fourier Series

FUNCTIONS OF SEVERAL VARIABLES CAN HAVE FOURIER SERIES REPRESENTATIONS as well. We motivate this discussion by looking at the vibrations of a rectangular membrane. You can think of this as a drumhead with a rectangular cross section as shown in Figure 2.22. We stretch the membrane over the drumhead and fasten the material to the boundary of the rectangle. The height of the vibrating membrane is described by its height from equilibrium, $u(x, y, t)$.

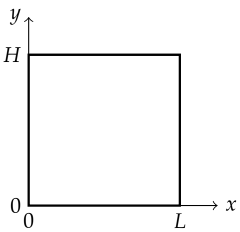


Figure 2.22: The rectangular membrane of length L and width H. There are fixed boundary conditions along the edges.

Example 2.10. *The vibrating rectangular membrane.*

The problem is given by the two-dimensional wave equation in Cartesian coordinates,

$$u_{tt} = c^2(u_{xx} + u_{yy}), \quad t > 0, 0 < x < L, 0 < y < H, \tag{2.81}$$

a set of boundary conditions,

$$\begin{aligned} u(0, y, t) = 0, \quad u(L, y, t) = 0, \quad t > 0, \quad 0 < y < H, \\ u(x, 0, t) = 0, \quad u(x, H, t) = 0, \quad t > 0, \quad 0 < x < L, \end{aligned} \tag{2.82}$$

and a pair of initial conditions (since the equation is second order in time),

$$u(x, y, 0) = f(x, y), \quad u_t(x, y, 0) = g(x, y). \tag{2.83}$$

The general solution is obtained in a course on partial differential equations using what is called the Method of Separation of Variables. One assumes solutions of the form $u(x, y, t) = X(x)Y(y)T(t)$ which satisfy the given boundary conditions, $u(0, y, t) = 0$, $u(L, y, t) = 0$, $u(x, 0, t) = 0$, and $u(x, H, t) = 0$. After some work, one finds the general solution is given by a linear superposition of these product solutions. The general solution is

The general solution for the vibrating rectangular membrane.

$$u(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (a_{nm} \cos \omega_{nm} t + b_{nm} \sin \omega_{nm} t) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}, \tag{2.84}$$

where

$$\omega_{nm} = c\sqrt{\left(\frac{n\pi}{L}\right)^2 + \left(\frac{m\pi}{H}\right)^2}. \tag{2.85}$$

Next, one imposes the initial conditions just like we had indicated in the side note at the beginning of this chapter for the one-dimensional wave equation. The first initial condition is $u(x, y, 0) = f(x, y)$. Setting $t = 0$ in the general solution, we obtain

$$f(x, y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}. \tag{2.86}$$

This is a double Fourier sine series. The goal is to find the unknown coefficients a_{nm}.

The coefficients a_{nm} can be found knowing what we already know about Fourier sine series. We can write the initial condition as the single sum

$$f(x, y) = \sum_{n=1}^{\infty} A_n(y) \sin \frac{n\pi x}{L}, \tag{2.87}$$

where

$$A_n(y) = \sum_{m=1}^{\infty} a_{nm} \sin \frac{m\pi y}{H}. \quad (2.88)$$

These are two Fourier sine series. Recalling from Chapter 2 that the coefficients of Fourier sine series can be computed as integrals, we have

$$\begin{aligned} A_n(y) &= \frac{2}{L} \int_0^L f(x,y) \sin \frac{n\pi x}{L}\, dx, \\ a_{nm} &= \frac{2}{H} \int_0^H A_n(y) \sin \frac{m\pi y}{H}\, dy. \end{aligned} \quad (2.89)$$

Inserting the integral for $A_n(y)$ into that for a_{nm}, we have an integral representation for the Fourier coefficients in the double Fourier sine series,

$$a_{nm} = \frac{4}{LH} \int_0^H \int_0^L f(x,y) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}\, dxdy. \quad (2.90)$$

The Fourier coefficients for the double Fourier sine series.

We can carry out the same process for satisfying the second initial condition, $u_t(x,y,0) = g(x,y)$ for the initial velocity of each point. Inserting the general solution into this initial condition, we obtain

$$g(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_{nm}\omega_{nm} \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}. \quad (2.91)$$

Again, we have a double Fourier sine series. But, now we can quickly determine the Fourier coefficients using the above expression for a_{nm} to find that

$$b_{nm} = \frac{4}{\omega_{nm} LH} \int_0^H \int_0^L g(x,y) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}\, dxdy. \quad (2.92)$$

This completes the full solution of the vibrating rectangular membrane problem. Namely, we have obtained the solution

The full solution of the vibrating rectangular membrane.

$$u(x,y,t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (a_{nm} \cos \omega_{nm} t + b_{nm} \sin \omega_{nm} t) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}, \quad (2.93)$$

where

$$a_{nm} = \frac{4}{LH} \int_0^H \int_0^L f(x,y) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}\, dxdy, \quad (2.94)$$

$$b_{nm} = \frac{4}{\omega_{nm} LH} \int_0^H \int_0^L g(x,y) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}\, dxdy, \quad (2.95)$$

and the angular frequencies are given by

$$\omega_{nm} = c\sqrt{\left(\frac{n\pi}{L}\right)^2 + \left(\frac{m\pi}{H}\right)^2}. \quad (2.96)$$

In this example we encountered a double Fourier sine series. This suggests a function $f(x,y)$ defined on the rectangular region $[0,L] \times [0,H]$ has a double Fourier sine series representation,

$$f(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_{nm} \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H}, \quad (2.97)$$

where

$$b_{nm} = \frac{4}{LH} \int_0^H \int_0^L f(x,y) \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{H} \, dxdy \quad n,m = 1,2,\ldots. \tag{2.98}$$

Of course, we would expect some of the same convergence problems already seen with Fourier series.

Example 2.11. *Find the double Fourier sine series representation of $f(x,y) = xy$ on the unit square.*

For this example, we seek the series representation

$$f(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_{nm} \sin n\pi x \sin m\pi y. \tag{2.99}$$

We compute the Fourier coefficients:

$$\begin{aligned} b_{nm} &= 4 \int_0^1 \int_0^1 f(x,y) \sin n\pi x \sin m\pi y \, dxdy \\ &= 4 \int_0^1 \int_0^1 xy \sin n\pi x \sin m\pi y \, dxdy \\ &= 4 \left(\int_0^1 x \sin n\pi x \, dx \right) \left(\int_0^1 \sin m\pi y \, dy \right) \\ &= 4 \left[-\frac{\cos n\pi}{n\pi} \right] \left[-\frac{\cos m\pi}{m\pi} \right] \\ &= \frac{4(-1)^{n+m}}{nm\pi^2}. \end{aligned}$$

Therefore,

$$xy \sim 4 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{(-1)^{n+m}}{nm\pi^2} \sin n\pi x \sin m\pi y. \tag{2.100}$$

We could just as well seek a double Fourier cosine series on $[0, L] \times [0, H]$,

$$\begin{aligned} f(x,y) \sim & \frac{a_{00}}{4} + \frac{1}{2} \sum_{n=1}^{\infty} a_{n0} \cos \frac{n\pi x}{L} + \frac{1}{2} \sum_{m=1}^{\infty} a_{0m} \cos \frac{m\pi y}{H} \\ & + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} \cos \frac{n\pi x}{L} \cos \frac{m\pi y}{H}, \end{aligned} \tag{2.101}$$

where the Fourier coefficients are given by

$$a_{nm} = \frac{4}{LH} \int_0^H \int_0^L f(x,y) \cos \frac{n\pi x}{L} \cos \frac{m\pi y}{H} \, dxdy, \quad n,m = 0,1,\ldots. \tag{2.102}$$

The more general double Fourier trigonometric series on $[0, L] \times [0, H]$ would take the form

$$\begin{aligned} f(x,y) \sim & \frac{a_{00}}{4} + \frac{1}{2} \sum_{n=1}^{\infty} \left[a_{n0} \cos \frac{2n\pi x}{L} + b_{n0} \sin \frac{2n\pi x}{L} \right] \\ & + \frac{1}{2} \sum_{m=1}^{\infty} \left[a_{0m} \cos \frac{2m\pi y}{H} + b_{0m} \sin \frac{2m\pi y}{H} \right] \\ & + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} \cos \frac{2n\pi x}{L} \cos \frac{2m\pi y}{H}, \end{aligned}$$

$$+\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} b_{nm}\sin\frac{2n\pi x}{L}\sin\frac{2m\pi y}{H},$$
$$+\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} c_{nm}\cos\frac{2n\pi x}{L}\sin\frac{2m\pi y}{H},$$
$$+\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} d_{nm}\sin\frac{2n\pi x}{L}\cos\frac{2m\pi y}{H}. \tag{2.103}$$

The corresponding double Fourier coefficients would take the form you might expect.

Problems

1. Write $y(t) = 3\cos 2t - 4\sin 2t$ in the form $y(t) = A\cos(2\pi f t + \phi)$.

2. Determine the period of the following functions:

a. $f(x) = \cos\frac{x}{3}$.

b. $f(x) = \sin 2\pi x$.

c. $f(x) = \sin 2\pi x - 0.1\cos 3\pi x$.

d. $f(x) = |\sin 5\pi x|$.

e. $f(x) = \cot 2\pi x$.

f. $f(x) = \cos^2\frac{x}{2}$.

g. $f(x) = 3\sin\frac{\pi x}{2} + 2\cos\frac{3\pi x}{4}$.

3. Derive the coefficients b_n in Equation (2.4).

4. Let $f(x)$ be defined for $x \in [-L, L]$. Parseval's identity is given by

$$\frac{1}{L}\int_{-L}^{L} f^2(x)\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2.$$

Assuming the the Fourier series of $f(x)$ converges uniformly in $(-L, L)$, prove Parseval's identity by multiplying the Fourier series representation by $f(x)$ and integrating from $x = -L$ to $x = L$. [In Section 5.6.3 we will encounter Parseval's equality for Fourier transforms which is a continuous version of this identity.]

5. Consider the square wave function

$$f(x) = \begin{cases} 1, & 0 < x < \pi, \\ -1, & \pi < x < 2\pi. \end{cases}$$

a. Find the Fourier series representation of this function and plot the first 50 terms.

b. Apply Parseval's identity in Problem 4 to the result in part a.

c. Use the result of part b to show $\frac{\pi^2}{8} = \sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}$.

6. For the following sets of functions: (i) show that each is orthogonal on the given interval, and (ii) determine the corresponding orthonormal set. [See page 43.]

a. $\{\sin 2nx\}, \quad n = 1, 2, 3, \ldots, \quad 0 \le x \le \pi.$

b. $\{\cos n\pi x\}, \quad n = 0, 1, 2, \ldots, \quad 0 \le x \le 2.$

c. $\{\sin \frac{n\pi x}{L}\}, \quad n = 1, 2, 3, \ldots, \quad x \in [-L, L].$

7. Consider $f(x) = 4\sin^3 2x$.

a. Derive the trigonometric identity giving $\sin^3\theta$ in terms of $\sin\theta$ and $\sin 3\theta$ using DeMoivre's Formula.

b. Find the Fourier series of $f(x) = 4\sin^3 2x$ on $[0, 2\pi]$ without computing any integrals.

8. Find the Fourier series of the following:

a. $f(x) = x$, $x \in [0, 2\pi]$.

b. $f(x) = \frac{x^2}{4}$, $|x| < \pi$.

c. $f(x) = \begin{cases} \frac{\pi}{2}, & 0 < x < \pi, \\ -\frac{\pi}{2}, & \pi < x < 2\pi. \end{cases}$

d. $f(x) = \begin{cases} x, & 0 < x < \pi, \\ \pi, & \pi < x < 2\pi. \end{cases}$

e. $f(x) = \begin{cases} \pi - x, & 0 < x < \pi, \\ 0, & \pi < x < 2\pi. \end{cases}$

9. Find the Fourier series of each function $f(x)$ of period 2π. For each series, plot the Nth partial sum,

$$S_N = \frac{a_0}{2} + \sum_{n=1}^{N} \left[a_n \cos nx + b_n \sin nx\right],$$

for $N = 5, 10, 50$ and describe the convergence (Is it fast? What is it converging to?, etc.) [Some simple Maple code for computing partial sums is shown in the notes.]

a. $f(x) = x$, $|x| < \pi$.

b. $f(x) = |x|$, $|x| < \pi$.

c. $f(x) = \cos x$, $|x| < \pi$.

d. $f(x) = \begin{cases} 0, & -\pi < x < 0, \\ 1, & 0 < x < \pi. \end{cases}$

10. Find the Fourier series of $f(x) = x$ on the given interval. Plot the Nth partial sums and describe what you see.

a. $0 < x < 2$.

b. $-2 < x < 2$.

c. $1 < x < 2$.

11. The result in Problem 8b above gives a Fourier series representation of $\frac{x^2}{4}$. By picking the right value for x and a little arrangement of the series, show that [See Example 2.5.]

a.
$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots.$$

b.
$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots.$$

Hint: Consider how the series in part a. can be used to do this.

c. Use the Fourier series representation result in Problem 8e to obtain the series in part b.

12. Sketch (by hand) the graphs of each of the following functions over four periods. Then sketch the extensions of each of the functions as both an even and odd periodic function. Determine the corresponding Fourier sine and cosine series and verify the convergence to the desired function using Maple.

a. $f(x) = x^2$, $0 < x < 1$.

b. $f(x) = x(2 - x)$, $0 < x < 2$.

c. $f(x) = \begin{cases} 0, & 0 < x < 1, \\ 1, & 1 < x < 2. \end{cases}$

d. $f(x) = \begin{cases} \pi, & 0 < x < \pi, \\ 2\pi - x, & \pi < x < 2\pi. \end{cases}$

13. Consider the function $f(x) = x$, $-\pi < x < \pi$.

a. Show that $x = 2\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\sin nx}{n}$.

b. Integrate the series in part a and show that

$$x^2 = \frac{\pi^2}{3} - 4\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\cos nx}{n^2}.$$

c. Find the Fourier cosine series of $f(x) = x^2$ on $[0, \pi]$ and compare it to the result in part b.

d. Apply Parseval's identity in Problem 4 to the series in part a for $f(x) = x$ on $-\pi < x < \pi$. This gives another means to finding the value $\zeta(2)$, where the Riemann zeta function is defined by

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s}.$$

14. Consider the function $f(x) = x$, $0 < x < 2$.

a. Find the Fourier sine series representation of this function and plot the first 50 terms.

b. Find the Fourier cosine series representation of this function and plot the first 50 terms.

c. Apply Parseval's identity in Problem 4 to the result in part b.

d. Use the result of part c to find the sum $\sum_{n=1}^{\infty} \frac{1}{n^4}$.

15. Differentiate the Fourier sine series term by term in Problem 14. Show that the result is not the derivative of $f(x) = x$.

16. The temperature, $u(x,t)$, of a one-dimensional rod of length L satisfies the heat equation,

$$\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}.$$

a. Show that the general solution,

$$u(x,t) = \sum_{n=0}^{\infty} b_n \sin \frac{n\pi x}{L} e^{-n^2\pi^2 kt/L^2},$$

satisfies the one-dimensional heat equation and the boundary conditions $u(0,t) = 0$ and $u(L,t) = 0$.

b. For $k = 1$ and $L = \pi$, find the solution satisfying the initial condition $u(x,0) = \sin x$. Plot six solutions on the same set of axes for $t \in [0,1]$.

c. For $k = 1$ and $L = 1$, find the solution satisfying the initial condition $u(x,0) = x(1-x)$. Plot six solutions on the same set of axes for $t \in [0,1]$.

17. The height, $u(x,t)$, of a one-dimensional vibrating string of length L satisfies the wave equation,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}.$$

a. Show that the general solution,

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} A_n \cos \frac{n\pi ct}{L} \sin \frac{n\pi x}{L} \\ &\quad + B_n \sin \frac{n\pi ct}{L} \sin \frac{n\pi x}{L}, \end{aligned}$$

satisfies the one-dimensional wave equation and the boundary conditions $u(0,t) = 0$ and $u(L,t) = 0$.

b. For $c = 1$ and $L = 1$, find the solution satisfying the initial conditions $u(x,0) = x(1-x)$ and $u_t(x,0) = x(1-x)$. Plot five solutions for $t \in [0,1]$.

c. For $c = 1$ and $L = 1$, find the solution satisfying the initial condition

$$u(x,0) = \begin{cases} 4x, & 0 \le x \le \frac{1}{4}, \\ \frac{4}{3}(1-x), & \frac{1}{4} \le x \le 1. \end{cases}$$

Plot five solutions for $t \in [0, 0.5]$.

18. Show that

$$u(x,y,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(a_{nm}\cos\omega_{nm}t + b_{nm}\sin\omega_{nm}t)\sin\frac{n\pi x}{L}\sin\frac{m\pi y}{H},$$

where

$$\omega_{nm} = c\sqrt{\left(\frac{n\pi}{L}\right)^2 + \left(\frac{m\pi}{H}\right)^2},$$

satisfies the two-dimensional wave equation

$$u_{tt} = c^2(u_{xx} + u_{yy}), \quad t > 0, 0 < x < L, 0 < y < H,$$

and the boundary conditions,

$$u(0,y,t) = 0, \quad u(L,y,t) = 0, \quad t > 0, \quad 0 < y < H,$$
$$u(x,0,t) = 0, \quad u(x,H,t) = 0, \quad t > 0, \quad 0 < x < L,$$

19. Find the double Fourier sine series representation of the following:

a. $f(x,y) = \sin\pi x\sin 2\pi y$ on $[0,1]\times[0,1]$.

b. $f(x,y) = x(2-x)\sin y$ on $[0,2]\times[0,\pi]$.

c. $f(x,y) = x^2y^3$ on $[0,1]\times[0,1]$.

20. Derive the Fourier coefficients in the double Fourier trigonometric series in Equation (2.103).

3
Generalized Fourier Series and Function Spaces

"Understanding is, after all, what science is all about and science is a great deal more than mindless computation." Sir Roger Penrose (1931-)

In this chapter we provide a glimpse into more general notions for generalized Fourier series and the convergence of Fourier series. It is useful to think about the general context in which one finds oneself when discussing Fourier series and transforms. We can view the sine and cosine functions in the Fourier trigonometric series representations as basis vectors in an infinite dimensional function space. A given function in that space may then be represented as a linear combination over this infinite basis. With this in mind, we might wonder

- Do we have enough basis vectors for the function space?
- Are the infinite series expansions convergent?
- For other other bases, what functions can be represented by such expansions?

In this chapter we touch a little on these ideas, leaving some of the deeper results for more advanced courses.

3.1 Finite Dimensional Vector Spaces

Much of the discussion and terminology that we will use comes from the theory of vector spaces . Until now you may only have dealt with finite dimensional vector spaces. Even then, you might only be comfortable with two and three dimensions. We will review a little of what we know about finite dimensional spaces so that we can introduce more general function spaces later.

The notion of a vector space is a generalization of three dimensional vectors and operations on them. In three dimensions, we have objects called vectors,[1] which are represented by arrows of a specific length and pointing

[1] In introductory physics one defines a vector as any quantity that has both magnitude and direction.

in a given direction. To each vector, we can associate a point in a three dimensional Cartesian system. We just attach the tail of the vector $\mathbf{v}$ to the origin and the head lands at the point (x, y, z).[2] We then use unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ along the coordinate axes to write

[2] In multivariate calculus one concentrates on the component form of vectors. These representations are easily generalized as we will see.

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Having defined vectors, we then learned how to add vectors and multiply vectors by numbers, or scalars. We then learned that there were two types of multiplication of vectors. We could multiply them to get a scalar or a vector. This led to dot products and cross products, respectively. The dot product is useful for determining the length of a vector, the angle between two vectors, if the vectors are perpendicular, or projections of one vector onto another. The cross product is used to produce orthogonal vectors, areas of parallelograms, and volumes of parallelepipeds.

In physics you first learned about vector products when you defined work, $W = \mathbf{F} \cdot \mathbf{r}$. Cross products were useful in describing things like torque, $\tau = \mathbf{r} \times \mathbf{F}$, or the force on a moving charge in a magnetic field, $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$. We will return to these more complicated vector operations later when we need them.

Properties and definition of vector spaces.

The properties three dimensional vectors are generalized to spaces of more than three dimensions in linear algebra courses. The properties roughly outlined above need to be preserved. So, we will start with a space of vectors and the operations of addition and scalar multiplication. We will need a set of scalars, which generally come from some field. However, in our applications the field will either be the set of real numbers or the set of complex numbers.

A field is a set together with two operations, usually addition and multiplication, such that we have

- Closure under addition and multiplication
- Associativity of addition and multiplication
- Commutativity of addition and multiplication
- Additive and multiplicative identity
- Additive and multiplicative inverses
- Distributivity of multiplication over addition

A vector space V over a field F is a set that is closed under addition and scalar multiplication and satisfies the following conditions:

For any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in F$

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
3. There exists a $\mathbf{0}$ such that $\mathbf{0} + \mathbf{v} = \mathbf{v}$.
4. There exists an additive inverse, $-\mathbf{v}$, such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

   There are also several distributive properties:

5. $a(b\mathbf{v}) = (ab)\mathbf{v}$.
6. $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$.
7. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.
8. There exists a multiplicative identity, 1, such that $1(\mathbf{v}) = \mathbf{v}$.

Basis vectors.

For now, we will restrict our examples to two and three dimensions and the field will consist of the set of real numbers.

In three dimensions the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ play an important role. Any vector in the three dimensional space can be written as a linear combination of these vectors,

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

In fact, given any three non-coplanar vectors, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, all vectors can be written as a linear combination of those vectors,

$$\mathbf{v} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3.$$

Such vectors are said to span the space and are called a basis for the space.

We can generalize these ideas. In an n-dimensional vector space any vector in the space can be represented as the sum over n linearly independent vectors (the equivalent of non-coplanar vectors). Such a linearly independent set of vectors $\{\mathbf{v}_j\}_{j=1}^n$ satisfies the condition

n-dimensional vector spaces.

Linearly independent vectors.

$$\sum_{j=1}^{n} c_j\mathbf{v}_j = \mathbf{0} \quad \Leftrightarrow \quad c_j = 0.$$

Note that we will often use summation notation instead of writing out all of the terms in the sum. Also, the symbol $\Leftrightarrow$ means "if and only if," or "is equivalent to." Each side of the symbol implies the other side.

Now we can define a basis for an n-dimensional vector space. We begin with the standard basis in an n-dimensional vector space. It is a generalization of the standard basis in three dimensions ($\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$).

The standard basis vectors, $\mathbf{e}_k$ are a natural generalization of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$.

We define the standard basis with the notation

$$\mathbf{e}_k = (0, \ldots, 0, \underbrace{1}_{k\text{th space}}, 0, \ldots, 0), \quad k = 1, \ldots, n. \tag{3.1}$$

We can expand any $\mathbf{v} \in V$ as

$$\mathbf{v} = \sum_{k=1}^{n} v_k\mathbf{e}_k, \tag{3.2}$$

where the v_k's are called the components of the vector in this basis. Sometimes we will write $\mathbf{v}$ as an n-tuple $(v_1, v_2, \ldots, v_n)$. This is similar to the ambiguous use of (x, y, z) to denote both vectors and points in the three dimensional space.

The only other thing we will need at this point is to generalize the dot product. Recall that there are two forms for the dot product in three dimensions. First, one has that

For more general vector spaces the term inner product is used to generalize the notions of dot and scalar products as we will see below.

$$\mathbf{u} \cdot \mathbf{v} = uv\cos\theta, \tag{3.3}$$

where u and v denote the length of the vectors. The other form is the component form:

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 = \sum_{k=1}^{3} u_kv_k. \tag{3.4}$$

Of course, this form is easier to generalize. So, we define the scalar product between two n-dimensional vectors as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^{n} u_k v_k. \tag{3.5}$$

Actually, there are a number of notations that are used in other texts. One can write the scalar product as $(\mathbf{u}, \mathbf{v})$ or even in the Dirac bra-ket notation[3] $\langle \mathbf{u} | \mathbf{v} \rangle$.

[3] The bra-ket notation was introduced by Paul Adrien Maurice Dirac (1902-1984) in order to facilitate computations of inner products in quantum mechanics. In the notation $\langle \mathbf{u}|\mathbf{v}\rangle$, $\langle \mathbf{u}|$ is the bra and $|\mathbf{v}\rangle$ is the ket. The kets live in a vector space and represented by column vectors with respect to a given basis. The bras live in the dual vector space and are represented by row vectors. The correspondence between bra and kets is $|\mathbf{v}\rangle = \overline{|\mathbf{v}\rangle^T}$. One can operate on kets, $A|\mathbf{v}\rangle$, and make sense out of operations like$\langle \mathbf{u}|A|\mathbf{v}\rangle$, which are used to obtain expectation values associated with the operator. Finally, the outer product, $|\mathbf{v}\rangle\langle \mathbf{v}|$ is used to perform vector space projections.

We note that the (real) scalar product satisfies some simple properties. For vectors $\mathbf{v}$, $\mathbf{w}$ and real scalar α we have

1. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
2. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$.
3. $\langle \alpha\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle$.

While it does not always make sense to talk about angles between general vectors in higher dimensional vector spaces, there is one concept that is useful. It is that of orthogonality, which in three dimensions is another way of saying the vectors are perpendicular to each other. So, we also say that vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. If $\{\mathbf{a}_k\}_{k=1}^n$, is a set of basis vectors such that

$$\langle \mathbf{a}_j, \mathbf{a}_k \rangle = 0, \quad k \neq j,$$

Orthogonal basis vectors.

then it is called an *orthogonal basis*.

If in addition each basis vector is a unit vector, then one has an orthonormal basis. This generalization of the unit basis can be expressed more compactly. We will denote such a basis of unit vectors by $\mathbf{e}_j$ for $j = 1 \ldots n$. Then,

$$\langle \mathbf{e}_j, \mathbf{e}_k \rangle = \delta_{jk}, \tag{3.6}$$

where we have introduced the Kronecker delta (named after Leopold Kronecker (1823-1891))

$$\delta_{jk} \equiv \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases} \tag{3.7}$$

Normalization of vectors.

The process of making vectors have unit length is called normalization. This is simply done by dividing by the length of the vector. Recall that the length of a vector, $\mathbf{v}$, is obtained as $v = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. So, if we want to find a unit vector in the direction of $\mathbf{v}$, then we simply normalize it as

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{v}.$$

Notice that we used a hat to indicate that we have a unit vector. Furthermore, if $\{\mathbf{a}_j\}_{j=1}^n$, is a set of orthogonal basis vectors, then

$$\hat{\mathbf{a}}_j = \frac{\mathbf{a}_i}{\sqrt{\langle \mathbf{a}_j, \mathbf{a}_j \rangle}}, \quad j = 1 \ldots n.$$

Example 3.1. *Find the angle between the vectors* $\mathbf{u} = (-2, 1, 3)$ *and* $\mathbf{v} = (1, 0, 2)$. *we need the lengths of each vector,*

$$u = \sqrt{(-2)^2 + 1^2 + 3^2} = \sqrt{14},$$

$$v = \sqrt{1^2 + 0^2 + 2^2} = \sqrt{5}.$$

We also need the scalar product of these vectors,

$$\mathbf{u} \cdot \mathbf{v} = -2 + 6 = 4.$$

This gives

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{uv} = \frac{4}{\sqrt{5}\sqrt{14}}.$$

So, $\theta = 61.4^\circ$.

Example 3.2. *Normalize the vector* $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

The length of the vector is $v = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$. *So, the unit vector in the direction of* $\mathbf{v}$ *is* $\hat{\mathbf{v}} = \frac{2}{3}\mathbf{i} + \frac{1}{3}\mathbf{j} - \frac{2}{3}\mathbf{k}$.

Let $\{\mathbf{a}_k\}_{k=1}^n$, be a set of orthogonal basis vectors for vector space V. We know that any vector $\mathbf{v}$ can be represented in terms of this basis, $\mathbf{v} = \sum_{k=1}^n v_k \mathbf{a}_k$. If we know the basis and vector, can we find the components, v_k? The answer is yes. We can use the scalar product of $\mathbf{v}$ with each basis element $\mathbf{a}_j$. Using the properties of the scalar product, we have for $j = 1, \ldots, n$

$$\begin{aligned} \langle \mathbf{a}_j, \mathbf{v} \rangle &= \langle \mathbf{a}_j, \sum_{k=1}^n v_k \mathbf{a}_k \rangle \\ &= \sum_{k=1}^n v_k \langle \mathbf{a}_j, \mathbf{a}_k \rangle. \end{aligned} \tag{3.8}$$

Since we know the basis elements, we can easily compute the numbers

$$A_{jk} \equiv \langle \mathbf{a}_j, \mathbf{a}_k \rangle$$

and

$$b_j \equiv \langle \mathbf{a}_j, \mathbf{v} \rangle.$$

Therefore, the system (3.8) for the v_k's is a linear algebraic system, which takes the form

$$b_j = \sum_{k=1}^n A_{jk} v_k. \tag{3.9}$$

We can write this set of equations in a more compact form. The set of numbers A_{jk}, $j, k = 1, \ldots n$ are the elements of an $n \times n$ matrix A with A_{jk} being an element in the jth row and kth column. We write such matrices with the n^2 entries A_{ij} as

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{pmatrix}. \tag{3.10}$$

Also, v_j and b_j can be written as column vectors, $\mathbf{v}$ and $\mathbf{b}$, respectively. Thus, system (3.8) can be written in matrix form as

$$A\mathbf{v} = \mathbf{b}.$$

However, if the basis is orthogonal, then the matrix $A_{jk} \equiv \langle \mathbf{a}_j, \mathbf{a}_k \rangle$ is diagonal,

$$A = \begin{pmatrix} A_{11} & 0 & \dots & \dots & 0 \\ 0 & A_{22} & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & A_{nn} \end{pmatrix}. \tag{3.11}$$

and the system is easily solvable. Recall that two vectors are orthogonal if and only if

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0, \quad i \neq j. \tag{3.12}$$

Thus, in this case we have that

$$\langle \mathbf{a}_j, \mathbf{v} \rangle = v_j \langle \mathbf{a}_j, \mathbf{a}_j \rangle, \quad j = 1, \dots, n. \tag{3.13}$$

or

$$v_j = \frac{\langle \mathbf{a}_j, \mathbf{v} \rangle}{\langle \mathbf{a}_j, \mathbf{a}_j \rangle}. \tag{3.14}$$

In fact, if the basis is orthonormal, i.e., the basis consists of an orthogonal set of unit vectors, then A is the identity matrix and the solution takes on a simpler form:

$$v_j = \langle \mathbf{a}_j, \mathbf{v} \rangle. \tag{3.15}$$

Example 3.3. *Consider the set of vectors* $\mathbf{a}_1 = \mathbf{i} + \mathbf{j}$ *and* $\mathbf{a}_2 = \mathbf{i} - 2\mathbf{j}$.

1. *Determine the matrix elements* $A_{jk} = \langle \mathbf{a}_j, \mathbf{a}_k \rangle$.
2. *Is this an orthogonal basis?*
3. *Expand the vector* $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$ *in the basis* $\{\mathbf{a}_1, \mathbf{a}_2\}$.

First, we compute the matrix elements of A:

$$\begin{aligned} A_{11} &= \langle \mathbf{a}_1, \mathbf{a}_1 \rangle = 2 \\ A_{12} &= \langle \mathbf{a}_1, \mathbf{a}_2 \rangle = -1 \\ A_{21} &= \langle \mathbf{a}_2, \mathbf{a}_1 \rangle = -1 \\ A_{22} &= \langle \mathbf{a}_2, \mathbf{a}_2 \rangle = 5 \end{aligned} \tag{3.16}$$

So,

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}.$$

Since $A_{12} = A_{21} \neq 0$, *the vectors are not orthogonal. However, they are linearly independent. Obviously, if* $c_1 = c_2 = 0$, *then the linear combination*

$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$. *Conversely, we assume that* $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$ *and solve for the coefficients. Inserting the given vectors, we have*

$$\begin{aligned} \mathbf{0} &= c_1(\mathbf{i}+\mathbf{j}) + c_2(\mathbf{i}-2\mathbf{j}) \\ &= (c_1+c_2)\mathbf{i} + (c_1-2c_2)\mathbf{j}. \end{aligned} \tag{3.17}$$

This implies that

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1 - 2c_2 &= 0. \end{aligned} \tag{3.18}$$

Solving this system, one has $c_1 = 0$, $c_2 = 0$. *Therefore, the two vectors are linearly independent.*

In order to determine the components of $\mathbf{v}$ *with respect to the new basis, we need to set up the system (3.8) and solve for the* v_k*'s. We have first,*

$$\begin{aligned} \mathbf{b} &= \begin{pmatrix} \langle \mathbf{a}_1, \mathbf{v}\rangle \\ \langle \mathbf{a}_2, \mathbf{v}\rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle \mathbf{i}+\mathbf{j}, 2\mathbf{i}+3\mathbf{j}\rangle \\ \langle \mathbf{i}-2\mathbf{j}, 2\mathbf{i}+3\mathbf{j}\rangle \end{pmatrix} \\ &= \begin{pmatrix} 5 \\ -4 \end{pmatrix}. \end{aligned} \tag{3.19}$$

So, now we have to solve the system $A\mathbf{v} = \mathbf{b}$ *for* $\mathbf{v}$:

$$\begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \end{pmatrix}. \tag{3.20}$$

We can solve this with matrix methods, $\mathbf{v} = A^{-1}\mathbf{b}$, *or rewrite it as a system of two equations and two unknowns as*

$$\begin{aligned} 2v_1 - v_2 &= 5 \\ -v_1 + 5v_2 &= -4. \end{aligned} \tag{3.21}$$

The solution of this set of algebraic equations is $v_1 = \frac{7}{3}$, $v_2 = -\frac{1}{3}$. *Thus,* $\mathbf{v} = \frac{7}{3}\mathbf{a}_1 - \frac{1}{3}\mathbf{a}_2$. *We will return later to using matrix methods to solve such systems.*

3.2 *Function Spaces*

EARLIER WE STUDIED FINITE DIMENSIONAL VECTOR SPACES. Given a set of basis vectors, $\{\mathbf{a}_k\}_{k=1}^n$, in vector space V, we showed that we can expand any vector $\mathbf{v} \in \mathbf{V}$ in terms of this basis, $\mathbf{v} = \sum_{k=1}^n v_k\mathbf{a}_k$. We then spent some time looking at the simple case of extracting the components v_k of the vector. The keys to doing this simply were to have a scalar product and an orthogonal basis set. These are also the key ingredients that we will need in the infinite dimensional case. In fact, we already did this when we studied Fourier series.

Recall when we found Fourier trigonometric series representations of functions, we started with a function (vector) that we wanted to expand in a set of trigonometric functions (basis) and we sought the Fourier coefficients (components). In this section we will extend our notions from finite dimensional spaces to infinite dimensional spaces and we will develop the needed background in which to think about more general Fourier series expansions. This conceptual framework is very important in other areas in mathematics (such as ordinary and partial differential equations) and physics (such as quantum mechanics and electrodynamics).

We note that the above determination of vector components for finite dimensional spaces is precisely what we did to compute the Fourier coefficients using trigonometric bases. Reading further, you will see how this works.

We will consider various infinite dimensional function spaces. Functions in these spaces would differ by their properties. For example, we could consider the space of continuous functions on [0,1], the space of differentiably continuous functions, or the set of functions integrable from a to b. As you will see, there are many types of function spaces . In order to view these spaces as vector spaces, we must be able to add functions and multiply them by scalars in such as way that they satisfy the definition of a vector space as defined in Chapter 3.

We will also need a scalar product defined on this space of functions. There are several types of scalar products, or inner products, that we can define. An inner product $\langle , \rangle$ on a real vector space V is a mapping from $V \times V$ into R such that for $u, v, w \in V$ and $\alpha \in R$, one has

1. $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0$ iff $v = 0$.
2. $\langle v, w\rangle = \langle w, v\rangle$.
3. $\langle \alpha v, w\rangle = \alpha\langle v, w\rangle$.
4. $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

A real vector space equipped with the above inner product leads to what is called a real inner product space. For complex inner product spaces, the above properties hold with the third property replaced with $\langle v, w\rangle = \overline{\langle w, v\rangle}$.

For the time being, we will only deal with real valued functions and, thus we will need an inner product appropriate for such spaces. One such definition is the following. Let $f(x)$ and $g(x)$ be functions defined on $[a, b]$ and introduce the weight function $\sigma(x) > 0$. Then, we define the inner product, if the integral exists, as

$$\langle f, g\rangle = \int_a^b f(x)g(x)\sigma(x)\,dx. \tag{3.22}$$

The space of square integrable functions.

Spaces in which $\langle f, f\rangle < \infty$ under this inner product are called the space of square integrable functions on (a, b) under weight σ and are denoted as $L^2_\sigma(a, b)$. In what follows, we will assume for simplicity that $\sigma(x) = 1$. This is possible to do using a change of variables.

Now that we have function spaces equipped with an inner product, we seek a basis for the space. For an n-dimensional space we need n basis vectors. For an infinite dimensional space, how many will we need? How do we know when we have enough? We will provide some answers to these questions later.

Let's assume that we have a basis of functions $\{\phi_n(x)\}_{n=1}^{\infty}$. Given a function $f(x)$, how can we go about finding the components of f in this basis? In other words, let

$$f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x).$$

How do we find the c_n's? Does this remind you of Fourier series expansions? Does it remind you of the problem we had earlier for finite dimensional spaces? [You may want to review the discussion at the end of Section 3.1 as you read the next derivation.]

Formally, we take the inner product of f with each ϕ_j and use the properties of the inner product to find

$$\begin{aligned} \langle \phi_j, f \rangle &= \langle \phi_j, \sum_{n=1}^{\infty} c_n \phi_n \rangle \\ &= \sum_{n=1}^{\infty} c_n \langle \phi_j, \phi_n \rangle. \end{aligned} \tag{3.23}$$

If the basis is an orthogonal basis, then we have

$$\langle \phi_j, \phi_n \rangle = N_j \delta_{jn}, \tag{3.24}$$

where δ_{jn} is the Kronecker delta. Recall from Chapter 3 that the Kronecker delta is defined as

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases} \tag{3.25}$$

Continuing with the derivation, we have

For the generalized Fourier series expansion $f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x)$, we have determined the generalized Fourier coefficients to be $c_j = \langle \phi_j, f \rangle / \langle \phi_j, \phi_j \rangle$.

$$\begin{aligned} \langle \phi_j, f \rangle &= \sum_{n=1}^{\infty} c_n \langle \phi_j, \phi_n \rangle \\ &= \sum_{n=1}^{\infty} c_n N_j \delta_{jn}. \end{aligned} \tag{3.26}$$

Expanding the sum, we see that the Kronecker delta picks out one nonzero term:

$$\begin{aligned} \langle \phi_j, f \rangle &= c_1 N_j \delta_{j1} + c_2 N_j \delta_{j2} + \ldots + c_j N_j \delta_{jj} + \ldots \\ &= c_j N_j. \end{aligned} \tag{3.27}$$

So, the expansion coefficients are

$$c_j = \frac{\langle \phi_j, f \rangle}{N_j} = \frac{\langle \phi_j, f \rangle}{\langle \phi_j, \phi_j \rangle} \quad j = 1, 2, \ldots.$$

We summarize this important result:

Generalized Basis Expansion

Let $f(x)$ be represented by an expansion over a basis of orthogonal functions, $\{\phi_n(x)\}_{n=1}^{\infty}$,

$$f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x).$$

Then, the expansion coefficients are formally determined as

$$c_n = \frac{\langle \phi_n, f \rangle}{\langle \phi_n, \phi_n \rangle}.$$

This will be referred to as the general Fourier series expansion and the c_j's are called the Fourier coefficients. Technically, equality only holds when the infinite series converges to the given function on the interval of interest.

Example 3.4. *Find the coefficients of the Fourier sine series expansion of $f(x)$, given by*

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx, \quad x \in [-\pi, \pi].$$

In the last chapter we established that the set of functions $\phi_n(x) = \sin nx$ for $n = 1, 2, \ldots$ is orthogonal on the interval $[-\pi, \pi]$. Recall that using trigonometric identities, we have for $n \neq m$

$$\langle \phi_n, \phi_m \rangle = \int_{-\pi}^{\pi} \sin nx \sin mx \, dx = \pi \delta_{nm}. \tag{3.28}$$

Therefore, the set $\phi_n(x) = \sin nx$ for $n = 1, 2, \ldots$ is an orthogonal set of functions on the interval $[-\pi, \pi]$.

We determine the expansion coefficients using

$$b_n = \frac{\langle \phi_n, f \rangle}{N_n} = \frac{\langle \phi_n, f \rangle}{\langle \phi_n, \phi_n \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx.$$

Does this result look familiar?

Just as with vectors in three dimensions, we can normalize these basis functions to arrive at an orthonormal basis. This is simply done by dividing by the length of the vector. Recall that the length of a vector is obtained as $v = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. In the same way, we define the norm of a function by

$$\|f\| = \sqrt{\langle f, f \rangle}.$$

Note that there are many types of norms, but this induced norm will be sufficient.[4]

For this example, the norms of the basis functions are $\|\phi_n\| = \sqrt{\pi}$. Defining $\psi_n(x) = \frac{1}{\sqrt{\pi}} \phi_n(x)$, we can normalize the ϕ_n's and have obtained an orthonormal basis of functions on $[-\pi, \pi]$.

We can also use the normalized basis to determine the expansion coefficients. In this case we have

$$b_n = \frac{\langle \psi_n, f \rangle}{N_n} = \langle \psi_n, f \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx.$$

[4] The norm defined here is the natural, or induced, norm on the inner product space. Norms are a generalization of the concept of lengths of vectors. Denoting $\|\mathbf{v}\|$ the norm of $\mathbf{v}$, it needs to satisfy the properties

1. $\|\mathbf{v}\| \geq 0$. $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
2. $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|$.
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Examples of common norms are

1. Euclidean norm:
$$\|\mathbf{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}.$$
2. Taxicab norm:
$$\|\mathbf{v}\| = |v_1| + \cdots + |v_n|.$$
3. L^p norm:
$$\|f\| = \left(\int [f(x)]^p \, dx \right)^{\frac{1}{p}}.$$

3.3 *Classical Orthogonal Polynomials*

There are other basis functions that can be used to develop series representations of functions. In this section we introduce the classical orthogonal polynomials. We begin by noting that the sequence of functions $\{1, x, x^2, \ldots\}$ is a basis of linearly independent functions. In fact, by the Stone-Weierstraß Approximation Theorem[5] this set is a basis of $L^2_\sigma(a, b)$, the space of square integrable functions over the interval $[a, b]$ relative to weight $\sigma(x)$. However, we will show that the sequence of functions $\{1, x, x^2, \ldots\}$ does not provide an orthogonal basis for these spaces. We will then proceed to find an appropriate orthogonal basis of functions.

[5] **Stone-Weierstraß Approximation Theorem** Suppose f is a continuous function defined on the interval $[a, b]$. For every $\epsilon > 0$, there exists a polynomial function $P(x)$ such that for all $x \in [a, b]$, we have $|f(x) - P(x)| < \epsilon$. Therefore, every continuous function defined on $[a, b]$ can be uniformly approximated as closely as we wish by a polynomial function.

We are familiar with being able to expand functions over a basis of powers of x, $\{1, x, x^2, \ldots\}$, since these expansions are just Maclaurin series representations of the functions about $x = 0$,

$$f(x) \sim \sum_{n=0}^{\infty} c_n x^n.$$

However, this basis is not an orthogonal set of basis functions. One can easily see this by integrating the product of two even, or two odd, basis functions with $\sigma(x) = 1$ and (a, b)=$(-1, 1)$. For example,

$$\int_{-1}^{1} x^0 x^2 \, dx = \frac{2}{3}.$$

The Gram-Schmidt Orthogonalization Process.

Since we have found that orthogonal bases have been useful in determining the coefficients for expansions of given functions, we might ask, "Given a set of linearly independent basis vectors, can one find an orthogonal basis of the given space?" The answer is yes. We recall from introductory linear algebra, which mostly covers finite dimensional vector spaces, that there is a method for carrying out this so-called Gram-Schmidt Orthogonalization Process. We will review this process for finite dimensional vectors and then generalize to function spaces.

Let's assume that we have three vectors that span the usual three-dimensional space, $\mathbf{R}^3$, given by $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ and shown in Figure 3.1. We seek an orthogonal basis $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, beginning one vector at a time.

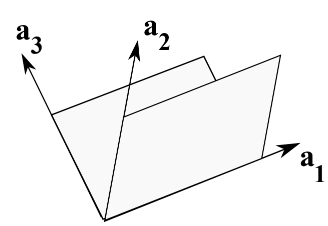


Figure 3.1: The basis $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, of $\mathbf{R}^3$.

First we take one of the original basis vectors, say $\mathbf{a}_1$, and define

$$\mathbf{e}_1 = \mathbf{a}_1.$$

It is sometimes useful to normalize these basis vectors, denoting such a normalized vector with a "hat":

$$\hat{\mathbf{e}}_1 = \frac{\mathbf{e}_1}{e_1},$$

where $e_1 = \sqrt{\mathbf{e}_1 \cdot \mathbf{e}_1}$.

Next, we want to determine an $\mathbf{e}_2$ that is orthogonal to $\mathbf{e}_1$. We take another element of the original basis, $\mathbf{a}_2$. In Figure 3.2 we show the orientation

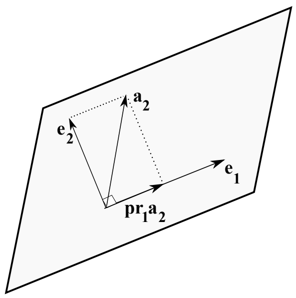


Figure 3.2: A plot of the vectors $\mathbf{e}_1$, $\mathbf{a}_2$, and $\mathbf{e}_2$ needed to find the projection of $\mathbf{a}_2$, on $\mathbf{e}_1$.

of the vectors. Note that the desired orthogonal vector is $\mathbf{e}_2$. We can now write $\mathbf{a}_2$ as the sum of $\mathbf{e}_2$ and the projection of $\mathbf{a}_2$ on $\mathbf{e}_1$. Denoting this projection by $\mathbf{pr}_1\mathbf{a}_2$, we then have

$$\mathbf{e}_2 = \mathbf{a}_2 - \mathbf{pr}_1\mathbf{a}_2. \tag{3.29}$$

Recall the projection of one vector onto another from your vector calculus class.

$$\mathbf{pr}_1\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1. \tag{3.30}$$

This is easily proven by writing the projection as a vector of length $a_2 \cos\theta$ in direction $\hat{\mathbf{e}}_1$, where θ is the angle between $\mathbf{e}_1$ and $\mathbf{a}_2$. Using the definition of the dot product, $\mathbf{a} \cdot \mathbf{b} = ab\cos\theta$, the projection formula follows.

Combining Equations (3.29) and (3.30), we find that

$$\mathbf{e}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1. \tag{3.31}$$

It is a simple matter to verify that $\mathbf{e}_2$ is orthogonal to $\mathbf{e}_1$:

$$\begin{aligned} \mathbf{e}_2 \cdot \mathbf{e}_1 &= \mathbf{a}_2 \cdot \mathbf{e}_1 - \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1 \cdot \mathbf{e}_1 \\ &= \mathbf{a}_2 \cdot \mathbf{e}_1 - \mathbf{a}_2 \cdot \mathbf{e}_1 = 0. \end{aligned} \tag{3.32}$$

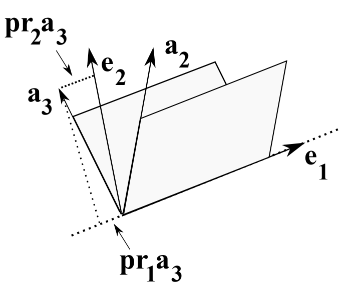


Figure 3.3: A plot of vectors for determining $\mathbf{e}_3$.

Next, we seek a third vector $\mathbf{e}_3$ that is orthogonal to both $\mathbf{e}_1$ and $\mathbf{e}_2$. Pictorially, we can write the given vector $\mathbf{a}_3$ as a combination of vector projections along $\mathbf{e}_1$ and $\mathbf{e}_2$ with the new vector. This is shown in Figure 3.3. Thus, we can see that

$$\mathbf{e}_3 = \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{e}_2}{e_2^2}\mathbf{e}_2. \tag{3.33}$$

Again, it is a simple matter to compute the scalar products with $\mathbf{e}_1$ and $\mathbf{e}_2$ to verify orthogonality.

We can easily generalize this procedure to the N-dimensional case. Let $\mathbf{a}_n$, $n = 1, ..., N$ be a set of linearly independent vectors in $\mathbf{R}^N$. Then, an orthogonal basis can be found by setting $\mathbf{e}_1 = \mathbf{a}_1$ and defining

$$\boxed{\mathbf{e}_n = \mathbf{a}_n - \sum_{j=1}^{n-1} \frac{\mathbf{a}_n \cdot \mathbf{e}_j}{e_j^2}\mathbf{e}_j, \quad n = 2, 3, \ldots, N. \qquad (3.34)}$$

Now we can generalize this idea to (real) function spaces. Let $f_n(x)$, $n \in N_0 = \{0, 1, 2, \ldots\}$, be a linearly independent sequence of continuous functions defined for $x \in [a, b]$. Then, an orthogonal basis of functions, $\phi_n(x)$, $n \in N_0$ can be found and is given by

$$\phi_0(x) = f_0(x)$$

and

$$\boxed{\phi_n(x) = f_n(x) - \sum_{j=0}^{n-1} \frac{\langle f_n, \phi_j \rangle}{\|\phi_j\|^2}\phi_j(x), \quad n = 1, 2, \ldots. \qquad (3.35)}$$

Here we are using inner products relative to weight $\sigma(x)$,

$$\langle f, g \rangle = \int_a^b f(x)g(x)\sigma(x)\,dx. \tag{3.36}$$

Note the similarity between the orthogonal basis in Equation (3.35) and the expression for the finite dimensional case in Equation (3.34).

Example 3.5. *Apply the Gram-Schmidt Orthogonalization Process to the set $f_n(x) = x^n$, $n \in N_0$, when $x \in (-1,1)$ and $\sigma(x) = 1$.*

First, we have $\phi_0(x) = f_0(x) = 1$. Note that

$$\int_{-1}^{1} \phi_0^2(x)\,dx = 2.$$

We could use this result to fix the normalization of the new basis, but we will hold off doing that for now.

Now we compute the second basis element:

$$\begin{aligned} \phi_1(x) &= f_1(x) - \frac{\langle f_1, \phi_0 \rangle}{\|\phi_0\|^2}\phi_0(x) \\ &= x - \frac{\langle x, 1 \rangle}{\|1\|^2} 1 = x, \end{aligned} \tag{3.37}$$

since $\langle x, 1 \rangle$ is the integral of an odd function over a symmetric interval.

For $\phi_2(x)$, we have

$$\begin{aligned} \phi_2(x) &= f_2(x) - \frac{\langle f_2, \phi_0 \rangle}{\|\phi_0\|^2}\phi_0(x) - \frac{\langle f_2, \phi_1 \rangle}{\|\phi_1\|^2}\phi_1(x) \\ &= x^2 - \frac{\langle x^2, 1 \rangle}{\|1\|^2} 1 - \frac{\langle x^2, x \rangle}{\|x\|^2} x \\ &= x^2 - \frac{\int_{-1}^{1} x^2\,dx}{\int_{-1}^{1} dx} \\ &= x^2 - \frac{1}{3}. \end{aligned} \tag{3.38}$$

So far, we have the orthogonal set $\{1, x, x^2 - \frac{1}{3}\}$. If one chooses to normalize these by forcing $\phi_n(1) = 1$, then one obtains the classical Legendre polynomials, $P_n(x)$. Thus,

$$P_2(x) = \frac{1}{2}(3x^2 - 1).$$

Note that this normalization is different from the usual one. In fact, we see that $P_2(x)$ does not have a unit norm,

$$\|P_2\|^2 = \int_{-1}^{1} P_2^2(x)\,dx = \frac{2}{5}.$$

The set of Legendre[6] polynomials is just one set of classical orthogonal polynomials that can be obtained in this way. Many of these special functions had originally appeared as solutions of important boundary value problems in physics. They all have similar properties and we will just elaborate some of these for the Legendre functions in the next section. Others in this group are shown in Table 3.1.

[6] Adrien-Marie Legendre (1752-1833) was a French mathematician who made many contributions to analysis and algebra.

Table 3.1: Common Classical Orthogonal Polynomials with the Interval and Weight Function Used to Define Them.

Polynomial	Symbol	Interval	$\sigma(x)$
Hermite	$H_n(x)$	$(-\infty, \infty)$	e^{-x^2}
Laguerre	$L_n^\alpha(x)$	$[0, \infty)$	e^{-x}
Legendre	$P_n(x)$	(-1,1)	1
Gegenbauer	$C_n^\lambda(x)$	(-1,1)	$(1-x^2)^{\lambda-1/2}$
Tchebychef of the 1st kind	$T_n(x)$	(-1,1)	$(1-x^2)^{-1/2}$
Tchebychef of the 2nd kind	$U_n(x)$	(-1,1)	$(1-x^2)^{-1/2}$
Jacobi	$P_n^{(\nu,\mu)}(x)$	(-1,1)	$(1-x)^\nu(1-x)^\mu$

3.4 *Fourier-Legendre Series*

In the last chapter we saw how useful Fourier series expansions were for solving the heat and wave equations. In the study of partial differential equations in higher dimensions and one finds that problems with spherical symmetry can lead to the series representations in terms of a basis of Legendre polynomials. For example, we could consider the steady-state temperature distribution inside a hemispherical igloo, which takes the form

$$\phi(r, \theta) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos\theta)$$

in spherical coordinates. Evaluating this function at the surface $r = a$ as $\phi(a, \theta) = f(\theta)$, leads to a Fourier-Legendre series expansion of function f :

$$f(\theta) = \sum_{n=0}^{\infty} c_n P_n(\cos\theta),$$

where $c_n = A_n a^n$.

In this section we would like to explore Fourier-Legendre series expansions of functions $f(x)$ defined on $(-1, 1)$:

$$f(x) \sim \sum_{n=0}^{\infty} c_n P_n(x). \tag{3.39}$$

As with Fourier trigonometric series, we can determine the expansion coefficients by multiplying both sides of Equation (3.39) by $P_m(x)$ and integrating for $x \in [-1, 1]$. Orthogonality gives the usual form for the generalized Fourier coefficients,

$$c_n = \frac{\langle f, P_n \rangle}{\|P_n\|^2}, n = 0, 1, \ldots.$$

We will later show that

$$\|P_n\|^2 = \frac{2}{2n+1}.$$

Therefore, the Fourier-Legendre coefficients are

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx. \tag{3.40}$$

3.4.1 *Properties of Legendre Polynomials*

We can do examples of Fourier-Legendre expansions given just a few facts about Legendre polynomials. The first property that the Legendre polynomials have is the Rodrigues Formula:

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n, \quad n \in N_0. \tag{3.41}$$

From the Rodrigues formula, one can show that $P_n(x)$ is an nth degree polynomial. Also, for n odd, the polynomial is an odd function and for n even, the polynomial is an even function.

Example 3.6. *Determine $P_2(x)$ from the Rodrigues Formula:*

$$\begin{aligned} P_2(x) &= \frac{1}{2^2 2!}\frac{d^2}{dx^2}(x^2-1)^2 \\ &= \frac{1}{8}\frac{d^2}{dx^2}(x^4-2x^2+1) \\ &= \frac{1}{8}\frac{d}{dx}(4x^3-4x) \\ &= \frac{1}{8}(12x^2-4) \\ &= \frac{1}{2}(3x^2-1). \end{aligned} \tag{3.42}$$

Note that we get the same result as we found in the last section using orthogonalization.

n	$(x^2-1)^n$	$\frac{d^n}{dx^n}(x^2-1)^n$	$\frac{1}{2^n n!}$	$P_n(x)$
0	1	1	1	1
1	x^2-1	$2x$	$\frac{1}{2}$	x
2	x^4-2x^2+1	$12x^2-4$	$\frac{1}{8}$	$\frac{1}{2}(3x^2-1)$
3	$x^6-3x^4+3x^2-1$	$120x^3-72x$	$\frac{1}{48}$	$\frac{1}{2}(5x^3-3x)$

Table 3.2: Tabular computation of the Legendre polynomials using the Rodrigues Formula.

The first several Legendre polynomials are given in Table 3.2. In Figure 3.4 we show plots of these Legendre polynomials.

The Three-Term Recursion Formula.

All of the classical orthogonal polynomials satisfy a three-term recursion formula (or, recurrence relation or formula). In the case of the Legendre polynomials, we have

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n = 1, 2, \ldots. \tag{3.43}$$

This can also be rewritten by replacing n with $n-1$ as

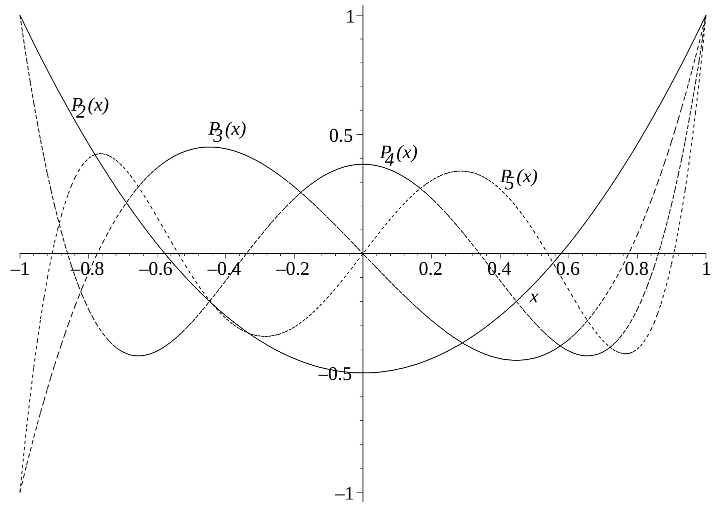
$$(2n-1)xP_{n-1}(x) = nP_n(x) + (n-1)P_{n-2}(x), \quad n = 1, 2, \ldots. \tag{3.44}$$

Example 3.7. *Use the recursion formula to find $P_2(x)$ and $P_3(x)$, given that $P_0(x) = 1$ and $P_1(x) = x$.*

We first begin by inserting $n = 1$ into Equation (3.43):

$$2P_2(x) = 3xP_1(x) - P_0(x) = 3x^2 - 1.$$

Figure 3.4: Plots of the Legendre polynomials $P_2(x)$, $P_3(x)$, $P_4(x)$, and $P_5(x)$.



So, $P_2(x) = \frac{1}{2}(3x^2 - 1)$.

For $n = 2$, *we have*

$$\begin{aligned} 3P_3(x) &= 5xP_2(x) - 2P_1(x) \\ &= \frac{5}{2}x(3x^2-1) - 2x \\ &= \frac{1}{2}(15x^3 - 9x). \end{aligned} \tag{3.45}$$

This gives $P_3(x) = \frac{1}{2}(5x^3 - 3x)$. *These expressions agree with the earlier results.*

The first proof of the three-term recursion formula is based upon the nature of the Legendre polynomials as an orthogonal basis, while the second proof is derived using generating functions.

We will prove the three-term recursion formula in two ways. First, we use the orthogonality properties of Legendre polynomials and the following lemma.

Lemma 3.1. *The leading coefficient of* x^n *in* $P_n(x)$ *is* $\frac{1}{2^n n!}\frac{(2n)!}{n!}$.

Proof. We can prove this using the Rodrigues Formula. First, we focus on the leading coefficient of $(x^2-1)^n$, which is x^{2n}. The first derivative of x^{2n} is $2nx^{2n-1}$. The second derivative is $2n(2n-1)x^{2n-2}$. The jth derivative is

$$\frac{d^j x^{2n}}{dx^j} = [2n(2n-1)\dots(2n-j+1)]x^{2n-j}.$$

Thus, the nth derivative is given by

$$\frac{d^n x^{2n}}{dx^n} = [2n(2n-1)\dots(n+1)]x^n.$$

This proves that $P_n(x)$ has degree n. The leading coefficient of $P_n(x)$ can now be written as

$$\begin{aligned} \frac{[2n(2n-1)\dots(n+1)]}{2^n n!} &= \frac{[2n(2n-1)\dots(n+1)]}{2^n n!}\frac{n(n-1)\dots 1}{n(n-1)\dots 1} \\ &= \frac{1}{2^n n!}\frac{(2n)!}{n!}. \end{aligned} \tag{3.46}$$

□

Theorem 3.1. *Legendre polynomials satisfy the three-term recursion formula*

$$(2n-1)xP_{n-1}(x) = nP_n(x) + (n-1)P_{n-2}(x), \quad n = 1, 2, \ldots. \tag{3.47}$$

Proof. In order to prove the three-term recursion formula, we consider the expression $(2n-1)xP_{n-1}(x) - nP_n(x)$. While each term is a polynomial of degree n, the leading order terms cancel. We need only look at the coefficient of the leading order term first expression. It is

$$\frac{2n-1}{2^{n-1}(n-1)!}\frac{(2n-2)!}{(n-1)!} = \frac{1}{2^{n-1}(n-1)!}\frac{(2n-1)!}{(n-1)!} = \frac{(2n-1)!}{2^{n-1}\left[(n-1)!\right]^2}.$$

The coefficient of the leading term for $nP_n(x)$ can be written as

$$n\frac{1}{2^n n!}\frac{(2n)!}{n!} = n\left(\frac{2n}{2n^2}\right)\left(\frac{1}{2^{n-1}(n-1)!}\right)\frac{(2n-1)!}{(n-1)!}\frac{(2n-1)!}{2^{n-1}\left[(n-1)!\right]^2}.$$

It is easy to see that the leading order terms in the expression $(2n-1)xP_{n-1}(x) - nP_n(x)$ cancel.

The next terms will be of degree $n-2$. This is because the P_n's are either even or odd functions, thus only containing even, or odd, powers of x. We conclude that

$$(2n-1)xP_{n-1}(x) - nP_n(x) = \text{ polynomial of degree } n-2.$$

Therefore, since the Legendre polynomials form a basis, we can write this polynomial as a linear combination of Legendre polynomials:

$$(2n-1)xP_{n-1}(x) - nP_n(x) = c_0P_0(x) + c_1P_1(x) + \ldots + c_{n-2}P_{n-2}(x). \tag{3.48}$$

Multiplying Equation (3.48) by $P_m(x)$ for $m = 0, 1, \ldots, n-3$, integrating from -1 to 1, and using orthogonality, we obtain

$$0 = c_m\|P_m\|^2, \quad m = 0, 1, \ldots, n-3.$$

[Note: $\int_{-1}^{1} x^k P_n(x)\,dx = 0$ for $k \le n-1$. Thus, $\int_{-1}^{1} xP_{n-1}(x)P_m(x)\,dx = 0$ for $m \le n-3$.]

Thus, all these c_m's are zero, leaving Equation (3.48) as

$$(2n-1)xP_{n-1}(x) - nP_n(x) = c_{n-2}P_{n-2}(x).$$

The final coefficient can be found using the normalization condition, $P_n(1) = 1$. Thus, $c_{n-2} = (2n-1) - n = n-1$. □

3.4.2 *The Generating Function for Legendre Polynomials*

A second proof of the three-term recursion formula can be obtained from the generating function of the Legendre polynomials. Many special functions have such generating functions. In this case, it is given by

$$g(x,t) = \frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \quad |x| \le 1, |t| < 1. \tag{3.49}$$

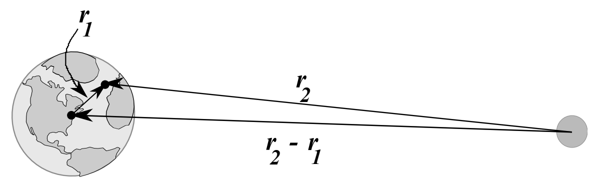


Figure 3.5: The position vectors used to describe the tidal force on the Earth due to the moon.

This generating function occurs often in applications. In particular, it arises in potential theory, such as electromagnetic or gravitational potentials. These potential functions are $\frac{1}{r}$ type functions.

For example, the gravitational potential between the Earth and the moon is proportional to the reciprocal of the magnitude of the difference between their positions relative to some coordinate system. An even better example would be to place the origin at the center of the Earth and consider the forces on the non-pointlike Earth due to the moon. Consider a piece of the Earth at position $\mathbf{r}_1$ and the moon at position $\mathbf{r}_2$ as shown in Figure 3.5. The tidal potential Φ is proportional to

$$\Phi \propto \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|} = \frac{1}{\sqrt{(\mathbf{r}_2 - \mathbf{r}_1)\cdot(\mathbf{r}_2 - \mathbf{r}_1)}} = \frac{1}{\sqrt{r_1^2 - 2r_1r_2\cos\theta + r_2^2}},$$

where θ is the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$.

Typically, one of the position vectors is much larger than the other. Let's assume that $r_1 \ll r_2$. Then, one can write

$$\Phi \propto \frac{1}{\sqrt{r_1^2 - 2r_1r_2\cos\theta + r_2^2}} = \frac{1}{r_2}\frac{1}{\sqrt{1 - 2\frac{r_1}{r_2}\cos\theta + \left(\frac{r_1}{r_2}\right)^2}}.$$

Now, define $x = \cos\theta$ and $t = \frac{r_1}{r_2}$. We then have that the tidal potential is proportional to the generating function for the Legendre polynomials! So, we can write the tidal potential as

$$\Phi \propto \frac{1}{r_2}\sum_{n=0}^{\infty} P_n(\cos\theta)\left(\frac{r_1}{r_2}\right)^n.$$

The first term in the expansion, $\frac{1}{r_2}$, is the gravitational potential that gives the usual force between the Earth and the moon. [Recall that the gravitational potential for mass m at distance r from M is given by $\Phi = -\frac{GMm}{r}$ and that the force is the gradient of the potential, $\mathbf{F} = -\nabla\Phi \propto \nabla\left(\frac{1}{r}\right)$.] The next terms will give expressions for the tidal effects.

Now that we have some idea as to where this generating function might have originated, we can proceed to use it. First of all, the generating function can be used to obtain special values of the Legendre polynomials.

Example 3.8. *Evaluate $P_n(0)$ using the generating function. $P_n(0)$ is found by considering $g(0,t)$. Setting $x = 0$ in Equation (3.49), we have*

$$g(0,t) = \frac{1}{\sqrt{1+t^2}}$$

$$\begin{aligned} &= \sum_{n=0}^{\infty} P_n(0)t^n \\ &= P_0(0) + P_1(0)t + P_2(0)t^2 + P_3(0)t^3 + \dots . \end{aligned} \tag{3.50}$$

We can use the binomial expansion to find the final answer. Namely, we have

$$\frac{1}{\sqrt{1+t^2}} = 1 - \frac{1}{2}t^2 + \frac{3}{8}t^4 + \dots .$$

Comparing these expansions, we have the $P_n(0) = 0$ for n odd and for even integers one can show (see Problem 12) that[7]

$$P_{2n}(0) = (-1)^n \frac{(2n-1)!!}{(2n)!!}, \tag{3.51}$$

where $n!!$ is the double factorial,

$$n!! = \begin{cases} n(n-2)\dots(3)1, & n > 0, odd, \\ n(n-2)\dots(4)2, & n > 0, even, \\ 1, & n = 0, -1. \end{cases} .$$

[7] This example can be finished by first proving that

$$(2n)!! = 2^n n!$$

and

$$(2n-1)!! = \frac{(2n)!}{(2n)!!} = \frac{(2n)!}{2^n n!}.$$

Example 3.9. *Evaluate $P_n(-1)$. This is a simpler problem. In this case we have*

$$g(-1, t) = \frac{1}{\sqrt{1+2t+t^2}} = \frac{1}{1+t} = 1 - t + t^2 - t^3 + \dots .$$

Therefore, $P_n(-1) = (-1)^n$.

Proof of the three-term recursion formula using the generating function.

Example 3.10. *Prove the three-term recursion formula,*

$$(k+1)P_{k+1}(x) - (2k+1)xP_k(x) + kP_{k-1}(x) = 0, \quad k = 1, 2, \dots,$$

using the generating function.

We can also use the generating function to find recurrence relations. To prove the three term recursion (3.43) that we introduced above, then we need only differentiate the generating function with respect to t in Equation (3.49) and rearrange the result. First note that

$$\frac{\partial g}{\partial t} = \frac{x-t}{(1-2xt+t^2)^{3/2}} = \frac{x-t}{1-2xt+t^2} g(x,t).$$

Combining this with

$$\frac{\partial g}{\partial t} = \sum_{n=0}^{\infty} nP_n(x)t^{n-1},$$

we have

$$(x-t)g(x,t) = (1-2xt+t^2) \sum_{n=0}^{\infty} nP_n(x)t^{n-1}.$$

Inserting the series expression for $g(x,t)$ and distributing the sum on the right side, we obtain

$$(x-t) \sum_{n=0}^{\infty} P_n(x)t^n = \sum_{n=0}^{\infty} nP_n(x)t^{n-1} - \sum_{n=0}^{\infty} 2nxP_n(x)t^n + \sum_{n=0}^{\infty} nP_n(x)t^{n+1}.$$

Multiplying out the $x - t$ factor and rearranging, leads to three separate sums:

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} - \sum_{n=0}^{\infty}(2n+1)xP_n(x)t^n + \sum_{n=0}^{\infty}(n+1)P_n(x)t^{n+1} = 0. \quad (3.52)$$

Each term contains powers of t that we would like to combine into a single sum. This is done by reindexing. For the first sum, we could use the new index $k = n-1$. Then, the first sum can be written

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} = \sum_{k=-1}^{\infty}(k+1)P_{k+1}(x)t^k.$$

Using different indices is just another way of writing out the terms. Note that

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} = 0 + P_1(x) + 2P_2(x)t + 3P_3(x)t^2 + \dots$$

and

$$\sum_{k=-1}^{\infty}(k+1)P_{k+1}(x)t^k = 0 + P_1(x) + 2P_2(x)t + 3P_3(x)t^2 + \dots$$

actually give the same sum. The indices are sometimes referred to as dummy indices because they do not show up in the expanded expression and can be replaced with another letter.

If we want to do so, we could now replace all the k's with n's. However, we will leave the k's in the first term and now reindex the next sums in Equation (3.52). The second sum just needs the replacement $n = k$ and the last sum we re-index using $k = n+1$. Therefore, Equation (3.52) becomes

$$\sum_{k=-1}^{\infty}(k+1)P_{k+1}(x)t^k - \sum_{k=0}^{\infty}(2k+1)xP_k(x)t^k + \sum_{k=1}^{\infty} kP_{k-1}(x)t^k = 0. \quad (3.53)$$

We can now combine all the terms, noting the $k = -1$ term is automatically zero and the $k = 0$ terms give

$$P_1(x) - xP_0(x) = 0. \quad (3.54)$$

Of course, we know this already. So, that leaves the $k > 0$ terms:

$$\sum_{k=1}^{\infty}\left[(k+1)P_{k+1}(x) - (2k+1)xP_k(x) + kP_{k-1}(x)\right]t^k = 0. \quad (3.55)$$

Since this is true for all t, the coefficients of the t^k's are zero, or

$$(k+1)P_{k+1}(x) - (2k+1)xP_k(x) + kP_{k-1}(x) = 0, \quad k = 1, 2, \dots.$$

While this is the standard form for the three-term recurrence relation, the earlier form is obtained by setting $k = n-1$.

There are other recursion relations that we list in the box below. Equation (3.56) was derived using the generating function. Differentiating it with respect to x, we find Equation (3.57). Equation (3.58) can be proven using the generating function by differentiating $g(x,t)$ with respect to x and rearranging the resulting infinite series just as in this last manipulation. This will be left as Problem 4. Combining this result with Equation (3.56), we can derive Equations (3.59) and (3.60). Adding and subtracting these equations yields Equations (3.61) and (3.62).

Recursion Formulae for Legendre Polynomials for $n = 1, 2, \ldots$.

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) \tag{3.56}$$

$$(n+1)P'_{n+1}(x) = (2n+1)[P_n(x) + xP'_n(x)] - nP'_{n-1}(x) \tag{3.57}$$

$$P_n(x) = P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x) \tag{3.58}$$

$$P'_{n-1}(x) = xP'_n(x) - nP_n(x) \tag{3.59}$$

$$P'_{n+1}(x) = xP'_n(x) + (n+1)P_n(x) \tag{3.60}$$

$$P'_{n+1}(x) + P'_{n-1}(x) = 2xP'_n(x) + P_n(x). \tag{3.61}$$

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x). \tag{3.62}$$

$$(x^2-1)P'_n(x) = nxP_n(x) - nP_{n-1}(x) \tag{3.63}$$

Finally, Equation (3.63) can be obtained using Equations (3.59) and (3.60). Just multiply Equation (3.59) by x,

$$x^2P'_n(x) - nxP_n(x) = xP'_{n-1}(x).$$

Now use Equation (3.60), but first replace n with $n-1$ to eliminate the $xP'_{n-1}(x)$ term:

$$x^2P'_n(x) - nxP_n(x) = P'_n(x) - nP_{n-1}(x).$$

Rearranging gives the Equation (3.63).

Example 3.11. *Use the generating function to prove*

$$\|P_n\|^2 = \int_{-1}^{1} P_n^2(x)\,dx = \frac{2}{2n+1}.$$

Another use of the generating function is to obtain the normalization constant. This can be done by first squaring the generating function in order to get the products $P_n(x)P_m(x)$, and then integrating over x.

The normalization constant.

Squaring the generating function must be done with care, as we need to make proper use of the dummy summation index. So, we first write

$$\begin{aligned} \frac{1}{1-2xt+t^2} &= \left[\sum_{n=0}^{\infty} P_n(x)t^n\right]^2 \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P_n(x)P_m(x)t^{n+m}. \end{aligned} \tag{3.64}$$

Integrating from $x=-1$ to $x=1$ and using the orthogonality of the Legendre polynomials, we have

$$\int_{-1}^{1} \frac{dx}{1-2xt+t^2} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} t^{n+m}\int_{-1}^{1} P_n(x)P_m(x)\,dx$$

$$= \sum_{n=0}^{\infty} t^{2n} \int_{-1}^{1} P_n^2(x)\,dx. \tag{3.65}$$

However, one can show that[8]

$$\int_{-1}^{1} \frac{dx}{1-2xt+t^2} = \frac{1}{t}\ln\left(\frac{1+t}{1-t}\right).$$

Expanding this expression about $t=0$, we obtain[9]

$$\frac{1}{t}\ln\left(\frac{1+t}{1-t}\right) = \sum_{n=0}^{\infty} \frac{2}{2n+1} t^{2n}.$$

Comparing this result with Equation (3.65), we find that

$$\|P_n\|^2 = \int_{-1}^{1} P_n^2(x)\,dx = \frac{2}{2n+1}. \tag{3.66}$$

[8] You will need the integral

$$\int \frac{dx}{a+bx} = \frac{1}{b}\ln(a+bx) + C.$$

[9] You will need the series expansion

$$\begin{aligned} \ln(1+x) &= \sum_{n=1}^{\infty} (-1)^{n+1}\frac{x^n}{n} \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots. \end{aligned}$$

3.4.3 *The Differential Equation for Legendre Polynomials*

The Legendre polynomials satisfy a second-order linear differential equation. This differential equation occurs naturally in the solution of initial-boundary value problems in three dimensions which possess some spherical symmetry. There are two approaches we could take in showing that the Legendre polynomials satisfy a particular differential equation. Either we can write down the equations and attempt to solve it, or we could use the above properties to obtain the equation. For now, we will seek the differential equation satisfied by $P_n(x)$ using the above recursion relations.

We begin by differentiating Equation (3.63) and using Equation (3.59) to simplify:

$$\begin{aligned} \frac{d}{dx}\left((x^2-1)P_n'(x)\right) &= nP_n(x) + nxP_n'(x) - nP_{n-1}'(x) \\ &= nP_n(x) + n^2P_n(x) \\ &= n(n+1)P_n(x). \end{aligned} \tag{3.67}$$

Therefore, Legendre polynomials, or Legendre functions of the first kind, are solutions of the differential equation

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0.$$

As this is a linear second-order differential equation, we expect two linearly independent solutions. The second solution, called the Legendre function of the second kind, is given by $Q_n(x)$ and is not well behaved at $x=\pm 1$. For example,

$$Q_0(x) = \frac{1}{2}\ln\frac{1+x}{1-x}.$$

We will not need these for physically interesting examples in this book.

A generalization of the Legendre equation is given by $(1-x^2)y'' - 2xy' + \left[n(n+1) - \frac{m^2}{1-x^2}\right]y = 0$. Solutions to this equation, $P_n^m(x)$ and $Q_n^m(x)$, are called the associated Legendre functions of the first and second kind.

3.4.4 Fourier-Legendre Series Examples

With these properties of Legendre functions, we are now prepared to compute the expansion coefficients for the Fourier-Legendre series representation of a given function.

Example 3.12. *Expand $f(x) = x^3$ in a Fourier-Legendre series.*

We simply need to compute

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} x^3 P_n(x)\, dx. \tag{3.68}$$

We first note that

$$\int_{-1}^{1} x^m P_n(x)\, dx = 0 \quad \text{for } m > n.$$

As a result, we have that $c_n = 0$ for $n > 3$. We could just compute $\int_{-1}^{1} x^3 P_m(x)\, dx$ for $m = 0, 1, 2, \ldots$ outright by looking up Legendre polynomials. We note that x^3 is an odd function. So, $c_0 = 0$ and $c_2 = 0$.

This leaves us with only two coefficients to compute. We refer to Table 3.2 and find that

$$c_1 = \frac{3}{2} \int_{-1}^{1} x^4\, dx = \frac{3}{5}$$

$$c_3 = \frac{7}{2} \int_{-1}^{1} x^3 \left[\frac{1}{2}(5x^3 - 3x)\right] dx = \frac{2}{5}.$$

Thus,

$$x^3 = \frac{3}{5} P_1(x) + \frac{2}{5} P_3(x).$$

Of course, this is simple to check using Table 3.2:

$$\frac{3}{5} P_1(x) + \frac{2}{5} P_3(x) = \frac{3}{5} x + \frac{2}{5}\left[\frac{1}{2}(5x^3 - 3x)\right] = x^3.$$

We could have obtained this result without doing any integration. Write x^3 as a linear combination of $P_1(x)$ and $P_3(x)$:

$$\begin{aligned} x^3 &= c_1 x + \frac{1}{2} c_2 (5x^3 - 3x) \\ &= (c_1 - \frac{3}{2} c_2) x + \frac{5}{2} c_2 x^3. \end{aligned} \tag{3.69}$$

Equating coefficients of like terms, we have that $c_2 = \frac{2}{5}$ and $c_1 = \frac{3}{2} c_2 = \frac{3}{5}$.

Example 3.13. *Expand the Heaviside[10] function in a Fourier-Legendre series.*

The Heaviside function is defined as

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{3.70}$$

In this case, we cannot find the expansion coefficients without some integration. We have to compute

$$\begin{aligned} c_n &= \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx \\ &= \frac{2n+1}{2} \int_{0}^{1} P_n(x)\, dx. \end{aligned} \tag{3.71}$$

[10] Oliver Heaviside (1850-1925) was an English mathematician, physicist, and engineer who used complex analysis to study circuits and was a co-founder of vector analysis. The Heaviside function is also called the step function.

We can make use of identity (3.62),

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x), \quad n > 0. \tag{3.72}$$

We have for $n > 0$

$$c_n = \frac{1}{2}\int_0^1 [P'_{n+1}(x) - P'_{n-1}(x)]\, dx = \frac{1}{2}[P_{n-1}(0) - P_{n+1}(0)].$$

For $n = 0$, we have

$$c_0 = \frac{1}{2}\int_0^1 dx = \frac{1}{2}.$$

This leads to the expansion

$$f(x) \sim \frac{1}{2} + \frac{1}{2}\sum_{n=1}^{\infty} [P_{n-1}(0) - P_{n+1}(0)]P_n(x).$$

We still need to evaluate the Fourier-Legendre coefficients

$$c_n = \frac{1}{2}[P_{n-1}(0) - P_{n+1}(0)].$$

Since $P_n(0) = 0$ for n odd, the c_n's vanish for n even. Letting $n = 2k-1$, we re-index the sum, obtaining

$$f(x) \sim \frac{1}{2} + \frac{1}{2}\sum_{k=1}^{\infty} [P_{2k-2}(0) - P_{2k}(0)]P_{2k-1}(x).$$

We can compute the nonzero Fourier coefficients, $c_{2k-1} = \frac{1}{2}[P_{2k-2}(0) - P_{2k}(0)]$, using a result from Problem 12:

$$P_{2k}(0) = (-1)^k \frac{(2k-1)!!}{(2k)!!}. \tag{3.73}$$

Namely, we have

$$\begin{aligned} c_{2k-1} &= \frac{1}{2}[P_{2k-2}(0) - P_{2k}(0)] \\ &= \frac{1}{2}\left[(-1)^{k-1}\frac{(2k-3)!!}{(2k-2)!!} - (-1)^k \frac{(2k-1)!!}{(2k)!!}\right] \\ &= -\frac{1}{2}(-1)^k \frac{(2k-3)!!}{(2k-2)!!}\left[1 + \frac{2k-1}{2k}\right] \\ &= -\frac{1}{2}(-1)^k \frac{(2k-3)!!}{(2k-2)!!}\frac{4k-1}{2k}. \end{aligned} \tag{3.74}$$

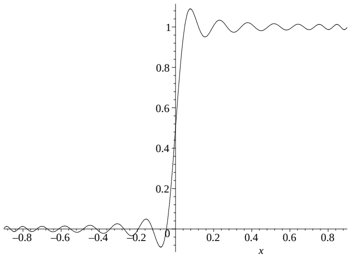


Figure 3.6: Sum of first 21 terms for Fourier-Legendre series expansion of Heaviside function.

Thus, the Fourier-Legendre series expansion for the Heaviside function is given by

$$f(x) \sim \frac{1}{2} - \frac{1}{2}\sum_{n=1}^{\infty} (-1)^n \frac{(2n-3)!!}{(2n-2)!!}\frac{4n-1}{2n} P_{2n-1}(x). \tag{3.75}$$

The sum of the first 21 terms of this series are shown in Figure 3.6. We note the slow convergence to the Heaviside function. Also, we see that the Gibbs phenomenon is present due to the jump discontinuity at $x = 0$. [See Section 2.5.]

3.5 *Gamma Function*

A FUNCTION THAT OFTEN OCCURS IN THE STUDY OF SPECIAL FUNCTIONS is the Gamma function. We will need the Gamma function in the next section on Fourier-Bessel series.

The name and symbol for the Gamma function were first given by Legendre in 1811. However, the search for a generalization of the factorial extends back to the 1720's when Euler provided the first representation of the factorial as an infinite product, later to be modified by others like Gauß, Weierstraß, and Legendre.

For $x > 0$ we define the Gamma function as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt, \quad x > 0. \tag{3.76}$$

The Gamma function is a generalization of the factorial function and a plot is shown in Figure 3.7. In fact, we have

$$\Gamma(1) = 1$$

and

$$\Gamma(x+1) = x\Gamma(x).$$

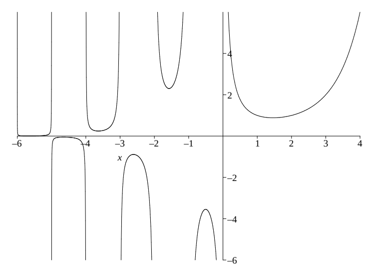


Figure 3.7: Plot of the Gamma function.

The reader can prove this identity by simply performing an integration by parts. (See Problem 7.) In particular, for integers $n \in Z^+$, we then have

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-2) = n(n-1)\cdots 2\Gamma(1) = n!.$$

We can also define the Gamma function for negative, non-integer values of x. We first note that by iteration on $n \in Z^+$, we have

$$\Gamma(x+n) = (x+n-1)\cdots(x+1)x\Gamma(x), \quad x > 0.$$

Solving for $\Gamma(x)$, we then find

$$\Gamma(x) = \frac{\Gamma(x+n)}{(x+n-1)\cdots(x+1)x}, \quad -n < x < 0.$$

Note that the Gamma function is undefined at zero and the negative integers.

Example 3.14. *We now prove that*

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

This is done by direct computation of the integral:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-\frac{1}{2}} e^{-t}\, dt.$$

Letting $t = z^2$, we have

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty e^{-z^2}\, dz.$$

Due to the symmetry of the integrand, we obtain the classic integral

$$\Gamma\left(\frac{1}{2}\right) = \int_{-\infty}^\infty e^{-z^2}\, dz,$$

which can be performed using a standard trick. Consider the integral

$$I = \int_{-\infty}^{\infty} e^{-x^2}\, dx.$$

Then,

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2}\, dx \int_{-\infty}^{\infty} e^{-y^2}\, dy.$$

Note that we changed the integration variable. This will allow us to write this product of integrals as a double integral:

In Example 5.5 we show the more general result:

$$\int_{-\infty}^{\infty} e^{-\beta y^2}\, dy = \sqrt{\frac{\pi}{\beta}}.$$

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\, dxdy.$$

This is an integral over the entire xy-plane. We can transform this Cartesian integration to an integration over polar coordinates. The integral becomes

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\, rdrd\theta.$$

This is simple to integrate and we have $I^2 = \pi$. So, the final result is found by taking the square root of both sides:

$$\Gamma\left(\frac{1}{2}\right) = I = \sqrt{\pi}.$$

In Problem 12, the reader will prove the identity

$$\Gamma(n + \frac{1}{2}) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}.$$

Another useful relation, which we only state, is

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}.$$

The are many other important relations, including infinite products, which we will not need at this point. The reader is encouraged to read about these elsewhere. In the meantime, we move on to the discussion of another important special function in physics and mathematics.

3.6 Fourier-Bessel Series

Bessel functions arise in many problems in physics possessing cylindrical symmetry, such as the vibrations of circular drumheads and the radial modes in optical fibers. They also provide us with another orthogonal set of basis functions.

Bessel functions have a long history and were named after Friedrich Wilhelm Bessel (1784-1846).

The first occurrence of Bessel functions (zeroth order) was in the work of Daniel Bernoulli on heavy chains (1738). More general Bessel functions were studied by Leonhard Euler in 1781 and in his study of the vibrating membrane in 1764. Joseph Fourier found them in the study of heat conduction in solid cylinders and Siméon Poisson (1781-1840) in heat conduction of spheres (1823).

The history of Bessel functions, did not just originate in the study of the wave and heat equations. These solutions originally came up in the study of the Kepler problem, describing planetary motion. According to G. N. Watson in his *Treatise on Bessel Functions*, the formulation and solution of Kepler's Problem was discovered by Joseph-Louis Lagrange (1736-1813), in 1770. Namely, the problem was to express the radial coordinate and what is called the eccentric anomaly, E, as functions of time. Lagrange found expressions for the coefficients in the expansions of r and E in trigonometric functions of time. However, he only computed the first few coefficients. In 1816, Friedrich Wilhelm Bessel (1784-1846) had shown that the coefficients in the expansion for r could be given an integral representation. In 1824, he presented a thorough study of these functions, which are now called Bessel functions.

You might have seen Bessel functions in a course on differential equations as solutions of the differential equation

$$x^2y'' + xy' + (x^2 - p^2)y = 0. \tag{3.77}$$

Solutions to this equation are obtained in the form of series expansions. Namely, one seeks solutions of the form

$$y(x) = \sum_{j=0}^{\infty} a_j x^{j+n}$$

by determining the form the coefficients must take. We will leave this for a homework exercise and simply report the results.

One solution of the differential equation is the *Bessel function of the first kind of order p*, given as

$$y(x) = J_p(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n+p+1)} \left(\frac{x}{2}\right)^{2n+p}. \tag{3.78}$$

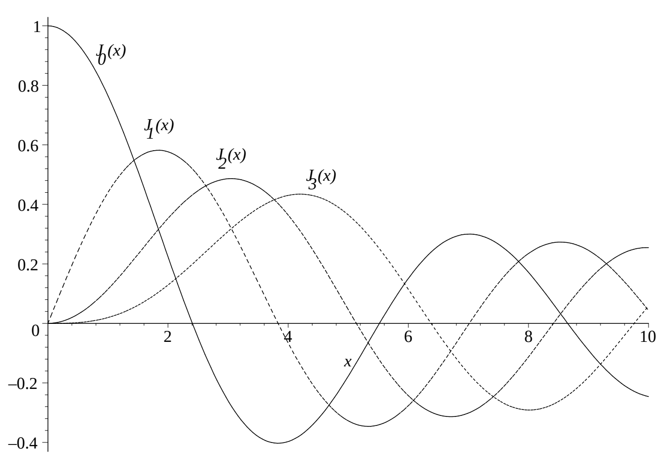


Figure 3.8: Plots of the Bessel functions $J_0(x)$, $J_1(x)$, $J_2(x)$, and $J_3(x)$.

In Figure 3.8, we display the first few Bessel functions of the first kind of integer order. Note that these functions can be described as decaying oscillatory functions.

A second linearly independent solution is obtained for p not an integer as $J_{-p}(x)$. However, for p an integer, the $\Gamma(n+p+1)$ factor leads to evaluations of the Gamma function at zero, or negative integers, when p is negative. Thus, the above series is not defined in these cases.

Another method for obtaining a second linearly independent solution is through a linear combination of $J_p(x)$ and $J_{-p}(x)$ as

$$N_p(x) = Y_p(x) = \frac{\cos \pi p J_p(x) - J_{-p}(x)}{\sin \pi p}. \tag{3.79}$$

These functions are called the Neumann functions, or Bessel functions of the second kind of order p.

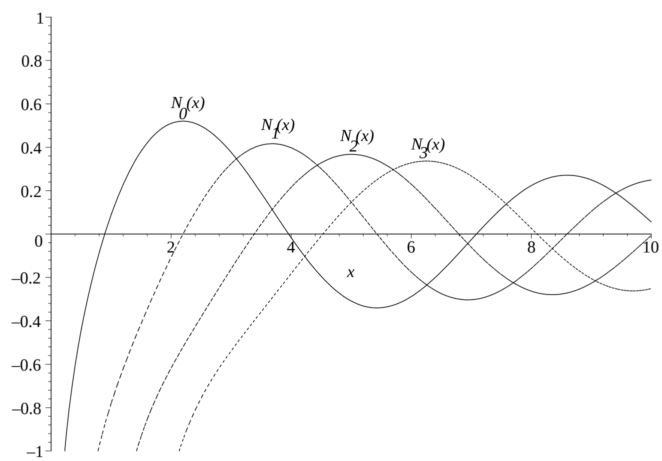


Figure 3.9: Plots of the Neumann functions $N_0(x)$, $N_1(x)$, $N_2(x)$, and $N_3(x)$.

In Figure 3.9, we display the first few Bessel functions of the second kind of integer order. Note that these functions are also decaying oscillatory functions. However, they are singular at $x = 0$.

In many applications, one desires bounded solutions at $x = 0$. These functions do not satisfy this boundary condition. For example, one standard problem is to describe the oscillations of a circular drumhead. For this problem one solves the two dimensional wave equation using separation of variables in cylindrical coordinates. The radial equation leads to a Bessel equation. The Bessel function solutions describe the radial part of the solution and one does not expect a singular solution at the center of the drum. The amplitude of the oscillation must remain finite. Thus, only Bessel functions of the first kind can be used.

Bessel functions satisfy a variety of properties, which we will only list at this time for Bessel functions of the first kind. The reader will have the opportunity to prove these for homework.

Derivative Identities These identities follow directly from the manipulation of the series solution.

$$\frac{d}{dx}\left[x^p J_p(x)\right] = x^p J_{p-1}(x). \tag{3.80}$$

$$\frac{d}{dx}\left[x^{-p} J_p(x)\right] = -x^{-p} J_{p+1}(x). \tag{3.81}$$

Recursion Formulae The next identities follow from adding, or subtracting, the derivative identities.

$$J_{p-1}(x) + J_{p+1}(x) = \frac{2p}{x} J_p(x). \tag{3.82}$$

$$J_{p-1}(x) - J_{p+1}(x) = 2J_p'(x). \tag{3.83}$$

Orthogonality One can recast the Bessel equation into an eigenvalue problem whose solutions form an orthogonal basis of functions on $L^2_x(0, a)$. Using Sturm-Liouville Theory, one can show that

$$\int_0^a x J_p(j_{pn}\frac{x}{a}) J_p(j_{pm}\frac{x}{a})\,dx = \frac{a^2}{2}\left[J_{p+1}(j_{pn})\right]^2 \delta_{n,m}, \tag{3.84}$$

where j_{pn} is the nth root of $J_p(x)$, $J_p(j_{pn}) = 0$, $n = 1, 2, \ldots$. A list of some of these roots is provided in Table 3.3.

n	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1	2.405	3.832	5.136	6.380	7.588	8.771
2	5.520	7.016	8.417	9.761	11.065	12.339
3	8.654	10.173	11.620	13.015	14.373	15.700
4	11.792	13.324	14.796	16.223	17.616	18.980
5	14.931	16.471	17.960	19.409	20.827	22.218
6	18.071	19.616	21.117	22.583	24.019	25.430
7	21.212	22.760	24.270	25.748	27.199	28.627
8	24.352	25.904	27.421	28.908	30.371	31.812
9	27.493	29.047	30.569	32.065	33.537	34.989

Table 3.3: The zeros of Bessel Functions, $J_m(j_{mn}) = 0$.

Generating Function

$$e^{x(t-\frac{1}{t})/2} = \sum_{n=-\infty}^{\infty} J_n(x) t^n, \quad x > 0, t \neq 0. \tag{3.85}$$

Integral Representation

$$J_n(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin\theta - n\theta)\,d\theta, \quad x > 0, n \in \mathrm{Z}. \tag{3.86}$$

Fourier-Bessel Series

Since the Bessel functions are an orthogonal set of functions of a Sturm-Liouville problem, we can expand square integrable functions in this basis. In fact, the Sturm-Liouville problem is given in the form

$$x^2 y'' + xy' + (\lambda x^2 - p^2)y = 0, \quad x \in [0, a], \tag{3.87}$$

satisfying the boundary conditions: $y(x)$ is bounded at $x = 0$ and $y(a) = 0$. The solutions are then of the form $J_p(\sqrt{\lambda}x)$, as can be shown by making the substitution $t = \sqrt{\lambda}x$ in the differential equation. Namely, we let $y(x) = u(t)$ and note that

$$\frac{dy}{dx} = \frac{dt}{dx}\frac{du}{dt} = \sqrt{\lambda}\frac{du}{dt}.$$

Then,

$$t^2u'' + tu' + (t^2 - p^2)u = 0,$$

which has a solution $u(t) = J_p(t)$.

In the study of boundary value problems in differential equations, Sturm-Liouville problems are a bountiful source of basis functions for the space of square integrable functions, as will be seen in the next section.

Using Sturm-Liouville theory, one can show that $J_p(j_{pn}\frac{x}{a})$ is a basis of eigenfunctions and the resulting *Fourier-Bessel series expansion* of $f(x)$ defined on $x \in [0, a]$ is

$$f(x) = \sum_{n=1}^{\infty} c_n J_p(j_{pn}\frac{x}{a}), \tag{3.88}$$

where the Fourier-Bessel coefficients are found using the orthogonality relation as

$$c_n = \frac{2}{a^2\left[J_{p+1}(j_{pn})\right]^2}\int_0^a x f(x) J_p(j_{pn}\frac{x}{a})\,dx. \tag{3.89}$$

Example 3.15. *Expand $f(x) = 1$ for $0 < x < 1$ in a Fourier-Bessel series of the form*

$$f(x) = \sum_{n=1}^{\infty} c_n J_0(j_{0n}x)$$

.

We need only compute the Fourier-Bessel coefficients in Equation (3.89):

$$c_n = \frac{2}{[J_1(j_{0n})]^2}\int_0^1 x J_0(j_{0n}x)\,dx. \tag{3.90}$$

From the identity

$$\frac{d}{dx}\left[x^p J_p(x)\right] = x^p J_{p-1}(x), \tag{3.91}$$

we have

$$\begin{aligned}
\int_0^1 x J_0(j_{0n}x)\,dx &= \frac{1}{j_{0n}^2}\int_0^{j_{0n}} y J_0(y)\,dy \\
&= \frac{1}{j_{0n}^2}\int_0^{j_{0n}} \frac{d}{dy}\left[yJ_1(y)\right]\,dy \\
&= \frac{1}{j_{0n}^2}\left[yJ_1(y)\right]_0^{j_{0n}} \\
&= \frac{1}{j_{0n}}J_1(j_{0n}).
\end{aligned} \tag{3.92}$$

As a result, the desired Fourier-Bessel expansion is given as

$$1 = 2\sum_{n=1}^{\infty}\frac{J_0(j_{0n}x)}{j_{0n}J_1(j_{0n})}, \quad 0 < x < 1. \tag{3.93}$$

In Figure 3.10, we show the partial sum for the first fifty terms of this series. Note once again the slow convergence due to the Gibbs phenomenon.

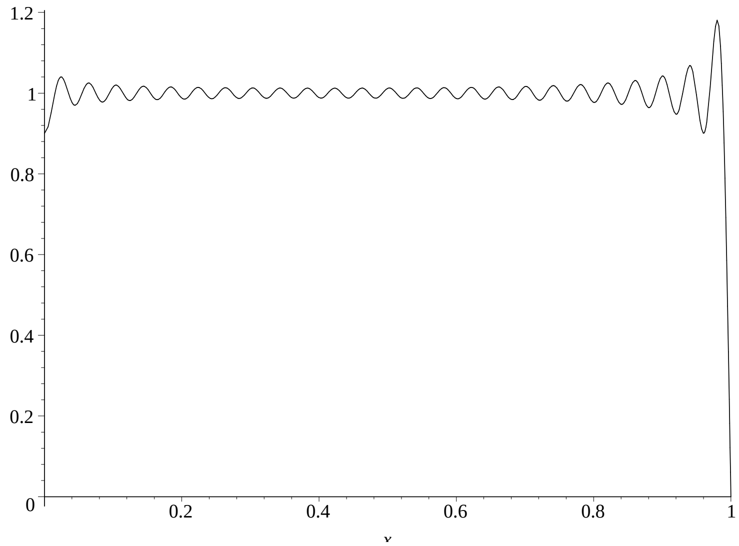


Figure 3.10: Plot of the first 50 terms of the Fourier-Bessel series in Equation (3.93) for $f(x) = 1$ on $0 < x < 1$.

3.7 Appendix: The Least Squares Approximation

In the first section of this chapter, we showed that we can expand functions over an infinite set of basis functions as

$$f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x)$$

and that the generalized Fourier coefficients are given by

$$c_n = \frac{< \phi_n, f >}{< \phi_n, \phi_n >}.$$

In this section we turn to a discussion of approximating $f(x)$ by the partial sums $\sum_{n=1}^{N} c_n \phi_n(x)$ and showing that the Fourier coefficients are the best coefficients minimizing the deviation of the partial sum from $f(x)$. This will lead us to a discussion of the convergence of Fourier series.

More specifically, we set the following goal:

Goal

To find the best approximation of $f(x)$ on $[a, b]$ by $S_N(x) = \sum_{n=1}^{N} c_n \phi_n(x)$ for a set of fixed functions $\phi_n(x)$; i.e., to find the expansion coefficients, c_n, such that $S_N(x)$ approximates $f(x)$ in the least squares sense.

We want to measure the deviation of the finite sum from the given function. Essentially, we want to look at the error made in the approximation. This is done by introducing the mean square deviation:

$$E_N = \int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx,$$

The mean square deviation.

where we have introduced the weight function $\rho(x) > 0$. It gives us a sense as to how close the Nth partial sum is to $f(x)$.

We want to minimize this deviation by choosing the right c_n's. We begin by inserting the partial sums and expand the square in the integrand:

$$\begin{aligned} E_N &= \int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx \\ &= \int_a^b \left[f(x) - \sum_{n=1}^N c_n \phi_n(x) \right]^2 \rho(x)\, dx \\ &= \int_a^b f^2(x)\rho(x)\, dx - 2\int_a^b f(x) \sum_{n=1}^N c_n \phi_n(x) \rho(x)\, dx \\ &\quad + \int_a^b \sum_{n=1}^N c_n \phi_n(x) \sum_{m=1}^N c_m \phi_m(x) \rho(x)\, dx. \end{aligned} \tag{3.94}$$

Looking at the three resulting integrals, we see that the first term is just the inner product of f with itself. The other integrations can be rewritten after interchanging the order of integration and summation. The double sum can be reduced to a single sum using the orthogonality of the ϕ_n's. Thus, we have

$$\begin{aligned} E_N &= < f, f > -2\sum_{n=1}^N c_n < f, \phi_n > + \sum_{n=1}^N \sum_{m=1}^N c_n c_m < \phi_n, \phi_m > \\ &= < f, f > -2\sum_{n=1}^N c_n < f, \phi_n > + \sum_{n=1}^N c_n^2 < \phi_n, \phi_n > . \end{aligned} \tag{3.95}$$

We are interested in finding the coefficients, so we will complete the square in c_n. Focusing on the last two terms, we have

$$\begin{aligned} & -2\sum_{n=1}^N c_n < f, \phi_n > + \sum_{n=1}^N c_n^2 < \phi_n, \phi_n > \\ &= \sum_{n=1}^N < \phi_n, \phi_n > c_n^2 - 2 < f, \phi_n > c_n \\ &= \sum_{n=1}^N < \phi_n, \phi_n > \left[c_n^2 - \frac{2 < f, \phi_n >}{< \phi_n, \phi_n >} c_n \right] \\ &= \sum_{n=1}^N < \phi_n, \phi_n > \left[\left(c_n - \frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 - \left(\frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 \right]. \end{aligned} \tag{3.96}$$

Up to this point, we have shown that the mean square deviation is given as

$$E_N = < f, f > + \sum_{n=1}^N < \phi_n, \phi_n > \left[\left(c_n - \frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 - \left(\frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 \right].$$

So, E_N is minimized by choosing

$$c_n = \frac{< f, \phi_n >}{< \phi_n, \phi_n >}.$$

However, these are the Fourier Coefficients. This minimization is often referred to as Minimization in Least Squares Sense.

Minimization in Least Squares Sense

Inserting the Fourier coefficients into the mean square deviation yields

Bessel's Inequality.

$$0 \leq E_N =< f, f > - \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > .$$

Thus, we obtain Bessel's Inequality:

$$< f, f > \geq \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > .$$

Convergence in the mean.

For convergence, we next let N get large and see if the partial sums converge to the function. In particular, we say that the infinite series converges in the mean if

$$\int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx \ \to 0 \text{ as } N \to \infty.$$

Letting N get large in Bessel's inequality shows that the sum $\sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n >$ converges if

$$(< f, f >= \int_a^b f^2(x) \rho(x)\, dx < \infty.$$

The space of all such f is denoted $L^2_\rho(a, b)$, the space of square integrable functions on (a, b) with weight $\rho(x)$.

From the nth term divergence test from calculus, we know that the convergence of $\sum a_n$ implies that $a_n \to 0$ as $n \to \infty$. Therefore, in this problem, the terms $c_n^2 < \phi_n, \phi_n >$ approach zero as n gets large. This is only possible if the c_n's go to zero as n gets large. Thus, if $\sum_{n=1}^{N} c_n \phi_n$ converges in the mean to f, then $\int_a^b [f(x) - \sum_{n=1}^{N} c_n \phi_n]^2 \rho(x)\, dx$ approaches zero as $N \to \infty$. This implies from the above derivation of Bessel's inequality that

$$< f, f > - \sum_{n=1}^{N} c_n^2 (\phi_n, \phi_n) \to 0.$$

This leads to Parseval's equality:

Parseval's equality.

$$< f, f >= \sum_{n=1}^{\infty} c_n^2 < \phi_n, \phi_n > .$$

Parseval's equality holds if and only if

$$\lim_{N \to \infty} \int_a^b \left(f(x) - \sum_{n=1}^{N} c_n \phi_n(x)\right)^2 \rho(x)\ dx = 0.$$

If this is true for every square integrable function in $L^2_\rho(a, b)$, then the set of functions $\{\phi_n(x)\}_{n=1}^{\infty}$ is said to be complete. One can view these functions as an infinite dimensional basis for the space of square integrable functions on (a, b) with weight $\rho(x) > 0$.

One can extend the above limit $c_n \to 0$ as $n \to \infty$, by assuming that $\frac{\phi_n(x)}{\|\phi_n\|}$ is uniformly bounded and that $\int_a^b |f(x)|\rho(x)\,dx < \infty$. This is the Riemann-Lebesgue Lemma, but will not be proven here.

3.8 Appendix: Convergence of Trigonometric Fourier Series

In this section we list definitions, lemmas and theorems needed to provide convergence arguments for trigonometric Fourier series. We will not attempt to discuss the derivations in depth, but provide enough for the interested reader to see what is involved in establishing convergence.

Definitions

1. For any nonnegative integer k, a function u is C^k if every k-th order partial derivative of u exists and is continuous.
2. For two functions $f$and g defined on an interval $[a, b]$, we will define the **inner product** as $< f, g >= \int_a^b f(x)g(x)\,dx$.
3. A function f is **periodic with period** $\boldsymbol{p}$ if $f(x + p) = f(x)$ for all x.
4. Let f be a function defined on $[-L, L]$ such that $f(-L) = f(L)$. The **periodic extension** $\tilde{f}$of f is the unique periodic function of period $2L$ such that $\tilde{f}(x) = f(x)$for all $x \in [-L, L]$.
5. The expression
$$D_N(x) = \frac{1}{2} + \sum_{n=1}^{N} \cos \frac{n\pi x}{L}$$
is called the ***N*-th Dirichlet Kernel.** [This will be summed later and the sequences of kernels converges to what is called the **Dirac Delta function**.]
6. A sequence of functions $\{s_1(x), s_2(x), \ldots\}$ is said to **converge pointwise** to $f(x)$ on the interval $[-L, L]$ if for each fixed x in the interval,
$$\lim_{N\to\infty} |f(x) - s_N(x)| = 0.$$
7. A sequence of functions $\{s_1(x), s_2(x), \ldots\}$ is said to **converge uniformly** to $f(x)$ on the interval $[-L, L]$ if
$$\lim_{N\to\infty} \left(\max_{|x|\le L} |f(x) - s_N(x)| \right) = 0.$$
8. **One-sided limits**: $f(x_0^+) = \lim_{x\downarrow x_0} f(x)$ and $f(x_0^-) = \lim_{x\uparrow x_0} f(x)$.
9. A function f is **piecewise continuous** on $[a, b]$ if the function satisfies
   a. f is defined and continuous at all but a finite number of points of $[a, b]$.

b. For all $x \in (a,b)$, the limits $f(x^+)$ and $f(x^-)$ exist.

c. $f(a^+)$ and $f(b^-)$ exist.

10. A function is **piecewise** C^1 on $[a,b]$ if $f(x)$ and $f'(x)$ are piecewise continuous on $[a,b]$.

Lemmas

1. **Bessel's Inequality:** Let $f(x)$ be defined on $[-L,L]$ and $\int \_{-L}^L f^2(x)\,dx < \infty$. If the trigonometric Fourier coefficients exist, then $a_0^2 + \sum_{n=1}^N (a_n^2 + b_n^2) \le \frac{1}{L}\int_{-L}^L f^2(x)\,dx$. This follows from the earlier section on the Least Squares Approximation.
2. **Riemann-Lebesgue Lemma:** Under the conditions of Bessel's Inequality, the Fourier coefficients approach zero as $n \to 0$. This is based upon some earlier convergence results seen in Calculus in which one learns for a series of nonnegative terms, $\sum c_n$with $c_n \ge 0$, *if* c_n*does not approach 0 as* $n \to \infty$, *then* $\sum c_n$*does not converge*. Therefore, the contrapositive holds, if $\sum c_n$ *converges, then* $c_n \to 0$ *as* $n \to \infty$. From Bessel's Inequality, we see that when f is square integrable, the series formed by the sums of squares of the Fourier coefficients converges. Therefore, the Fourier coefficients must go to zero as n increases. This is also referred to in the earlier section on the Least Squares Approximation. However, an extension to absolutely integrable functions exists, which is called the Riemann-Lebesgue Lemma.
3. **Green's Formula:** Let $f$and $g$be C^2 functions on $[a,b]$. Then $< f'', g > - < f, g'' > = [f'(x)g(x) - f(x)g'(x)]|_a^b$. [Note: This is just an iteration of integration by parts.]
4. **Special Case of Green's Formula:** Let $f$and $g$be C^2 functions on $[-L,L]$ and both functions satisfy the conditions $f(-L) = f(L)$ and $f'(-L) = f'(L)$. Then $< f'', g > = < f, g'' >$.
5. **Lemma 1:** If g is a periodic function of period $2L$and c any real number, then $\int_{-L+c}^{L+c} g(x)\,dx = \int_{-L}^{L} g(x)\,dx$.
6. **Lemma 2:** Let f be a C^2 function on $[-L,L]$ such that $f(-L) = f(L)$ and $f'(-L) = f'(L)$. Then for $M = \max_{|x|\le L} |f''(x)|$ and $n \ge 1$,

$$|a_n| = \left| \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}\,dx \right| \le \frac{2L^2 M}{n^2\pi^2} \tag{3.97}$$

$$|b_n| = \left| \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}\,dx \right| \le \frac{2L^2 M}{n^2\pi^2}. \tag{3.98}$$

7. **Lemma 3:** For any real θ such that $\sin\frac{\theta}{2} \ne 0$,

$$\frac{1}{2} + \cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{\sin((n+\frac{1}{2})\theta)}{2\sin\frac{\theta}{2}}$$

8. **Lemma 4:** Let $h(x)$ be C^1 on $[-L, L]$. Then

$$\lim_{n\to\infty} \frac{1}{L} \int_{-L}^{L} D_n(x)h(x)\, dx = h(0).$$

Convergence Theorems

1. **Theorem 1.** (Pointwise Convergence) Let f be C^1 on $[-L, L]$ with $f(-L) = f(L), f'(-L) = f'(L)$. Then $FS\, f(x) = f(x)$ for all x in $[-L, L]$.
2. **Theorem 2.** (Uniform Convergence) Let f be C^2 on $[-L, L]$ with $f(-L) = f(L), f'(-L) = f'(L)$. Then $FS\, f(x)$ converges uniformly to $f(x)$. In particular,

$$|f(x) - S_N(x)| \le \frac{4L^2M}{\pi^2 N}$$

for all x in $[-L, L]$, where $M = \max_{|x|\le L} |f''(x)|$.
3. **Theorem 3.** (Piecewise C^1 - Pointwise Convergence) Let f be a piecewise C^1 function on $[-L, L]$. Then $FS\, f(x)$ converges to the periodic extension of

$$f(x) = \begin{cases} \frac{1}{2}[f(x^+) + f(x^-)], & -L < x < L \\ \frac{1}{2}[f(L^+) + f(L^-)], & x = \pm L \end{cases}$$

for all x in $[-L, L]$.
4. **Theorem 4.** (Piecewise C^1 - Uniform Convergence) Let f be a piecewise C^1 function on $[-L, L]$ such that $f(-L) = f(L)$. Then $FS\, f(x)$ converges uniformly to $f(x)$.

Proof of Convergence

We are considering the Fourier series of $f(x)$:

$$FS\, f(x) = a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \right],$$

where the Fourier coefficients are given by

$$\begin{aligned} a_0 &= \tfrac{1}{2L} \int_{-L}^{L} f(x)\, dx, \\ a_n &= \tfrac{1}{L} \int_{-L}^{L} f(x) \cos\tfrac{n\pi x}{L}\, dx, \\ b_n &= \tfrac{1}{L} \int_{-L}^{L} f(x) \sin\tfrac{n\pi x}{L}\, dx. \end{aligned}$$

We are first interested in the pointwise convergence of the infinite series. Thus, we need to look at the partial sums for each x. Writing out the partial sums, inserting the Fourier coefficients and rearranging, we have

$$\begin{aligned}
S_N(x) &= a_0 + \sum_{n=1}^{N}\left[a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right] \\
&= \frac{1}{2L}\int_{-L}^{L} f(y)\,dy + \sum_{n=1}^{N}\left[\left(\frac{1}{L}\int_{-L}^{L} f(y)\cos\frac{n\pi y}{L}\,dy\right)\cos\frac{n\pi x}{L}\right. \\
&\quad \left. + \left(\frac{1}{L}\int_{-L}^{L} f(y)\sin\frac{n\pi y}{L}\,dy.\right)\sin\frac{n\pi x}{L}\right] \\
&= \frac{1}{L}\int_{-L}^{L}\left\{\frac{1}{2} + \sum_{n=1}^{N}\left(\cos\frac{n\pi y}{L}\cos\frac{n\pi x}{L} + \sin\frac{n\pi y}{L}\sin\frac{n\pi x}{L}\right)\right\} f(y)\,dy \\
&= \frac{1}{L}\int_{-L}^{L}\left\{\frac{1}{2} + \sum_{n=1}^{N}\cos\frac{n\pi(y-x)}{L}\right\} f(y)\,dy \\
&\equiv \frac{1}{L}\int_{-L}^{L} D_N(y-x) f(y)\,dy \qquad (3.99)
\end{aligned}$$

Here

$$D_N(x) = \frac{1}{2} + \sum_{n=1}^{N}\cos\frac{n\pi x}{L}$$

is called the ***N*-th Dirichlet Kernel**. What we seek to prove is (**Lemma 4**) that

$$\lim_{N\to\infty}\frac{1}{L}\int_{-L}^{L} D_N(y-x) f(y)\,dy = f(x).$$

[Technically, we need the periodic extension of f.] So, we need to consider the Dirichlet kernel. Then pointwise convergence follows, as $\lim_{N\to\infty} S_N(x) = f(x)$.

Proposition:

$$D_n(x) = \begin{cases} \frac{\sin((n+\frac{1}{2})\frac{\pi x}{L})}{2\sin\frac{\pi x}{2L}}, & \sin\frac{\pi x}{2L} \neq 0 \\ n+\frac{1}{2}, & \sin\frac{\pi x}{2L} = 0 \end{cases}.$$

Proof: Actually, this follows from **Lemma 3**. Let $\theta = \frac{\pi x}{L}$ and multiply $D_n(x)$ by $2\sin\frac{\theta}{2}$ to obtain:

$$\begin{aligned}
2\sin\frac{\theta}{2}D_n(x) &= 2\sin\frac{\theta}{2}\left[\frac{1}{2} + \cos\theta + \cdots + \cos n\theta\right] \\
&= \sin\frac{\theta}{2} + 2\cos\theta\sin\frac{\theta}{2} + 2\cos 2\theta\sin\frac{\theta}{2} + \cdots + 2\cos n\theta\sin\frac{\theta}{2} \\
&= \sin\frac{\theta}{2} + \left(\sin\frac{3\theta}{2} - \sin\frac{\theta}{2}\right) + \left(\sin\frac{5\theta}{2} - \sin\frac{3\theta}{2}\right) + \cdots \\
&\quad + \left(\sin\left((n+\frac{1}{2})\theta\right) - \sin\left((n-\frac{1}{2})\theta\right)\right) \\
&= \sin\left((n+\frac{1}{2})\theta\right). \qquad (3.100)
\end{aligned}$$

Thus,

$$2\sin\frac{\theta}{2}D_n(x) = \sin\left((n+\frac{1}{2})\theta\right),$$

or if $\sin\frac{\theta}{2} \neq 0$,

$$D_n(x) = \frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\frac{\theta}{2}},\ \theta = \frac{\pi x}{L}.$$

If $\sin\frac{\theta}{2} = 0$, then one needs to apply L'Hospital's Rule:

$$\begin{aligned}\lim_{\theta\to 2m\pi}\frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\frac{\theta}{2}} &= \lim_{\theta\to 2m\pi}\frac{(n+\frac{1}{2})\cos\left((n+\frac{1}{2})\theta\right)}{\cos\frac{\theta}{2}} \\ &= \frac{(n+\frac{1}{2})\cos\left(2mn\pi+m\pi\right)}{\cos m\pi} \\ &= n+\frac{1}{2}. \end{aligned} \tag{3.101}$$

As $n \to \infty$, $D_n(x) \to \delta(x)$, the **Dirac delta function,** on the interval $[-L, L]$. In Figures 5.13-5.14 are some plots for $L = \pi$and $n = 25, 50, 100$. Note how the central peaks of $D_N(x)$ grow as N gets large and the values of $D_N(x)$ tend towards zero for nonzero x.

Figure 3.11: Nth Dirichlet Kernel for N=25.

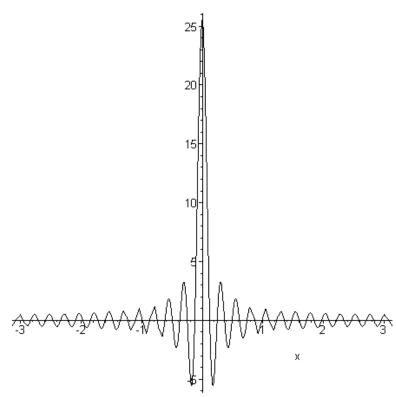
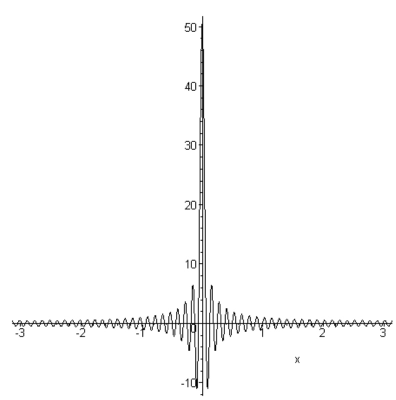


Figure 3.12: Nth Dirichlet Kernel for N=50.



The Dirac delta function can be defined as that quantity satisfying

a. $\delta(x) = 0,\ x \neq 0;$

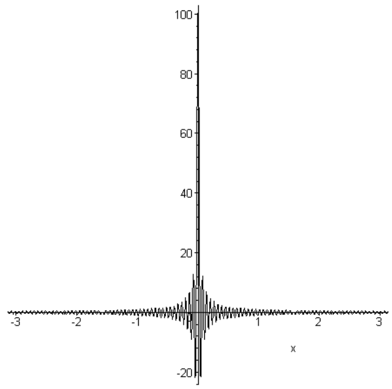


Figure 3.13: Nth Dirichlet Kernel for N=100.

b. $\int_{-\infty}^{\infty} \delta(x)\, dx = 1.$

This generalized function, or **distribution**, also has the property:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\, dx = f(a).$$

Thus, under the appropriate conditions on f, one can show

$$\lim_{N\to\infty} \frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\, dy = f(x).$$

We need to prove **Lemma 4** first.

Proof: Since $\frac{1}{L}\int_{-L}^{L} D_N(x)\, dx = \frac{1}{2L}\int_{-L}^{L} dx = 1$, we have that

$$\begin{aligned} \frac{1}{L}\int_{-L}^{L} D_N(x)h(x)\, dx - h(0) &= \frac{1}{L}\int_{-L}^{L} D_N(x)\left[h(x)-h(0)\right]\, dx \\ &= \frac{1}{2L}\int_{-L}^{L}\left[\cos\frac{n\pi x}{L} + \cot\frac{\pi x}{L}\sin\frac{n\pi x}{L}\right]\left[h(x)-h(0)\right]\, dx. \end{aligned} \tag{3.102}$$

The two terms look like the Fourier coefficients. An application of the Riemann-L:ebesgue Lemma indicates that these coefficients tend to zero as $n\to\infty$, provided the functions being expanded are square integrable and the integrals above exist. The cosine integral follows, but a little work is needed for the sine integral. One can use L'Hospital's Rule with $h \in C^1$.

Now we apply **Lemma 4** to get the convergence from

$$\lim_{N\to\infty} \frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\, dy = f(x).$$

Due to periodicity, we have

$$\frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\, dy = \frac{1}{L}\int_{-L}^{L} D_N(y-x)\tilde{f}(y)\, dy$$

$$
\begin{aligned}
&= \frac{1}{L}\int_{-L+x}^{L+x} D_N(y-x)\tilde{f}(y)\,dy \\
&= \frac{1}{L}\int_{-L}^{L} D_N(z)\tilde{f}(x+z)\,dz. \qquad (3.103)
\end{aligned}
$$

We can apply **Lemma 4** providing $\tilde{f}(z+x)$ is C^1 in z, which is true since f is C^1 and behaves well at $\pm L$.

To prove **Theorem 2** on uniform convergence, we need only combine **Theorem 1** with **Lemma 2**. Then we have,

$$
\begin{aligned}
|f(x)-S_N(x)| &= |f(x)-S_N(x)| \\
&\le \sum_{n=N+1}^{\infty}\left[\left|a_n\cos\frac{n\pi x}{L}\right| + \left|b_n\sin\frac{n\pi x}{L}\right|\right] \\
&\le \sum_{n=N+1}^{\infty}[|a_n|+|b_n|] \qquad (3.104) \\
&\le \frac{4L^2M}{\pi^2}\sum_{n=N+1}^{\infty}\frac{1}{n^2} \\
&\le \frac{4L^2M}{\pi^2 N}. \qquad (3.105)
\end{aligned}
$$

This gives the uniform convergence.

These Theorems can be relaxed to include piecewise C^1 functions. **Lemma 4** needs to be changed for such functions to the result that

$$
\lim_{n\to\infty}\frac{1}{L}\int_{-L}^{L} D_n(x)h(x)\,dx = \frac{1}{2}[h(0^+)+h(0^-)]
$$

by splitting the integral into integrals over $[-L,0]$, $[0,L]$ and applying a one-sided L'Hospital's Rule. Proving uniform convergence under the conditions in **Theorem 4** takes a little more effort, but it can be done.

Problems

1. Consider the set of vectors $(-1,1,1)$, $(1,-1,1)$, $(1,1,-1)$.

a. Use the Gram-Schmidt process to find an orthonormal basis for R^3 using this set in the given order.

b. What do you get if you do reverse the order of these vectors?

2. Use the Gram-Schmidt process to find the first four orthogonal polynomials satisfying the following:

a. Interval: $(-\infty,\infty)$ Weight Function: e^{-x^2}.

b. Interval: $(0,\infty)$ Weight Function: e^{-x}.

3. Find $P_4(x)$ using

a. The Rodrigues Formula in Equation (3.41).

b. The three-term recursion formula in Equation (3.43).

4. In Equations (3.56) through (3.63) we provide several identities for Legendre polynomials. Derive the results in Equations (3.57) through (3.63) as described in the text. Namely,

a. Differentiating Equation (3.56) with respect to x, derive Equation (3.57).

b. Derive Equation (3.58) by differentiating $g(x,t)$ with respect to x and rearranging the resulting infinite series.

c. Combining the previous result with Equation (3.56), derive Equations (3.59) and (3.60).

d. Adding and subtracting Equations (3.59) and (3.60), obtain Equations (3.61) and (3.62).

e. Derive Equation (3.63) using some of the other identities.

5. Use the recursion relation (3.43) to evaluate $\int_{-1}^{1} xP_n(x)P_m(x)\,dx$, $n \le m$.

6. Expand the following in a Fourier-Legendre series for $x \in (-1,1)$.

a. $f(x) = x^2$.

b. $f(x) = 5x^4 + 2x^3 - x + 3$.

c. $f(x) = \begin{cases} -1, & -1 < x < 0, \\ 1, & 0 < x < 1. \end{cases}$

d. $f(x) = \begin{cases} x, & -1 < x < 0, \\ 0, & 0 < x < 1. \end{cases}$

7. Use integration by parts to show $\Gamma(x+1) = x\Gamma(x)$.

8. Prove the double factorial identities:

$$(2n)!! = 2^n n!$$

and

$$(2n-1)!! = \frac{(2n)!}{2^n n!}.$$

9. Express the following as Gamma functions. Namely, noting the form $\Gamma(x+1) = \int_0^\infty t^x e^{-t}\,dt$ and using an appropriate substitution, each expression can be written in terms of a Gamma function.

a. $\int_0^\infty x^{2/3} e^{-x}\,dx$.

b. $\int_0^\infty x^5 e^{-x^2}\,dx$.

c. $\int_0^1 \left[\ln\left(\frac{1}{x}\right)\right]^n dx$.

10. The coefficients C_k^p in the binomial expansion for $(1+x)^p$ are given by

$$C_k^p = \frac{p(p-1)\cdots(p-k+1)}{k!}.$$

a. Write C_k^p in terms of Gamma functions.

b. For $p = 1/2$, use the properties of Gamma functions to write $C_k^{1/2}$ in terms of factorials.

c. Confirm your answer in part b by deriving the Maclaurin series expansion of $(1+x)^{1/2}$.

11. The Hermite polynomials, $H_n(x)$, satisfy the following:

i. $< H_n, H_m >= \int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx = \sqrt{\pi} 2^n n! \delta_{n,m}$.

ii. $H_n'(x) = 2nH_{n-1}(x)$.

iii. $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$.

iv. $H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right)$.

Using these, show that

a. $H_n'' - 2xH_n' + 2nH_n = 0$. [Use properties ii. and iii.]

b. $\int_{-\infty}^{\infty} xe^{-x^2} H_n(x) H_m(x)\, dx = \sqrt{\pi} 2^{n-1} n! \left[\delta_{m,n-1} + 2(n+1)\delta_{m,n+1}\right]$. [Use properties i. and iii.]

c. $H_n(0) = \begin{cases} 0, & n \text{ odd}, \\ (-1)^m \frac{(2m)!}{m!}, & n = 2m. \end{cases}$ [Let $x = 0$ in iii. and iterate. Note from iv. that $H_0(x) = 1$ and $H_1(x) = 2x$.]

12. In Maple one can type **simplify(LegendreP(2*n-2,0)-LegendreP(2*n,0));** to find a value for $P_{2n-2}(0) - P_{2n}(0)$. It gives the result in terms of Gamma functions. However, in Example 3.13 for Fourier-Legendre series, the value is given in terms of double factorials! So, we have

$$P_{2n-2}(0) - P_{2n}(0) = \frac{\sqrt{\pi}(4n-1)}{2\Gamma(n+1)\Gamma\left(\frac{3}{2}-n\right)} = (-1)^{n-1} \frac{(2n-3)!!}{(2n-2)!!} \frac{4n-1}{2n}.$$

You will verify that both results are the same by doing the following:

a. Prove that $P_{2n}(0) = (-1)^n \frac{(2n-1)!!}{(2n)!!}$ using the generating function and a binomial expansion.

b. Prove that $\Gamma\left(n+\frac{1}{2}\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}$ using $\Gamma(x) = (x-1)\Gamma(x-1)$ and iteration.

c. Verify the result from Maple that $P_{2n-2}(0) - P_{2n}(0) = \frac{\sqrt{\pi}(4n-1)}{2\Gamma(n+1)\Gamma\left(\frac{3}{2}-n\right)}$.

d. Can either expression for $P_{2n-2}(0) - P_{2n}(0)$ be simplified further?

13. A solution of Bessel's equation, $x^2y'' + xy' + (x^2 - n^2)y = 0$, , can be found using the guess $y(x) = \sum_{j=0}^{\infty} a_j x^{j+n}$. One obtains the recurrence relation $a_j = \frac{-1}{j(2n+j)} a_{j-2}$. Show that for $a_0 = (n!2^n)^{-1}$, we get the Bessel function of the first kind of order n from the even values $j = 2k$:

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{x}{2}\right)^{n+2k}.$$

14. Use the infinite series in Problem 13 to derive the derivative identities (3.80) and (3.81):

a. $\frac{d}{dx}\left[x^n J_n(x)\right] = x^n J_{n-1}(x)$.
b. $\frac{d}{dx}\left[x^{-n} J_n(x)\right] = -x^{-n} J_{n+1}(x)$.

15. Prove the following identities based on those in Problem 14.

a. $J_{p-1}(x) + J_{p+1}(x) = \frac{2p}{x} J_p(x)$.
b. $J_{p-1}(x) - J_{p+1}(x) = 2J_p'(x)$.

16. Use the derivative identities of Bessel functions, (3.80) and (3.81), and integration by parts to show that

$$\int x^3 J_0(x)\, dx = x^3 J_1(x) - 2x^2 J_2(x) + C.$$

17. Use the generating function to find $J_n(0)$ and $J_n'(0)$.

18. Bessel functions $J_p(\lambda x)$ are solutions of $x^2 y'' + xy' + (\lambda^2 x^2 - p^2)y = 0$. Assume that $x \in (0,1)$ and that $J_p(\lambda) = 0$ and $J_p(0)$ is finite.

a. Show that this equation can be written in the form

$$\frac{d}{dx}\left(x\frac{dy}{dx}\right) + (\lambda^2 x - \frac{p^2}{x})y = 0.$$

This is the standard Sturm-Liouville form for Bessel's equation.

b. Prove that

$$\int_0^1 x J_p(\lambda x) J_p(\mu x)\, dx = 0, \quad \lambda \neq \mu$$

by considering

$$\int_0^1 \left[J_p(\mu x)\frac{d}{dx}\left(x\frac{d}{dx}J_p(\lambda x)\right) - J_p(\lambda x)\frac{d}{dx}\left(x\frac{d}{dx}J_p(\mu x)\right)\right] dx.$$

Thus, the solutions corresponding to different eigenvalues (λ, μ) are orthogonal.

c. Prove that

$$\int_0^1 x\left[J_p(\lambda x)\right]^2\, dx = \frac{1}{2}J_{p+1}^2(\lambda) = \frac{1}{2}J_p'^2(\lambda).$$

19. We can rewrite Bessel functions, $J_\nu(x)$, in a form which will allow the order to be non-integer by using the gamma function. You will need the results from Problem 12b for $\Gamma\left(k+\frac{1}{2}\right)$.

a. Extend the series definition of the Bessel function of the first kind of order ν, $J_\nu(x)$, for $\nu \geq 0$ by writing the series solution for $y(x)$ in Problem 13 using the gamma function.
b. Extend the series to $J_{-\nu}(x)$, for $\nu \geq 0$. Discuss the resulting series and what happens when ν is a positive integer.

c. Use these results to obtain the closed form expressions

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x,$$

$$J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}} \cos x.$$

d. Use the results in part c with the recursion formula for Bessel functions to obtain a closed form for $J_{3/2}(x)$.

20. In this problem you will derive the expansion

$$x^2 = \frac{c^2}{2} + 4\sum_{j=2}^{\infty} \frac{J_0(\alpha_j x)}{\alpha_j^2 J_0(\alpha_j c)}, \quad 0 < x < c,$$

where the $\alpha_j's$ are the positive roots of $J_1(\alpha c) = 0$, by following the below steps.

a. List the first five values of α for $J_1(\alpha c) = 0$ using Table 3.3 and Figure 3.8. [Note: Be careful in determining α_1.]

b. Show that $\|J_0(\alpha_1 x)\|^2 = \frac{c^2}{2}$. Recall,

$$\|J_0(\alpha_j x)\|^2 = \int_0^c x J_0^2(\alpha_j x)\, dx.$$

c. Show that $\|J_0(\alpha_j x)\|^2 = \frac{c^2}{2}\left[J_0(\alpha_j c)\right]^2$, $j = 2, 3, \ldots$. (This is the most involved step.) First note from Problem 18 that $y(x) = J_0(\alpha_j x)$ is a solution of

$$x^2 y'' + xy' + \alpha_j^2 x^2 y = 0.$$

i. Verify the Sturm-Liouville form of this differential equation: $(xy')' = -\alpha_j^2 xy$.

ii. Multiply the equation in part i. by $y(x)$ and integrate from $x = 0$ to $x = c$ to obtain

$$\begin{aligned} \int_0^c (xy')' y\, dx &= -\alpha_j^2 \int_0^c xy^2\, dx \\ &= -\alpha_j^2 \int_0^c x J_0^2(\alpha_j x)\, dx. \end{aligned} \quad (3.106)$$

iii. Noting that $y(x) = J_0(\alpha_j x)$, integrate the left hand side by parts and use the following to simplify the resulting equation.

1. $J_0'(x) = -J_1(x)$ from Equation (3.81).
2. Equation (3.84).
3. $J_2(\alpha_j c) + J_0(\alpha_j c) = 0$ from Equation (3.82).

iv. Now you should have enough information to complete this part.

d. Use the results from parts b and c and Problem 16 to derive the expansion coefficients for

$$x^2 = \sum_{j=1}^{\infty} c_j J_0(\alpha_j x)$$

in order to obtain the desired expansion.

4 Complex Analysis

"He is not a true man of science who does not bring some sympathy to his studies, and expect to learn something by behavior as well as by application. It is childish to rest in the discovery of mere coincidences, or of partial and extraneous laws. The study of geometry is a petty and idle exercise of the mind, if it is applied to no larger system than the starry one. Mathematics should be mixed not only with physics but with ethics; that is mixed mathematics. The fact which interests us most is the life of the naturalist. The purest science is still biographical." Henry David Thoreau (1817 - 1862)

We have seen that we can seek the frequency content of a signal $f(t)$ defined on an interval $[0, T]$ by looking for the the Fourier coefficients in the Fourier series expansion

In this chapter we introduce complex numbers and complex functions. We will later see that the rich structure of complex functions will lead to a deeper understanding of analysis, interesting techniques for computing integrals, and a natural way to express analog and discrete signals.

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{2\pi nt}{T} + b_n \sin\frac{2\pi nt}{T}.$$

The coefficients can be written as integrals such as

$$a_n = \frac{2}{T}\int_0^T f(t)\cos\frac{2\pi nt}{T}\,dt.$$

However, we have also seen that, using Euler's Formula, trigonometric functions can be written in a complex exponential form,

$$\cos\frac{2\pi nt}{T} = \frac{e^{2\pi int/T} + e^{-2\pi int/T}}{2}.$$

We can use these ideas to rewrite the trigonometric Fourier series as a sum over complex exponentials in the form

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi int/T},$$

where the Fourier coefficients now take the form

$$c_n = \int_0^T f(t)e^{-2\pi int/T}\,dt.$$

This representation will be useful in the analysis of analog signals, which are ideal signals defined on an infinite interval and containing a continuum of frequencies. We will see the above sum become an integral and will naturally find ourselves needing to work with functions of complex variables and performing integrals of complex functions.

With this ultimate goal in mind, we will now take a tour of complex analysis. We will begin with a review of some facts about complex numbers and then introduce complex functions. This will lead us to the calculus of functions of a complex variable, including differentiation and integration of complex functions.

4.1 *Complex Numbers*

Complex numbers were first introduced in order to solve some simple problems. The history of complex numbers only extends about five hundred years. In essence, it was found that we need to find the roots of equations such as $x^2 + 1 = 0$. The solution is $x = \pm\sqrt{-1}$. Due to the usefulness of this concept, which was not realized at first, a special symbol was introduced - the imaginary unit, $i = \sqrt{-1}$. In particular, Girolamo Cardano $(1501 - 1576)$ was one of the first to use square roots of negative numbers when providing solutions of cubic equations. However, complex numbers did not become an important part of mathematics or science until the late seventh century after people like Abraham de Moivre (1667 - 1754), the Bernoulli[1] family, and Leonhard Euler (1707 - 1783) took them seriously.

[1] The Bernoulli's were a family of Swiss mathematicians spanning three generations. It all started with Jacob Bernoulli (1654 - 1705) and his brother Johann Bernoulli (1667 - 1748). Jacob had a son, Nicolaus Bernoulli (1687 - 1759) and Johann (1667 - 1748) had three sons, Nicolaus Bernoulli II (1695 - 1726), Daniel Bernoulli (1700 - 1872), and Johann Bernoulli II (1710 - 1790). The last generation consisted of Johann II's sons, Johann Bernoulli III (1747 - 1807) and Jacob Bernoulli II (1759 - 1789). Johann, Jacob and Daniel Bernoulli were the most famous of the Bernoulli's. Jacob studied with Leibniz, Johann studied under his older brother and later taught Leonhard Euler and Daniel Bernoulli, who is known for his work in hydrodynamics.

A complex number is a number of the form $z = x + iy$, where x and y are real numbers. x is called the real part of z and y is the imaginary part of z. Examples of such numbers are $3 + 3i$, $-1i = -i$, $4i$ and 5. Note that $5 = 5 + 0i$ and $4i = 0 + 4i$.

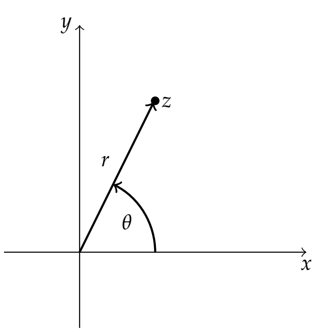


Figure 4.1: The Argand diagram for plotting complex numbers in the complex z-plane.

There is a geometric representation of complex numbers in a two-dimensional plane, known as the complex plane C. This is given by the Argand diagram as shown in Figure 4.1. Here we can think of the complex number $z = x + iy$ as a point (x, y) in the z-complex plane or as a vector. The magnitude, or length, of this vector is called the complex modulus of z, denoted by $|z| = \sqrt{x^2 + y^2}$. We can also use the geometric picture to develop a polar representation of complex numbers. From Figure 4.1 we can see that in terms of r and θ, we have that

The complex modulus, $|z| = \sqrt{x^2 + y^2}$.

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta. \end{aligned} \tag{4.1}$$

Thus, using Euler's Formula (Example 1.34), we have

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}. \tag{4.2}$$

Complex numbers can be represented in rectangular (Cartesian), $z = x + iy$, or polar form, $z = re^{i\theta}$. Here we define the argument, θ, and modulus, $|z| = r$, of complex numbers.

So, given r and θ we have $z = re^{i\theta}$. However, given the Cartesian form, $z = x + iy$, we can also determine the polar form, since

$$\begin{aligned} r &= \sqrt{x^2 + y^2}, \\ \tan\theta &= \frac{y}{x}. \end{aligned} \tag{4.3}$$

Note that $r = |z|$.

Locating $1+i$ in the complex plane, it is possible to immediately determine the polar form from the angle and length of the "complex vector." This is shown in Figure 4.2. It is obvious that $\theta = \frac{\pi}{4}$ and $r = \sqrt{2}$.

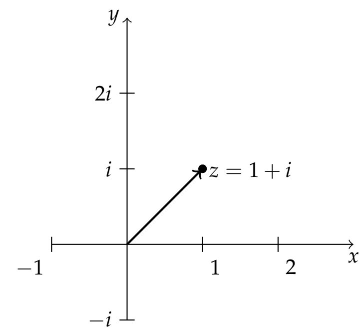


Figure 4.2: Locating $1+i$ in the complex z-plane.

Example 4.1. *Write $z = 1+i$ in polar form.*

If one did not see the polar form from the plot in the z-plane, then one could systematically determine the results. First, write $z = 1+i$ in polar form, $z = re^{i\theta}$, for some r and θ.

Using the above relations between polar and Cartesian representations, we have $r = \sqrt{x^2+y^2} = \sqrt{2}$ and $\tan\theta = \frac{y}{x} = 1$. This gives $\theta = \frac{\pi}{4}$. So, we have found that

$$1+i = \sqrt{2}e^{i\pi/4}.$$

We can also define binary operations of addition, subtraction, multiplication, and division of complex numbers to produce a new complex number.

The addition of two complex numbers is simply done by adding the real and imaginary parts of each number. So,

We can easily add, subtract, multiply, and divide complex numbers.

$$(3+2i) + (1-i) = 4+i.$$

Subtraction is just as easy,

$$(3+2i) - (1-i) = 2+3i.$$

We can multiply two complex numbers just like we multiply any binomials, though we now can use the fact that $i^2 = -1$. For example, we have

$$(3+2i)(1-i) = 3+2i-3i+2i(-i) = 5-i.$$

We can even divide one complex number into another one and get a complex number as the quotient. Before we do this, we need to introduce the complex conjugate, $\bar{z}$, of a complex number. The complex conjugate of $z = x+iy$, where x and y are real numbers, is given as

The complex conjugate of $z = x+iy$ is given as $\bar{z} = x - iy$.

$$\bar{z} = x - iy.$$

Complex conjugates satisfy the following relations for complex numbers z and w and real number x.

$$\begin{aligned} \overline{z+w} &= \bar{z} + \bar{w}. \\ \overline{zw} &= \bar{z}\bar{w}. \\ \bar{\bar{z}} &= z. \\ \bar{x} &= x. \end{aligned} \tag{4.4}$$

One consequence is that the complex conjugate of $re^{i\theta}$ is

$$\overline{re^{i\theta}} = \overline{\cos\theta + i\sin\theta} = \cos\theta - i\sin\theta = re^{-i\theta}.$$

Another consequence is that

$$z\bar{z} = re^{i\theta}re^{-i\theta} = r^2.$$

Thus, the product of a complex number with its complex conjugate is a real number. We can also prove this result using the Cartesian form

$$z\bar{z} = (x+iy)(x-iy) = x^2+y^2 = |z|^2.$$

Now we are in a position to write the quotient of two complex numbers in the standard form of a real plus an imaginary number.

Example 4.2. *Simplify the expression* $z = \frac{3+2i}{1-i}$.

This simplification is accomplished by multiplying the numerator and denominator of this expression by the complex conjugate of the denominator:

$$z = \frac{3+2i}{1-i} = \frac{3+2i}{1-i}\frac{1+i}{1+i} = \frac{1+5i}{2}.$$

Therefore, the quotient is a complex number and in standard form is given by $z = \frac{1}{2} + \frac{5}{2}i$.

We can also consider powers of complex numbers. For example,

$$(1+i)^2 = 2i,$$

$$(1+i)^3 = (1+i)(2i) = 2i-2.$$

But, what is $(1+i)^{1/2} = \sqrt{1+i}$?

In general, we want to find the nth root of a complex number. Let $t = z^{1/n}$. To find t in this case is the same as asking for the solution of

$$z = t^n$$

given z. But, this is the root of an nth degree equation, for which we expect n roots. If we write z in polar form, $z = re^{i\theta}$, then we would naively compute

$$\begin{aligned} z^{1/n} &= \left(re^{i\theta}\right)^{1/n} \\ &= r^{1/n}e^{i\theta/n} \\ &= r^{1/n}\left[\cos\frac{\theta}{n} + i\sin\frac{\theta}{n}\right]. \end{aligned} \tag{4.5}$$

For example,

$$(1+i)^{1/2} = \left(\sqrt{2}e^{i\pi/4}\right)^{1/2} = 2^{1/4}e^{i\pi/8}.$$

The function $f(z) = z^{1/n}$ is multivalued: $z^{1/n} = r^{1/n}e^{i(\theta+2k\pi)/n}$, $k = 0, 1, \ldots, n-1$.

But this is only one solution. We expected two solutions for $n = 2$.

The reason we only found one solution is that the polar representation for z is not unique. We note that

$$\boxed{e^{2k\pi i} = 1, \quad k = 0, \pm 1, \pm 2, \ldots.}$$

So, we can rewrite z as $z = re^{i\theta}e^{2k\pi i} = re^{i(\theta+2k\pi)}$. Now we have that

$$z^{1/n} = r^{1/n} e^{i(\theta+2k\pi)/n}, \quad k = 0, 1, \ldots, n-1.$$

Note that these are the only distinct values for the roots. We can see this by considering the case $k = n$. Then we find that

$$e^{i(\theta+2\pi i n)/n} = e^{i\theta/n} e^{2\pi i} = e^{i\theta/n}.$$

So, we have recovered the $n = 0$ value. Similar results can be shown for the other k values larger than n.

Now we can finish the example we started.

Example 4.3. *Determine the square roots of $1+i$, or $\sqrt{1+i}$.*

As we have seen, we first write $1+i$ in polar form: $1+i = \sqrt{2}e^{i\pi/4}$. Then we introduce $e^{2k\pi i} = 1$ and find the roots:

$$\begin{aligned} (1+i)^{1/2} &= \left(\sqrt{2}e^{i\pi/4}e^{2k\pi i}\right)^{1/2}, \quad k = 0,1, \\ &= 2^{1/4}e^{i(\pi/8+k\pi)}, \quad k = 0,1, \\ &= 2^{1/4}e^{i\pi/8}, 2^{1/4}e^{9\pi i/8}. \end{aligned} \tag{4.6}$$

Finally, what is $\sqrt[n]{1}$? Our first guess would be $\sqrt[n]{1} = 1$. But we now know that there should be n roots. These roots are called the nth roots of unity. Using the above result with $r = 1$ and $\theta = 0$, we have that

The nth roots of unity, $\sqrt[n]{1}$.

$$\sqrt[n]{1} = \left[\cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}\right], \quad k = 0, \ldots, n-1.$$

For example, we have

$$\sqrt[3]{1} = \left[\cos\frac{2\pi k}{3} + i\sin\frac{2\pi k}{3}\right], \quad k = 0,1,2.$$

These three roots can be written out as

$$\sqrt[3]{1} = 1, -\frac{1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i.$$

We can locate these cube roots of unity in the complex plane. In Figure 4.3, we see that these points lie on the unit circle and are at the vertices of an equilateral triangle. In fact, all nth roots of unity lie on the unit circle and are the vertices of a regular n-gon with one vertex at $z = 1$.

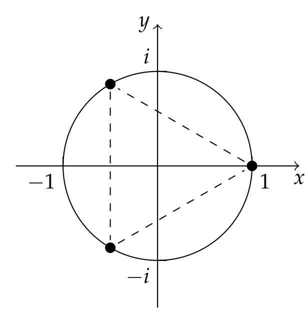


Figure 4.3: Locating the cube roots of unity in the complex z-plane.

4.2 *Complex Valued Functions*

We would like to next explore complex functions and the calculus of complex functions. We begin by defining a function that takes complex numbers into complex numbers, $f : C \to C$. It is difficult to visualize such functions. For real functions of one variable, $f : R \to R$, we graph these functions by first drawing two intersecting copies of R and then proceed to map the domain into the range of f.

It would be more difficult to do this for complex functions. Imagine placing together two orthogonal copies of the complex plane, C. One would need a four-dimensional space in order to complete the visualization. Instead, one typically uses two copies of the complex plane side by side in order to indicate how such functions behave. Over the years there have been several ways to visualize complex functions. We will describe a few of these in this chapter.

We will assume that the domain lies in the z-plane and the image lies in the w-plane. We will then write the complex function as $w = f(z)$. We show these planes in Figure 4.4 as well as the mapping between the planes.

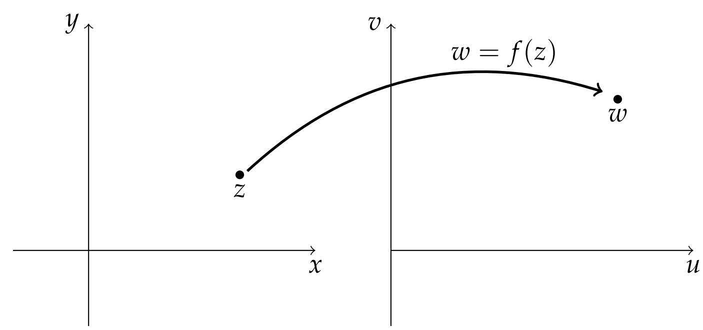


Figure 4.4: Defining a complex valued function, $w = f(z)$, on C for $z = x + iy$ and $w = u + iv$.

Letting $z = x + iy$ and $w = u + iv$, we can write the real and imaginary parts of $f(z)$:

$$w = f(z) = f(x + iy) = u(x, y) + iv(x, y).$$

We see that one can view this function as a function of z or a function of x and y. Often, we have an interest in writing out the real and imaginary parts of the function, $u(x, y)$ and $v(x, y)$, which are functions of two real variables, x and y. We will look at several functions to determine the real and imaginary parts.

Example 4.4. *Find the real and imaginary parts of $f(z) = z^2$.*

For example, we can look at the simple function $f(z) = z^2$. It is a simple matter to determine the real and imaginary parts of this function. Namely, we have

$$z^2 = (x + iy)^2 = x^2 - y^2 + 2ixy.$$

Therefore, we have that

$$u(x, y) = x^2 - y^2, \quad v(x, y) = 2xy.$$

In Figure 4.5 we show how a grid in the z-plane is mapped by $f(z) = z^2$ into the w-plane. For example, the horizontal line $x = 1$ is mapped to $u(1, y) = 1 - y^2$ and $v(1, y) = 2y$. Eliminating the "parameter" y between these two equations, we have $u = 1 - v^2/4$. This is a parabolic curve. Similarly, the horizontal line $y = 1$ results in the curve $u = v^2/4 - 1$.

If we look at several curves, $x =$const and $y =$const, then we get a family of intersecting parabolae, as shown in Figure 4.5.

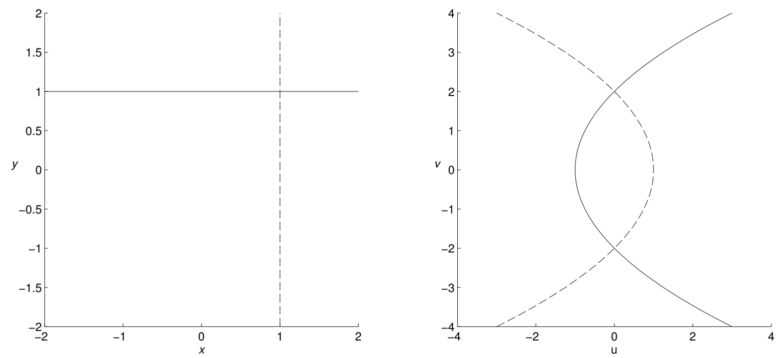


Figure 4.5: 2D plot showing how the function $f(z) = z^2$ maps the lines $x = 1$ and $y = 1$ in the z-plane into parabolae in the w-plane.

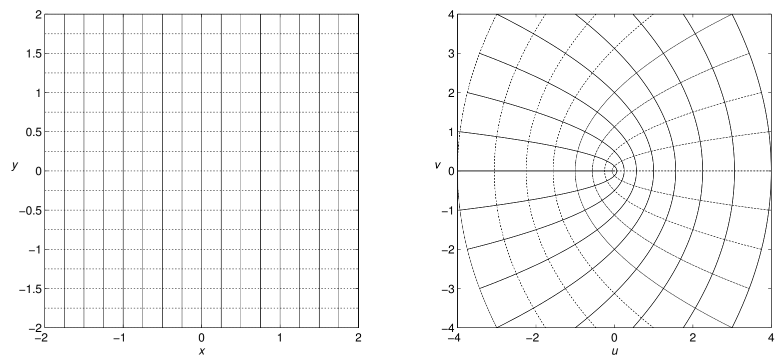


Figure 4.6: 2D plot showing how the function $f(z) = z^2$ maps a grid in the z-plane into the w-plane.

Example 4.5. *Find the real and imaginary parts of $f(z) = e^z$.*

For this case, we make use of Euler's Formula (from Example 1.34):

$$\begin{aligned} e^z &= e^{x+iy} \\ &= e^x e^{iy} \\ &= e^x(\cos y + i \sin y). \end{aligned} \tag{4.7}$$

Thus, $u(x, y) = e^x \cos y$ and $v(x, y) = e^x \sin y$. In Figure 4.7 we show how a grid in the z-plane is mapped by $f(z) = e^z$ into the w-plane.

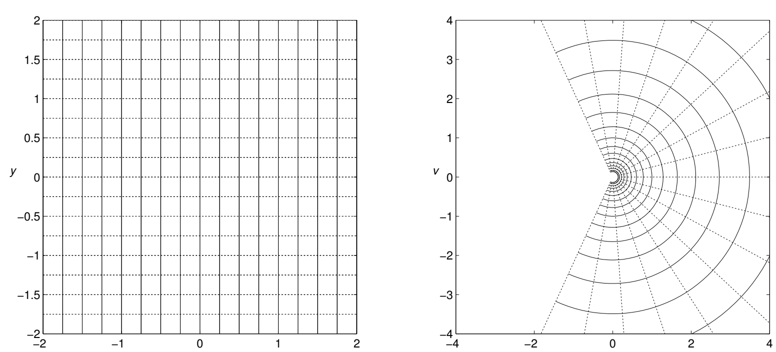


Figure 4.7: 2D plot showing how the function $f(z) = e^z$ maps a grid in the z-plane into the w-plane.

Example 4.6. *Find the real and imaginary parts of $f(z) = z^{1/2}$.*

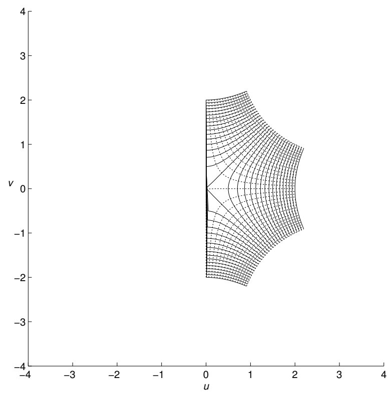


Figure 4.8: 2D plot showing how the function $f(z) = \sqrt{z}$ maps a grid in the z-plane into the w-plane.

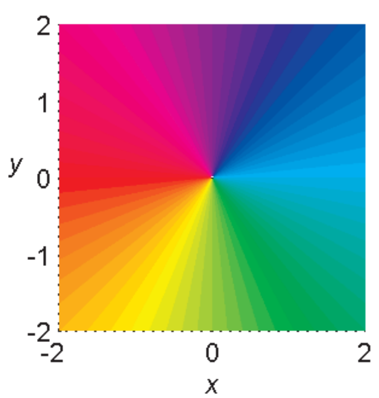


Figure 4.9: Domain coloring of the complex z-plane assigning colors to $\arg(z)$.

We have that

$$z^{1/2} = \sqrt{x^2+y^2}\left(\cos\left(\theta + k\pi\right) + i\sin\left(\theta + k\pi\right)\right), \quad k = 0, 1. \tag{4.8}$$

Thus,

$$u = |z|\cos\left(\theta + k\pi\right), u = |z|\cos\left(\theta + k\pi\right),$$

for $|z| = \sqrt{x^2+y^2}$ *and* $\theta = \tan^{-1}(y/x)$. *For each k-value, one has a different surface and curves of constant* θ *give* $u/v = c_1$*; and, curves of constant nonzero complex modulus give concentric circles,* $u^2 + v^2 = c_2$*, for* c_1 *and* c_2 *constants.*

Example 4.7. *Find the real and imaginary parts of* $f(z) = \ln z$.

In this case, we make use of the polar form of a complex number, $z = re^{i\theta}$. *Our first thought would be to simply compute*

$$\ln z = \ln r + i\theta.$$

However, the natural logarithm is multivalued, just like the square root function. Recalling that $e^{2\pi ik} = 1$ *for k an integer, we have* $z = re^{i(\theta+2\pi k)}$. *Therefore,*

$$\boxed{\ln z = \ln r + i(\theta + 2\pi k), \qquad k = \text{ integer.}}$$

The natural logarithm is a multivalued function. In fact, there are an infinite number of values for a given z. Of course, this contradicts the definition of a function that you were first taught.

Thus, one typically will only report the principal value, Log $z = \ln r + i\theta$*, for* θ *restricted to some interval of length* 2π*, such as* $[0, 2\pi)$. *In order to account for the multivaluedness, one introduces a way to extend the complex plane so as to include all of the branches. This is done by assigning a plane to each branch, using (branch) cuts along lines, and then gluing the planes together at the branch cuts to form what is called a Riemann surface. We will not elaborate upon this any further here and refer the interested reader to more advanced texts. Comparing the multivalued logarithm to the principal value logarithm, we have*

$$\ln z = Log\ z + 2n\pi i.$$

We should note that some books use $\log z$ *instead of* $\ln z$. *It should not be confused with the common logarithm.*

4.2.1 *Complex Domain Coloring*

ANOTHER METHOD FOR VISUALIZING COMPLEX FUNCTIONS is domain coloring. The idea was described by Frank A. Farris. There are a few approaches to this method. The main idea is that one colors each point of the z-plane (the domain) according to $\arg(z)$ as shown in Figure 4.9. The modulus, $|f(z)|$ is then plotted as a surface. Examples are shown for $f(z) = z^2$ in Figure 4.10 and $f(z) = 1/z(1-z)$ in Figure 4.11.

We would like to put all of this information in one plot. We can do this by adjusting the brightness of the colored domain using the modulus of the

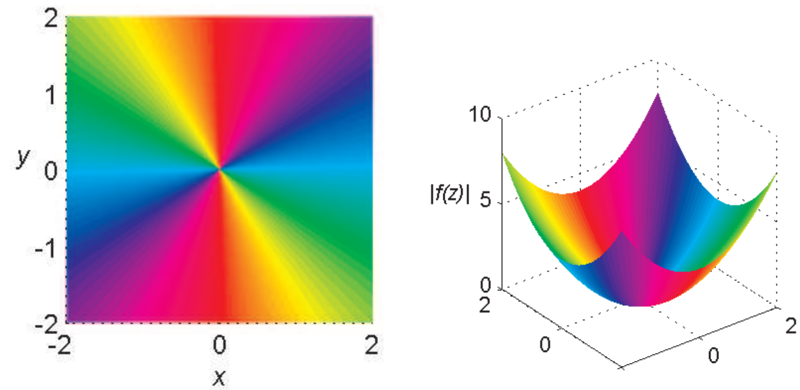


Figure 4.10: Domain coloring for $f(z) = z^2$. The left figure shows the phase coloring. The right figure shows the colored surface with height $|f(z)|$.

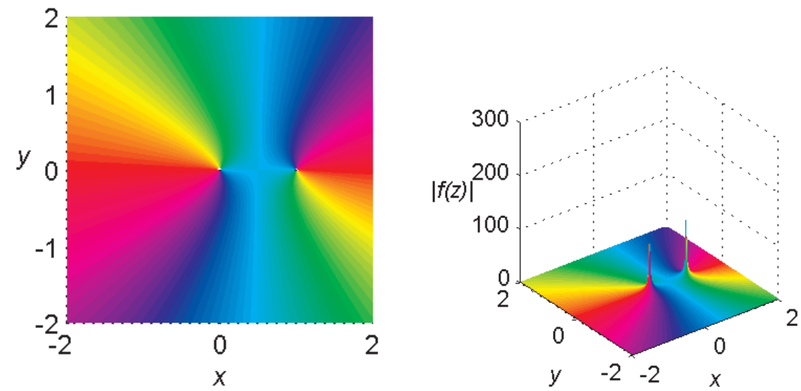


Figure 4.11: Domain coloring for $f(z) = 1/z(1-z)$. The left figure shows the phase coloring. The right figure shows the colored surface with height $|f(z)|$.

function. In the plots that follow we use the fractional part of $\ln|z|$. In Figure 4.12 we show the effect for the z-plane using $f(z) = z$. In the figures that follow, we look at several other functions. In these plots, we have chosen to view the functions in a circular window.

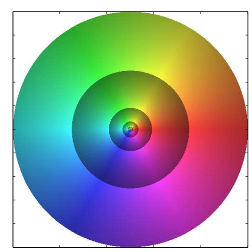


Figure 4.12: Domain coloring for the function $f(z) = z$ showing a coloring for $\arg(z)$ and brightness based on $|f(z)|$.

One can see the rich behavior hidden in these figures. As you progress in your reading, especially after the next chapter, you should return to these figures and locate the zeros, poles, branch points, and branch cuts. A search online will lead you to other colorings and superposition of the uv grid on these figures.

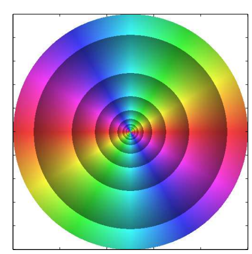


Figure 4.13: Domain coloring for the function $f(z) = z^2$.

As a final picture, we look at iteration in the complex plane. Consider the function $f(z) = z^2 - 0.75 - 0.2i$. Interesting figures result when studying the iteration in the complex plane. In Figure 4.15 we show $f(z)$ and $f^{20}(z)$, which is the iteration of f twenty times. It leads to an interesting coloring. What happens when one keeps iterating? Such iterations lead to the study of Julia and Mandelbrot sets . In Figure 4.16 we show six iterations of $f(z) = (1 - i/2)\sin x$.

The following code was used in MATLAB to produce these figures.

```
fn = @(x) (1-i/2)*sin(x);
xmin=-2; xmax=2; ymin=-2; ymax=2;
Nx=500;
Ny=500;
```

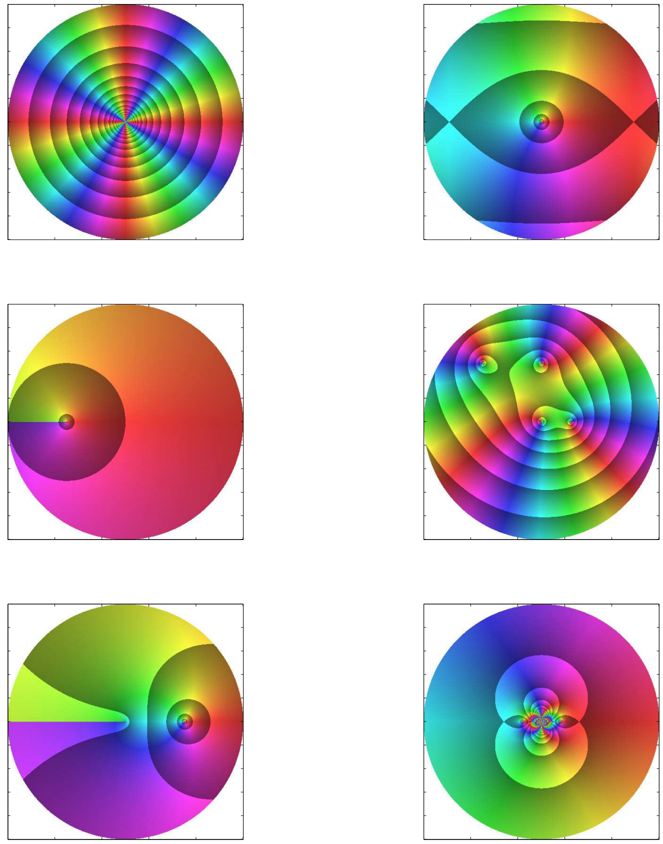


Figure 4.14: Domain coloring for several functions. On the top row, the domain coloring is shown for $f(z) = z^4$ and $f(z) = \sin z$. On the second row, plots for $f(z) = \sqrt{1+z}$ and $f(z) = \frac{1}{z(1/2-z)(z-i)(z-i+1)}$ are shown. In the last row, domain colorings for $f(z) = \ln z$ and $f(z) = \sin(1/z)$ are shown.

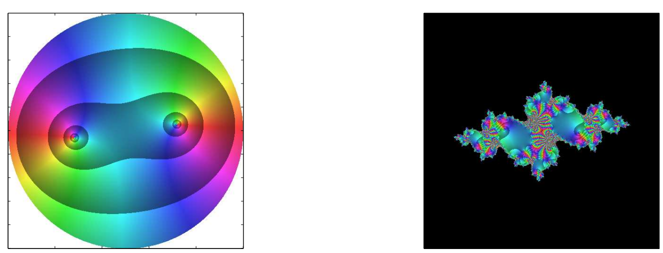


Figure 4.15: Domain coloring for $f(z) = z^2 - 0.75 - 0.2i$. The left figure shows the phase coloring. On the right is the plot for $f^{20}(z)$.

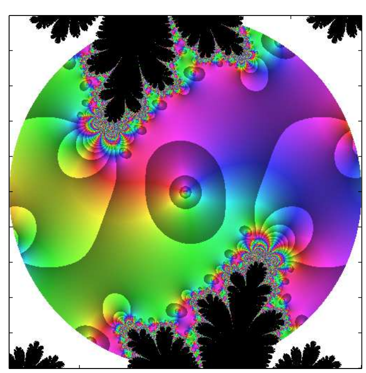


Figure 4.16: Domain coloring for six iterations of $f(z) = (1 - i/2)\sin x$.

```
x=linspace(xmin,xmax,Nx);
y=linspace(ymin,ymax,Ny);
[X,Y] = meshgrid(x,y); z = complex(X,Y);
tmp=z; for n=1:6
    tmp = fn(tmp);
end Z=tmp;
XX=real(Z);
YY=imag(Z);
R2=max(max(X.^2));
R=max(max(XX.^2+YY.^2));

circle(:,:,1) = X.^2+Y.^2 < R2;
circle(:,:,2)=circle(:,:,1);
circle(:,:,3)=circle(:,:,1);

addcirc(:,:,1)=circle(:,:,1)==0;
addcirc(:,:,2)=circle(:,:,1)==0;
addcirc(:,:,3)=circle(:,:,1)==0;

warning off MATLAB:divideByZero;
hsvCircle=ones(Nx,Ny,3);
hsvCircle(:,:,1)=atan2(YY,XX)*180/pi+(atan2(YY,XX)*180/pi<0)*360;
hsvCircle(:,:,1)=hsvCircle(:,:,1)/360; lgz=log(XX.^2+YY.^2)/2;
hsvCircle(:,:,2)=0.75; hsvCircle(:,:,3)=1-(lgz-floor(lgz))/2;
hsvCircle(:,:,1) = flipud((hsvCircle(:,:,1)));

hsvCircle(:,:,2) = flipud((hsvCircle(:,:,2)));

hsvCircle(:,:,3) =flipud((hsvCircle(:,:,3)));

rgbCircle=hsv2rgb(hsvCircle);
rgbCircle=rgbCircle.*circle+addcirc;

image(rgbCircle)
axis square
set(gca,'XTickLabel',{})
set(gca,'YTickLabel',{})
```

4.3 *Complex Differentiation*

NEXT WE WANT TO DIFFERENTIATE COMPLEX FUNCTIONS. We generalize the definition from single variable calculus,

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}, \tag{4.9}$$

provided this limit exists.

The computation of this limit is similar to what one sees in multivariable calculus for limits of real functions of two variables. Letting $z = x + iy$ and $\delta z = \delta x + i\delta y$, then

$$z + \delta x = (x + \delta x) + i(y + \delta y).$$

Letting $\Delta z \to 0$ means that we get closer to z. There are many paths that one can take that will approach z. [See Figure 4.17.]

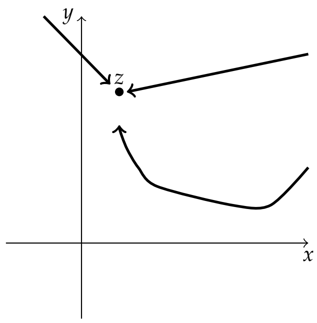


Figure 4.17: There are many paths that approach z as $\Delta z \to 0$.

It is sufficient to look at two paths in particular. We first consider the path $y =$ constant. This horizontal path is shown in Figure 4.18. For this path, $\Delta z = \Delta x + i\Delta y = \Delta x$, since y does not change along the path. The derivative, if it exists, is then computed as

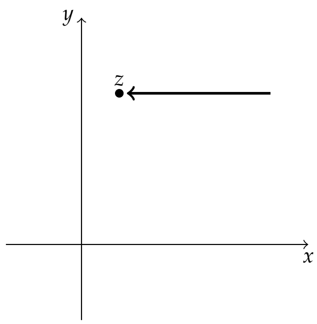


Figure 4.18: A path that approaches z with $y =$ constant.

$$\begin{aligned} f'(z) &= \lim_{\Delta z \to 0} \frac{f(z+\Delta z) - f(z)}{\Delta z} \\ &= \lim_{\Delta x \to 0} \frac{u(x+\Delta x, y) + iv(x+\Delta x, y) - (u(x,y) + iv(x,y))}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{u(x+\Delta x, y) - u(x,y)}{\Delta x} + \lim_{\Delta x \to 0} i\frac{v(x+\Delta x, y) - v(x,y)}{\Delta x}. \end{aligned} \tag{4.10}$$

The last two limits are easily identified as partial derivatives of real valued functions of two variables. Thus, we have shown that when $f'(z)$ exists,

$$f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}. \tag{4.11}$$

A similar computation can be made if, instead, we take the vertical path, $x =$ constant, in Figure 4.17). In this case, $\Delta z = i\Delta y$ and

$$\begin{aligned} f'(z) &= \lim_{\Delta z \to 0} \frac{f(z+\Delta z) - f(z)}{\Delta z} \\ &= \lim_{\Delta y \to 0} \frac{u(x, y+\Delta y) + iv(x, y+\Delta y) - (u(x,y) + iv(x,y))}{i\Delta y} \\ &= \lim_{\Delta y \to 0} \frac{u(x, y+\Delta y) - u(x,y)}{i\Delta y} + \lim_{\Delta y \to 0} \frac{v(x, y+\Delta y) - v(x,y)}{\Delta y}. \end{aligned} \tag{4.12}$$

Therefore,

$$f'(z) = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}. \tag{4.13}$$

We have found two different expressions for $f'(z)$ by following two different paths to z. If the derivative exists, then these two expressions must be the same. Equating the real and imaginary parts of these expressions, we have

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$$
$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \quad (4.14)$$

These are known as the Cauchy-Riemann Equations[2].

[2] The Cauchy-Riemann Equations. Augustin-Louis Cauchy (1789 - 1857) was a French mathematician well known for his work in analysis. Georg Friedrich Bernhard Riemann (1826 - 1866) was a German mathematician who made major contributions to geometry and analysis.

Theorem 4.1. *$f(z)$ is holomorphic (differentiable) if and only if the Cauchy-Riemann Equations are satisfied.*

Example 4.8. $f(z) = z^2$.

In this case we have already seen that $z^2 = x^2 - y^2 + 2ixy$. Therefore, $u(x, y) = x^2 - y^2$ and $v(x, y) = 2xy$. We first check the Cauchy-Riemann Equations.

$$\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y}$$
$$\frac{\partial v}{\partial x} = 2y = -\frac{\partial u}{\partial y}. \quad (4.15)$$

Therefore, $f(z) = z^2$ is differentiable.

We can further compute the derivative using either Equation (4.11) or Equation (4.13). Thus,

$$f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = 2x + i(2y) = 2z.$$

This result is not surprising.

Example 4.9. $f(z) = \bar{z}$.

In this case we have $f(z) = x - iy$. Therefore, $u(x, y) = x$ and $v(x, y) = -y$. But, $\frac{\partial u}{\partial x} = 1$ and $\frac{\partial v}{\partial y} = -1$. Thus, the Cauchy-Riemann Equations are not satisfied and we conclude that $f(z) = \bar{z}$ is not differentiable.

Harmonic functions satisfy Laplace's Equation.

Another consequence of the Cauchy-Riemann Equations is that both $u(x, y)$ and $v(x, y)$ are harmonic functions. A real-valued function $u(x, y)$ is harmonic if it satisfies Laplace's Equation in 2D, $\nabla^2 u = 0$, or

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Theorem 4.2. *$f(z) = u(x, y) + iv(x, y)$ is differentiable only if u and v are harmonic functions.*

This is easily proven using the Cauchy-Riemann Equations.

$$\begin{aligned}
\frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial u}{\partial x} \\
&= \frac{\partial}{\partial x}\frac{\partial v}{\partial y} \\
&= \frac{\partial}{\partial y}\frac{\partial v}{\partial x} \\
&= -\frac{\partial}{\partial y}\frac{\partial u}{\partial y} \\
&= -\frac{\partial^2 u}{\partial y^2}. \quad (4.16)
\end{aligned}$$

Example 4.10. *Is $u(x,y) = x^2 + y^2$ harmonic?*

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 2 + 2 \neq 0.$$

No, it is not.

Example 4.11. *Is $u(x,y) = x^2 - y^2$ harmonic?*

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 2 - 2 = 0.$$

Yes, it is.

The harmonic conjugate function.

Given a harmonic function $u(x,y)$, can one find a function, $v(x,y)$, such $f(z) = u(x,y) + iv(x,y)$ is differentiable? In this case, v are called the harmonic conjugate of u.

Example 4.12. *Find the harmonic conjugate of $u(x,y) = x^2 - y^2$ and determine $f(z) = u + iv$ such that $u + iv$ is differentiable.*

The Cauchy-Riemann Equations tell us the following about the unknown function, $v(x,y)$:

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} = 2y,$$

$$\frac{\partial v}{\partial y} = \frac{\partial u}{\partial x} = 2x.$$

We can integrate the first of these equations to obtain

$$v(x,y) = \int 2y\, dx = 2xy + c(y).$$

Here, $c(y)$ is an arbitrary function of y. One can check to see that this works by simply differentiating the result with respect to x.

However, the second equation must also hold. So, we differentiate the result with respect to y to find that

$$\frac{\partial v}{\partial y} = 2x + c'(y).$$

Since we were supposed to get $2x$, we have that $c'(y) = 0$. Thus, $c(y) = k$ is a constant.

We have just shown that we get an infinite number of functions,

$$v(x,y) = 2xy + k,$$

such that

$$f(z) = x^2 - y^2 + i(2xy + k)$$

is differentiable. In fact, for $k = 0$, this is nothing other than $f(z) = z^2$.

4.4 Complex Integration

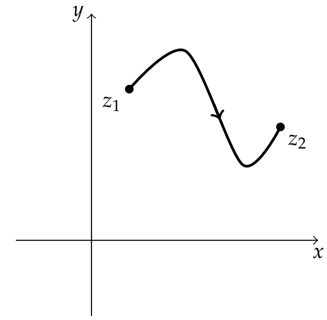


Figure 4.19: We would like to integrate a complex function $f(z)$ over the path Γ in the complex plane.

WE HAVE INTRODUCED FUNCTIONS OF A COMPLEX VARIABLE. We also established when functions are differentiable as complex functions, or holomorphic. In this chapter we will turn to integration in the complex plane. We will learn how to compute complex path integrals, or contour integrals. We will see that contour integral methods are also useful in the computation of some of the real integrals that we will face when exploring Fourier transforms in the next chapter.

4.4.1 Complex Path Integrals

IN THIS SECTION WE WILL INVESTIGATE the computation of complex path integrals. Given two points in the complex plane, connected by a path Γ as shown in Figure 4.19, we would like to define the integral of $f(z)$ along Γ,

$$\int_\Gamma f(z)\, dz.$$

A natural procedure would be to work in real variables, by writing

$$\int_\Gamma f(z)\, dz = \int_\Gamma [u(x,y) + iv(x,y)]\,(dx + idy),$$

since $z = x + iy$ and $dz = dx + idy$.

In order to carry out the integration, we then have to find a parametrization of the path and use methods from a multivariate calculus class. Namely, let u and v be continuous in domain D, and Γ a piecewise smooth curve in D. Let $(x(t), y(t))$ be a parametrization of Γ for $t_0 \le t \le t_1$ and $f(z) = u(x,y) + iv(x,y)$ for $z = x + iy$. Then

$$\int_\Gamma f(z)\, dz = \int_{t_0}^{t_1} [u(x(t),y(t)) + iv(x(t),y(t))]\,(\frac{dx}{dt} + i\frac{dy}{dt})dt. \tag{4.17}$$

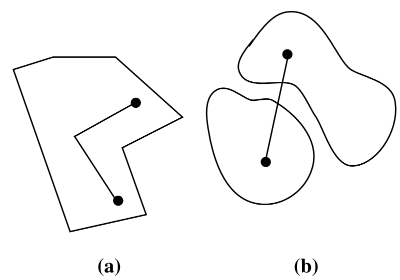


Figure 4.20: Examples of (a) a connected set and (b) a disconnected set.

Here we have used

$$dz = dx + idy = \left(\frac{dx}{dt} + i\frac{dy}{dt}\right) dt.$$

Furthermore, a set D is called a domain if it is both open and connected.

Before continuing, we first define open and connected. A set D is connected if and only if for all z_1, and z_2 in D, there exists a piecewise smooth curve connecting z_1 to z_2 and lying in D. Otherwise it is called disconnected. Examples are shown in Figure 4.20

A set D is open if and only if for all z_0 in D, there exists an open disk $|z - z_0| < \rho$ in D. In Figure 4.21 we show a region with two disks.

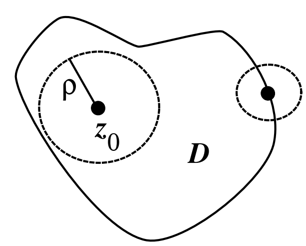


Figure 4.21: Locations of open disks inside and on the boundary of a region.

For all points on the interior of the region, one can find at least one disk contained entirely in the region. The closer one is to the boundary, the smaller the radii of such disks. However, for a point on the boundary, every such disk would contain points inside and outside the disk. Thus, an open set in the complex plane would not contain any of its boundary points.

We now have a prescription for computing path integrals. Let's see how this works with a couple of examples.

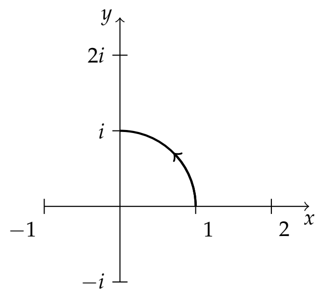


Figure 4.22: Contour for Example 4.13.

Example 4.13. *Evaluate* $\int_C z^2\, dz$*, where* C *= the arc of the unit circle in the first quadrant as shown in Figure 4.22.*

There are two ways we could carry out the parametrization. First, we note that the standard parametrization of the unit circle is

$$(x(\theta), y(\theta)) = (\cos\theta, \sin\theta), \quad 0 \le \theta \le 2\pi.$$

For a quarter circle in the first quadrant, $0 \le \theta \le \frac{\pi}{2}$*, we let* $z = \cos\theta + i\sin\theta$*. Therefore,* $dz = (-\sin\theta + i\cos\theta)\, d\theta$ *and the path integral becomes*

$$\int_C z^2\, dz = \int_0^{\frac{\pi}{2}} (\cos\theta + i\sin\theta)^2(-\sin\theta + i\cos\theta)\, d\theta.$$

We can expand the integrand and integrate, having to perform some trigonometric integrations.

$$\int_0^{\frac{\pi}{2}} [\sin^3\theta - 3\cos^2\theta\sin\theta + i(\cos^3\theta - 3\cos\theta\sin^2\theta)]\, d\theta.$$

The reader should work out these trigonometric integrations and confirm the result. For example, you can use

$$\sin^3\theta = \sin\theta(1 - \cos^2\theta))$$

to write the real part of the integrand as

$$\sin\theta - 4\cos^2\theta\sin\theta.$$

The resulting antiderivative becomes

$$-\cos\theta + \frac{4}{3}\cos^3\theta.$$

The imaginary integrand can be integrated in a similar fashion.

While this integral is doable, there is a simpler procedure. We first note that $z = e^{i\theta}$ *on C. So,* $dz = ie^{i\theta}d\theta$*. The integration then becomes*

$$\begin{aligned} \int_C z^2\, dz &= \int_0^{\frac{\pi}{2}} (e^{i\theta})^2 ie^{i\theta}\, d\theta \\ &= i\int_0^{\frac{\pi}{2}} e^{3i\theta}\, d\theta \\ &= \left.\frac{ie^{3i\theta}}{3i}\right|_0^{\pi/2} \\ &= -\frac{1+i}{3}. \end{aligned} \tag{4.18}$$

Example 4.14. *Evaluate $\int_\Gamma z\,dz$, for the path $\Gamma = \gamma_1 \cup \gamma_2$ shown in Figure 4.23.*

In this problem we have a path that is a piecewise smooth curve. We can compute the path integral by computing the values along the two segments of the path and adding the results. Let the two segments be called γ_1 and γ_2 as shown in Figure 4.23 and parametrize each path separately.

Over γ_1 we note that $y = 0$. Thus, $z = x$ for $x \in [0,1]$. It is natural to take x as the parameter. So, we let $dz = dx$ to find

$$\int_{\gamma_1} z\,dz = \int_0^1 x\,dx = \frac{1}{2}.$$

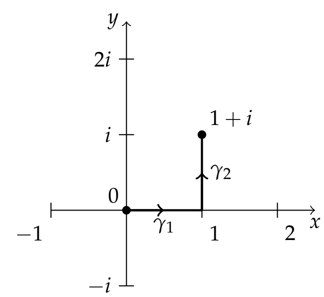


Figure 4.23: Contour for Example 4.14 with $\Gamma = \gamma_1 \cup \gamma_2$.

For path γ_2, we have that $z = 1 + iy$ for $y \in [0,1]$ and $dz = i\,dy$. Inserting this parametrization into the integral, the integral becomes

$$\int_{\gamma_2} z\,dz = \int_0^1 (1 + iy)\,idy = i - \frac{1}{2}.$$

Combining the results for the paths γ_1 and γ_2, we have $\int_\Gamma z\,dz = \frac{1}{2} + (i - \frac{1}{2}) = i$.

Example 4.15. *Evaluate $\int_{\gamma_3} z\,dz$, where γ_3 is the path shown in Figure 4.24.*

In this case we take a path from $z = 0$ to $z = 1 + i$ along a different path than in the last example. Let $\gamma_3 = \{(x,y)|y = x^2, x \in [0,1]\} = \{z|z = x + ix^2, x \in [0,1]\}$. Then, $dz = (1 + 2ix)\,dx$.

The integral becomes

$$\begin{aligned}\int_{\gamma_3} z\,dz &= \int_0^1 (x + ix^2)(1 + 2ix)\,dx \\ &= \int_0^1 (x + 3ix^2 - 2x^3)\,dx = \\ &= \left[\frac{1}{2}x^2 + ix^3 - \frac{1}{2}x^4\right]_0^1 = i. \qquad (4.19)\end{aligned}$$

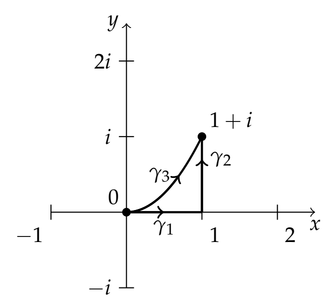


Figure 4.24: Contour for Example 4.15.

In the last case we found the same answer as obtained in Example 4.14. But we should not take this as a general rule for all complex path integrals. In fact, it is not true that integrating over different paths always yields the same results. However, when this is true, then we refer to this property as path independence. In particular, the integral $\int f(z)\,dz$ is path independent if

$$\int_{\Gamma_1} f(z)\,dz = \int_{\Gamma_2} f(z)\,dz$$

for all paths from z_1 to z_2 as shown in Figure 4.25.

We can show that if $\int f(z)\,dz$ is path independent, then the integral of $f(z)$ over all closed loops is zero:

$$\int_{\text{closed loops}} f(z)\,dz = 0.$$

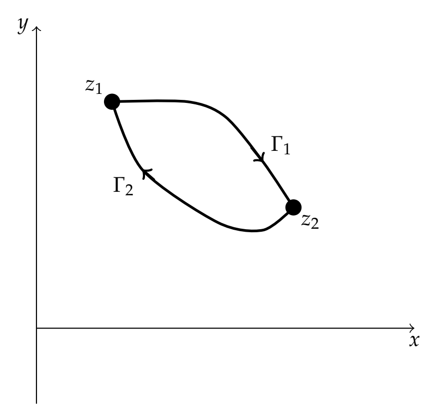


Figure 4.25: $\int_{\Gamma_1} f(z)\,dz = \int_{\Gamma_2} f(z)\,dz$ for all paths from z_1 to z_2 when the integral of $f(z)$ is path independent.

A simple closed contour.

A common notation for integrating over closed loops is $\oint_C f(z)\,dz$. But first we have to define what we mean by a closed loop. A simple closed contour is a path satisfying

a. The end point is the same as the beginning point. (This makes the loop closed.)

b. The are no self-intersections. (This makes the loop simple.)

A loop in the shape of a figure eight is closed, but it is not simple.

Now consider an integral over the closed loop C shown in Figure 4.26. We pick two points on the loop, breaking it into two contours, C_1 and C_2. Then we make use of the path independence by defining C_2^- to be the path along C_2 but in the opposite direction. Then,

$\oint_C f(z)\,dz = 0$ if the integral is path independent.

$$\begin{aligned}\oint_C f(z)\,dz &= \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz \\ &= \int_{C_1} f(z)\,dz - \int_{C_2^-} f(z)\,dz. \qquad (4.20)\end{aligned}$$

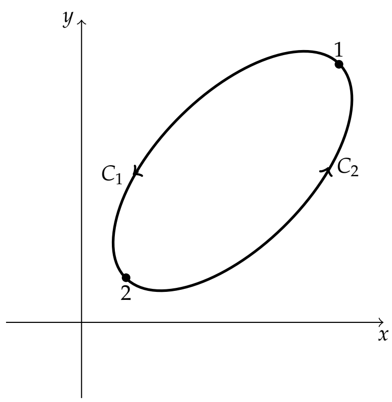


Figure 4.26: The integral $\oint_C f(z)\,dz$ around C is zero if the integral $\int_\Gamma f(z)\,dz$ is path independent.

Assuming that the integrals from point 1 to point 2 are path independent, then the integrals over C_1 and C_2^- are equal. Therefore, we have $\oint_C f(z)\,dz = 0$.

Example 4.16. *Consider the integral $\oint_C z\,dz$ for C the closed contour shown in Figure 4.24 starting at $z = 0$ following path γ_1, then γ_2 and returning to $z = 0$. Based on the earlier examples and the fact that going backward on γ_3 introduces a negative sign, we have*

$$\oint_C z\,dz = \int_{\gamma_1} z\,dz + \int_{\gamma_2} z\,dz - \int_{\gamma_3} z\,dz = \frac{1}{2} + \left(i - \frac{1}{2}\right) - i = 0.$$

4.4.2 *Cauchy's Theorem*

NEXT WE WANT TO INVESTIGATE if we can determine that integrals over simple closed contours vanish without doing all the work of parametrizing the contour. First, we need to establish the direction about which we traverse the contour. We can define the orientation of a curve by referring to the normal of the curve.

Recall that the normal is a perpendicular to the curve. There are two such perpendiculars. The above normal points outward and the other normal points toward the interior of a closed curve. We will define a positively oriented contour as one that is traversed with the outward normal pointing to the right. As one follows loops, the interior would then be on the left.

A curve with parametrization $(x(t), y(t))$ has a normal $(n_x, n_y) = (-\frac{dx}{dt}, \frac{dy}{dt})$.

We now consider $\oint_C (u + iv)\,dz$ over a simple closed contour. This can be written in terms of two real integrals in the xy-plane.

$$\begin{aligned}\oint_C (u+iv)\,dz &= \int_C (u+iv)(dx + i\,dy) \\ &= \int_C u\,dx - v\,dy + i\int_C v\,dx + u\,dy. \qquad (4.21)\end{aligned}$$

These integrals in the plane can be evaluated using Green's Theorem in the Plane. Recall this theorem from your last semester of calculus:

Green's Theorem in the Plane.

Theorem 4.3. *Let $P(x,y)$ and $Q(x,y)$ be continuously differentiable functions on and inside the simple closed curve C as shown in Figure 4.27. Denoting the enclosed region S, we have*

$$\int_C P\,dx + Q\,dy = \int\int_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy. \tag{4.22}$$

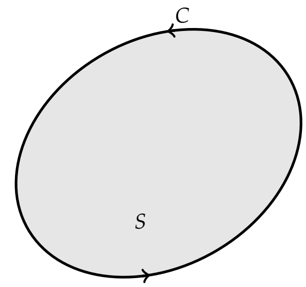


Figure 4.27: Region used in Green's Theorem.

Green's Theorem in the Plane is one of the major integral theorems of vector calculus. It was discovered by George Green (1793 - 1841) and published in 1828, about four years before he entered Cambridge as an undergraduate.

Using Green's Theorem to rewrite the first integral in Equation (4.21), we have

$$\int_C u\,dx - v\,dy = \int\int_S \left(\frac{-\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dxdy.$$

If u and v satisfy the Cauchy-Riemann Equations (4.14), then the integrand in the double integral vanishes. Therefore,

$$\int_C u\,dx - v\,dy = 0.$$

In a similar fashion, one can show that

$$\int_C v\,dx + u\,dy = 0.$$

We have thus proven the following theorem:

Cauchy's Theorem

Theorem 4.4. *If u and v satisfy the Cauchy-Riemann Equations (4.14) inside and on the simple closed contour C, then*

$$\oint_C (u + iv)\,dz = 0. \tag{4.23}$$

Corollary $\oint_C f(z)\,dz = 0$ *when f is differentiable in domain D with $C \subset D$.*

Either one of these is referred to as **Cauchy's Theorem**.

Example 4.17. *Evaluate $\oint_{|z-1|=3} z^4\,dz$.*

Since $f(z) = z^4$ is differentiable inside the circle $|z-1| = 3$, this integral vanishes.

We can use Cauchy's Theorem to show that we can deform one contour into another, perhaps simpler, contour.

One can deform contours into simpler ones.

Theorem 4.5. *If $f(z)$ is holomorphic between two simple closed contours, C and C', then $\oint_C f(z)\,dz = \oint_{C'} f(z)\,dz$.*

Proof. We consider the two curves C and C' as shown in Figure 4.28. Connecting the two contours with contours Γ_1 and Γ_2 (as shown in the figure), C is seen to split into contours C_1 and C_2 and C' into contours C_1' and C_2'. Note that $f(z)$ is differentiable inside the newly formed regions between the curves. Also, the boundaries of these regions are now simple closed curves.

Therefore, Cauchy's Theorem tells us that the integrals of $f(z)$ over these regions are zero.

Noting that integrations over contours opposite the positive orientation are the negative of integrals that are positively oriented, we have from Cauchy's Theorem that

$$\int_{C_1} f(z)\,dz + \int_{\Gamma_1} f(z)\,dz - \int_{C_1'} f(z)\,dz + \int_{\Gamma_2} f(z)\,dz = 0$$

and

$$\int_{C_2} f(z)\,dz - \int_{\Gamma_2} f(z)\,dz - \int_{C_2'} f(z)\,dz - \int_{\Gamma_1} f(z)\,dz = 0.$$

In the first integral, we have traversed the contours in the following order: C_1, Γ_1, C_1' backward, and Γ_2. The second integral denotes the integration over the lower region, but going backward over all contours except for C_2.

Combining these results by adding the two equations above, we have

$$\int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz - \int_{C_1'} f(z)\,dz - \int_{C_2'} f(z)\,dz = 0.$$

Noting that $C = C_1 + C_2$ and $C' = C_1' + C_2'$, we have

$$\oint_C f(z)\,dz = \oint_{C'} f(z)\,dz,$$

as was to be proven. □

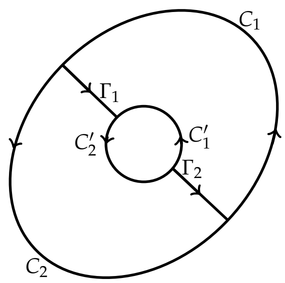


Figure 4.28: The contours needed to prove that $\oint_C f(z)\,dz = \oint_{C'} f(z)\,dz$ when $f(z)$ is holomorphic between the contours C and C'.

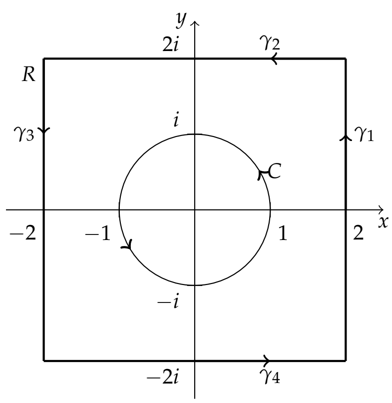


Figure 4.29: The contours used to compute $\oint_R \frac{dz}{z}$. Note that to compute the integral around R we can deform the contour to the circle C since $f(z)$ is differentiable in the region between the contours.

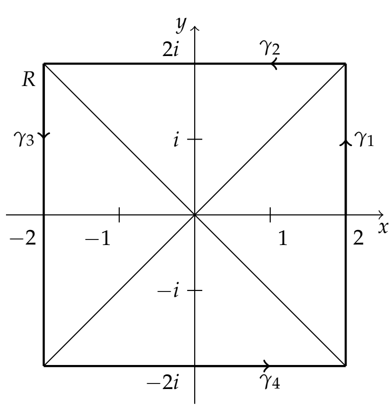


Figure 4.30: The contours used to compute $\oint_R \frac{dz}{z}$. The added diagonals are for the reader to easily see the arguments used in the evaluation of the limits when integrating over the segments of the square R.

Example 4.18. *Compute* $\oint_R \frac{dz}{z}$ *for* R *the rectangle* $[-2,2]\times[-2i,2i]$.

We can compute this integral by looking at four separate integrals over the sides of the rectangle in the complex plane. One simply parametrizes each line segment, perform the integration and sum the four separate results. From the deformation theorem, we can instead integrate over a simpler contour by deforming the rectangle into a circle as long as $f(z) = \frac{1}{z}$ *is differentiable in the region bounded by the rectangle and the circle. So, using the unit circle, as shown in Figure 4.29, the integration might be easier to perform.*

More specifically, the the deformation theorem tells us that

$$\oint_R \frac{dz}{z} = \oint_{|z|=1} \frac{dz}{z}$$

The latter integral can be computed using the parametrization $z = e^{i\theta}$ *for* $\theta \in [0, 2\pi]$. *Thus,*

$$\begin{aligned}\oint_{|z|=1} \frac{dz}{z} &= \int_0^{2\pi} \frac{ie^{i\theta}\,d\theta}{e^{i\theta}} \\ &= i\int_0^{2\pi} d\theta = 2\pi i. \qquad (4.24)\end{aligned}$$

Therefore, we have found that $\oint_R \frac{dz}{z} = 2\pi i$ *by deforming the original simple closed contour.*

For fun, let's do this the long way to see how much effort was saved. We will label the contour as shown in Figure 4.30. The lower segment, γ_4, *of the square can be* *simple parametrized by noting that along this segment $z = x - 2i$ for $x \in [-2, 2]$. Then, we have*

$$\begin{aligned}
\oint_{\gamma_4} \frac{dz}{z} &= \int_{-2}^{2} \frac{dx}{x-2i} \\
&= \ln\left|x-2i\right|_{-2}^{2} \\
&= \left(\ln(2\sqrt{2}) - \frac{\pi i}{4}\right) - \left(\ln(2\sqrt{2}) - \frac{3\pi i}{4}\right) \\
&= \frac{\pi i}{2}. \qquad (4.25)
\end{aligned}$$

We note that the arguments of the logarithms are determined from the angles made by the diagonals provided in Figure 4.30.

Similarly, the integral along the top segment, $z = x + 2i$, $x \in [-2, 2]$, is computed as

$$\begin{aligned}
\oint_{\gamma_2} \frac{dz}{z} &= \int_{2}^{-2} \frac{dx}{x+2i} \\
&= \ln\left|x+2i\right|_{2}^{-2} \\
&= \left(\ln(2\sqrt{2}) + \frac{3\pi i}{4}\right) - \left(\ln(2\sqrt{2}) + \frac{\pi i}{4}\right) \\
&= \frac{\pi i}{2}. \qquad (4.26)
\end{aligned}$$

The integral over the right side, $z = 2 + iy$, $y \in [-2, 2]$, is

$$\begin{aligned}
\oint_{\gamma_1} \frac{dz}{z} &= \int_{-2}^{2} \frac{idy}{2+iy} \\
&= \ln\left|2+iy\right|_{-2}^{2} \\
&= \left(\ln(2\sqrt{2}) + \frac{\pi i}{4}\right) - \left(\ln(2\sqrt{2}) - \frac{\pi i}{4}\right) \\
&= \frac{\pi i}{2}. \qquad (4.27)
\end{aligned}$$

Finally, the integral over the left side, $z = -2 + iy$, $y \in [-2, 2]$, is

$$\begin{aligned}
\oint_{\gamma_3} \frac{dz}{z} &= \int_{2}^{-2} \frac{idy}{-2+iy} \\
&= \ln\left|-2+iy\right|_{-2}^{2} \\
&= \left(\ln(2\sqrt{2}) + \frac{5\pi i}{4}\right) - \left(\ln(2\sqrt{2}) + \frac{3\pi i}{4}\right) \\
&= \frac{\pi i}{2}. \qquad (4.28)
\end{aligned}$$

Therefore, we have that

$$\oint_R \frac{dz}{z} = \int_{\gamma_1} \frac{dz}{z} + \int_{\gamma_2} \frac{dz}{z} + \int_{\gamma_3} \frac{dz}{z} + \int_{\gamma_4} \frac{dz}{z}$$

$$= \frac{\pi i}{2} + \frac{\pi i}{2} + \frac{\pi i}{2} + \frac{\pi i}{2}$$
$$= 4\left(\frac{\pi i}{2}\right) = 2\pi i. \quad (4.29)$$

This gives the same answer as we found using a simple contour deformation.

Morera's Theorem.

The converse of Cauchy's Theorem is not true, namely $\oint_C f(z)\,dz = 0$ does not always imply that $f(z)$ is differentiable. What we do have is Morera's Theorem (Giacinto Morera, 1856 - 1909):

Theorem 4.6. *Let f be continuous in a domain D. Suppose that for every simple closed contour C in D, $\oint_C f(z)\,dz = 0$. Then f is differentiable in D.*

The proof is a bit more detailed than we need to go into here. However, this theorem is useful in the next section.

4.4.3 Analytic Functions and Cauchy's Integral Formula

In the previous section we saw that Cauchy's Theorem was useful for computing particular integrals without having to parametrize the contours or for deforming contours into simpler contours. The integrand needs to possess certain differentiability properties. In this section, we will generalize the functions that we can integrate slightly so that we can integrate a larger family of complex functions. This will lead us to the Cauchy's Integral Formula, which extends Cauchy's Theorem to functions analytic in an annulus. However, first we need to explore the concept of analytic functions.

A function $f(z)$ is analytic in domain D if for every open disk $|z - z_0| < \rho$ lying in D, $f(z)$ can be represented as a power series in z_0. Namely,

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n.$$

This series converges uniformly and absolutely inside the circle of convergence, $|z - z_0| < R$, with radius of convergence R. [See the Appendix for a review of convergence.]

There are various types of complex-valued functions.

A holomorphic function is (complex) differentiable in a neighborhood of every point in its domain.

An analytic function has a convergent Taylor series expansion in a neighborhood of each point in its domain. We see here that analytic functions are holomorphic and vice versa.

If a function is holomorphic throughout the complex plane, then it is called an entire function.

Finally, a function which is holomorphic on all of its domain except at a set of isolated poles (to be defined later), then it is called a meromorphic function.

Since $f(z)$ can be written as a uniformly convergent power series, we can integrate it term by term over any simple closed contour in D containing z_0. In particular, we have to compute integrals like $\oint_C (z - z_0)^n\,dz$. As we will see in the homework problems, these integrals evaluate to zero for most n. Thus, we can show that for $f(z)$ analytic in D and on any closed contour C lying in D, $\oint_C f(z)\,dz = 0$. Also, f is a uniformly convergent sum of continuous functions, so $f(z)$ is also continuous. Thus, by Morera's Theorem, we have that $f(z)$ is differentiable if it is analytic. Often terms like analytic, differentiable, and holomorphic are used interchangeably, though there is a subtle distinction due to their definitions.

As examples of series expansions about a given point, we will consider series expansions and regions of convergence for $f(z) = \frac{1}{1+z}$.

Example 4.19. *Find the series expansion of* $f(z) = \frac{1}{1+z}$ *about* $z_0 = 0$.

This case is simple. From Chapter 1 we recall that $f(z)$ *is the sum of a geometric series for* $|z| < 1$. *We have*

$$f(z) = \frac{1}{1+z} = \sum_{n=0}^{\infty} (-z)^n.$$

Thus, this series expansion converges inside the unit circle ($|z| < 1$) *in the complex plane.*

Example 4.20. *Find the series expansion of* $f(z) = \frac{1}{1+z}$ *about* $z_0 = \frac{1}{2}$.

We now look into an expansion about a different point. We could compute the expansion coefficients using Taylor's formula for the coefficients. However, we can also make use of the formula for geometric series after rearranging the function. We seek an expansion in powers of $z - \frac{1}{2}$. *So, we rewrite the function in a form that is a function of* $z - \frac{1}{2}$. *Thus,*

$$f(z) = \frac{1}{1+z} = \frac{1}{1 + (z - \frac{1}{2} + \frac{1}{2})} = \frac{1}{\frac{3}{2} + (z - \frac{1}{2})}.$$

This is not quite in the form we need. It would be nice if the denominator were of the form of one plus something. [Note: This is similar to what we had seen in Example 1.33.] We can get the denominator into such a form by factoring out the $\frac{3}{2}$. *Then we would have*

$$f(z) = \frac{2}{3} \frac{1}{1 + \frac{2}{3}(z - \frac{1}{2})}.$$

The second factor now has the form $\frac{1}{1-r}$, *which would be the sum of a geometric series with first term* $a = 1$ *and ratio* $r = -\frac{2}{3}(z - \frac{1}{2})$ *provided that* $|r| < 1$. *Therefore, we have found that*

$$f(z) = \frac{2}{3} \sum_{n=0}^{\infty} \left[-\frac{2}{3}(z - \frac{1}{2}) \right]^n$$

for

$$\left| -\frac{2}{3}(z - \frac{1}{2}) \right| < 1.$$

This convergence interval can be rewritten as

$$\left| z - \frac{1}{2} \right| < \frac{3}{2},$$

which is a circle centered at $z = \frac{1}{2}$ *with radius* $\frac{3}{2}$.

In Figure 4.31 we show the regions of convergence for the power series expansions of $f(z) = \frac{1}{1+z}$ about $z = 0$ and $z = \frac{1}{2}$. We note that the first expansion gives that $f(z)$ is at least analytic inside the region $|z| < 1$. The second expansion shows that $f(z)$ is analytic in a larger region, $|z - \frac{1}{2}| < \frac{3}{2}$. We will see later that there are expansions which converge outside these regions and that some yield expansions involving negative powers of $z - z_0$.

We now present the main theorem of this section:

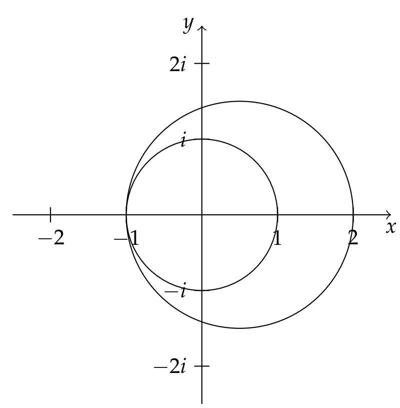


Figure 4.31: Regions of convergence for expansions of $f(z) = \frac{1}{1+z}$ about $z = 0$ and $z = \frac{1}{2}$.

Cauchy Integral Formula

Theorem 4.7. *Let $f(z)$ be analytic in $|z - z_0| < \rho$ and let C be the boundary (circle) of this disk. Then,*

$$f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0}\, dz. \tag{4.30}$$

Proof. In order to prove this, we first make use of the analyticity of $f(z)$. We insert the power series expansion of $f(z)$ about z_0 into the integrand. Then we have

$$\begin{aligned} \frac{f(z)}{z-z_0} &= \frac{1}{z-z_0}\left[\sum_{n=0}^{\infty} c_n (z-z_0)^n\right] \\ &= \frac{1}{z-z_0}\left[c_0 + c_1(z-z_0) + c_2(z-z_0)^2 + \dots\right] \\ &= \frac{c_0}{z-z_0} + \underbrace{c_1 + c_2(z-z_0) + \dots}_{\text{analytic function}}. \end{aligned} \tag{4.31}$$

As noted, the integrand can be written as

$$\frac{f(z)}{z-z_0} = \frac{c_0}{z-z_0} + h(z),$$

where $h(z)$ is an analytic function, since $h(z)$ is representable as a series expansion about z_0. We have already shown that analytic functions are differentiable, so by Cauchy's Theorem $\oint_C h(z)\, dz = 0$.

Noting also that $c_0 = f(z_0)$ is the first term of a Taylor series expansion about $z = z_0$, we have

$$\oint_C \frac{f(z)}{z-z_0}\, dz = \oint_C \left[\frac{c_0}{z-z_0} + h(z)\right] dz = f(z_0) \oint_C \frac{1}{z-z_0}\, dz.$$

We need only compute the integral $\oint_C \frac{1}{z-z_0}\, dz$ to finish the proof of Cauchy's Integral Formula. This is done by parametrizing the circle, $|z - z_0| = \rho$, as shown in Figure 4.32. This is simply done by letting

$$z - z_0 = \rho e^{i\theta}.$$

(Note that this has the right complex modulus since $|e^{i\theta}| = 1$. Then $dz = i\rho e^{i\theta} d\theta$. Using this parametrization, we have

$$\oint_C \frac{dz}{z-z_0} = \int_0^{2\pi} \frac{i\rho e^{i\theta}\, d\theta}{\rho e^{i\theta}} = i\int_0^{2\pi} d\theta = 2\pi i.$$

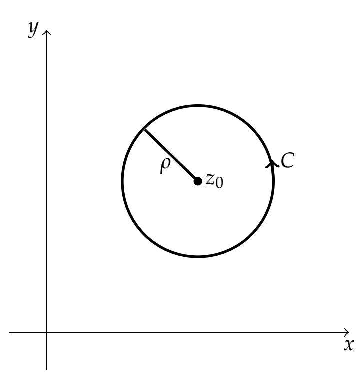


Figure 4.32: Circular contour used in proving the Cauchy Integral Formula.

Therefore,

$$\oint_C \frac{f(z)}{z-z_0}\, dz = f(z_0) \oint_C \frac{1}{z-z_0}\, dz = 2\pi i f(z_0),$$

as was to be shown. □

Example 4.21. *Compute* $\oint_{|z|=4} \frac{\cos z}{z^2-6z+5}\, dz$.

In order to apply the Cauchy Integral Formula, we need to factor the denominator, $z^2 - 6z + 5 = (z-1)(z-5)$. *We next locate the zeros of the denominator. In Figure 4.33 we show the contour and the points* $z = 1$ *and* $z = 5$. *The only point inside the region bounded by the contour is* $z = 1$. *Therefore, we can apply the Cauchy Integral Formula for* $f(z) = \frac{\cos z}{z-5}$ *to the integral*

$$\int_{|z|=4} \frac{\cos z}{(z-1)(z-5)}\, dz = \int_{|z|=4} \frac{f(z)}{(z-1)}\, dz = 2\pi i f(1).$$

Therefore, we have

$$\int_{|z|=4} \frac{\cos z}{(z-1)(z-5)}\, dz = -\frac{\pi i \cos(1)}{2}.$$

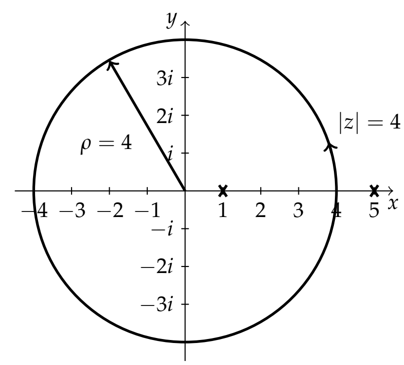


Figure 4.33: Circular contour used in computing $\oint_{|z|=4} \frac{\cos z}{z^2-6z+5}\, dz$.

We have shown that $f(z_0)$ has an integral representation for $f(z)$ analytic in $|z - z_0| < \rho$. In fact, all derivatives of an analytic function have an integral representation. This is given by

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz. \tag{4.32}$$

This can be proven following a derivation similar to that for the Cauchy Integral Formula. Inserting the Taylor series expansion for $f(z)$ into the integral on the right hand side, we have

$$\begin{aligned} \oint_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz &= \sum_{m=0}^{\infty} c_m \oint_C \frac{(z-z_0)^m}{(z-z_0)^{n+1}}\, dz \\ &= \sum_{m=0}^{\infty} c_m \oint_C \frac{dz}{(z-z_0)^{n-m+1}}. \end{aligned} \tag{4.33}$$

Picking $k = n - m$, the integrals in the sum can be computed using the following result:

$$\oint_C \frac{dz}{(z-z_0)^{k+1}} = \begin{cases} 0, & k \neq 0, \\ 2\pi i, & k = 0. \end{cases} \tag{4.34}$$

The proof is left for the exercises.

The only nonvanishing integrals, $\oint_C \frac{dz}{(z-z_0)^{n-m+1}}$, occur when $k = n - m = 0$, or $m = n$. Therefore, the series of integrals collapses to one term and we have

$$\oint_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz = 2\pi i c_n.$$

We finish the proof by recalling that the coefficients of the Taylor series expansion for $f(z)$ are given by

$$c_n = \frac{f^{(n)}(z_0)}{n!}.$$

Then,

$$\oint_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz = \frac{2\pi i}{n!} f^{(n)}(z_0)$$

and the result follows.

4.4.4 Laurent Series

Until this point we have only talked about series whose terms have nonnegative powers of $z - z_0$. It is possible to have series representations in which there are negative powers. In the last section we investigated expansions of $f(z) = \frac{1}{1+z}$ about $z = 0$ and $z = \frac{1}{2}$. The regions of convergence for each series was shown in Figure 4.31. Let us reconsider each of these expansions, but for values of z outside the region of convergence previously found.

Example 4.22. $f(z) = \frac{1}{1+z}$ *for* $|z| > 1$.

As before, we make use of the geometric series . Since $|z| > 1$*, we instead rewrite the function as*

$$f(z) = \frac{1}{1+z} = \frac{1}{z}\frac{1}{1+\frac{1}{z}}.$$

We now have the function in a form of the sum of a geometric series with first term $a = 1$ *and ratio* $r = -\frac{1}{z}$*. We note that* $|z| > 1$ *implies that* $|r| < 1$*. Thus, we have the geometric series*

$$f(z) = \frac{1}{z}\sum_{n=0}^{\infty}\left(-\frac{1}{z}\right)^n.$$

This can be re-indexed[3] *as*

$$f(z) = \sum_{n=0}^{\infty}(-1)^n z^{-n-1} = \sum_{j=1}^{\infty}(-1)^{j-1}z^{-j}.$$

Note that this series, which converges outside the unit circle, $|z| > 1$*, has negative powers of* z*.*

[3] Re-indexing a series is often useful in series manipulations. In this case, we have the series

$$\sum_{n=0}^{\infty}(-1)^n z^{-n-1} = z^{-1} - z^{-2} + z^{-3} + \ldots.$$

The index is n. You can see that the index does not appear when the sum is expanded showing the terms. The summation index is sometimes referred to as a dummy index for this reason. Reindexing allows one to rewrite the shorthand summation notation while capturing the same terms. In this example, the exponents are $-n-1$. We can simplify the notation by letting $-n-1 = -j$, or $j = n+1$. Noting that $j = 1$ when $n = 0$, we get the sum $\sum_{j=1}^{\infty}(-1)^{j-1}z^{-j}$.

Example 4.23. $f(z) = \frac{1}{1+z}$ *for* $|z - \frac{1}{2}| > \frac{3}{2}$.

As before, we express this in a form in which we can use a geometric series expansion. We seek powers of $z - \frac{1}{2}$*. So, we add and subtract* $\frac{1}{2}$ *to the* z *to obtain:*

$$f(z) = \frac{1}{1+z} = \frac{1}{1 + (z - \frac{1}{2} + \frac{1}{2})} = \frac{1}{\frac{3}{2} + (z - \frac{1}{2})}.$$

Instead of factoring out the $\frac{3}{2}$ *as we had done in Example 4.20, we factor out the* $(z - \frac{1}{2})$ *term. Then, we obtain*

$$f(z) = \frac{1}{1+z} = \frac{1}{(z-\frac{1}{2})}\frac{1}{\left[1 + \frac{3}{2}(z - \frac{1}{2})^{-1}\right]}.$$

Now we identify $a = 1$ *and* $r = -\frac{3}{2}(z - \frac{1}{2})^{-1}$*. This leads to the series*

$$\begin{aligned} f(z) &= \frac{1}{z - \frac{1}{2}}\sum_{n=0}^{\infty}\left(-\frac{3}{2}(z - \frac{1}{2})^{-1}\right)^n \\ &= \sum_{n=0}^{\infty}\left(-\frac{3}{2}\right)^n\left(z - \frac{1}{2}\right)^{-n-1}. \end{aligned} \tag{4.35}$$

This converges for $|z - \frac{1}{2}| > \frac{3}{2}$ *and can also be re-indexed to verify that this series involves negative powers of* $z - \frac{1}{2}$*.*

This leads to the following theorem:

Theorem 4.8. *Let $f(z)$ be analytic in an annulus, $R_1 < |z - z_0| < R_2$, with C a positively oriented simple closed curve around z_0 and inside the annulus as shown in Figure 4.34. Then,*

$$f(z) = \sum_{j=0}^{\infty} a_j (z - z_0)^j + \sum_{j=1}^{\infty} b_j (z - z_0)^{-j},$$

with

$$a_j = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{j+1}}\, dz,$$

and

$$b_j = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{-j+1}}\, dz.$$

The above series can be written in the more compact form

$$f(z) = \sum_{j=-\infty}^{\infty} c_j (z - z_0)^j.$$

Such a series expansion is called a Laurent series expansion, named after its discoverer Pierre Alphonse Laurent (1813 - 1854).

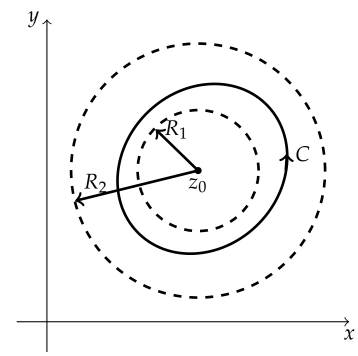


Figure 4.34: This figure shows an annulus, $R_1 < |z - z_0| < R_2$, with C a positively oriented simple closed curve around z_0 and inside the annulus.

Example 4.24. *Expand $f(z) = \frac{1}{(1-z)(2+z)}$ in the annulus $1 < |z| < 2$.*

Using partial fractions, we can write this as

$$f(z) = \frac{1}{3}\left[\frac{1}{1-z} + \frac{1}{2+z}\right].$$

We can expand the first fraction, $\frac{1}{1-z}$, as an analytic function in the region $|z| > 1$ and the second fraction, $\frac{1}{2+z}$, as an analytic function in $|z| < 2$. This is done as follows. First, we write

$$\frac{1}{2+z} = \frac{1}{2[1 - (-\frac{z}{2})]} = \frac{1}{2}\sum_{n=0}^{\infty}\left(-\frac{z}{2}\right)^n.$$

Then, we write

$$\frac{1}{1-z} = -\frac{1}{z[1 - \frac{1}{z}]} = -\frac{1}{z}\sum_{n=0}^{\infty}\frac{1}{z^n}.$$

Therefore, in the common region, $1 < |z| < 2$, we have that

$$\begin{aligned}\frac{1}{(1-z)(2+z)} &= \frac{1}{3}\left[\frac{1}{2}\sum_{n=0}^{\infty}\left(-\frac{z}{2}\right)^n - \sum_{n=0}^{\infty}\frac{1}{z^{n+1}}\right] \\ &= \sum_{n=0}^{\infty}\frac{(-1)^n}{6(2^n)}z^n + \sum_{n=1}^{\infty}\frac{(-1)}{3}z^{-n}. \qquad (4.36)\end{aligned}$$

We note that this is not a Taylor series expansion due to the existence of terms with negative powers in the second sum.

Example 4.25. *Find series representations of* $f(z) = \frac{1}{(1-z)(2+z)}$ *throughout the complex plane.*

In the last example we found series representations of $f(z) = \frac{1}{(1-z)(2+z)}$ *in the annulus* $1 < |z| < 2$. *However, we can also find expansions which converge for other regions. We first write*

$$f(z) = \frac{1}{3}\left[\frac{1}{1-z} + \frac{1}{2+z}\right].$$

We then expand each term separately.

The first fraction, $\frac{1}{1-z}$, *can be written as the sum of the geometric series*

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n, \quad |z| < 1.$$

This series converges inside the unit circle. We indicate this by region 1 in Figure 4.35.

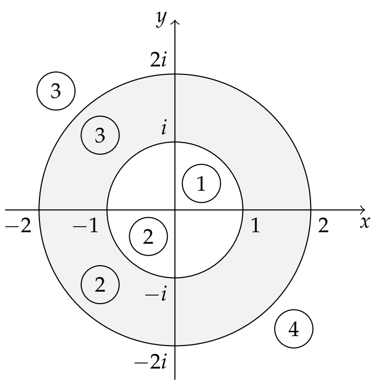


Figure 4.35: Regions of convergence for Laurent expansions of $f(z) = \frac{1}{(1-z)(2+z)}$.

In the last example, we showed that the second fraction, $\frac{1}{2+z}$, *has the series expansion*

$$\frac{1}{2+z} = \frac{1}{2[1-(-\frac{z}{2})]} = \frac{1}{2}\sum_{n=0}^{\infty}\left(-\frac{z}{2}\right)^n.$$

which converges in the circle $|z| < 2$. *This is labeled as region 2 in Figure 4.35.*

Regions 1 and 2 intersect for $|z| < 1$, *so, we can combine these two series representations to obtain*

$$\frac{1}{(1-z)(2+z)} = \frac{1}{3}\left[\sum_{n=0}^{\infty} z^n + \frac{1}{2}\sum_{n=0}^{\infty}\left(-\frac{z}{2}\right)^n\right], \quad |z| < 1.$$

In the annulus, $1 < |z| < 2$, *we had already seen in the last example that we needed a different expansion for the fraction* $\frac{1}{1-z}$. *We looked for an expansion in powers of* $1/z$ *which would converge for large values of* z. *We had found that*

$$\frac{1}{1-z} = -\frac{1}{z\left(1-\frac{1}{z}\right)} = -\frac{1}{z}\sum_{n=0}^{\infty}\frac{1}{z^n}, \quad |z| > 1.$$

This series converges in region 3 in Figure 4.35. Combining this series with the one for the second fraction, we obtain a series representation that converges in the overlap of regions 2 and 3. Thus, in the annulus $1 < |z| < 2$, *we have*

$$\frac{1}{(1-z)(2+z)} = \frac{1}{3}\left[\frac{1}{2}\sum_{n=0}^{\infty}\left(-\frac{z}{2}\right)^n - \sum_{n=0}^{\infty}\frac{1}{z^{n+1}}\right].$$

So far, we have series representations for $|z| < 2$. *The only region not covered yet is outside this disk,* $|z| > 2$. *In Figure 4.35 we see that series 3, which converges in region 3, will converge in the last section of the complex plane. We just need one more series expansion for* $1/(2+z)$ *for large* z. *Factoring out a* z *in the denominator, we can write this as a geometric series with* $r = 2/z$:

$$\frac{1}{2+z} = \frac{1}{z[\frac{2}{z}+1]} = \frac{1}{z}\sum_{n=0}^{\infty}\left(-\frac{2}{z}\right)^n.$$

This series converges for $|z| > 2$. Therefore, it converges in region 4 and the final series representation is

$$\frac{1}{(1-z)(2+z)} = \frac{1}{3}\left[\frac{1}{z}\sum_{n=0}^{\infty}\left(-\frac{2}{z}\right)^n - \sum_{n=0}^{\infty}\frac{1}{z^{n+1}}\right].$$

4.4.5 Singularities and The Residue Theorem

In the last section we found, that we could integrate functions satisfying some analyticity properties along contours without using detailed parametrizations around the contours. We can deform contours if the function is analytic in the region between the original and new contour. In this section we will extend our tools for performing contour integrals.

The integrand in the Cauchy Integral Formula was of the form $g(z) = \frac{f(z)}{z-z_0}$, where $f(z)$ is well behaved at z_0. The point $z = z_0$ is called a singularity of $g(z)$, as $g(z)$ is not defined there. More specifically, a singularity of $f(z)$ is a point at which $f(z)$ fails to be analytic.

Singularities of complex functions.

We can also classify these singularities. Typically, these are isolated singularities. As we saw from the proof of the Cauchy Integral Formula, $g(z) = \frac{f(z)}{z-z_0}$ has a Laurent series expansion about $z = z_0$, given by

$$g(z) = \frac{f(z_0)}{z-z_0} + f'(z_0) + \frac{1}{2}f''(z_0)(z-z_0) + \ldots.$$

It is the nature of the first term that gives information about the type of singularity that $g(z)$ has. Namely, in order to classify the singularities of $f(z)$, we look at the principal part of the Laurent series of $f(z)$ about $z = z_0$, $\sum_{j-1}^{\infty} b_j(z-z_0)^{-j}$, which consists of the negative powers of $z - z_0$.

Classification of singularities.

There are three types of singularities: removable, poles, and essential singularities. They are defined as follows:

1. If $f(z)$ is bounded near z_0, then z_0 is a removable singularity.
2. If there are a finite number of terms in the principal part of the Laurent series of $f(z)$ about $z = z_0$, then z_0 is called a pole.
3. If there are an infinite number of terms in the principal part of the Laurent series of $f(z)$ about $z = z_0$, then z_0 is called an essential singularity.

Example 4.26. $f(z) = \frac{\sin z}{z}$ *has a removable singularity at* $z = 0$.

At first it looks like there is a possible singularity at $z = 0$, since the denominator is zero at $z = 0$. However, we know from the first semester of calculus that $\lim_{z\to 0}\frac{\sin z}{z} = 1$. *Furthermore, we can expand* $\sin z$ *about* $z = 0$ *and see that*

$$\frac{\sin z}{z} = \frac{1}{z}(z - \frac{z^3}{3!} + \ldots) = 1 - \frac{z^2}{3!} + \ldots.$$

Thus, there are only nonnegative powers in the series expansion. So, $z = 0$ is a removable singularity.

Example 4.27. *$f(z) = \frac{e^z}{(z-1)^n}$ has poles at $z = 1$ for n a positive integer.*

For $n = 1$, we have $f(z) = \frac{e^z}{z-1}$. This function has a singularity at $z = 1$. The series expansion is found by expanding e^z about $z = 1$:

$$f(z) = \frac{e}{z-1}e^{z-1} = \frac{e}{z-1} + e + \frac{e}{2!}(z-1) + \ldots .$$

Note that the principal part of the Laurent series expansion about $z = 1$ only has one term, $\frac{e}{z-1}$. Therefore, $z = 1$ is a pole. Since the leading term has an exponent of -1, $z = 1$ is called a pole of order one, or a simple pole.

Simple pole.

For $n = 2$ we have $f(z) = \frac{e^z}{(z-1)^2}$. The series expansion is again found by expanding e^z about $z = 1$:

$$f(z) = \frac{e}{(z-1)^2}e^{z-1} = \frac{e}{(z-1)^2} + \frac{e}{z-1} + \frac{e}{2!} + \frac{e}{3!}(z-1) + \ldots .$$

Note that the principal part of the Laurent series has two terms involving $(z-1)^{-2}$ and $(z-1)^{-1}$. Since the leading term has an exponent of -2, $z = 1$ is called a pole of order 2, or a double pole.

Double pole.

Example 4.28. *$f(z) = e^{\frac{1}{z}}$ has an essential singularity at $z = 0$.*

In this case, we have the series expansion about $z = 0$ given by

$$f(z) = e^{\frac{1}{z}} = \sum_{n=0}^{\infty} \frac{\left(\frac{1}{z}\right)^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} z^{-n}.$$

We see that there are an infinite number of terms in the principal part of the Laurent series. So, this function has an essential singularity at $z = 0$.

Poles of order k.

In the above examples we have seen poles of order one (a simple pole) and two (a double pole). In general, we can say that $f(z)$ has a pole of order k at z_0 if and only if $(z-z_0)^k f(z)$ has a removable singularity at z_0, but $(z-z_0)^{k-1} f(z)$ for $k > 0$ does not.

Example 4.29. *Determine the order of the pole of $f(z) = \cot z \csc z$ at $z = 0$.*

First we rewrite $f(z)$ in terms of sines and cosines.

$$f(z) = \cot z \csc z = \frac{\cos z}{\sin^2 z}.$$

We note that the denominator vanishes at $z = 0$.

How do we know that the pole is not a simple pole? Well, we check to see if $(z-0)f(z)$ has a removable singularity at $z = 0$:

$$\begin{aligned} \lim_{z\to 0}(z-0)f(z) &= \lim_{z\to 0} \frac{z\cos z}{\sin^2 z} \\ &= \left(\lim_{z\to 0} \frac{z}{\sin z}\right)\left(\lim_{z\to 0} \frac{\cos z}{\sin z}\right) \\ &= \lim_{z\to 0} \frac{\cos z}{\sin z}. \end{aligned} \tag{4.37}$$

We see that this limit is undefined. So now we check to see if $(z-0)^2 f(z)$ has a removable singularity at $z = 0$:

$$\lim_{z\to 0}(z-0)^2 f(z) = \lim_{z\to 0} \frac{z^2 \cos z}{\sin^2 z}$$

$$= \left(\lim_{z\to 0}\frac{z}{\sin z}\right)\left(\lim_{z\to 0}\frac{z\cos z}{\sin z}\right)$$
$$= \lim_{z\to 0}\frac{z}{\sin z}\cos(0) = 1. \qquad (4.38)$$

In this case, we have obtained a finite, nonzero, result. So, $z = 0$ is a pole of order 2.

We could have also relied on series expansions. Expanding both the sine and cosine functions in a Taylor series expansion, we have

$$f(z) = \frac{\cos z}{\sin^2 z} = \frac{1 - \frac{1}{2!}z^2 + \ldots}{(z - \frac{1}{3!}z^3 + \ldots)^2}.$$

Factoring a z from the expansion in the denominator,

$$f(z) = \frac{1}{z^2}\frac{1 - \frac{1}{2!}z^2 + \ldots}{(1 - \frac{1}{3!}z + \ldots)^2} = \frac{1}{z^2}\left(1 + O(z^2)\right),$$

we can see that the leading term will be a $1/z^2$, indicating a pole of order 2.

We will see how knowledge of the poles of a function can aid in the computation of contour integrals. We now show that if a function, $f(z)$, has a pole of order k, then

Integral of a function with a simple pole inside C.

$$\oint_C f(z)\,dz = 2\pi i \operatorname{Res}[f(z); z_0],$$

Residues of a function with poles of order k.

where we have defined $\operatorname{Res}[f(z); z_0]$ as the residue of $f(z)$ at $z = z_0$. In particular, for a pole of order k the residue is given by

Residues for Poles of order k

$$\operatorname{Res}[f(z); z_0] = \lim_{z\to z_0}\frac{1}{(k-1)!}\frac{d^{k-1}}{dz^{k-1}}\left[(z-z_0)^k f(z)\right]. \qquad (4.39)$$

Proof. Let $\phi(z) = (z-z_0)^k f(z)$ be an analytic function. Then $\phi(z)$ has a Taylor series expansion about z_0. As we had seen in the last section, we can write the integral representation of any derivative of ϕ as

$$\phi^{(k-1)}(z_0) = \frac{(k-1)!}{2\pi i}\oint_C \frac{\phi(z)}{(z-z_0)^k}\,dz.$$

Inserting the definition of $\phi(z)$, we then have

$$\phi^{(k-1)}(z_0) = \frac{(k-1)!}{2\pi i}\oint_C f(z)\,dz.$$

Solving for the integral, we have the following result:

$$\oint_C f(z)\,dz = \frac{2\pi i}{(k-1)!}\frac{d^{k-1}}{dz^{k-1}}\left[(z-z_0)^k f(z)\right]_{z=z_0}$$
$$\equiv 2\pi i \operatorname{Res}[f(z); z_0] \qquad (4.40)$$

□

Note: If z_0 is a simple pole, the residue is easily computed as

$$Res[f(z); z_0] = \lim_{z\to z_0} (z - z_0) f(z).$$

The residue for a simple pole.

In fact, one can show (Problem 18) that for g and h analytic functions at z_0, with $g(z_0) \neq 0$, $h(z_0) = 0$, and $h'(z_0) \neq 0$,

$$\text{Res}\left[\frac{g(z)}{h(z)}; z_0\right] = \frac{g(z_0)}{h'(z_0)}.$$

Example 4.30. *Find the residues of* $f(z) = \frac{z-1}{(z+1)^2(z^2+4)}$.

$f(z)$ has poles at $z = -1$, $z = 2i$, and $z = -2i$. The pole at $z = -1$ is a double pole (pole of order 2). The other poles are simple poles. We compute those residues first:

$$\begin{aligned} Res[f(z); 2i] &= \lim_{z\to 2i} (z - 2i) \frac{z-1}{(z+1)^2(z+2i)(z-2i)} \\ &= \lim_{z\to 2i} \frac{z-1}{(z+1)^2(z+2i)} \\ &= \frac{2i-1}{(2i+1)^2(4i)} = -\frac{1}{50} - \frac{11}{100}i. \end{aligned} \tag{4.41}$$

$$\begin{aligned} Res[f(z); -2i] &= \lim_{z\to -2i} (z + 2i) \frac{z-1}{(z+1)^2(z+2i)(z-2i)} \\ &= \lim_{z\to -2i} \frac{z-1}{(z+1)^2(z-2i)} \\ &= \frac{-2i-1}{(-2i+1)^2(-4i)} = -\frac{1}{50} + \frac{11}{100}i. \end{aligned} \tag{4.42}$$

For the double pole, we have to do a little more work.

$$\begin{aligned} Res[f(z); -1] &= \lim_{z\to -1} \frac{d}{dz}\left[(z+1)^2 \frac{z-1}{(z+1)^2(z^2+4)}\right] \\ &= \lim_{z\to -1} \frac{d}{dz}\left[\frac{z-1}{z^2+4}\right] \\ &= \lim_{z\to -1} \frac{d}{dz}\left[\frac{z^2+4-2z(z-1)}{(z^2+4)^2}\right] \\ &= \lim_{z\to -1} \frac{d}{dz}\left[\frac{-z^2+2z+4}{(z^2+4)^2}\right] \\ &= \frac{1}{25}. \end{aligned} \tag{4.43}$$

Example 4.31. *Find the residue of $f(z) = \cot z$ at $z = 0$.*

We write $f(z) = \cot z = \frac{\cos z}{\sin z}$ and note that $z = 0$ is a simple pole. Thus,

$$Res[\cot z; z = 0] = \lim_{z\to 0} \frac{z \cos z}{\sin z} = \cos(0) = 1.$$

The residue of $f(z)$ at z_0 is the coefficient of the $(z - z_0)^{-1}$ term, $c_{-1} = b_1$, of the Laurent series expansion about z_0.

Another way to find the residue of a function $f(z)$ at a singularity z_0 is to look at the Laurent series expansion about the singularity. This is because the residue of $f(z)$ at z_0 is the coefficient of the $(z - z_0)^{-1}$ term, or $c_{-1} = b_1$.

Example 4.32. *Find the residue of* $f(z) = \frac{1}{z(3-z)}$ *at* $z = 0$ *using a Laurent series expansion.*

First, we need the Laurent series expansion about $z = 0$ *of the form* $\sum_{-\infty}^{\infty} c_n z^n$. *A partial fraction expansion gives*

$$f(z) = \frac{1}{z(3-z)} = \frac{1}{3}\left(\frac{1}{z} + \frac{1}{3-z}\right).$$

The first term is a power of z. *The second term needs to be written as a convergent series for small* z. *This is given by*

$$\begin{aligned} \frac{1}{3-z} &= \frac{1}{3(1-z/3)} \\ &= \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{z}{3}\right)^n. \end{aligned} \tag{4.44}$$

Thus, we have found

$$f(z) = \frac{1}{3}\left(\frac{1}{z} + \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{z}{3}\right)^n\right).$$

The coefficient of z^{-1} *can be read off to give* $Res[f(z); z = 0] = \frac{1}{3}$.

Example 4.33. *Find the residue of* $f(z) = z\cos\frac{1}{z}$ *at* $z = 0$ *using a Laurent series expansion.*

Finding the residue at an essential singularity.

In this case, $z = 0$ *is an essential singularity. The only way to find residues at essential singularities is to use Laurent series. Since*

$$\cos z = 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \frac{1}{6!}z^6 + \ldots,$$

then we have

$$\begin{aligned} f(z) &= z\left(1 - \frac{1}{2!z^2} + \frac{1}{4!z^4} - \frac{1}{6!z^6} + \ldots\right) \\ &= z - \frac{1}{2!z} + \frac{1}{4!z^3} - \frac{1}{6!z^5} + \ldots. \end{aligned} \tag{4.45}$$

From the second term we have that $Res[f(z); z = 0] = -\frac{1}{2}$.

We are now ready to use residues in order to evaluate integrals.

Example 4.34. *Evaluate* $\oint_{|z|=1} \frac{dz}{\sin z}$.

We begin by looking for the singularities of the integrand. These are located at values of z *for which* $\sin z = 0$. *Thus,* $z = 0, \pm\pi, \pm 2\pi, \ldots,$ *are the singularities. However, only* $z = 0$ *lies inside the contour, as shown in Figure 4.36. We note further that* $z = 0$ *is a simple pole, since*

$$\lim_{z\to 0}(z-0)\frac{1}{\sin z} = 1.$$

Therefore, the residue is one and we have

$$\oint_{|z|=1} \frac{dz}{\sin z} = 2\pi i.$$

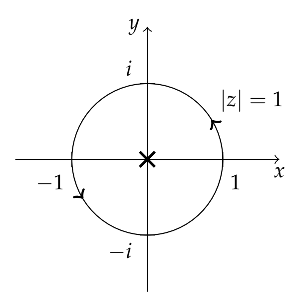


Figure 4.36: Contour for computing $\oint_{|z|=1} \frac{dz}{\sin z}$.

In general, we could have several poles of different orders. For example, we will be computing

$$\oint_{|z|=2} \frac{dz}{z^2-1}.$$

The integrand has singularities at $z^2 - 1 = 0$, or $z = \pm 1$. Both poles are inside the contour, as seen in Figure 4.38. One could do a partial fraction decomposition and have two integrals with one pole each integral. Then, the result could be found by adding the residues from each pole.

In general, when there are several poles, we can use the Residue Theorem:

The Residue Theorem.

The Residue Theorem

Theorem 4.9. *Let $f(z)$ be a function which has poles z_j, $j = 1, \ldots, N$ inside a simple closed contour C and no other singularities in this region. Then,*

$$\oint_C f(z)\,dz = 2\pi i \sum_{j=1}^{N} Res[f(z); z_j], \tag{4.46}$$

where the residues are computed using Equation (4.39),

$$Res[f(z); z_0] = \lim_{z\to z_0} \frac{1}{(k-1)!} \frac{d^{k-1}}{dz^{k-1}} \left[(z-z_0)^k f(z) \right].$$

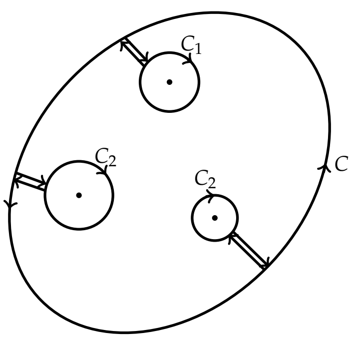


Figure 4.37: A depiction of how one cuts out poles to prove that the integral around C is the sum of the integrals around circles with the poles at the center of each.

The proof of this theorem is based upon the contours shown in Figure 4.37. One constructs a new contour C' by encircling each pole, as show in the figure. Then one connects a path from C to each circle. In the figure two separated paths along the cut are shown only to indicate the direction followed on the cut. The new contour is then obtained by following C and crossing each cut as it is encountered. Then one goes around a circle in the negative sense and returns along the cut to proceed around C. The sum of the contributions to the contour integration involve two integrals for each cut, which will cancel due to the opposing directions. Thus, we are left with

$$\oint_{C'} f(z)\,dz = \oint_C f(z)\,dz - \oint_{C_1} f(z)\,dz - \oint_{C_2} f(z)\,dz - \oint_{C_3} f(z)\,dz = 0.$$

Of course, the sum is zero because $f(z)$ is analytic in the enclosed region, since all singularities have been cut out. Solving for $\oint_C f(z)\,dz$, one has that this integral is the sum of the integrals around the separate poles, which can be evaluated with single residue computations. Thus, the result is that $\oint_C f(z)\,dz$ is $2\pi i$ times the sum of the residues.

Example 4.35. *Evaluate $\oint_{|z|=2} \frac{dz}{z^2-1}$.*

We first note that there are two poles in this integral since

$$\frac{1}{z^2-1} = \frac{1}{(z-1)(z+1)}.$$

In Figure 4.38 we plot the contour and the two poles, denoted by an "x." Since both poles are inside the contour, we need to compute the residues for each one. They are each simple poles, so we have

$$
\begin{aligned}
Res\left[\frac{1}{z^2-1}; z=1\right] &= \lim_{z\to 1}(z-1)\frac{1}{z^2-1} \\
&= \lim_{z\to 1}\frac{1}{z+1} = \frac{1}{2}, \qquad (4.47)
\end{aligned}
$$

and

$$
\begin{aligned}
Res\left[\frac{1}{z^2-1}; z=-1\right] &= \lim_{z\to -1}(z+1)\frac{1}{z^2-1} \\
&= \lim_{z\to -1}\frac{1}{z-1} = -\frac{1}{2}. \qquad (4.48)
\end{aligned}
$$

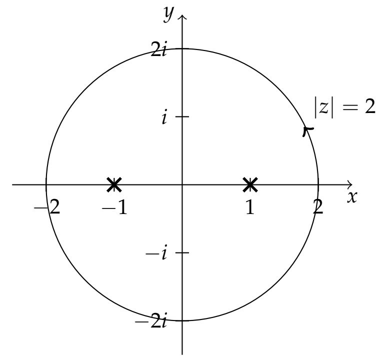


Figure 4.38: Contour for computing $\oint_{|z|=2} \frac{dz}{z^2-1}$.

Then,

$$\oint_{|z|=2} \frac{dz}{z^2-1} = 2\pi i(\frac{1}{2}-\frac{1}{2}) = 0.$$

Example 4.36. *Evaluate* $\oint_{|z|=3} \frac{z^2+1}{(z-1)^2(z+2)}\,dz$.

In this example, there are two poles $z=1,-2$ inside the contour. [See Figure 4.39.] $z=1$ is a second-order pole and $z=-2$ is a simple pole. Therefore, we need to compute the residues at each pole of $f(z)=\frac{z^2+1}{(z-1)^2(z+2)}$:

$$
\begin{aligned}
Res[f(z); z=1] &= \lim_{z\to 1}\frac{1}{1!}\frac{d}{dz}\left[(z-1)^2\frac{z^2+1}{(z-1)^2(z+2)}\right] \\
&= \lim_{z\to 1}\left(\frac{z^2+4z-1}{(z+2)^2}\right) \\
&= \frac{4}{9}. \qquad (4.49)
\end{aligned}
$$

$$
\begin{aligned}
Res[f(z); z=-2] &= \lim_{z\to -2}(z+2)\frac{z^2+1}{(z-1)^2(z+2)} \\
&= \lim_{z\to -2}\frac{z^2+1}{(z-1)^2} \\
&= \frac{5}{9}. \qquad (4.50)
\end{aligned}
$$

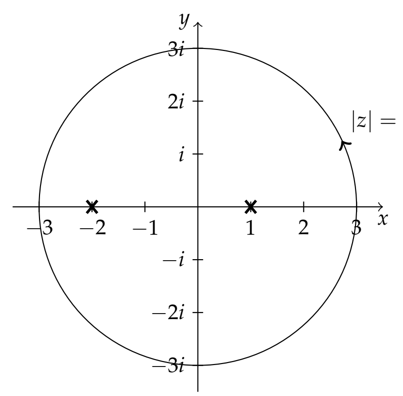


Figure 4.39: Contour for computing $\oint_{|z|=3} \frac{z^2+1}{(z-1)^2(z+2)}\,dz$.

The evaluation of the integral is found by computing $2\pi i$ times the sum of the residues:

$$\oint_{|z|=3} \frac{z^2+1}{(z-1)^2(z+2)}\,dz = 2\pi i\left(\frac{4}{9}+\frac{5}{9}\right) = 2\pi i.$$

Example 4.37. *Compute* $\oint_{|z|=2} z^3 e^{2/z}\,dz$.

In this case, $z=0$ is an essential singularity and is inside the contour. A Laurent series expansion about $z=0$ gives

$$
\begin{aligned}
z^3 e^{2/z} &= z^3\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{2}{z}\right)^n \\
&= \sum_{n=0}^{\infty}\frac{2^n}{n!}z^{3-n} \\
&= z^3+\frac{2}{2!}z^2+\frac{4}{3!}z+\frac{8}{4!}+\frac{16}{5!z}+\dots. \qquad (4.51)
\end{aligned}
$$

The residue is the coefficient of z^{-1}*, or* $Res[z^3e^{2/z}; z=0] = -\frac{2}{15}$*. Therefore,*

$$\oint_{|z|=2} z^3 e^{2/z}\,dz = \frac{4}{15}\pi i.$$

Example 4.38. *Evaluate* $\int_0^{2\pi} \frac{d\theta}{2+\cos\theta}$*.*

Here we have a real integral in which there are no signs of complex functions. In fact, we could apply simpler methods from a calculus course to do this integral, attempting to write $1+\cos\theta = 2\cos^2\frac{\theta}{2}$*. However, we do not get very far.*

One trick, useful in computing integrals whose integrand is in the form $f(\cos\theta, \sin\theta)$*, is to transform the integration to the complex plane through the transformation* $z = e^{i\theta}$*. Then,*

Computation of integrals of functions of sines and cosines, $f(\cos\theta, \sin\theta)$.

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2} = \frac{1}{2}\left(z+\frac{1}{z}\right),$$

$$\sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i} = -\frac{i}{2}\left(z-\frac{1}{z}\right).$$

Under this transformation, $z = e^{i\theta}$*, the integration now takes place around the unit circle in the complex plane. Noting that* $dz = ie^{i\theta}\,d\theta = iz\,d\theta$*, we have*

$$\begin{aligned}\int_0^{2\pi} \frac{d\theta}{2+\cos\theta} &= \oint_{|z|=1} \frac{\frac{dz}{iz}}{2+\frac{1}{2}\left(z+\frac{1}{z}\right)} \\ &= -i\oint_{|z|=1} \frac{dz}{2z+\frac{1}{2}(z^2+1)} \\ &= -2i\oint_{|z|=1} \frac{dz}{z^2+4z+1}. \qquad (4.52)\end{aligned}$$

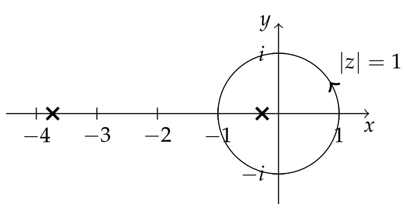


Figure 4.40: Contour for computing $\int_0^{2\pi} \frac{d\theta}{2+\cos\theta}$.

We can apply the Residue Theorem to the resulting integral. The singularities occur at the roots of $z^2+4z+1=0$*. Using the quadratic formula, we have the roots* $z = -2\pm\sqrt{3}$*.*

The location of these poles are shown in Figure 4.40. Only $z = -2+\sqrt{3}$ *lies inside the integration contour. We will therefore need the residue of* $f(z) = \frac{-2i}{z^2+4z+1}$ *at this simple pole:*

$$\begin{aligned}Res[f(z); z=-2+\sqrt{3}] &= \lim_{z\to-2+\sqrt{3}} (z-(-2+\sqrt{3}))\frac{-2i}{z^2+4z+1} \\ &= -2i \lim_{z\to-2+\sqrt{3}} \frac{z-(-2+\sqrt{3})}{(z-(-2+\sqrt{3}))(z-(-2-\sqrt{3}))} \\ &= -2i \lim_{z\to-2+\sqrt{3}} \frac{1}{z-(-2-\sqrt{3})} \\ &= \frac{-2i}{-2+\sqrt{3}-(-2-\sqrt{3})} \\ &= \frac{-i}{\sqrt{3}} \\ &= \frac{-i\sqrt{3}}{3}. \qquad (4.53)\end{aligned}$$

Therefore, we have

$$\int_0^{2\pi} \frac{d\theta}{2+\cos\theta} = -2i \oint_{|z|=1} \frac{dz}{z^2+4z+1} = 2\pi i \left(\frac{-i\sqrt{3}}{3} \right) = \frac{2\pi\sqrt{3}}{3}. \quad (4.54)$$

Before moving on to further applications, we note that there is another way to compute the integral in the last example. Karl Theodor Wilhelm Weierstraß (1815 - 1897) introduced a substitution method for computing integrals involving rational functions of sine and cosine. One makes the substitution $t = \tan\frac{\theta}{2}$ and converts the integrand into a rational function of t. One can show that this substitution implies that

The Weierstraß substitution method.

$$\sin\theta = \frac{2t}{1+t^2}, \quad \cos\theta = \frac{1-t^2}{1+t^2},$$

and

$$d\theta = \frac{2dt}{1+t^2}.$$

The details are left for Problem 8. In order to see how it works, we will apply the Weierstraß substitution method to the last example.

Example 4.39. *Apply the Weierstraß substitution method to compute* $\int_0^{2\pi} \frac{d\theta}{2+\cos\theta}$.

$$\begin{aligned} \int_0^{2\pi} \frac{d\theta}{2+\cos\theta} &= \int_{-\infty}^{\infty} \frac{1}{2+\frac{1-t^2}{1+t^2}} \frac{2dt}{1+t^2} \\ &= 2\int_{-\infty}^{\infty} \frac{dt}{t^2+3} \\ &= \frac{2}{3}\sqrt{3} \left[\tan^{-1}\left(\frac{\sqrt{3}}{3}t \right) \right]_{-\infty}^{\infty} = \frac{2\pi\sqrt{3}}{3}. \end{aligned} \quad (4.55)$$

4.4.6 *Infinite Integrals*

INFINITE INTEGRALS OF THE FORM $\int_{-\infty}^{\infty} f(x)\,dx$ occur often in physics. They can represent wave packets, wave diffraction, Fourier transforms, and arise in other applications. In this section, we will see that such integrals may be computed by extending the integration to a contour in the complex plane.

Recall from your calculus experience that these integrals are improper integrals. The way that one determines if improper integrals exist, or converge, is to carefully compute these integrals using limits such as

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{R\to\infty} \int_{-R}^{R} f(x)\,dx.$$

For example, we evaluate the integral of $f(x) = x$ as

$$\int_{-\infty}^{\infty} x\,dx = \lim_{R\to\infty} \int_{-R}^{R} x\,dx = \lim_{R\to\infty} \left(\frac{R^2}{2} - \frac{(-R)^2}{2} \right) = 0.$$

One might also be tempted to carry out this integration by splitting the integration interval, $(-\infty, 0] \cup [0, \infty)$. However, the integrals $\int_0^\infty x\,dx$ and $\int_{-\infty}^0 x\,dx$ do not exist. A simple computation confirms this.

$$\int_0^\infty x\,dx = \lim_{R\to\infty} \int_0^R x\,dx = \lim_{R\to\infty} \left(\frac{R^2}{2}\right) = \infty.$$

Therefore,

$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^0 f(x)\,dx + \int_0^\infty f(x)\,dx$$

does not exist while $\lim_{R\to\infty} \int_{-R}^R f(x)\,dx$ does exist. We will be interested in computing the latter type of integral. Such an integral is called the Cauchy Principal Value Integral and is denoted with either a P, PV, or a bar through the integral:

The Cauchy principal value integral.

$$P\int_{-\infty}^\infty f(x)\,dx = PV\int_{-\infty}^\infty f(x)\,dx = ⨍_{-\infty}^\infty f(x)\,dx = \lim_{R\to\infty} \int_{-R}^R f(x)\,dx. \quad (4.56)$$

If there is a discontinuity in the integral, one can further modify this definition of principal value integral to bypass the singularity. For example, if $f(x)$ is continuous on $a \le x \le b$ and not defined at $x = x_0 \in [a, b]$, then

$$\int_a^b f(x)\,dx = \lim_{\epsilon\to 0} \left(\int_a^{x_0-\epsilon} f(x)\,dx + \int_{x_0+\epsilon}^b f(x)\,dx \right).$$

In our discussions we will be computing integrals over the real line in the Cauchy principal value sense.

Example 4.40. *Compute $\int_{-1}^1 \frac{dx}{x^3}$ in the Cauchy Principal Value sense.*
In this case, $f(x) = \frac{1}{x^3}$ is not defined at $x = 0$. So, we have

$$\begin{aligned} \int_{-1}^1 \frac{dx}{x^3} &= \lim_{\epsilon\to 0} \left(\int_{-1}^{-\epsilon} \frac{dx}{x^3} + \int_\epsilon^1 \frac{dx}{x^3} \right) \\ &= \lim_{\epsilon\to 0} \left(-\frac{1}{2x^2}\Big|_{-1}^{-\epsilon} - \frac{1}{2x^2}\Big|_\epsilon^1 \right) = 0. \end{aligned} \quad (4.57)$$

Computation of real integrals by embedding the problem in the complex plane.

We now proceed to the evaluation of principal value integrals using complex integration methods. We want to evaluate the integral $\int_{-\infty}^\infty f(x)\,dx$. We will extend this into an integration in the complex plane. We extend $f(x)$ to $f(z)$ and assume that $f(z)$ is analytic in the upper half plane ($\text{Im}(z) > 0$) except at isolated poles. We then consider the integral $\int_{-R}^R f(x)\,dx$ as an integral over the interval $(-R, R)$. We view this interval as a piece of a larger contour C_R obtained by completing the contour with a semicircle Γ_R of radius R extending into the upper half plane as shown in Figure 4.41. Note that a similar construction is sometimes needed extending the integration into the lower half plane ($\text{Im}(z) < 0$) as we will later see.

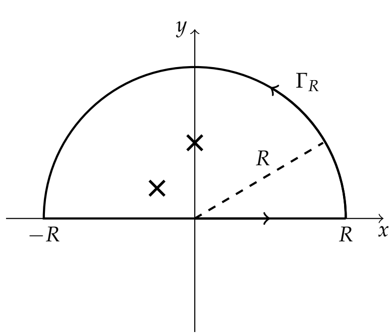


Figure 4.41: Contours for computing $P\int_{-\infty}^\infty f(x)\,dx$.

The integral around the entire contour C_R can be computed using the Residue Theorem and is related to integrations over the pieces of the contour by

$$\oint_{C_R} f(z)\,dz = \int_{\Gamma_R} f(z)\,dz + \int_{-R}^R f(z)\,dz. \quad (4.58)$$

Taking the limit $R \to \infty$ and noting that the integral over $(-R, R)$ is the desired integral, we have

$$P \int_{-\infty}^{\infty} f(x)\,dx = \oint_C f(z)\,dz - \lim_{R\to\infty} \int_{\Gamma_R} f(z)\,dz, \qquad (4.59)$$

where we have identified C as the limiting contour as R gets large.

Now the key to carrying out the integration is that the second integral vanishes in the limit. This is true if $R|f(z)| \to 0$ along Γ_R as $R \to \infty$. This can be seen by the following argument. We parametrize the contour Γ_R using $z = Re^{i\theta}$. Then, when $|f(z)| < M(R)$,

$$\begin{aligned} \left| \int_{\Gamma_R} f(z)\,dz \right| &= \left| \int_0^{2\pi} f(Re^{i\theta}) Re^{i\theta}\,d\theta \right| \\ &\leq R \int_0^{2\pi} \left| f(Re^{i\theta}) \right| d\theta \\ &< RM(R) \int_0^{2\pi} d\theta \\ &= 2\pi RM(R). \end{aligned} \qquad (4.60)$$

So, if $\lim_{R\to\infty} RM(R) = 0$, then $\lim_{R\to\infty} \int_{\Gamma_R} f(z)\,dz = 0$.

We now demonstrate how to use complex integration methods in evaluating integrals over real valued functions.

Example 4.41. *Evaluate* $\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$.

We already know how to do this integral using calculus without complex analysis. We have that

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \lim_{R\to\infty} \left(2\tan^{-1} R \right) = 2\left(\frac{\pi}{2}\right) = \pi.$$

We will apply the methods of this section and confirm this result. The needed contours are shown in Figure 4.42 and the poles of the integrand are at $z = \pm i$. We first write the integral over the bounded contour C_R as the sum of an integral from $-R$ to R along the real axis plus the integral over the semicircular arc in the upper half complex plane,

$$\int_{C_R} \frac{dz}{1+z^2} = \int_{-R}^{R} \frac{dx}{1+x^2} + \int_{\Gamma_R} \frac{dz}{1+z^2}.$$

Next, we let R get large.

We first note that $f(z) = \frac{1}{1+z^2}$ goes to zero fast enough on Γ_R as R gets large.

$$R|f(z)| = \frac{R}{|1 + R^2 e^{2i\theta}|} = \frac{R}{\sqrt{1 + 2R^2\cos\theta + R^4}}.$$

Thus, as $R \to \infty$, $R|f(z)| \to 0$ and $C_R \to C$. So,

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \oint_C \frac{dz}{1+z^2}.$$

We need only compute the residue at the enclosed pole, $z = i$.

$$Res[f(z); z = i] = \lim_{z\to i} (z - i) \frac{1}{1+z^2} = \lim_{z\to i} \frac{1}{z+i} = \frac{1}{2i}.$$

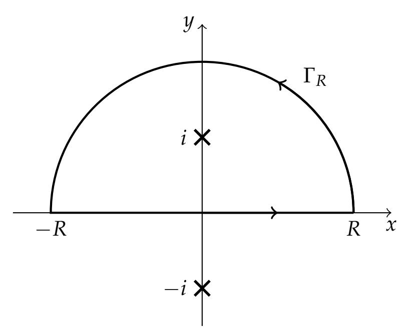


Figure 4.42: Contour for computing $\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$.

Then, using the Residue Theorem, we have

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = 2\pi i \left(\frac{1}{2i}\right) = \pi.$$

Example 4.42. *Evaluate* $P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx$.

For this example, the integral is unbounded at $z = 0$. Constructing the contours as before we are faced for the first time with a pole lying on the contour. We cannot ignore this fact. We can proceed with the computation by carefully going around the pole with a small semicircle of radius ϵ, as shown in Figure 4.43. Then the principal value integral computation becomes

$$P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx = \lim_{\epsilon\to 0, R\to\infty} \left(\int_{-R}^{-\epsilon} \frac{\sin x}{x}\, dx + \int_{\epsilon}^{R} \frac{\sin x}{x}\, dx \right). \tag{4.61}$$

We will also need to rewrite the sine function in terms of exponentials in this integral. There are two approaches that we could take. First, we could employ the definition of the sine function in terms of complex exponentials. This gives two integrals to compute:

$$P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx = \frac{1}{2i}\left(P\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\, dx - P\int_{-\infty}^{\infty} \frac{e^{-ix}}{x}\, dx \right). \tag{4.62}$$

The other approach would be to realize that the sine function is the imaginary part of an exponential, $\operatorname{Im} e^{ix} = \sin x$. Then, we would have

$$P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx = \operatorname{Im}\left(P\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\, dx \right). \tag{4.63}$$

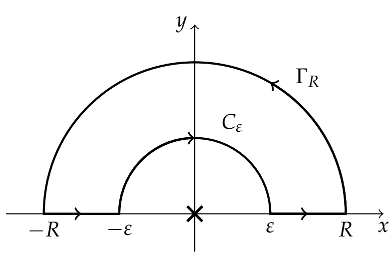


Figure 4.43: Contour for computing $P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx$.

We first consider $P\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\, dx$, which is common to both approaches. We use the contour in Figure 4.43. Then we have

$$\oint_{C_R} \frac{e^{iz}}{z}\, dz = \int_{\Gamma_R} \frac{e^{iz}}{z}\, dz + \int_{-R}^{-\epsilon} \frac{e^{iz}}{z}\, dz + \int_{C_\epsilon} \frac{e^{iz}}{z}\, dz + \int_{\epsilon}^{R} \frac{e^{iz}}{z}\, dz.$$

The integral $\oint_{C_R} \frac{e^{iz}}{z}\, dz$ vanishes since there are no poles enclosed in the contour! The sum of the second and fourth integrals gives the integral we seek as $\epsilon \to 0$ and $R \to \infty$. The integral over Γ_R will vanish as R gets large according to Jordan's Lemma.

Jordan's Lemma gives conditions for the vanishing of integrals over Γ_R as R gets large. We state a version of Jordan's Lemma here for reference and give a proof at the end of this chapter.

Jordan's Lemma

If $f(z)$ converges uniformly to zero as $z \to \infty$, then

$$\lim_{R\to\infty} \int_{C_R} f(z) e^{ikz}\, dz = 0,$$

where $k > 0$ and C_R is the upper half of the circle $|z| = R$.

A similar result applies for $k < 0$, but one closes the contour in the lower half plane. [See Section 4.4.8 for the proof of Jordan's Lemma.]

The remaining integral around the small semicircular arc must be done separately. We have

$$\int_{C_\epsilon} \frac{e^{iz}}{z}\, dz = \int_\pi^0 \frac{\exp(i\epsilon e^{i\theta})}{\epsilon e^{i\theta}} i\epsilon e^{i\theta}\, d\theta = -\int_0^\pi i \exp(i\epsilon e^{i\theta})\, d\theta.$$

Taking the limit as ϵ goes to zero, the integrand goes to i and we have

$$\int_{C_\epsilon} \frac{e^{iz}}{z}\, dz = -\pi i.$$

So far, we have that

$$P\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\, dx = -\lim_{\epsilon\to 0}\int_{C_\epsilon} \frac{e^{iz}}{z}\, dz = \pi i.$$

Note that we have not previously done integrals in which a singularity lies on the contour. One can show, as in this example, that points on the contour can be accounted for using half of a residue (times $2\pi i$). For the semicircle C_ϵ, the reader can verify this. The negative sign comes from going clockwise around the semicircle.

At this point, we can get the answer using the second approach in Equation (4.63). Namely,

$$P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx = \mathrm{Im}\left(P\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\, dx\right) = \mathrm{Im}(\pi i) = \pi. \tag{4.64}$$

It is instructive to carry out the first approach in Equation (4.62). We will need to compute $P\int_{-\infty}^{\infty} \frac{e^{-ix}}{x}\, dx$. This is done in a similar maaner to the above computation, being careful with the sign changes due to the orientations of the contours as shown in Figure 4.44.

We note that the contour is closed in the lower half plane. This is because $k < 0$ in the application of Jordan's Lemma. One can understand why this is the case from the following observation. Consider the exponential in Jordan's Lemma. Let $z = z_R + iz_I$. Then,

$$e^{ikz} = e^{ik(z_R + iz_I)} = e^{-kz_I} e^{ikz_R}.$$

As $|z|$ gets large, the second factor just oscillates. The first factor would go to zero if $kz_I > 0$. So, if $k > 0$, we would close the contour in the upper half plane. If $k < 0$, then we would close the contour in the lower half plane. In the current computation, $k = -1$, so we use the lower half plane.

Working out the details, we find the same value for

$$P\int_{-\infty}^{\infty} \frac{e^{-ix}}{x}\, dx = \pi i.$$

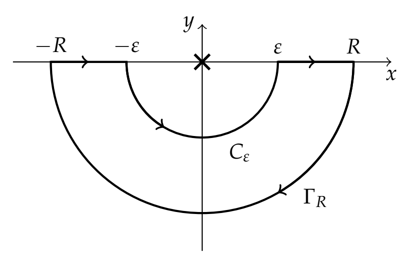


Figure 4.44: Contour in the lower half plane for computing $P\int_{-\infty}^{\infty} \frac{e^{-ix}}{x}\, dx$.

Finally, we can compute the original integral as

$$\begin{aligned} P\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx &= \frac{1}{2i}\left(P\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\, dx - P\int_{-\infty}^{\infty} \frac{e^{-ix}}{x}\, dx\right) \\ &= \frac{1}{2i}(\pi i + \pi i) \\ &= \pi. \end{aligned} \tag{4.65}$$

This is the same result as we obtained using Equation(4.63).

Example 4.43. *Evaluate* $\oint_{|z|=1} \frac{dz}{z^2+1}$.

In this example, there are two simple poles, $z = \pm i$ lying on the contour, as seen in Figure 4.45. This problem is similar to Problem 1c, except we will do it using contour integration instead of a parametrization. We bypass the two poles by drawing small semicircles around them. Since the poles are not included in the closed contour, the Residue Theorem tells us that the integral over the path vanishes. We can write the full integration as a sum over three paths: $C_\pm$ for the semicircles and C for the original contour with the poles cut out. Then we take the limit as the semicircle radii go to zero. So,

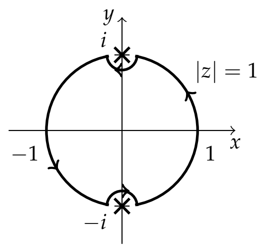


Figure 4.45: Example with poles on contour.

$$0 = \int_C \frac{dz}{z^2+1} + \int_{C_+} \frac{dz}{z^2+1} + \int_{C_-} \frac{dz}{z^2+1}.$$

The integral over the semicircle around i can be done using the parametrization $z = i + \epsilon e^{i\theta}$. Then $z^2 + 1 = 2i\epsilon e^{i\theta} + \epsilon^2 e^{2i\theta}$. This gives

$$\int_{C_+} \frac{dz}{z^2+1} = \lim_{\epsilon\to 0} \int_0^{-\pi} \frac{i\epsilon e^{i\theta}}{2i\epsilon e^{i\theta} + \epsilon^2 e^{2i\theta}}\, d\theta = \frac{1}{2}\int_0^{-\pi} d\theta = -\frac{\pi}{2}.$$

As in the last example, we note that this is just πi times the residue, $\operatorname{Res}\left[\frac{1}{z^2+1}; z = i\right] = \frac{1}{2i}$. Since the path is traced clockwise, we find that the contribution is $-\pi i \operatorname{Res} = -\frac{\pi}{2}$, which is what we obtained above. A Similar computation will give the contribution from $z = -i$ as $\frac{\pi}{2}$. Adding these values gives the total contribution from $C_\pm$ as zero. So, the final result is that

$$\oint_{|z|=1} \frac{dz}{z^2+1} = 0.$$

Example 4.44. *Evaluate $\int_{-\infty}^{\infty} \frac{e^{ax}}{1+e^x}\, dx$, for $0 < a < 1$.*

In dealing with integrals involving exponentials or hyperbolic functions it is sometimes useful to use different types of contours. This example is one such case. We will replace x with z and integrate over the contour in Figure 4.46. Letting $R \to \infty$, the integral along the real axis is the integral that we desire. The integral along the path for $y = 2\pi$ leads to a multiple of this integral since $z = x + 2\pi i$ along this path. Integration along the vertical paths vanishes as $R \to \infty$. This is captured in the following integrals:

$$\begin{aligned}\oint_{C_R} \frac{e^{az}}{1+e^z} dz &= \int_{-R}^{R} \frac{e^{ax}}{1+e^x}\, dx + \int_0^{2\pi} \frac{e^{a(R+iy)}}{1+e^{R+iy}}\, dy \\ &\quad + \int_R^{-R} \frac{e^{a(x+2\pi i)}}{1+e^{x+2\pi i}}\, dx + \int_{2\pi}^{0} \frac{e^{a(-R+iy)}}{1+e^{-R+iy}}\, dy \qquad (4.66)\end{aligned}$$

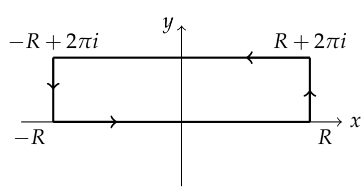


Figure 4.46: Example using a rectangular contour.

We can now let $R \to \infty$. For large R, the second integral decays as $e^{(a-1)R}$ and the fourth integral decays as e^{-aR}. Thus, we are left with

$$\begin{aligned}\oint_C \frac{e^{az}}{1+e^z} dz &= \lim_{R\to\infty}\left(\int_{-R}^{R} \frac{e^{ax}}{1+e^x} dx - e^{2\pi i a}\int_{-R}^{R} \frac{e^{ax}}{1+e^x} dx\right) \\ &= (1 - e^{2\pi i a})\int_{-\infty}^{\infty} \frac{e^{ax}}{1+e^x} dx. \qquad (4.67)\end{aligned}$$

We need only evaluate the left contour integral using the Residue Theorem. The poles are found from

$$1 + e^z = 0.$$

Within the contour, this is satisfied by $z = i\pi$. So,

$$Res\left[\frac{e^{az}}{1+e^z}; z = i\pi\right] = \lim_{z\to i\pi}(z - i\pi)\frac{e^{az}}{1+e^z} = -e^{i\pi a}.$$

Applying the Residue Theorem, we have

$$(1 - e^{2\pi i a})\int_{-\infty}^{\infty}\frac{e^{ax}}{1+e^x}dx = -2\pi i e^{i\pi a}.$$

Therefore, we have found that

$$\int_{-\infty}^{\infty}\frac{e^{ax}}{1+e^x}dx = \frac{-2\pi i e^{i\pi a}}{1 - e^{2\pi i a}} = \frac{\pi}{\sin \pi a}, \quad 0 < a < 1.$$

4.4.7 *Integration over Multivalued Functions*

We have seen that some complex functions inherently possess multivaluedness; that is, such "functions" do not evaluate to a single value, but have many values. The key examples were $f(z) = z^{1/n}$ and $f(z) = \ln z$. The nth roots have n distinct values and logarithms have an infinite number of values as determined by the range of the resulting arguments. We mentioned that the way to handle multivaluedness is to assign different branches to these functions, introduce a branch cut, and glue them together at the branch cuts to form Riemann surfaces. In this way we can draw continuous paths along the Riemann surfaces as we move from one Riemann sheet to another.

Before we do examples of contour integration involving multivalued functions, let's first try to get a handle on multivaluedness in a simple case. We will consider the square root function,

$$w = z^{1/2} = r^{1/2}e^{i(\frac{\theta}{2}+k\pi)}, \quad k = 0, 1.$$

There are two branches, corresponding to each k value. If we follow a path not containing the origin, then we stay in the same branch, so the final argument (θ) will be equal to the initial argument. However, if we follow a path that encloses the origin, this will not be true. In particular, for an initial point on the unit circle, $z_0 = e^{i\theta_0}$, we have its image as $w_0 = e^{i\theta_0/2}$. However, if we go around a full revolution, $\theta = \theta_0 + 2\pi$, then

$$z_1 = e^{i\theta_0 + 2\pi i} = e^{i\theta_0},$$

but

$$w_1 = e^{(i\theta_0+2\pi i)/2} = e^{i\theta_0/2}e^{\pi i} \neq w_0.$$

Here we obtain a final argument (θ) that is not equal to the initial argument! Somewhere, we have crossed from one branch to another. Points, such as the origin in this example, are called branch points. Actually, there are two branch points, because we can view the closed path around the origin as a closed path around complex infinity in the compactified complex plane. However, we will not go into that at this time.

We can demonstrate this in the following figures. In Figure 4.47 we show how the points A through E are mapped from the z-plane into the w-plane under the square root function for the principal branch, $k = 0$. As we trace out the unit circle in the z-plane, we only trace out a semicircle in the w-plane. If we consider the branch $k = 1$, we then trace out a semicircle in the lower half plane, as shown in Figure 4.48 following the points from F to J.

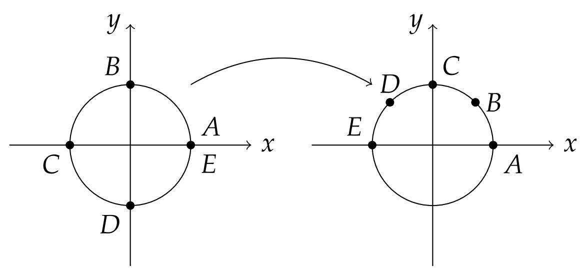


Figure 4.47: In this figure we show how points on the unit circle in the z-plane are mapped to points in the w-plane under the principal square root function.

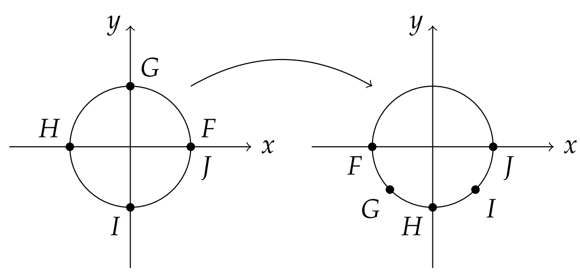


Figure 4.48: In this figure we show how points on the unit circle in the z-plane are mapped to points in the w-plane under the square root function for the second branch, $k = 1$.

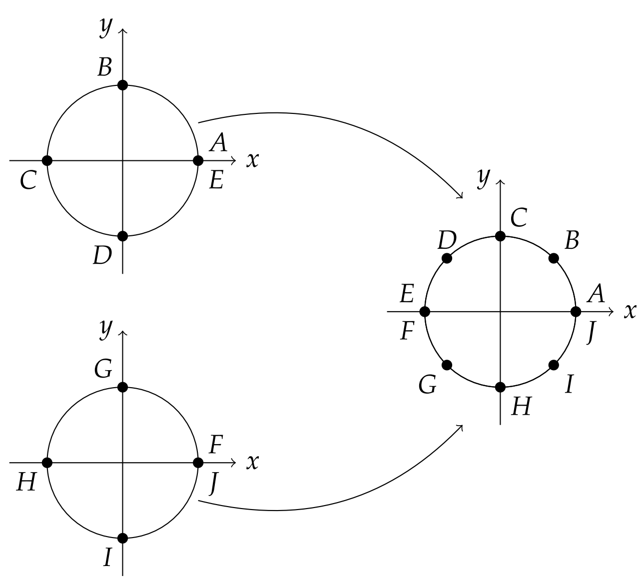


Figure 4.49: In this figure we show the combined mapping using two branches of the square root function.

We can combine these into one mapping depicting how the two complex planes corresponding to each branch provide a mapping to the w-plane. This is shown in Figure 4.49.

A common way to draw this domain, which looks like two separate complex planes, would be to glue them together. Imagine cutting each plane along the positive x-axis, extending between the two branch points, $z=0$ and $z=\infty$. As one approaches the cut on the principal branch, one can move onto the glued second branch. Then one continues around the origin on this branch until one once again reaches the cut. This cut is glued to the principal branch in such a way that the path returns to its starting point. The resulting surface we obtain is the Riemann surface shown in Figure 4.50. Note that there is nothing that forces us to place the branch cut at a particular place. For example, the branch cut could be along the positive real axis, the negative real axis, or any path connecting the origin and complex infinity.

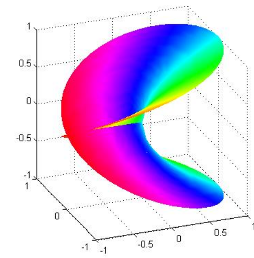


Figure 4.50: Riemann surface for $f(z)=z^{1/2}$.

We now look at examples involving integrals of multivalued functions.

Example 4.45. *Evaluate* $\int_0^\infty \frac{\sqrt{x}}{1+x^2}\,dx$.

We consider the contour integral $\oint_C \frac{\sqrt{z}}{1+z^2}\,dz$.

The first thing we can see in this problem is the square root function in the integrand. Being that there is a multivalued function, we locate the branch point and determine where to draw the branch cut. In Figure 4.51 we show the contour that we will use in this problem. Note that we picked the branch cut along the positive x-axis.

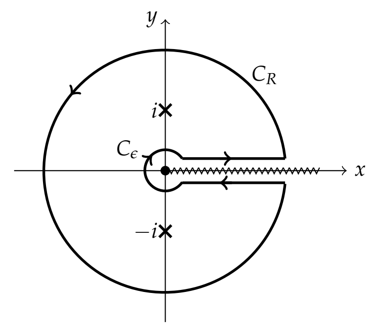


Figure 4.51: An example of a contour which accounts for a branch cut.

We take the contour C to be positively oriented, being careful to enclose the two poles and to hug the branch cut. It consists of two circles. The outer circle C_R *is a circle of radius R and the inner circle* C_ϵ *will have a radius of* ϵ*. The sought-after answer will be obtained by letting* $R\to\infty$ *and* $\epsilon\to 0$*. On the large circle we have that the integrand goes to zero fast enough as* $R\to\infty$*. The integral around the small circle vanishes as* $\epsilon\to 0$*. We can see this by parametrizing the circle as* $z=\epsilon e^{i\theta}$ *for* $\theta\in[0,2\pi]$:

$$\begin{aligned}\oint_{C_\epsilon}\frac{\sqrt{z}}{1+z^2}\,dz &= \int_0^{2\pi}\frac{\sqrt{\epsilon e^{i\theta}}}{1+(\epsilon e^{i\theta})^2}\,i\epsilon e^{i\theta}\,d\theta \\ &= i\epsilon^{3/2}\int_0^{2\pi}\frac{e^{3i\theta/2}}{1+(\epsilon^2 e^{2i\theta})}\,d\theta. \qquad (4.68)\end{aligned}$$

It should now be easy to see that as $\epsilon\to 0$*, this integral vanishes.*

The integral above the branch cut is the one we are seeking, $\int_0^\infty \frac{\sqrt{x}}{1+x^2}\,dx$*. The integral under the branch cut, where* $z=re^{2\pi i}$*, is*

$$\begin{aligned}\int\frac{\sqrt{z}}{1+z^2}\,dz &= \int_\infty^0\frac{\sqrt{re^{2\pi i}}}{1+r^2e^{4\pi i}}\,dr \\ &= \int_0^\infty\frac{\sqrt{r}}{1+r^2}\,dr. \qquad (4.69)\end{aligned}$$

We note that this is the same as that above the cut.

Up to this point, we have that the contour integral, as $R\to\infty$ *and* $\epsilon\to 0$*, is*

$$\oint_C\frac{\sqrt{z}}{1+z^2}\,dz = 2\int_0^\infty\frac{\sqrt{x}}{1+x^2}\,dx.$$

In order to finish this problem, we need the residues at the two simple poles.

$$Res\left[\frac{\sqrt{z}}{1+z^2}; z=i\right] = \frac{\sqrt{i}}{2i} = \frac{\sqrt{2}}{4}(1+i),$$

$$Res\left[\frac{\sqrt{z}}{1+z^2}; z=-i\right] = \frac{\sqrt{-i}}{-2i} = \frac{\sqrt{2}}{4}(1-i).$$

So,

$$2\int_0^\infty \frac{\sqrt{x}}{1+x^2}dx = 2\pi i\left(\frac{\sqrt{2}}{4}(1+i) + \frac{\sqrt{2}}{4}(1-i)\right) = \pi\sqrt{2}.$$

Finally, we have the value of the integral that we were seeking,

$$\int_0^\infty \frac{\sqrt{x}}{1+x^2}dx = \frac{\pi\sqrt{2}}{2}.$$

Example 4.46. *Compute* $\int_a^\infty f(x)\,dx$ *using contour integration involving logarithms.*[4]

[4] This approach was originally published in Neville, E. H., 1945, Indefinite integration by means of residues. *The Mathematical Student*, **13**, 16-35, and discussed in Duffy, D. G., *Transform Methods for Solving Partial Differential Equations*, 1994.

In this example, we will apply contour integration to the integral

$$\oint_C f(z)\ln(a-z)\,dz$$

for the contour shown in Figure 4.52.

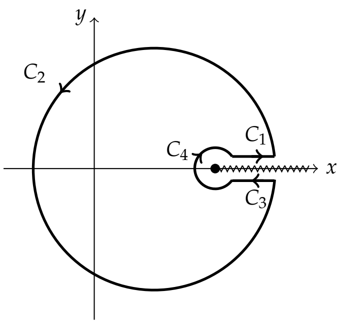


Figure 4.52: Contour needed to compute $\oint_C f(z)\ln(a-z)\,dz$.

We will assume that $f(z)$ is single valued and vanishes as $|z| \to \infty$. We will choose the branch cut to span from the origin along the positive real axis. Employing the Residue Theorem and breaking up the integrals over the pieces of the contour in Figure 4.52, we have schematically that

$$2\pi i \sum Res[f(z)\ln(a-z)] = \left(\int_{C_1} + \int_{C_2} + \int_{C_3} + \int_{C_4}\right) f(z)\ln(a-z)\,dz.$$

First of all, we assume that $f(z)$ is well behaved at $z = a$ and vanishes fast enough as $|z| = R \to \infty$. Then, the integrals over C_2 and C_4 will vanish. For example, for the path C_4, we let $z = a + \epsilon e^{i\theta}$, $0 < \theta < 2\pi$. Then,

$$\int_{C_4} f(z)\ln(a-z)\,dz. = \lim_{\epsilon\to 0}\int_{2\pi}^0 f(a+\epsilon e^{i\theta})\ln(\epsilon e^{i\theta})i\epsilon e^{i\theta}\,d\theta.$$

If $f(a)$ is well behaved, then we only need to show that $\lim_{\epsilon\to 0}\epsilon\ln\epsilon = 0$. This is left to the reader.

Similarly, we consider the integral over C_2 as R gets large,

$$\int_{C_2} f(z)\ln(a-z)\,dz = \lim_{R\to\infty}\int_0^{2\pi} f(Re^{i\theta})\ln(Re^{i\theta})iRe^{i\theta}\,d\theta.$$

Thus, we need only require that

$$\lim_{R\to\infty} R\ln R|f(Re^{i\theta})| = 0.$$

Next, we consider the two straight line pieces. For C_1, the integration along the real axis occurs for $z = x$, so

$$\int_{C_1} f(z)\ln(a-z)\,dz = \int_a^\infty f(x)\ln(a-x)\,dz.$$

However, integration over C_3 requires noting that we need the branch for the logarithm such that $\ln z = \ln(a-x) + 2\pi i$. Then,

$$\int_{C_3} f(z)\ln(a-z)\,dz = \int_\infty^a f(x)[\ln(a-x)+2\pi i]\,dz.$$

Combining these results, we have

$$\begin{aligned} 2\pi i \sum Res[f(z)\ln(a-z)] &= \int_a^\infty f(x)\ln(a-x)\,dz \\ &\quad + \int_\infty^a f(x)[\ln(a-x)+2\pi i]\,dz. \\ &= -2\pi i \int_a^\infty f(x)\,dz. \end{aligned} \tag{4.70}$$

Therefore,

$$\int_a^\infty f(x)\,dx = -\sum Res[f(z)\ln(a-z)].$$

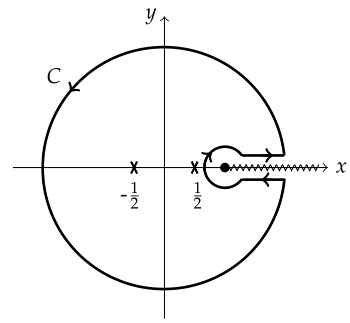


Figure 4.53: Contour needed to compute $\int_1^\infty \frac{dx}{4x^2-1}$.

Example 4.47. *Compute* $\int_1^\infty \frac{dx}{4x^2-1}$.

We can apply the last example to this case. We see from Figure 4.53 that the two poles at $z = \pm\frac{1}{2}$ are inside contour C. So, we compute the residues of $\frac{\ln(1-z)}{4z^2-1}$ at these poles and find that

$$\begin{aligned} \int_1^\infty \frac{dx}{4x^2-1} &= -Res\left[\frac{\ln(1-z)}{4z^2-1};\frac{1}{2}\right] - Res\left[\frac{\ln(1-z)}{4z^2-1};-\frac{1}{2}\right] \\ &= -\frac{\ln\frac{1}{2}}{4} + \frac{\ln\frac{3}{2}}{4} = \frac{\ln 3}{4}. \end{aligned} \tag{4.71}$$

4.4.8 *Appendix: Jordan's Lemma*

FOR COMPLETENESS, WE PROVE JORDAN'S LEMMA.

Theorem 4.10. *If $f(z)$ converges uniformly to zero as $z \to \infty$, then*

$$\lim_{R\to\infty} \int_{C_R} f(z)e^{ikz}\,dz = 0,$$

where $k > 0$ and C_R is the upper half of the circle $|z| = R$.

Proof. We consider the integral

$$I_R = \int_{C_R} f(z)e^{ikz}\,dz,$$

where $k > 0$ and C_R is the upper half of the circle $|z| = R$ in the complex plane. Let $z = Re^{i\theta}$ be a parametrization of C_R. Then,

$$I_R = \int_0^\pi f(Re^{i\theta})e^{ikR\cos\theta - aR\sin\theta}\, iRe^{i\theta}\,d\theta.$$

Since

$$\lim_{|z|\to\infty} f(z) = 0, \quad 0 \le \arg z \le \pi,$$

then for large $|R|$, $|f(z)| < \epsilon$ for some $\epsilon > 0$. Then,

$$\begin{aligned} |I_R| &= \left| \int_0^{\pi} f(Re^{i\theta}) e^{ikR\cos\theta - aR\sin\theta}\, iRe^{i\theta}\, d\theta \right| \\ &\leq \int_0^{\pi} \left| f(Re^{i\theta}) \right| \left| e^{ikR\cos\theta} \right| \left| e^{-aR\sin\theta} \right| \left| iRe^{i\theta} \right| d\theta \\ &\leq \epsilon R \int_0^{\pi} e^{-aR\sin\theta}\, d\theta \\ &= 2\epsilon R \int_0^{\pi/2} e^{-aR\sin\theta}\, d\theta. \end{aligned} \tag{4.72}$$

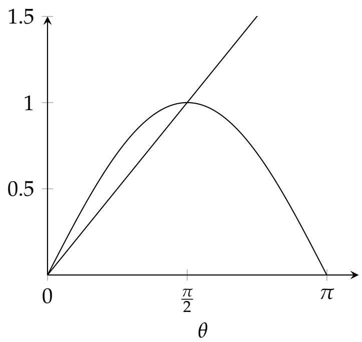


Figure 4.54: Plots of $y = \sin\theta$ and $y = \frac{2}{\pi}\theta$ to show where $\sin\theta \geq \frac{2}{\pi}\theta$.

The resulting integral still cannot be computed, but we can get a bound on it over the range $\theta \in [0, \pi/2]$. Note from Figure 4.54 that

$$\sin\theta \geq \frac{2}{\pi}\theta, \quad \theta \in [0, \pi/2].$$

Therefore, we have

$$|I_R| \leq 2\epsilon R \int_0^{\pi/2} e^{-2aR\theta/\pi}\, d\theta = \frac{2\epsilon R}{2aR/\pi}(1 - e^{-aR}).$$

For large R, we have

$$\lim_{R\to\infty} |I_R| \leq \frac{\pi\epsilon}{a}.$$

So, as $\epsilon \to 0$, the integral vanishes.

□

Problems

1. Write the following in standard form.

a. $(4 + 5i)(2 - 3i)$.

b. $(1 + i)^3$.

c. $\frac{5+3i}{1-i}$.

2. Write the following in polar form, $z = re^{i\theta}$.

a. $i - 1$.

b. $-2i$.

c. $\sqrt{3} + 3i$.

3. Write the following in rectangular form, $z = a + ib$.

a. $4e^{i\pi/6}$.

b. $\sqrt{2}e^{5i\pi/4}$.

c. $(1 - i)^{100}$.

4. Find all z such that $z^4 = 16i$. Write the solutions in rectangular form, $z = a + ib$, with no decimal approximation or trig functions.

5. Show that $\sin(x + iy) = \sin x \cosh y + i \cos x \sinh y$ using trigonometric identities and the exponential forms of these functions.

6. Find all z such that $\cos z = 2$, or explain why there are none. You will need to consider $\cos(x + iy)$ and equate real and imaginary parts of the resulting expression similar to Problem 5.

7. Find the principal value of i^i. Rewrite the base, i, as an exponential first.

8. Consider the circle $|z - 1| = 1$.

a. Rewrite the equation in rectangular coordinates by setting $z = x + iy$.

b. Sketch the resulting circle using part a.

c. Consider the image of the circle under the mapping $f(z) = z^2$, given by $|z^2 - 1| = 1$.

i. By inserting $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$, find the equation of the image curve in polar coordinates.

ii. Sketch the image curve. You may need to refer to your Calculus II text for polar plots. [Maple might help.]

9. Find the real and imaginary parts of the functions:

a. $f(z) = z^3$.

b. $f(z) = \sinh(z)$.

c. $f(z) = \cos \bar{z}$.

10. Find the derivative of each function in Problem 9 when the derivative exists. Otherwise, show that the derivative does not exist.

11. Let $f(z) = u + iv$ be differentiable. Consider the vector field given by $\mathbf{F} = v\mathbf{i} + u\mathbf{j}$. Show that the equations $\nabla \cdot \mathbf{F} = \mathbf{0}$ and $\nabla \times \mathbf{F} = \mathbf{0}$ are equivalent to the Cauchy-Riemann Equations. [You will need to recall from multivariable calculus the del operator, $\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$.]

12. What parametric curve is described by the function

$$\gamma(t) = (t - 3) + i(2t + 1),$$

$0 \le t \le 2$? [Hint: What would you do if you were instead considering the parametric equations $x = t - 3$ and $y = 2t + 1$?]

13. Write the equation that describes the circle of radius 3 which is centered at $z = 2 - i$ in (a) Cartesian form (in terms of x and y); (b) polar form (in terms of θ and r); (c) complex form (in terms of z, r, and $e^{i\theta}$).

14. Consider the function $u(x, y) = x^3 - 3xy^2$.

a. Show that $u(x, y)$ is harmonic; that is, $\nabla^2 u = 0$.

b. Find its harmonic conjugate, $v(x, y)$.

c. Find a differentiable function, $f(z)$, for which $u(x, y)$ is the real part.

d. Determine $f'(z)$ for the function in part c. [Use $f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$ and rewrite your answer as a function of z.]

15. Evaluate the following integrals:

a. $\int_C \bar{z}\, dz$, where C is the parabola $y = x^2$ from $z = 0$ to $z = 1 + i$.

b. $\int_C f(z)\, dz$, where $f(z) = 2z - \bar{z}$ and C is the path from $z = 0$ to $z = 2 + i$ consisting of two line segments from $z = 0$ to $z = 2$ and then $z = 2$ to $z = 2 + i$.

c. $\int_C \frac{1}{z^2+4}\, dz$ for C the positively oriented circle, $|z| = 2$. [Hint: Parametrize the circle as $z = 2e^{i\theta}$, multiply numerator and denominator by $e^{-i\theta}$, and put in trigonometric form.]

16. Let C be the positively oriented ellipse $3x^2 + y^2 = 9$. Define

$$F(z_0) = \int_C \frac{z^2 + 2z}{z - z_0}\, dz.$$

Find $F(2i)$ and $F(2)$. [Hint: Sketch the ellipse in the complex plane. Use the Cauchy Integral Theorem with an appropriate $f(z)$, or Cauchy's Theorem if z_0 is outside the contour.]

17. Show that

$$\int_C \frac{dz}{(z - 1 - i)^{n+1}} = \begin{cases} 0, & n \neq 0, \\ 2\pi i, & n = 0, \end{cases}$$

for C the boundary of the square $0 \leq x \leq 2$, $0 \leq y \leq 2$ taken counterclockwise. [Hint: Use the fact that contours can be deformed into simpler shapes (like a circle) as long as the integrand is analytic in the region between them. After picking a simpler contour, integrate using parametrization.]

18. Show that for g and h analytic functions at z_0, with $g(z_0) \neq 0$, $h(z_0) = 0$, and $h'(z_0) \neq 0$,

$$\text{Res}\left[\frac{g(z)}{h(z)}; z_0\right] = \frac{g(z_0)}{h'(z_0)}.$$

19. For the following, determine if the given point is a removable singularity, an essential singularity, or a pole (indicate its order).

a. $\frac{1-\cos z}{z^2}$, $\quad z = 0$.

b. $\frac{\sin z}{z^2}$, $\quad z = 0$.

c. $\frac{z^2-1}{(z-1)^2}$, $\quad z = 1$.

d. $ze^{1/z}$, $\quad z = 0$.

e. $\cos\frac{\pi}{z-\pi}$, $\quad z = \pi$.

20. Find the Laurent series expansion for $f(z) = \frac{\sinh z}{z^3}$ about $z = 0$. [Hint: You need to first do a MacLaurin series expansion for the hyperbolic sine.]

21. Find series representations for all indicated regions.

a. $f(z) = \frac{z}{z-1}$, $|z| < 1$, $|z| > 1$.

b. $f(z) = \frac{1}{(z-i)(z+2)}$, $|z| < 1$, $1 < |z| < 2$, $|z| > 2$. [Hint: Use partial fractions to write this as a sum of two functions first.]

22. Find the residues at the given points:

a. $\frac{2z^2+3z}{z-1}$ at $z = 1$.

b. $\frac{\ln(1+2z)}{z}$ at $z = 0$.

c. $\frac{\cos z}{(2z-\pi)^3}$ at $z = \frac{\pi}{2}$.

23. Consider the integral $\int_0^{2\pi} \frac{d\theta}{5-4\cos\theta}$.

a. Evaluate this integral by making the substitution $2\cos\theta = z + \frac{1}{z}$, $z = e^{i\theta}$, and using complex integration methods.

b. In the 1800's Weierstrass introduced a method for computing integrals involving rational functions of sine and cosine. One makes the substitution $t = \tan\frac{\theta}{2}$ and converts the integrand into a rational function of t. Note that the integration around the unit circle corresponds to $t \in (-\infty, \infty)$.

i. Show that

$$\sin\theta = \frac{2t}{1+t^2}, \quad \cos\theta = \frac{1-t^2}{1+t^2}.$$

ii. Show that

$$d\theta = \frac{2dt}{1+t^2}.$$

iii. Use the Weierstrass substitution to compute the above integral.

24. Do the following integrals:

a.

$$\oint_{|z-i|=3} \frac{e^z}{z^2+\pi^2}\, dz.$$

b.

$$\oint_{|z-i|=3} \frac{z^2-3z+4}{z^2-4z+3}\, dz.$$

c.

$$\int_{-\infty}^{\infty} \frac{\sin x}{x^2+4}\, dx.$$

[Hint: This is $\operatorname{Im}\int_{-\infty}^{\infty} \frac{e^{ix}}{x^2+4}\, dx$.]

25. Evaluate the integral $\int_0^{\infty} \frac{(\ln x)^2}{1+x^2}\, dx$.
[Hint: Replace x with $z = e^t$ and use the rectangular contour in Figure 4.55 with $R \to \infty$.]

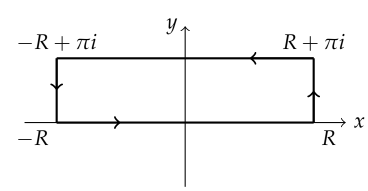


Figure 4.55: Rectangular contour for Problem 25.

26. Do the following integrals for fun!

a. For C the boundary of the square $|x| \le 2$, $|y| \le 2$,

$$\oint_C \frac{dz}{z(z-1)(z-3)^2}.$$

b.
$$\int_0^{\pi} \frac{\sin^2\theta}{13-12\cos\theta}\,d\theta.$$

c.
$$\int_{-\infty}^{\infty} \frac{dx}{x^2+5x+6}.$$

d.
$$\int_0^{\infty} \frac{\cos\pi x}{1-9x^2}\,dx.$$

e.
$$\int_{-\infty}^{\infty} \frac{dx}{(x^2+9)(1-x)^2}.$$

f.
$$\int_0^{\infty} \frac{\sqrt{x}}{(1+x)^2}\,dx.$$

5
Fourier and Laplace Transforms

"There is no branch of mathematics, however abstract, which may not some day be applied to phenomena of the real world.", Nikolai Lobatchevsky (1792-1856)

5.1 Introduction

In this chapter we turn to the study of Fourier transforms, which provide integral representations of functions defined on the entire real line. Such functions can represent analog signals. Recall that analog signals are continuous signals which are sums over a continuous set of frequencies. Our starting point will be to rewrite Fourier trigonometric series as Fourier exponential series. The sums over discrete frequencies will lead to a sum (integral) over continuous frequencies. The resulting integrals will be complex integrals, which can be evaluated using contour methods. We will investigate the properties of these Fourier transforms and get prepared to ask how the analog signal representations are related to the Fourier series expansions over discrete frequencies which we had seen in Chapter 2. Fourier series represented functions which were defined over finite domains such as $x \in [0, L]$. Our explorations will lead us into a discussion of the sampling of signals in the next chapter.

We will also discuss a related integral transform, the Laplace transform. Laplace transforms are useful in solving initial value problems in differential equations and can be used to relate the input to the output of a linear system. Both transforms provide an introduction to a more general theory of transforms, which are used to transform specific problems to simpler ones.

In this chapter we will explore the use of integral transforms. Given a function $f(x)$, we define an integral transform to a new function $F(k)$ as

$$F(k) = \int_a^b f(x)K(x,k)\,dx.$$

Here $K(x,k)$ is called the kernel of the transform. We will concentrate specifically on Fourier transforms,

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{ikx}\,dx,$$

and Laplace transforms

$$F(s) = \int_0^{\infty} f(t)e^{-st}\,dt.$$

In Figure 5.1 we summarize the transform scheme for solving an initial value problem. One can solve the differential equation directly, evolving the initial condition $y(0)$ into the solution $y(t)$ at a later time.

However, the transform method can be used to solve the problem indirectly. Starting with the differential equation and an initial condition, one computes its Transform (T) using

$$Y(s) = \int_0^{\infty} y(t)e^{-st}\,dt.$$

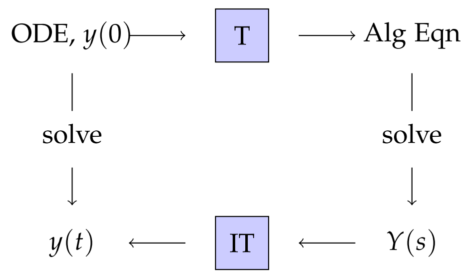


Figure 5.1: Schematic of using transforms to solve a linear ordinary differential equation.

Applying the transform to the differential equation, one obtains a simpler (algebraic) equation satisfied by $Y(s)$, which is simpler to solve than the original differential equation. Once $Y(s)$ has been found, then one applies the Inverse Transform (IT) to $Y(s)$ in order to get the desired solution, $y(t)$. We will see how all of this plays out by the end of the chapter.

We will begin by introducing the Fourier transform. First, we need to see how one can rewrite a trigonometric Fourier series as complex exponential series. Then we can extend the new representation of such series to analog signals, which typically have infinite periods. In later chapters we will highlight the connection between these analog signals and their associated digital signals.

5.2 *Complex Exponential Fourier Series*

Before deriving the Fourier transform, we will need to rewrite the trigonometric Fourier series representation as a complex exponential Fourier series. We first recall from Chapter 2 the trigonometric Fourier series representation of a function defined on $[-\pi, \pi]$ with period 2π. The Fourier series is given by

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos nx + b_n \sin nx\right), \tag{5.1}$$

where the Fourier coefficients were found as

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \quad n = 0, 1, \ldots, \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx, \quad n = 1, 2, \ldots. \end{aligned} \tag{5.2}$$

In order to derive the exponential Fourier series, we replace the trigonometric functions with exponential functions and collect like exponential terms. This gives

$$\begin{aligned} f(x) &\sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\left(\frac{e^{inx}+e^{-inx}}{2}\right) + b_n\left(\frac{e^{inx}-e^{-inx}}{2i}\right)\right] \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(\frac{a_n - ib_n}{2}\right)e^{inx} + \sum_{n=1}^{\infty}\left(\frac{a_n + ib_n}{2}\right)e^{-inx}. \end{aligned} \tag{5.3}$$

The coefficients of the complex exponentials can be rewritten by defining

$$c_n = \frac{1}{2}(a_n + ib_n), \quad n = 1, 2, \ldots. \tag{5.4}$$

This implies that

$$\bar{c}_n = \frac{1}{2}(a_n - ib_n), \quad n = 1, 2, \ldots. \tag{5.5}$$

So far, the representation is rewritten as

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \bar{c}_n e^{inx} + \sum_{n=1}^{\infty} c_n e^{-inx}.$$

Re-indexing the first sum, by introducing $k = -n$, we can write

$$f(x) \sim \frac{a_0}{2} + \sum_{k=-1}^{-\infty} \bar{c}_{-k} e^{-ikx} + \sum_{n=1}^{\infty} c_n e^{-inx}.$$

Since k is a dummy index, we replace it with a new n as

$$f(x) \sim \frac{a_0}{2} + \sum_{n=-1}^{-\infty} \bar{c}_{-n} e^{-inx} + \sum_{n=1}^{\infty} c_n e^{-inx}.$$

We can now combine all the terms into a simple sum. We first define c_n for negative n's by

$$c_n = \bar{c}_{-n}, \quad n = -1, -2, \ldots.$$

Letting $c_0 = \frac{a_0}{2}$, we can write the complex exponential Fourier series representation as

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{-inx}, \tag{5.6}$$

where

$$\begin{aligned} c_n &= \frac{1}{2}(a_n + ib_n), \quad n = 1, 2, \ldots, \\ c_n &= \frac{1}{2}(a_{-n} - ib_{-n}), \quad n = -1, -2, \ldots, \\ c_0 &= \frac{a_0}{2}. \end{aligned} \tag{5.7}$$

Given such a representation, we would like to write out the integral forms of the coefficients, c_n. So, we replace the a_n's and b_n's with their integral representations and replace the trigonometric functions with complex exponential functions. Doing this, we have for $n = 1, 2, \ldots,$

$$\begin{aligned} c_n &= \frac{1}{2}(a_n + ib_n) \\ &= \frac{1}{2}\left[\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx + \frac{i}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx\right] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\left(\frac{e^{inx} + e^{-inx}}{2}\right)dx + \frac{i}{2\pi}\int_{-\pi}^{\pi} f(x)\left(\frac{e^{inx} - e^{-inx}}{2i}\right)dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}\,dx. \end{aligned} \tag{5.8}$$

It is a simple matter to determine the c_n's for other values of n. For $n = 0$, we have that

$$c_0 = \frac{a_0}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx.$$

For $n = -1, -2, \ldots$, we find that

$$c_n = \bar{c}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{e^{-inx}}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}\,dx.$$

Therefore, we have obtained the complex exponential Fourier series coefficients for all n. Now we can define the complex exponential Fourier series for the function $f(x)$ defined on $[-\pi, \pi]$ as shown below.

Complex Exponential Series for $f(x)$ Defined on $[-\pi, \pi]$

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{-inx}, \tag{5.9}$$

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}\,dx. \tag{5.10}$$

We can easily extend the above analysis to other intervals. For example, for $x \in [-L, L]$ the Fourier trigonometric series is

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right)$$

with Fourier coefficients

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}\,dx, \quad n = 0, 1, \ldots,$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}\,dx, \quad n = 1, 2, \ldots.$$

This can be rewritten as an exponential Fourier series of the form

Complex Exponential Series for $f(x)$ Defined on $[-L, L]$

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{-in\pi x/L}, \tag{5.11}$$

$$c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{in\pi x/L}\,dx. \tag{5.12}$$

We can now use this complex exponential Fourier series for function defined on $[-L, L]$ to derive the Fourier transform by letting L get large. This will lead to a sum over a continuous set of frequencies, as opposed to the sum over discrete frequencies, which Fourier series represent.

5.3 *Exponential Fourier Transform*

Both the trigonometric and complex exponential Fourier series provide us with representations of a class of functions of finite period in terms of sums over a discrete set of frequencies. In particular, for functions defined on $x \in [-L, L]$, the period of the Fourier series representation is $2L$. We can write the arguments in the exponentials, $e^{-in\pi x/L}$, in terms of the angular frequency, $\omega_n = n\pi/L$, as $e^{-i\omega_n x}$. We note that the frequencies, ν_n, are then defined through $\omega_n = 2\pi\nu_n = \frac{n\pi}{L}$. Therefore, the complex exponential series is seen to be a sum over a discrete, or countable, set of frequencies.

We would now like to extend the finite interval to an infinite interval, $x \in (-\infty, \infty)$, and to extend the discrete set of (angular) frequencies to a continuous range of frequencies, $\omega \in (-\infty, \infty)$. One can do this rigorously. It amounts to letting L and n get large and keeping $\frac{n}{L}$ fixed.

We first define $\Delta\omega = \frac{\pi}{L}$, so that $\omega_n = n\Delta\omega$. Inserting the Fourier coefficients (5.12) into Equation (5.11), we have

$$\begin{aligned} f(x) &\sim \sum_{n=-\infty}^{\infty} c_n e^{-in\pi x/L} \\ &= \sum_{n=-\infty}^{\infty} \left(\frac{1}{2L} \int_{-L}^{L} f(\xi) e^{in\pi\xi/L}\, d\xi \right) e^{-in\pi x/L} \\ &= \sum_{n=-\infty}^{\infty} \left(\frac{\Delta\omega}{2\pi} \int_{-L}^{L} f(\xi) e^{i\omega_n \xi}\, d\xi \right) e^{-i\omega_n x}. \end{aligned} \tag{5.13}$$

Now, we let L get large, so that $\Delta\omega$ becomes small and ω_n approaches the angular frequency ω. Then,

$$\begin{aligned} f(x) &\sim \lim_{\Delta\omega \to 0, L \to \infty} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left(\int_{-L}^{L} f(\xi) e^{i\omega_n \xi}\, d\xi \right) e^{-i\omega_n x} \Delta\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\xi) e^{i\omega\xi}\, d\xi \right) e^{-i\omega x}\, d\omega. \end{aligned} \tag{5.14}$$

Looking at this last result, we formally arrive at the definition of the Fourier transform. It is embodied in the inner integral and can be written as

Definitions of the Fourier transform and the inverse Fourier transform.

$$F[f] = \hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x}\, dx. \tag{5.15}$$

This is a generalization of the Fourier coefficients (5.12).

Once we know the Fourier transform, $\hat{f}(\omega)$, we can reconstruct the original function, $f(x)$, using the inverse Fourier transform, which is given by the outer integration,

$$F^{-1}[\hat{f}] = f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-i\omega x}\, d\omega. \tag{5.16}$$

We note that it can be proven that the Fourier transform exists when $f(x)$ is absolutely integrable, that is,

$$\int_{-\infty}^{\infty} |f(x)|\, dx < \infty.$$

Such functions are said to be L_1.

We combine these results below, defining the Fourier and inverse Fourier transforms and indicating that they are inverse operations of each other.

We will then prove the first of the equations, Equation (5.19). [The second equation, Equation (5.20), follows in a similar way.]

The **Fourier transform** and **inverse Fourier transform** are inverse operations. Defining the Fourier transform as

$$F[f] = \hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x}\, dx. \tag{5.17}$$

and the inverse Fourier transform as

$$F^{-1}[\hat{f}] = f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega) e^{-i\omega x}\, d\omega, \tag{5.18}$$

then

$$F^{-1}[F[f]] = f(x) \tag{5.19}$$

and

$$F[F^{-1}[\hat{f}]] = \hat{f}(\omega). \tag{5.20}$$

Proof. The proof is carried out by inserting the definition of the Fourier transform, Equation (5.17), into the inverse transform definition, Equation (5.18), and then interchanging the orders of integration. Thus, we have

$$\begin{aligned} F^{-1}[F[f]] &= \frac{1}{2\pi}\int_{-\infty}^{\infty} F[f] e^{-i\omega x}\, d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(\xi) e^{i\omega\xi}\, d\xi\right] e^{-i\omega x}\, d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi) e^{i\omega(\xi - x)}\, d\xi d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{i\omega(\xi-x)}\, d\omega\right] f(\xi)\, d\xi. \end{aligned} \tag{5.21}$$

In order to complete the proof, we need to evaluate the inside integral, which does not depend upon $f(x)$. This is an improper integral, so we first define

$$D_\Omega(x) = \int_{-\Omega}^{\Omega} e^{i\omega x}\, d\omega$$

and compute the inner integral as

$$\int_{-\infty}^{\infty} e^{i\omega(\xi - x)}\, d\omega = \lim_{\Omega\to\infty} D_\Omega(\xi - x).$$

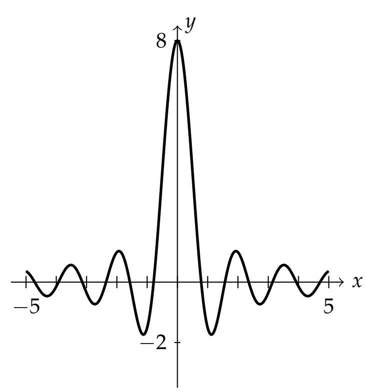


Figure 5.2: A plot of the function $D_\Omega(x)$ for $\Omega = 4$.

We can compute $D_\Omega(x)$. A simple evaluation yields

$$\begin{aligned} D_\Omega(x) &= \int_{-\Omega}^{\Omega} e^{i\omega x}\, d\omega \\ &= \left.\frac{e^{i\omega x}}{ix}\right|_{-\Omega}^{\Omega} \\ &= \frac{e^{ix\Omega} - e^{-ix\Omega}}{2ix} \\ &= \frac{2\sin x\Omega}{x}. \end{aligned} \tag{5.22}$$

A plot of this function is given in Figure 5.2 for $\Omega = 4$. For large Ω, the peak grows and the values of $D_\Omega(x)$ for $x \neq 0$ tend to zero as shown in Figure 5.3. In fact, as x approaches 0, $D_\Omega(x)$ approaches 2Ω. For $x \neq 0$, the $D_\Omega(x)$ function tends to zero.

We further note that

$$\lim_{\Omega\to\infty} D_\Omega(x) = 0, \quad x \neq 0,$$

and $\lim_{\Omega\to\infty} D_\Omega(x)$ is infinite at $x = 0$. However, the area is constant for each Ω. In fact,

$$\int_{-\infty}^{\infty} D_\Omega(x)\,dx = 2\pi.$$

We can show this by recalling the computation in Example 4.42,

$$\int_{-\infty}^{\infty} \frac{\sin x}{x}\,dx = \pi.$$

Then,

$$\begin{aligned} \int_{-\infty}^{\infty} D_\Omega(x)\,dx &= \int_{-\infty}^{\infty} \frac{2\sin x\Omega}{x}\,dx \\ &= \int_{-\infty}^{\infty} 2\frac{\sin y}{y}\,dy \\ &= 2\pi. \end{aligned} \tag{5.23}$$

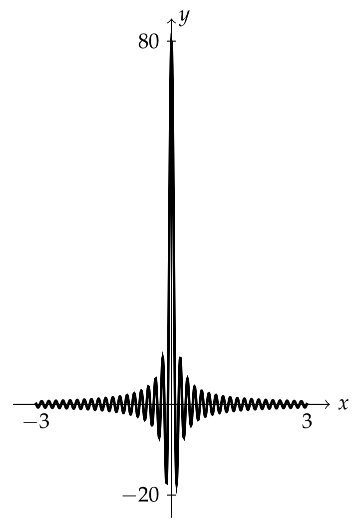


Figure 5.3: A plot of the function $D_\Omega(x)$ for $\Omega = 40$.

Another way to look at $D_\Omega(x)$ is to consider the sequence of functions $f_n(x) = \frac{\sin nx}{\pi x}$, $n = 1, 2, \ldots$. Thus we have shown that this sequence of functions satisfies the two properties,

$$\lim_{n\to\infty} f_n(x) = 0, \quad x \neq 0,$$

$$\int_{-\infty}^{\infty} f_n(x)\,dx = 1.$$

This is a key representation of such generalized functions. The limiting value vanishes at all but one point, but the area is finite.

Such behavior can be seen for the limit of other sequences of functions. For example, consider the sequence of functions

$$f_n(x) = \begin{cases} 0, & |x| > \frac{1}{n}, \\ \frac{n}{2}, & |x| \leq \frac{1}{n}. \end{cases}$$

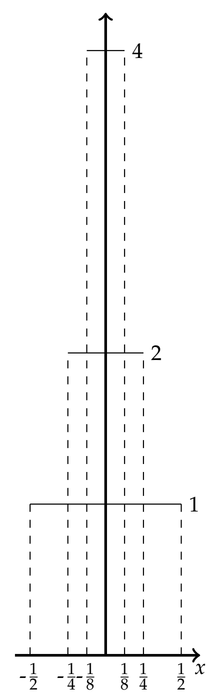


Figure 5.4: A plot of the functions $f_n(x)$ for $n = 2, 4, 8$.

This is a sequence of functions as shown in Figure 5.4. As $n \to \infty$, we find the limit is zero for $x \neq 0$ and is infinite for $x = 0$. However, the area under each member of the sequences is one. Thus, the limiting function is zero at most points but has area one.

The limit is not really a function. It is a generalized function. It is called the Dirac delta function, which is defined by

1. $\delta(x) = 0$ for $x \neq 0$.
2. $\int_{-\infty}^{\infty} \delta(x)\,dx = 1$.

Before returning to the proof that the inverse Fourier transform of the Fourier transform is the identity, we state one more property of the Dirac delta function, which we will prove in the next section. Namely, we will show that

$$\int_{-\infty}^{\infty} \delta(x-a) f(x)\, dx = f(a).$$

Returning to the proof, we now have that

$$\int_{-\infty}^{\infty} e^{i\omega(\xi - x)}\, d\omega = \lim_{\Omega\to\infty} D_\Omega(\xi - x) = 2\pi\delta(\xi - x).$$

Inserting this into Equation (5.21), we have

$$\begin{aligned} F^{-1}[F[f]] &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{i\omega(\xi-x)}\, d\omega\right] f(\xi)\, d\xi \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} 2\pi\delta(\xi - x) f(\xi)\, d\xi \\ &= f(x). \end{aligned} \tag{5.24}$$

Thus, we have proven that the inverse transform of the Fourier transform of f is f. □

5.4 *The Dirac Delta Function*

P. A. M. Dirac (1902 - 1984) introduced the δ function in his book, *The Principles of Quantum Mechanics*, 4th Ed., Oxford University Press, 1958, originally published in 1930, as part of his orthogonality statement for a basis of functions in a Hilbert space, $< \xi'|\xi'' >= c\delta(\xi' - \xi'')$ in the same way we introduced discrete orthogonality using the Kronecker delta.

In the last section we introduced the Dirac delta function, $\delta(x)$. As noted above, this is one example of what is known as a generalized function, or a distribution. Dirac had introduced this function in the 1930's in his study of quantum mechanics as a useful tool. It was later studied in a general theory of distributions and found to be more than a simple tool used by physicists. The Dirac delta function, as any distribution, only makes sense under an integral.

Two properties were used in the last section. First, one has that the area under the delta function is one:

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1.$$

Integration over more general intervals gives

$$\int_a^b \delta(x)\, dx = \begin{cases} 1, & 0 \in [a,b], \\ 0, & 0 \notin [a,b]. \end{cases} \tag{5.25}$$

The other property that was used was the sifting property:

$$\int_{-\infty}^{\infty} \delta(x-a) f(x)\, dx = f(a).$$

This can be seen by noting that the delta function is zero everywhere except at $x = a$. Therefore, the integrand is zero everywhere and the only contribution from $f(x)$ will be from $x = a$. So, we can replace $f(x)$ with $f(a)$ under the integral. Since $f(a)$ is a constant, we have that

$$\begin{aligned}\int_{-\infty}^{\infty} \delta(x-a) f(x)\, dx &= \int_{-\infty}^{\infty} \delta(x-a) f(a)\, dx \\ &= f(a) \int_{-\infty}^{\infty} \delta(x-a)\, dx = f(a). \qquad (5.26)\end{aligned}$$

Another property results from using a scaled argument, ax. In this case, we show that

$$\delta(ax) = |a|^{-1}\delta(x). \qquad (5.27)$$

As usual, this only has meaning under an integral sign. So, we place $\delta(ax)$ inside an integral and make a substitution $y = ax$:

$$\begin{aligned}\int_{-\infty}^{\infty} \delta(ax)\, dx &= \lim_{L\to\infty} \int_{-L}^{L} \delta(ax)\, dx \\ &= \lim_{L\to\infty} \frac{1}{a} \int_{-aL}^{aL} \delta(y)\, dy. \qquad (5.28)\end{aligned}$$

If $a > 0$ then

$$\int_{-\infty}^{\infty} \delta(ax)\, dx = \frac{1}{a} \int_{-\infty}^{\infty} \delta(y)\, dy.$$

However, if $a < 0$ then

$$\int_{-\infty}^{\infty} \delta(ax)\, dx = \frac{1}{a} \int_{\infty}^{-\infty} \delta(y)\, dy = -\frac{1}{a} \int_{-\infty}^{\infty} \delta(y)\, dy.$$

The overall difference in a multiplicative minus sign can be absorbed into one expression by changing the factor $1/a$ to $1/|a|$. Thus,

$$\boxed{\int_{-\infty}^{\infty} \delta(ax)\, dx = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(y)\, dy. \qquad (5.29)}$$

Example 5.1. *Evaluate* $\int_{-\infty}^{\infty} (5x+1)\delta(4(x-2))\, dx$. *This is a straight-forward integration:*

$$\int_{-\infty}^{\infty} (5x+1)\delta(4(x-2))\, dx = \frac{1}{4} \int_{-\infty}^{\infty} (5x+1)\delta(x-2)\, dx = \frac{11}{4}.$$

The first step is to write $\delta(4(x-2)) = \frac{1}{4}\delta(x-2)$. *Then, the final evaluation is given by*

$$\frac{1}{4} \int_{-\infty}^{\infty} (5x+1)\delta(x-2)\, dx = \frac{1}{4}(5(2)+1) = \frac{11}{4}.$$

Even more general than $\delta(ax)$ is the delta function $\delta(f(x))$. The integral of $\delta(f(x))$ can be evaluated, depending upon the number of zeros of $f(x)$. If there is only one zero, $f(x_1) = 0$, then one has that

$$\int_{-\infty}^{\infty} \delta(f(x))\, dx = \int_{-\infty}^{\infty} \frac{1}{|f'(x_1)|} \delta(x - x_1)\, dx.$$

This can be proven using the substitution $y = f(x)$ and is left as an exercise for the reader. This result is often written as

$$\delta(f(x)) = \frac{1}{|f'(x_1)|} \delta(x - x_1),$$

again keeping in mind that this only has meaning when placed under an integral.

Properties of the Dirac delta function:

$$\int_{-\infty}^{\infty} \delta(x-a) f(x)\, dx = f(a).$$

$$\int_{-\infty}^{\infty} \delta(ax)\, dx = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(y)\, dy.$$

$$\int_{-\infty}^{\infty} \delta(f(x))\, dx = \int_{-\infty}^{\infty} \sum_{j=1}^{n} \frac{\delta(x-x_j)}{|f'(x_j)|}\, dx.$$

(For n simple roots.)

These and other properties are often written outside the integral:

$$\delta(ax) = \frac{1}{|a|}\delta(x).$$

$$\delta(-x) = \delta(x).$$

$$\delta((x-a)(x-b)) = \frac{[\delta(x-a) + \delta(x-a)]}{|a-b|}.$$

$$\delta(f(x)) = \sum_j \frac{\delta(x-x_j)}{|f'(x_j)|},$$

for $f(x_j) = 0, f'(x_j) \neq 0$.

Example 5.2. *Evaluate* $\int_{-\infty}^{\infty} \delta(3x-2)x^2\,dx$.

This is not a simple $\delta(x-a)$*. So, we need to find the zeros of* $f(x) = 3x-2$*. There is only one,* $x = \frac{2}{3}$*. Also,* $|f'(x)| = 3$*. Therefore, we have*

$$\int_{-\infty}^{\infty} \delta(3x-2)x^2\,dx = \int_{-\infty}^{\infty} \frac{1}{3}\delta(x-\frac{2}{3})x^2\,dx = \frac{1}{3}\left(\frac{2}{3}\right)^2 = \frac{4}{27}.$$

Note that this integral can be evaluated the long way using the substitution $y = 3x-2$*. Then,* $dy = 3\,dx$ *and* $x = (y+2)/3$*. This gives*

$$\int_{-\infty}^{\infty} \delta(3x-2)x^2\,dx = \frac{1}{3}\int_{-\infty}^{\infty} \delta(y)\left(\frac{y+2}{3}\right)^2 dy = \frac{1}{3}\left(\frac{4}{9}\right) = \frac{4}{27}.$$

More generally, one can show that when $f(x_j) = 0$ and $f'(x_j) \neq 0$ for $j = 1, 2, \ldots, n$, (i.e., when one has n simple zeros), then

$$\delta(f(x)) = \sum_{j=1}^{n} \frac{1}{|f'(x_j)|}\delta(x-x_j).$$

Example 5.3. *Evaluate* $\int_0^{2\pi} \cos x\,\delta(x^2-\pi^2)\,dx$.

In this case, the argument of the delta function has two simple roots. Namely, $f(x) = x^2-\pi^2 = 0$ *when* $x = \pm\pi$*. Furthermore,* $f'(x) = 2x$*. Therefore,* $|f'(\pm\pi)| = 2\pi$*. This gives*

$$\delta(x^2-\pi^2) = \frac{1}{2\pi}[\delta(x-\pi)+\delta(x+\pi)].$$

Inserting this expression into the integral and noting that $x = -\pi$ *is not in the integration interval, we have*

$$\begin{aligned} \int_0^{2\pi} \cos x\,\delta(x^2-\pi^2)\,dx &= \frac{1}{2\pi}\int_0^{2\pi} \cos x\,[\delta(x-\pi)+\delta(x+\pi)]\,dx \\ &= \frac{1}{2\pi}\cos\pi = -\frac{1}{2\pi}. \end{aligned} \tag{5.30}$$

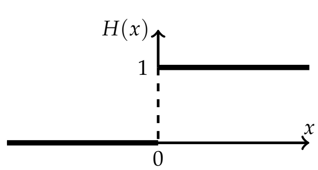


Figure 5.5: The Heaviside step function, $H(x)$.

Example 5.4. *Show* $H'(x) = \delta(x)$*, where the Heaviside function (or, step function) is defined as*

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases}$$

and is shown in Figure 5.5.

Looking at the plot, it is easy to see that $H'(x) = 0$ *for* $x \neq 0$*. In order to check that this gives the delta function, we need to compute the area integral. Therefore, we have*

$$\int_{-\infty}^{\infty} H'(x)\,dx = H(x)\Big|_{-\infty}^{\infty} = 1-0 = 1.$$

Thus, $H'(x)$ *satisfies the two properties of the Dirac delta function.*

5.5 Properties of the Fourier Transform

WE NOW RETURN TO THE FOURIER TRANSFORM. Before actually computing the Fourier transform of some functions, we prove a few of the properties of the Fourier transform.

First we note that there are several forms that one may encounter for the Fourier transform. In applications, functions can either be functions of time, $f(t)$, or space, $f(x)$. The corresponding Fourier transforms are then written as

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}\, dt, \tag{5.31}$$

or

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{ikx}\, dx. \tag{5.32}$$

ω is called the angular frequency and is related to the frequency ν by $\omega = 2\pi\nu$. The units of frequency are typically given in Hertz (Hz). Sometimes the frequency is denoted by f when there is no confusion. k is called the wavenumber. It has units of inverse length and is related to the wavelength, λ, by $k = \frac{2\pi}{\lambda}$.

We explore a few basic properties of the Fourier transform and use them in examples in the next section.

1. **Linearity:** For any functions $f(x)$ and $g(x)$ for which the Fourier transform exists and constant a, we have

$$F[f+g] = F[f] + F[g]$$

and

$$F[af] = aF[f].$$

These simply follow from the properties of integration and establish the linearity of the Fourier transform.

2. **Transform of a Derivative:** $F\left[\frac{df}{dx}\right] = -ik\hat{f}(k)$

Here we compute the Fourier transform (5.17) of the derivative by inserting the derivative in the Fourier integral and using integration by parts:

$$\begin{aligned} F\left[\frac{df}{dx}\right] &= \int_{-\infty}^{\infty} \frac{df}{dx} e^{ikx}\, dx \\ &= \lim_{L\to\infty} \left[f(x)e^{ikx}\right]_{-L}^{L} - ik\int_{-\infty}^{\infty} f(x)e^{ikx}\, dx. \end{aligned} \tag{5.33}$$

The limit will vanish if we assume that $\lim_{x\to\pm\infty} f(x) = 0$. This last integral is recognized as the Fourier transform of f, proving the given property.

3. **Higher Order Derivatives:** $F\left[\frac{d^n f}{dx^n}\right] = (-ik)^n \hat{f}(k)$

   The proof of this property follows from the last result, or doing several integration by parts. We will consider the case when $n = 2$. Noting that the second derivative is the derivative of $f'(x)$ and applying the last result, we have

$$\begin{aligned} F\left[\frac{d^2 f}{dx^2}\right] &= F\left[\frac{d}{dx} f'\right] \\ &= -ikF\left[\frac{df}{dx}\right] = (-ik)^2 \hat{f}(k). \end{aligned} \tag{5.34}$$

   This result will be true if

$$\lim_{x\to\pm\infty} f(x) = 0 \text{ and } \lim_{x\to\pm\infty} f'(x) = 0.$$

   The generalization to the transform of the nth derivative easily follows.

4. **Multiplication by x:** $F[xf(x)] = -i\frac{d}{dk}\hat{f}(k)$

   This property can be shown by using the fact that $\frac{d}{dk}e^{ikx} = ixe^{ikx}$ and the ability to differentiate an integral with respect to a parameter.

$$\begin{aligned} F[xf(x)] &= \int_{-\infty}^{\infty} xf(x)e^{ikx}\,dx \\ &= \int_{-\infty}^{\infty} f(x)\frac{d}{dk}\left(\frac{1}{i}e^{ikx}\right)dx \\ &= -i\frac{d}{dk}\int_{-\infty}^{\infty} f(x)e^{ikx}\,dx \\ &= -i\frac{d}{dk}\hat{f}(k). \end{aligned} \tag{5.35}$$

   This result can be generalized to $F[x^n f(x)]$ as an exercise.

5. **Shifting Properties:** For constant a, we have the following shifting properties:

These are the first and second shifting properties, or First and Second Shift Theorems.

$$f(x-a) \leftrightarrow e^{ika}\hat{f}(k), \tag{5.36}$$

$$f(x)e^{-iax} \leftrightarrow \hat{f}(k-a). \tag{5.37}$$

   Here we have denoted the Fourier transform pairs using a double arrow as $f(x) \leftrightarrow \hat{f}(k)$. These are easily proved by inserting the desired forms into the definition of the Fourier transform (5.17), or inverse Fourier transform (5.18). The first shift property (5.36) is shown by the following argument. We evaluate the Fourier transform:

$$F[f(x-a)] = \int_{-\infty}^{\infty} f(x-a)e^{ikx}\,dx.$$

   Now perform the substitution $y = x - a$. Then,

$$\begin{aligned} F[f(x-a)] &= \int_{-\infty}^{\infty} f(y)e^{ik(y+a)}\,dy \\ &= e^{ika}\int_{-\infty}^{\infty} f(y)e^{iky}\,dy \\ &= e^{ika}\hat{f}(k). \end{aligned} \tag{5.38}$$

The second shift property (5.37) follows in a similar way.

6. **Convolution of Functions:** We define the convolution of two functions $f(x)$ and $g(x)$ as

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\,dx. \tag{5.39}$$

Then, the Fourier transform of the convolution is the product of the Fourier transforms of the individual functions:

$$F[f * g] = \hat{f}(k)\hat{g}(k). \tag{5.40}$$

We will return to the proof of this property in Section 5.6.

5.5.1 *Fourier Transform Examples*

IN THIS SECTION WE WILL COMPUTE the Fourier transforms of several functions.

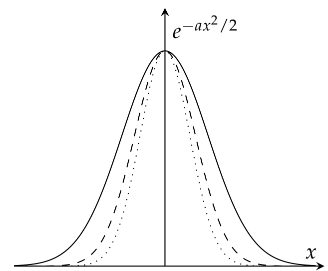


Figure 5.6: Plots of the Gaussian function $f(x) = e^{-ax^2/2}$ for $a = 1, 2, 3$.

Example 5.5. *Find the Fourier transform of a Gaussian, $f(x) = e^{-ax^2/2}$.*

This function, shown in Figure 5.6, is called the Gaussian function. It has many applications in areas such as quantum mechanics, molecular theory, probability, and heat diffusion. We will compute the Fourier transform of this function and show that the Fourier transform of a Gaussian is a Gaussian. In the derivation, we will introduce classic techniques for computing such integrals.

We begin by applying the definition of the Fourier transform,

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{ikx}\,dx = \int_{-\infty}^{\infty} e^{-ax^2/2+ikx}\,dx. \tag{5.41}$$

The first step in computing this integral is to complete the square in the argument of the exponential. Our goal is to rewrite this integral so that a simple substitution will lead to a classic integral of the form $\int_{-\infty}^{\infty} e^{\beta y^2}\,dy$, which we can integrate. The completion of the square follows as usual:

$$\begin{aligned}
-\frac{a}{2}x^2 + ikx &= -\frac{a}{2}\left[x^2 - \frac{2ik}{a}x\right] \\
&= -\frac{a}{2}\left[x^2 - \frac{2ik}{a}x + \left(-\frac{ik}{a}\right)^2 - \left(-\frac{ik}{a}\right)^2\right] \\
&= -\frac{a}{2}\left(x - \frac{ik}{a}\right)^2 - \frac{k^2}{2a}.
\end{aligned} \tag{5.42}$$

We now put this expression into the integral and make the substitutions $y = x - \frac{ik}{a}$ and $\beta = \frac{a}{2}$.

$$\begin{aligned}
\hat{f}(k) &= \int_{-\infty}^{\infty} e^{-ax^2/2+ikx}\,dx \\
&= e^{-\frac{k^2}{2a}} \int_{-\infty}^{\infty} e^{-\frac{a}{2}\left(x-\frac{ik}{a}\right)^2}\,dx \\
&= e^{-\frac{k^2}{2a}} \int_{-\infty-\frac{ik}{a}}^{\infty-\frac{ik}{a}} e^{-\beta y^2}\,dy.
\end{aligned} \tag{5.43}$$

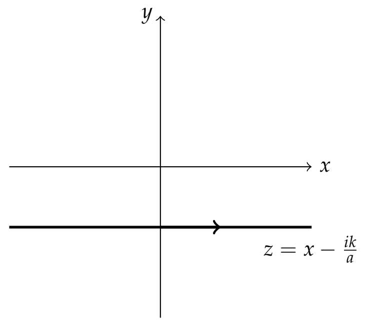


Figure 5.7: Simple horizontal contour.

One would be tempted to absorb the $-\frac{ik}{a}$ terms in the limits of integration. However, we know from our previous study that the integration takes place over a contour in the complex plane as shown in Figure 5.7.

In this case, we can deform this horizontal contour to a contour along the real axis since we will not cross any singularities of the integrand. So, we now safely write

$$\hat{f}(k) = e^{-\frac{k^2}{2a}} \int_{-\infty}^{\infty} e^{-\beta y^2}\, dy.$$

The resulting integral is a classic integral and can be performed using a standard trick. Define I by[1]

$$I = \int_{-\infty}^{\infty} e^{-\beta y^2}\, dy.$$

[1] Here we show

$$\int_{-\infty}^{\infty} e^{-\beta y^2}\, dy = \sqrt{\frac{\pi}{\beta}}.$$

Note that we solved the $\beta = 1$ case in Example 3.14, so a simple variable transformation $z = \sqrt{\beta} y$ is all that is needed to get the answer. However, it cannot hurt to see this classic derivation again.

Then,

$$I^2 = \int_{-\infty}^{\infty} e^{-\beta y^2}\, dy \int_{-\infty}^{\infty} e^{-\beta x^2}\, dx.$$

Note that we needed to change the integration variable so that we can write this product as a double integral:

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\beta(x^2+y^2)}\, dxdy.$$

This is an integral over the entire xy-plane. We now transform to polar coordinates to obtain

$$\begin{aligned} I^2 &= \int_0^{2\pi} \int_0^{\infty} e^{-\beta r^2}\, rdrd\theta \\ &= 2\pi \int_0^{\infty} e^{-\beta r^2}\, rdr \\ &= -\frac{\pi}{\beta}\left[e^{-\beta r^2}\right]_0^{\infty} = \frac{\pi}{\beta}. \end{aligned} \tag{5.44}$$

The final result is obtained by taking the square root, yielding

$$I = \sqrt{\frac{\pi}{\beta}}.$$

We can now insert this result to give the Fourier transform of the Gaussian function:

$$\hat{f}(k) = \sqrt{\frac{2\pi}{a}} e^{-k^2/2a}. \tag{5.45}$$

The Fourier transform of a Gaussian is a Gaussian.

Therefore, we have shown that the Fourier transform of a Gaussian is a Gaussian.

Example 5.6. *Find the Fourier transform of the box, or gate, function,*

$$f(x) = \begin{cases} b, & |x| \le a, \\ 0, & |x| > a. \end{cases}$$

This function is called the box function, or gate function. It is shown in Figure 5.8. The Fourier transform of the box function is relatively easy to compute. It is given by

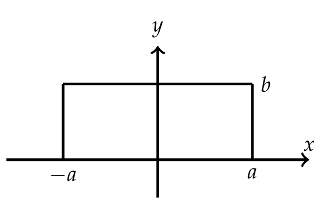


Figure 5.8: A plot of the box function in Example 5.6.

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x) e^{ikx}\, dx$$

$$
\begin{aligned}
&= \int_{-a}^{a} b e^{ikx}\, dx \\
&= \frac{b}{ik} e^{ikx}\Big|_{-a}^{a} \\
&= \frac{2b}{k} \sin ka. \qquad (5.46)
\end{aligned}
$$

We can rewrite this as

$$\hat{f}(k) = 2ab\frac{\sin ka}{ka} \equiv 2ab \operatorname{sinc} ka.$$

Here we introduced the sinc function,

$$\operatorname{sinc} x = \frac{\sin x}{x}.$$

A plot of this function is shown in Figure 5.9. This function appears often in signal analysis and it plays a role in the study of diffraction.

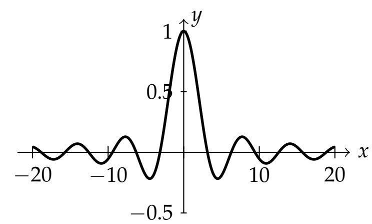


Figure 5.9: A plot of the Fourier transform of the box function in Example 5.6. This is the general shape of the sinc function.

We will now consider special limiting values for the box function and its transform. This will lead us to the Uncertainty Principle for signals, connecting the relationship between the localization properties of a signal and its transform.

1. $a \to \infty$ *and* b *fixed.*

   In this case, as a *gets large, the box function approaches the constant function* $f(x) = b$*. At the same time, we see that the Fourier transform approaches a Dirac delta function. We had seen this function earlier when we first defined the Dirac delta function. Compare Figure 5.9 with Figure 5.2. In fact,* $\hat{f}(k) = bD_a(k)$*. [Recall the definition of* $D_\Omega(x)$ *in Equation (5.22).] So, in the limit, we obtain* $\hat{f}(k) = 2\pi b\delta(k)$*. This limit implies the fact that the Fourier transform of* $f(x) = 1$ *is* $\hat{f}(k) = 2\pi\delta(k)$*. As the width of the box becomes wider, the Fourier transform becomes more localized. In fact, we have arrived at the important result that*

$$\int_{-\infty}^{\infty} e^{ikx}\, dx = 2\pi\delta(k).$$

$$\int_{-\infty}^{\infty} e^{ikx}\, dx = 2\pi\delta(k). \qquad (5.47)$$

2. $b \to \infty$*,* $a \to 0$*, and* $2ab = 1$*.*

   In this case, the box narrows and becomes steeper while maintaining a constant area of one. This is the way we had found a representation of the Dirac delta function previously. The Fourier transform approaches a constant in this limit. As a *approaches zero, the sinc function approaches one, leaving* $\hat{f}(k) \to 2ab = 1$*. Thus, the Fourier transform of the Dirac delta function is one. Namely, we have*

$$\int_{-\infty}^{\infty} \delta(x) e^{ikx}\, dx = 1. \qquad (5.48)$$

   In this case, we have that the more localized the function $f(x)$ *is, the more spread out the Fourier transform,* $\hat{f}(k)$*, is. We will summarize these notions in the next item by relating the widths of the function and its Fourier transform.*

3. *The Uncertainty Principle:* $\Delta x \Delta k = 4\pi$.

The widths of the box function and its Fourier transform are related, as we have seen in the last two limiting cases. It is natural to define the width, Δx, *of the box function as*

$$\Delta x = 2a.$$

The width of the Fourier transform is a little trickier. This function actually extends along the entire k-axis. However, as $\hat{f}(k)$ *became more localized, the central peak in Figure 5.9 became narrower. So, we define the width of this function,* Δk *as the distance between the first zeros on either side of the main lobe as shown in Figure 5.10. This gives*

$$\Delta k = \frac{2\pi}{a}.$$

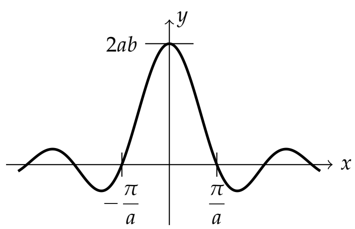


Figure 5.10: The width of the function $2ab\frac{\sin ka}{ka}$ is defined as the distance between the smallest magnitude zeros.

Combining these two relations, we find that

$$\Delta x \Delta k = 4\pi.$$

Thus, the more localized a signal, the less localized its transform and vice versa. This notion is referred to as the Uncertainty Principle. For general signals, one needs to define the effective widths more carefully, but the main idea holds:

$$\Delta x \Delta k \geq c > 0.$$

More formally, the Uncertainty Principle for signals is about the relation between duration and bandwidth, which are defined by $\Delta t = \frac{\|tf\|_2}{\|f\|_2}$ and $\Delta\omega = \frac{\|\omega\hat{f}\|_2}{\|\hat{f}\|_2}$, respectively, where $\|f\|_2 = \int_{-\infty}^{\infty} |f(t)|^2\,dt$ and $\|\hat{f}\|_2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\,d\omega$. Under appropriate conditions, one can prove that $\Delta t \Delta\omega \geq \frac{1}{2}$. Equality holds for Gaussian signals. Werner Heisenberg (1901 - 1976) introduced the Uncertainty Principle into quantum physics in 1926, relating uncertainties in the position (Δx) and momentum (Δp_x) of particles. In this case, $\Delta x \Delta p_x \geq \frac{1}{2}\hbar$. Here, the uncertainties are defined as the positive square roots of the quantum mechanical variances of the position and momentum.

We now turn to other examples of Fourier transforms.

Example 5.7. *Find the Fourier transform of* $f(x) = \begin{cases} e^{-ax}, & x \geq 0 \\ 0, & x < 0 \end{cases}$, $a > 0$.

The Fourier transform of this function is

$$\begin{aligned} \hat{f}(k) &= \int_{-\infty}^{\infty} f(x) e^{ikx}\,dx \\ &= \int_0^{\infty} e^{ikx-ax}\,dx \\ &= \frac{1}{a-ik}. \end{aligned} \tag{5.49}$$

Next, we will compute the inverse Fourier transform of this result and recover the original function.

Example 5.8. *Find the inverse Fourier transform of* $\hat{f}(k) = \frac{1}{a-ik}$.

The inverse Fourier transform of this function is

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(k) e^{-ikx}\,dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ikx}}{a-ik}\,dk.$$

This integral can be evaluated using contour integral methods. We evaluate the integral

$$I = \int_{-\infty}^{\infty} \frac{e^{-ixz}}{a-iz}\,dz,$$

using Jordan's Lemma from Section 4.4.8. According to Jordan's Lemma, we need to enclose the contour with a semicircle in the upper half plane for $x < 0$ and in the lower half plane for $x > 0$, as shown in Figure 5.11.

The integrations along the semicircles will vanish and we will have

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ikx}}{a-ik}\,dk \\ &= \pm\frac{1}{2\pi}\oint_C \frac{e^{-ixz}}{a-iz}\,dz \\ &= \begin{cases} 0, & x<0 \\ -\frac{1}{2\pi}2\pi i \; Res\,[z=-ia], & x>0 \end{cases} \\ &= \begin{cases} 0, & x<0 \\ e^{-ax}, & x>0 \end{cases}. \end{aligned} \tag{5.50}$$

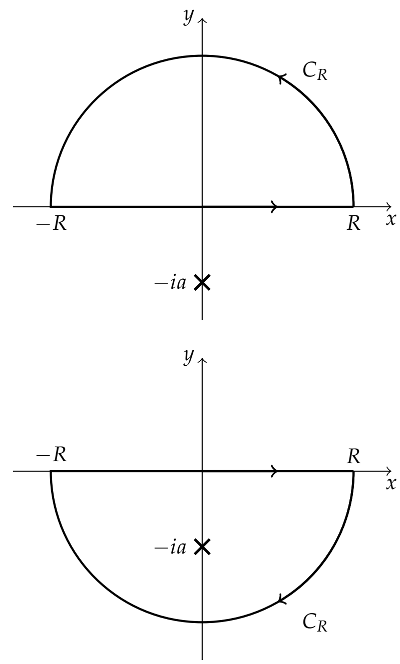


Figure 5.11: Contours for inverting $\hat{f}(k) = \frac{1}{a-ik}$.

Note that without paying careful attention to Jordan's Lemma, one might not retrieve the function from the last example.

Example 5.9. *Find the inverse Fourier transform of $\hat{f}(\omega) = \pi\delta(\omega+\omega_0) + \pi\delta(\omega-\omega_0)$.*

We would like to find the inverse Fourier transform of this function. Instead of carrying out any integration, we will make use of the properties of Fourier transforms. Since the transforms of sums are the sums of transforms, we can look at each term individually. Consider $\delta(\omega-\omega_0)$. This is a shifted function. From the shift theorems in Equations (5.36) and (5.37) we have the Fourier transform pair

$$e^{i\omega_0 t} f(t) \leftrightarrow \hat{f}(\omega-\omega_0).$$

Recalling from Example 5.6 that

$$\int_{-\infty}^{\infty} e^{i\omega t}\,dt = 2\pi\delta(\omega),$$

we have from the shift property that

$$F^{-1}[\delta(\omega-\omega_0)] = \frac{1}{2\pi}e^{-i\omega_0 t}.$$

The second term can be transformed similarly. Therefore, we have

$$F^{-1}[\pi\delta(\omega+\omega_0)+\pi\delta(\omega-\omega_0] = \frac{1}{2}e^{i\omega_0 t} + \frac{1}{2}e^{-i\omega_0 t} = \cos\omega_0 t.$$

Example 5.10. *Find the Fourier transform of the finite wave train.*

$$f(t) = \begin{cases} \cos\omega_0 t, & |t| \le a, \\ 0, & |t| > a. \end{cases}$$

For the last example, we consider the finite wave train, which will reappear in the last chapter on signal analysis. In Figure 5.12 we show a plot of this function.

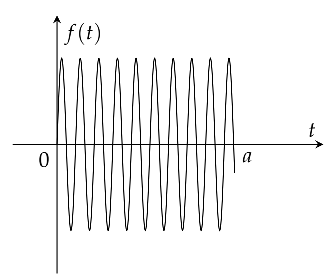


Figure 5.12: A plot of the finite wave train.

A straight-forward computation gives

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}\,dt$$

$$\begin{aligned}
&= \int_{-a}^{a} [\cos\omega_0 t + i\sin\omega_0 t] e^{i\omega t}\, dt \\
&= \int_{-a}^{a} \cos\omega_0 t \cos\omega t\, dt + i\int_{-a}^{a} \sin\omega_0 t \sin\omega t\, dt \\
&= \frac{1}{2}\int_{-a}^{a} \left[\cos((\omega+\omega_0)t) + \cos((\omega-\omega_0)t)\right]\, dt \\
&= \frac{\sin((\omega+\omega_0)a)}{\omega+\omega_0} + \frac{\sin((\omega-\omega_0)a)}{\omega-\omega_0}. \qquad (5.51)
\end{aligned}$$

5.6 The Convolution Operation

In the list of properties of the Fourier transform, we defined the convolution of two functions, $f(x)$ and $g(x)$, to be the integral

$$(f*g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\, dt. \qquad (5.52)$$

In some sense one is looking at a sum of the overlaps of one of the functions and all of the shifted versions of the other function. The German word for convolution is *faltung*, which means "folding" and in old texts this is referred to as the Faltung Theorem. In this section we will look into the convolution operation and its Fourier transform.

Before we get too involved with the convolution operation, it should be noted that there are really two things you need to take away from this discussion. The rest is detail. First, the convolution of two functions is a new functions as defined by Equation (5.52) when dealing with the Fourier transform. The second and most relevant is that the Fourier transform of the convolution of two functions is the product of the transforms of each function. The rest is all about the use and consequences of these two statements. In this section we will show how the convolution works and how it is useful.

The convolution is commutative.

First, we note that the convolution is commutative: $f*g = g*f$. This is easily shown by replacing $x-t$ with a new variable, $y = x-t$ and $dy = -dt$.

$$\begin{aligned}
(g*f)(x) &= \int_{-\infty}^{\infty} g(t)f(x-t)\, dt \\
&= -\int_{\infty}^{-\infty} g(x-y)f(y)\, dy \\
&= \int_{-\infty}^{\infty} f(y)g(x-y)\, dy \\
&= (f*g)(x). \qquad (5.53)
\end{aligned}$$

The best way to understand the folding of the functions in the convolution is to take two functions and convolve them. The next example gives a graphical rendition followed by a direct computation of the convolution. The reader is encouraged to carry out these analyses for other functions.

Example 5.11. *Graphical convolution of the box function and a triangle function.*

In order to understand the convolution operation, we need to apply it to specific *functions. We will first do this graphically for the box function*

$$f(x) = \begin{cases} 1, & |x| \le 1, \\ 0, & |x| > 1, \end{cases}$$

and the triangular function

$$g(x) = \begin{cases} x, & 0 \le x \le 1, \\ 0, & \textit{otherwise}, \end{cases}$$

as shown in Figure 5.13.

Next, we determine the contributions to the integrand. We consider the shifted and reflected function $g(t-x)$ in Equation (5.52) for various values of t. For $t=0$, we have $g(x-0) = g(-x)$. This function is a reflection of the triangle function, $g(x)$, as shown in Figure 5.14.

We then translate the triangle function performing horizontal shifts by t. In Figure 5.15 we show such a shifted and reflected $g(x)$ for $t = 2$, or $g(2-x)$.

In Figure 5.15 we show several plots of other shifts, $g(x-t)$, superimposed on $f(x)$.

The integrand is the product of $f(t)$ and $g(x-t)$ and the integral of the product $f(t)g(x-t)$ is given by the sum of the shaded areas for each value of x.

In the first plot of Figure 5.16, the area is zero, as there is no overlap of the functions. Intermediate shift values are displayed in the other plots in Figure 5.16. The value of the convolution at x is shown by the area under the product of the two functions for each value of x.

Plots of the areas of the convolution of the box and triangle functions for several values of x are given in Figure 5.15. We see that the value of the convolution integral builds up and then quickly drops to zero as a function of x. In Figure 5.17 the values of these areas is shown as a function of x.

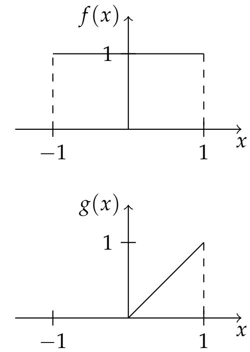


Figure 5.13: A plot of the box function $f(x)$ and the triangle function $g(x)$.

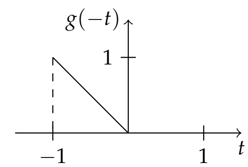


Figure 5.14: A plot of the reflected triangle function, $g(-t)$.

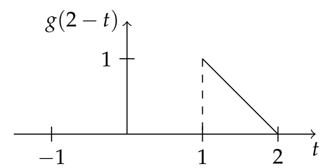


Figure 5.15: A plot of the reflected triangle function shifted by two units, $g(2-t)$.

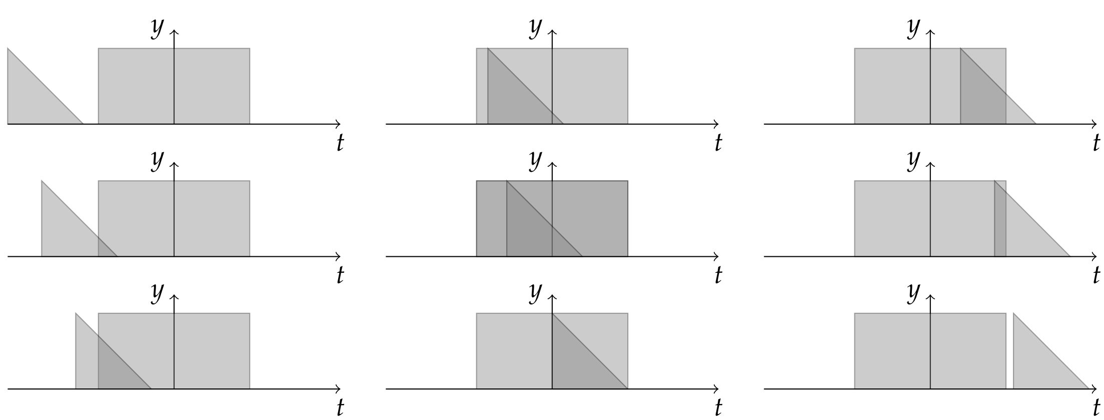


Figure 5.16: A plot of the box and triangle functions with the overlap indicated by the shaded area.

The plot of the convolution in Figure 5.17 is not easily determined using the graphical method. However, we can directly compute the convolution as shown in the next example.

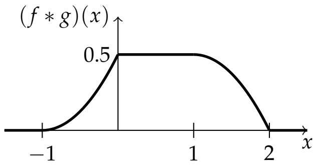


Figure 5.17: A plot of the convolution of the box and triangle functions.

Example 5.12. *Analytically find the convolution of the box function and the triangle function.*

The nonvanishing contributions to the convolution integral are when both $f(t)$ and $g(x-t)$ do not vanish. $f(t)$ is nonzero for $|t| \leq 1$, or $-1 \leq t \leq 1$. $g(x-t)$ is nonzero for $0 \leq x-t \leq 1$, or $x-1 \leq t \leq x$. These two regions are shown in Figure 5.18. On this region, $f(t)g(x-t) = x-t$.

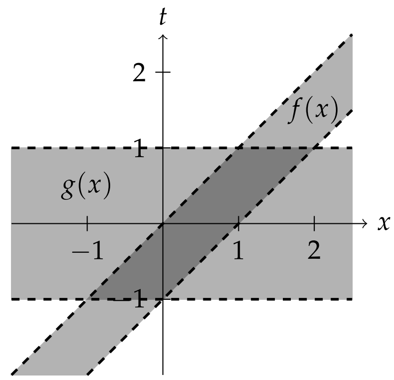


Figure 5.18: Intersection of the support of $g(x)$ and $f(x)$.

Isolating the intersection in Figure 5.19, we see in Figure 5.19 that there are three regions as shown by different shadings. These regions lead to a piecewise defined function with three different branches of nonzero values for $-1 < x < 0$, $0 < x < 1$, and $1 < x < 2$.

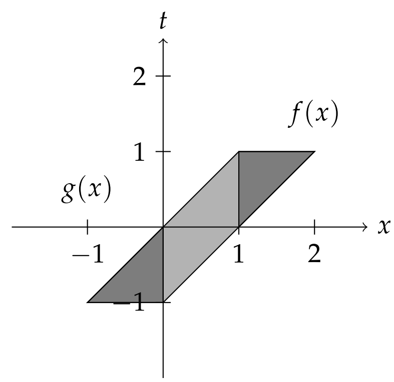


Figure 5.19: Intersection of the support of $g(x)$ and $f(x)$ showing the integration regions.

The values of the convolution can be determined through careful integration. The resulting integrals are given as

$$\begin{aligned}
(f*g)(x) &= \int_{-\infty}^{\infty} f(t)g(x-t)\,dt \\
&= \begin{cases} \int_{-1}^{x}(x-t)\,dt, & -1 < x < 0, \\ \int_{x-1}^{x}(x-t)\,dt, & 0 < x < 1, \\ \int_{x-1}^{1}(x-t)\,dt, & 1 < x < 2 \end{cases} \\
&= \begin{cases} \frac{1}{2}(x+1)^2, & -1 < x < 0, \\ \frac{1}{2}, & 0 < x < 1, \\ \frac{1}{2}\left[1-(x-1)^2\right] & 1 < x < 2. \end{cases}
\end{aligned} \tag{5.54}$$

A plot of this function is shown in Figure 5.17.

5.6.1 Convolution Theorem for Fourier Transforms

In this section we compute the Fourier transform of the convolution integral and show that the Fourier transform of the convolution is the product of the transforms of each function,

$$F[f * g] = \hat{f}(k)\hat{g}(k). \tag{5.55}$$

First, we use the definitions of the Fourier transform and the convolution to write the transform as

$$\begin{aligned} F[f * g] &= \int_{-\infty}^{\infty} (f * g)(x) e^{ikx}\, dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(t) g(x-t)\, dt \right) e^{ikx}\, dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(x-t) e^{ikx}\, dx \right) f(t)\, dt. \end{aligned} \tag{5.56}$$

We now substitute $y = x - t$ on the inside integral and separate the integrals:

$$\begin{aligned} F[f * g] &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(x-t) e^{ikx}\, dx \right) f(t)\, dt \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(y) e^{ik(y+t)}\, dy \right) f(t)\, dt \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(y) e^{iky}\, dy \right) f(t) e^{ikt}\, dt \\ &= \left(\int_{-\infty}^{\infty} f(t) e^{ikt}\, dt \right) \left(\int_{-\infty}^{\infty} g(y) e^{iky}\, dy \right). \end{aligned} \tag{5.57}$$

We see that the two integrals are just the Fourier transforms of f and g. Therefore, the Fourier transform of a convolution is the product of the Fourier transforms of the functions involved:

$$F[f * g] = \hat{f}(k)\hat{g}(k).$$

Example 5.13. *Compute the convolution of the box function of height one and width two with itself.*

Let $\hat{f}(k)$ be the Fourier transform of $f(x)$. Then, the Convolution Theorem says that $F[f * f](k) = \hat{f}^2(k)$, or

$$(f * f)(x) = F^{-1}[\hat{f}^2(k)].$$

For the box function, we have already found that

$$\hat{f}(k) = \frac{2}{k} \sin k.$$

So, we need to compute

$$\begin{aligned} (f * f)(x) &= F^{-1}[\frac{4}{k^2} \sin^2 k] \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{4}{k^2} \sin^2 k \right) e^{-ikx}\, dk. \end{aligned} \tag{5.58}$$

One way to compute this integral is to extend the computation into the complex k-plane. We first need to rewrite the integrand. Thus,

$$\begin{aligned}(f*f)(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{4}{k^2}\sin^2 k e^{-ikx}\,dk \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{k^2}[1-\cos 2k]e^{-ikx}\,dk \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{k^2}\left[1-\frac{1}{2}(e^{ik}+e^{-ik})\right]e^{-ikx}\,dk \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{k^2}\left[e^{-ikx}-\frac{1}{2}(e^{-i(1-k)}+e^{-i(1+k)})\right]dk. \qquad (5.59)\end{aligned}$$

We can compute the above integrals if we know how to compute the integral

$$I(y)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{-iky}}{k^2}\,dk.$$

Then, the result can be found in terms of $I(y)$ as

$$(f*f)(x)=I(x)-\frac{1}{2}[I(1-k)+I(1+k)].$$

We consider the integral

$$\oint_C \frac{e^{-iyz}}{\pi z^2}\,dz$$

over the contour in Figure 5.20.

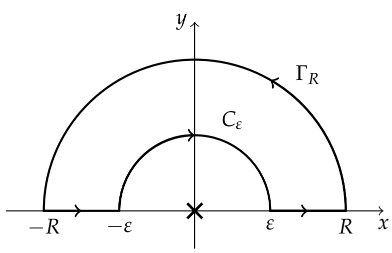


Figure 5.20: Contour for computing $P\int_{-\infty}^{\infty}\frac{e^{-iyz}}{\pi z^2}\,dz$.

We can see that there is a double pole at $z=0$. The pole is on the real axis. So, we will need to cut out the pole as we seek the value of the principal value integral.

Recall from Chapter 4 that

$$\oint_{C_R}\frac{e^{-iyz}}{\pi z^2}\,dz=\int_{\Gamma_R}\frac{e^{-iyz}}{\pi z^2}\,dz+\int_{-R}^{-\epsilon}\frac{e^{-iyz}}{\pi z^2}\,dz+\int_{C_\epsilon}\frac{e^{-iyz}}{\pi z^2}\,dz+\int_{\epsilon}^{R}\frac{e^{-iyz}}{\pi z^2}\,dz.$$

The integral $\oint_{C_R}\frac{e^{-iyz}}{\pi z^2}\,dz$ vanishes since there are no poles enclosed in the contour! The sum of the second and fourth integrals gives the integral we seek as $\epsilon\to 0$ and $R\to\infty$. The integral over Γ_R will vanish as R gets large according to Jordan's Lemma provided $y<0$. That leaves the integral over the small semicircle.

As before, we can show that

$$\lim_{\epsilon\to 0}\int_{C_\epsilon} f(z)\,dz=-\pi i\,\text{Res}[f(z);z=0].$$

Therefore, we find

$$I(y)=P\int_{-\infty}^{\infty}\frac{e^{-iyz}}{\pi z^2}\,dz=\pi i\,\text{Res}\left[\frac{e^{-iyz}}{\pi z^2};z=0\right].$$

A simple computation of the residue gives $I(y)=-y$, for $y<0$.

When $y>0$, we need to close the contour in the lower half plane in order to apply Jordan's Lemma. Carrying out the computation, one finds $I(y)=y$, for $y>0$. Thus,

$$I(y)=\begin{cases}-y, & y>0,\\ y, & y<0,\end{cases} \qquad (5.60)$$

We are now ready to finish the computation of the convolution. We have to combine the integrals $I(y)$, $I(y+1)$, *and* $I(y-1)$, *since* $(f*f)(x) = I(x) - \frac{1}{2}[I(1-k) + I(1+k)]$. *This gives different results in four intervals:*

$$\begin{aligned}(f*f)(x) &= x - \frac{1}{2}[(x-2)+(x+2)] = 0, \quad x < -2,\\ &= x - \frac{1}{2}[(x-2)-(x+2)] = 2+x \quad -2 < x < 0,\\ &= -x - \frac{1}{2}[(x-2)-(x+2)] = 2-x, \quad 0 < x < 2,\\ &= -x - \frac{1}{2}[-(x-2)-(x+2)] = 0, \quad x > 2. \end{aligned} \tag{5.61}$$

A plot of this solution is the triangle function:

$$(f*f)(x) = \begin{cases} 0, & x < -2 \\ 2+x, & -2 < x < 0 \\ 2-x, & 0 < x < 2 \\ 0, & x > 2, \end{cases} \tag{5.62}$$

which was shown in the last example.

Example 5.14. *Find the convolution of the box function of height one and width two with itself using a direct computation of the convolution integral.*

The nonvanishing contributions to the convolution integral are when both $f(t)$ *and* $f(x-t)$ *do not vanish.* $f(t)$ *is nonzero for* $|t| \le 1$, *or* $-1 \le t \le 1$. $f(x-t)$ *is nonzero for* $|x-t| \le 1$, *or* $x-1 \le t \le x+1$. *These two regions are shown in Figure 5.22. On this region,* $f(t)g(x-t) = 1$.

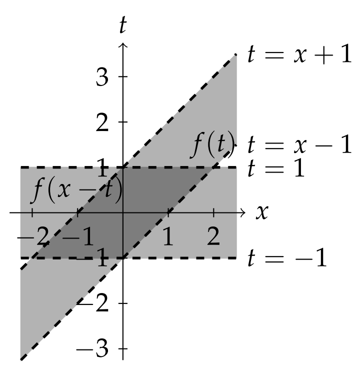


Figure 5.21: Plot of the regions of support for $f(t)$ and $f(x-t)$..

Thus, the nonzero contributions to the convolution are

$$(f*f)(x) = \begin{cases} \int_{-1}^{x+1} dt, & 0 \le x \le 2, \\ \int_{x-1}^{1} dt, & -2 \le x \le 0, \end{cases} = \begin{cases} 2+x, & 0 \le x \le 2, \\ 2-x, & -2 \le x \le 0. \end{cases}$$

Once again, we arrive at the triangle function.

In the last section we showed the graphical convolution. For completeness, we do the same for this example. In Figure 5.22 we show the results. We see that the convolution of two box functions is a triangle function.

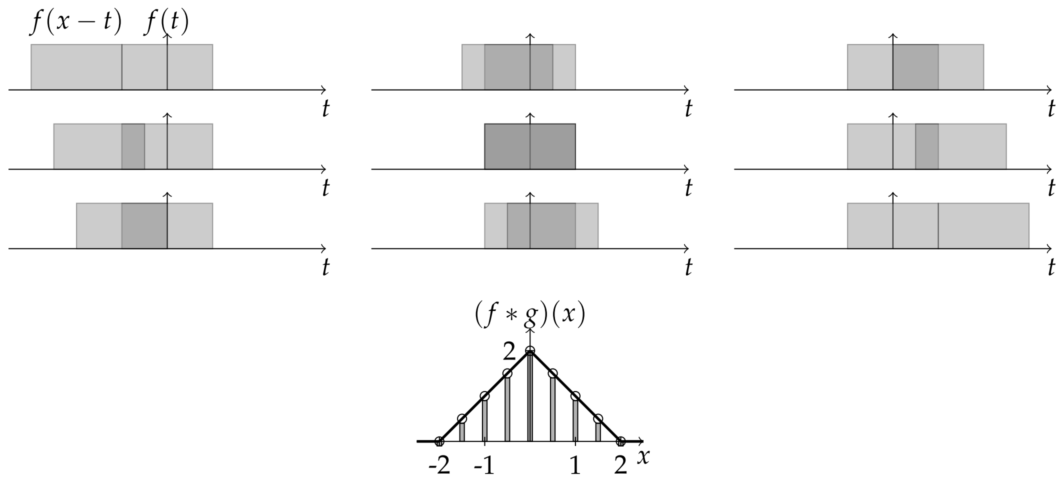


Figure 5.22: A plot of the convolution of a box function with itself. The areas of the overlaps of as $f(x-t)$ is translated across $f(t)$ are shown as well. The result is the triangular function.

Example 5.15. *Show the graphical convolution of the box function of height one and width two with itself.*

Let's consider a slightly more complicated example, the convolution of two Gaussian functions.

Example 5.16. *Convolution of two Gaussian functions $f(x) = e^{-ax^2}$.*

In this example we will compute the convolution of two Gaussian functions with different widths. Let $f(x) = e^{-ax^2}$ and $g(x) = e^{-bx^2}$. A direct evaluation of the integral would be to compute

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\, dt = \int_{-\infty}^{\infty} e^{-at^2-b(x-t)^2}\, dt.$$

This integral can be rewritten as

$$(f * g)(x) = e^{-bx^2} \int_{-\infty}^{\infty} e^{-(a+b)t^2+2bxt}\, dt.$$

One could proceed to complete the square and finish carrying out the integration. However, we will use the Convolution Theorem to evaluate the convolution and leave the evaluation of this integral to Problem 12.

Recalling the Fourier transform of a Gaussian from Example 5.5, we have

$$\hat{f}(k) = F[e^{-ax^2}] = \sqrt{\frac{\pi}{a}}e^{-k^2/4a} \tag{5.63}$$

and

$$\hat{g}(k) = F[e^{-bx^2}] = \sqrt{\frac{\pi}{b}}e^{-k^2/4b}.$$

Denoting the convolution function by $h(x) = (f * g)(x)$, the Convolution Theorem gives

$$\hat{h}(k) = \hat{f}(k)\hat{g}(k) = \frac{\pi}{\sqrt{ab}}e^{-k^2/4a}e^{-k^2/4b}.$$

This is another Gaussian function, as seen by rewriting the Fourier transform of $h(x)$ as

$$\hat{h}(k) = \frac{\pi}{\sqrt{ab}} e^{-\frac{1}{4}\left(\frac{1}{a}+\frac{1}{b}\right)k^2} = \frac{\pi}{\sqrt{ab}} e^{-\frac{a+b}{4ab}k^2}. \tag{5.64}$$

In order to complete the evaluation of the convolution of these two Gaussian functions, we need to find the inverse transform of the Gaussian in Equation (5.64). We can do this by looking at Equation (5.63). We have first that

$$F^{-1}\left[\sqrt{\frac{\pi}{a}} e^{-k^2/4a}\right] = e^{-ax^2}.$$

Moving the constants, we then obtain

$$F^{-1}[e^{-k^2/4a}] = \sqrt{\frac{a}{\pi}} e^{-ax^2}.$$

We now make the substitution $\alpha = \frac{1}{4a}$,

$$F^{-1}[e^{-\alpha k^2}] = \sqrt{\frac{1}{4\pi\alpha}} e^{-x^2/4\alpha}.$$

This is in the form needed to invert Equation (5.64). Thus, for $\alpha = \frac{a+b}{4ab}$, we find

$$(f * g)(x) = h(x) = \sqrt{\frac{\pi}{a+b}} e^{-\frac{ab}{a+b}x^2}.$$

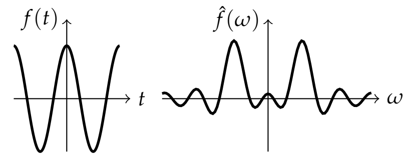


Figure 5.23: Schematic plot of a signal $f(t)$ and its Fourier transform $\hat{f}(\omega)$.

5.6.2 Application to Signal Analysis

THERE ARE MANY APPLICATIONS of the convolution operation. One of these areas is the study of analog signals. An analog signal is a continuous signal and may contain either a finite or continuous set of frequencies. Fourier transforms can be used to represent such signals as a sum over the frequency content of these signals. In this section we will describe how convolutions can be used in studying signal analysis.

Filtering signals.

The first application is filtering. For a given signal, there might be some noise in the signal, or some undesirable high frequencies. For example, a device used for recording an analog signal might naturally not be able to record high frequencies. Let $f(t)$ denote the amplitude of a given analog signal and $\hat{f}(\omega)$ be the Fourier transform of this signal such as the example provided in Figure 5.23. Recall that the Fourier transform gives the frequency content of the signal.

There are many ways to filter out unwanted frequencies. The simplest would be to just drop all the high (angular) frequencies. For example, for some cutoff frequency ω_0, frequencies $|\omega| > \omega_0$ will be removed. The Fourier transform of the filtered signal would then be zero for $|\omega| > \omega_0$. This could be accomplished by multiplying the Fourier transform of the signal by a function that vanishes for $|\omega| > \omega_0$. For example, we could use the gate function

$$p_{\omega_0}(\omega) = \begin{cases} 1, & |\omega| \le \omega_0, \\ 0, & |\omega| > \omega_0 \end{cases} \tag{5.65}$$

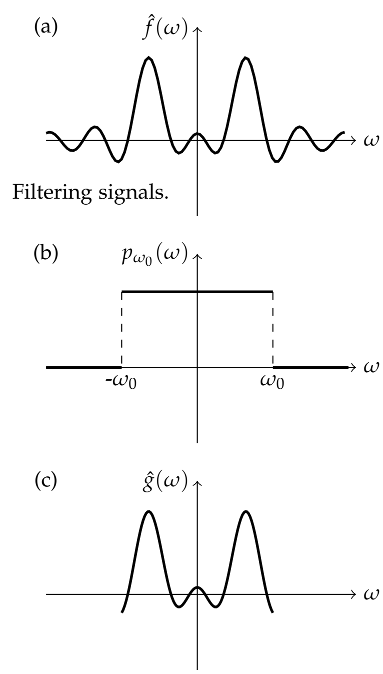


Figure 5.24: (a) Plot of the Fourier transform $\hat{f}(\omega)$ of a signal. (b) The gate function $p_{\omega_0}(\omega)$ used to filter out high frequencies. (c) The product of the functions, $\hat{g}(\omega) = \hat{f}(\omega)p_{\omega_0}(\omega)$, in (a) and (b) shows how the filters cuts out high frequencies, $|\omega| > \omega_0$.

as shown in Figure 5.24.

In general, we multiply the Fourier transform of the signal by some filtering function $\hat{h}(t)$ to get the Fourier transform of the filtered signal,

$$\hat{g}(\omega) = \hat{f}(\omega)\hat{h}(\omega).$$

The new signal, $g(t)$ is then the inverse Fourier transform of this product, giving the new signal as a convolution:

$$g(t) = F^{-1}[\hat{f}(\omega)\hat{h}(\omega)] = \int_{-\infty}^{\infty} h(t-\tau)f(\tau)\, d\tau. \tag{5.66}$$

Such processes occur often in systems theory as well. One thinks of $f(t)$ as the input signal into some filtering device, which in turn produces the output, $g(t)$. The function $h(t)$ is called the impulse response. This is because it is a response to the impulse function, $\delta(t)$. In this case, one has

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

Windowing signals.

Another application of the convolution is in windowing. This represents what happens when one measures a real signal. Real signals cannot be recorded for all values of time. Instead, data is collected over a finite time interval. If the length of time the data is collected is T, then the resulting signal is zero outside this time interval. This can be modeled in the same way as with filtering, except the new signal will be the product of the old signal with the windowing function. The resulting Fourier transform of the new signal will be a convolution of the Fourier transforms of the original signal and the windowing function.

Example 5.17. Finite Wave Train, Revisited.

We return to the finite wave train in Example 5.10 given by

$$h(t) = \begin{cases} \cos\omega_0 t, & |t| \le a, \\ 0, & |t| > a. \end{cases}$$

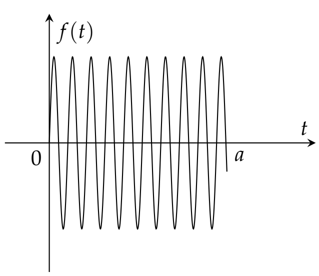


Figure 5.25: A plot of the finite wave train.

We can view this as a windowed version of $f(t) = \cos\omega_0 t$ obtained by multiplying $f(t)$ by the gate function

$$g_a(t) = \begin{cases} 1, & |x| \le a, \\ 0, & |x| > a. \end{cases} \tag{5.67}$$

This is shown in Figure 5.25. Then, the Fourier transform is given as a convolution,

$$\begin{aligned} \hat{h}(\omega) &= (\hat{f} * \hat{g}_a)(\omega) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega - \nu)\hat{g}_a(\nu)\, d\nu. \end{aligned} \tag{5.68}$$

The convolution in spectral space is defined with an extra factor of $1/2\pi$ so as to preserve the idea that the inverse Fourier transform of a convolution is the product of the corresponding signals.

Note that the convolution in frequency space requires the extra factor of $1/(2\pi)$.

We need the Fourier transforms of f and g_a in order to finish the computation. The Fourier transform of the box function was found in Example 5.6 as

$$\hat{g}_a(\omega) = \frac{2}{\omega}\sin\omega a.$$

The Fourier transform of the cosine function, $f(t) = \cos\omega_0 t$, is

$$\begin{aligned}\hat{f}(\omega) &= \int_{-\infty}^{\infty} \cos(\omega_0 t) e^{i\omega t}\, dt \\ &= \int_{-\infty}^{\infty} \frac{1}{2}\left(e^{i\omega_0 t} + e^{-i\omega_0 t}\right) e^{i\omega t}\, dt \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \left(e^{i(\omega+\omega_0)t} + e^{i(\omega-\omega_0)t}\right)\, dt \\ &= \pi\left[\delta(\omega+\omega_0) + \delta(\omega-\omega_0)\right]. \end{aligned} \tag{5.69}$$

Note that we had earlier computed the inverse Fourier transform of this function in Example 5.9.

Inserting these results in the convolution integral, we have

$$\begin{aligned}\hat{h}(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega-\nu)\hat{g}_a(\nu)\, d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \pi\left[\delta(\omega-\nu+\omega_0) + \delta(\omega-\nu-\omega_0)\right]\frac{2}{\nu}\sin\nu a\, d\nu \\ &= \frac{\sin(\omega+\omega_0)a}{\omega+\omega_0} + \frac{\sin(\omega-\omega_0)a}{\omega-\omega_0}. \end{aligned} \tag{5.70}$$

This is the same result we had obtained in Example 5.10.

5.6.3 Parseval's Equality

The integral/sum of the (modulus) square of a function is the integral/sum of the (modulus) square of the transform.

AS ANOTHER EXAMPLE OF THE CONVOLUTION THEOREM, we derive Parseval's Equality (named after Marc-Antoine Parseval (1755 - 1836)):

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\, d\omega. \tag{5.71}$$

This equality has a physical meaning for signals. The integral on the left side is a measure of the energy content of the signal in the time domain. The right side provides a measure of the energy content of the transform of the signal. Parseval's Equality, is simply a statement that the energy is invariant under the Fourier transform. Parseval's Equality is a special case of Plancherel's Formula (named after Michel Plancherel, 1885 - 1967).

Let's rewrite the Convolution Theorem in its inverse form

$$F^{-1}[\hat{f}(k)\hat{g}(k)] = (f * g)(t). \tag{5.72}$$

Then, by the definition of the inverse Fourier transform, we have

$$\int_{-\infty}^{\infty} f(t-u)g(u)\, du = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)\hat{g}(\omega)e^{-i\omega t}\, d\omega.$$

Setting $t = 0$,

$$\int_{-\infty}^{\infty} f(-u)g(u)\, du = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)\hat{g}(\omega)\, d\omega. \tag{5.73}$$

Now, let $g(t) = \overline{f(-t)}$, or $f(-t) = \overline{g(t)}$. We note that the Fourier transform of $g(t)$ is related to the Fourier transform of $f(t)$:

$$\begin{aligned} \hat{g}(\omega) &= \int_{-\infty}^{\infty} \overline{f(-t)} e^{i\omega t}\, dt \\ &= -\int_{\infty}^{-\infty} \overline{f(\tau)} e^{-i\omega\tau}\, d\tau \\ &= \overline{\int_{-\infty}^{\infty} f(\tau) e^{i\omega\tau}\, d\tau} = \overline{\hat{f}(\omega)}. \end{aligned} \tag{5.74}$$

So, inserting this result into Equation (5.73), we find that

$$\int_{-\infty}^{\infty} f(-u)\overline{f(-u)}\, du = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\, d\omega,$$

which yields Parseval's Equality in the form in Equation (5.71) after substituting $t = -u$ on the left.

As noted above, the forms in Equations (5.71) and (5.73) are often referred to as the Plancherel Formula or Parseval Formula. A more commonly defined Parseval equation is that given for Fourier series. For example, for a function $f(x)$ defined on $[-\pi, \pi]$, which has a Fourier series representation, we have

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} [f(x)]^2\, dx.$$

In general, there is a Parseval identity for functions that can be expanded in a complete sets of orthonormal functions, $\{\phi_n(x)\}$, $n = 1, 2, \ldots$, which is given by

$$\sum_{n=1}^{\infty} < f, \phi_n >^2 = \|f\|^2.$$

Here, $\|f\|^2 =< f, f >$. The Fourier series example is just a special case of this formula.

5.7 The Laplace Transform

The Laplace transform is named after Pierre-Simon de Laplace (1749 - 1827). Laplace made major contributions, especially to celestial mechanics, tidal analysis, and probability.

Up to this point we have only explored Fourier exponential transforms as one type of integral transform. The Fourier transform is useful on infinite domains. However, students are often introduced to another integral transform, called the Laplace transform, in their introductory differential equations class. These transforms are defined over semi-infinite domains and are useful for solving initial value problems for ordinary differential equations.

Integral transform on $[a, b]$ with respect to the integral kernel, $K(x, k)$.

The Fourier and Laplace transforms are examples of a broader class of transforms known as integral transforms. For a function $f(x)$ defined on an interval (a, b), we define the integral transform

$$F(k) = \int_a^b K(x, k) f(x)\, dx,$$

where $K(x,k)$ is a specified kernel of the transform. Looking at the Fourier transform, we see that the interval is stretched over the entire real axis and the kernel is of the form, $K(x,k) = e^{ikx}$. In Table 5.1 we show several types of integral transforms.

Laplace Transform	$F(s) = \int_0^\infty e^{-sx} f(x)\,dx$
Fourier Transform	$F(k) = \int_{-\infty}^\infty e^{ikx} f(x)\,dx$
Fourier Cosine Transform	$F(k) = \int_0^\infty \cos(kx) f(x)\,dx$
Fourier Sine Transform	$F(k) = \int_0^\infty \sin(kx) f(x)\,dx$
Mellin Transform	$F(k) = \int_0^\infty x^{k-1} f(x)\,dx$
Hankel Transform	$F(k) = \int_0^\infty x J_n(kx) f(x)\,dx$

Table 5.1: A Table of Common Integral Transforms.

It should be noted that these integral transforms inherit the linearity of integration. Namely, let $h(x) = \alpha f(x) + \beta g(x)$, where α and β are constants. Then,

$$\begin{aligned} H(k) &= \int_a^b K(x,k) h(x)\,dx, \\ &= \int_a^b K(x,k)(\alpha f(x) + \beta g(x))\,dx, \\ &= \alpha \int_a^b K(x,k) f(x)\,dx + \beta \int_a^b K(x,k) g(x)\,dx, \\ &= \alpha F(x) + \beta G(x). \end{aligned} \tag{5.75}$$

Therefore, we have shown linearity of the integral transforms. We have seen the linearity property used for Fourier transforms and we will use linearity in the study of Laplace transforms.

The Laplace transform of f, $F = \mathcal{L}[f]$.

We now turn to Laplace transforms. The Laplace transform of a function $f(t)$ is defined as

$$F(s) = \mathcal{L}[f](s) = \int_0^\infty f(t) e^{-st}\,dt, \quad s > 0. \tag{5.76}$$

This is an improper integral and one needs

$$\lim_{t\to\infty} f(t) e^{-st} = 0$$

to guarantee convergence.

Laplace transforms also have proven useful in engineering for solving circuit problems and doing systems analysis. In Figure 5.26 it is shown that a signal $x(t)$ is provided as input to a linear system, indicated by $h(t)$. One is interested in the system output, $y(t)$, which is given by a convolution of the input and system functions. By considering the transforms of $x(t)$ and $h(t)$, the transform of the output is given as a product of the Laplace transforms in the s-domain. In order to obtain the output, one needs to compute a convolution product for Laplace transforms similar to the convolution operation we had seen for Fourier transforms earlier in the chapter. Of course, for us to do this in practice, we have to know how to compute Laplace transforms.

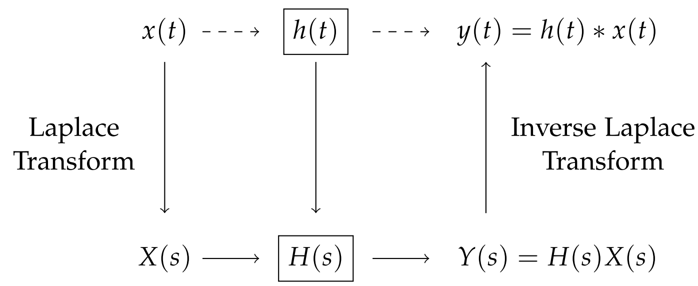


Figure 5.26: A schematic depicting the use of Laplace transforms in systems theory.

5.7.1 *Properties and Examples of Laplace Transforms*

It is typical that one makes use of Laplace transforms by referring to a Table of transform pairs. A sample of such pairs is given in Table 5.2. Combining some of these simple Laplace transforms with the properties of the Laplace transform, as shown in Table 5.3, we can deal with many applications of the Laplace transform. We will first prove a few of the given Laplace transforms and show how they can be used to obtain new transform pairs. In the next section we will show how these transforms can be used to sum infinite series and to solve initial value problems for ordinary differential equations.

Table 5.2: Table of Selected Laplace Transform Pairs.

$f(t)$	$F(s)$	$f(t)$	$F(s)$
c	$\frac{c}{s}$	e^{at}	$\frac{1}{s-a}, \quad s > a$
t^n	$\frac{n!}{s^{n+1}}, \quad s > 0$	$t^n e^{at}$	$\frac{n!}{(s-a)^{n+1}}$
$\sin \omega t$	$\frac{\omega}{s^2+\omega^2}$	$e^{at} \sin \omega t$	$\frac{\omega}{(s-a)^2+\omega^2}$
$\cos \omega t$	$\frac{s}{s^2+\omega^2}$	$e^{at} \cos \omega t$	$\frac{s-a}{(s-a)^2+\omega^2}$
$t \sin \omega t$	$\frac{2\omega s}{(s^2+\omega^2)^2}$	$t \cos \omega t$	$\frac{s^2-\omega^2}{(s^2+\omega^2)^2}$
$\sinh at$	$\frac{a}{s^2-a^2}$	$\cosh at$	$\frac{s}{s^2-a^2}$
$H(t-a)$	$\frac{e^{-as}}{s}, \quad s > 0$	$\delta(t-a)$	$e^{-as}, \quad a \geq 0, s > 0$

We begin with some simple transforms. These are found by simply using the definition of the Laplace transform.

Example 5.18. *Show that* $\mathcal{L}[1] = \frac{1}{s}$.

For this example, we insert $f(t) = 1$ *into the definition of the Laplace transform:*

$$\mathcal{L}[1] = \int_0^\infty e^{-st}\, dt.$$

This is an improper integral and the computation is understood by introducing an upper limit of a *and then letting* $a \to \infty$. *We will not always write this limit, but it will be understood that this is how one computes such improper integrals.*

Proceeding with the computation, we have

$$\begin{aligned}
\mathcal{L}[1] &= \int_0^\infty e^{-st}\,dt \\
&= \lim_{a\to\infty}\int_0^a e^{-st}\,dt \\
&= \lim_{a\to\infty}\left(-\frac{1}{s}e^{-st}\right)_0^a \\
&= \lim_{a\to\infty}\left(-\frac{1}{s}e^{-sa}+\frac{1}{s}\right) = \frac{1}{s}.
\end{aligned} \tag{5.77}$$

Thus, we have found that the Laplace transform of 1 is $\frac{1}{s}$. This result can be extended to any constant c, using the linearity of the transform, $\mathcal{L}[c] = c\mathcal{L}[1]$. Therefore,

$$\mathcal{L}[c] = \frac{c}{s}.$$

Example 5.19. *Show that* $\mathcal{L}[e^{at}] = \frac{1}{s-a}$, *for* $s > a$.

For this example, we can easily compute the transform. Again, we only need to compute the integral of an exponential function.

$$\begin{aligned}
\mathcal{L}[e^{at}] &= \int_0^\infty e^{at}e^{-st}\,dt \\
&= \int_0^\infty e^{(a-s)t}\,dt \\
&= \left(\frac{1}{a-s}e^{(a-s)t}\right)_0^\infty \\
&= \lim_{t\to\infty}\frac{1}{a-s}e^{(a-s)t} - \frac{1}{a-s} = \frac{1}{s-a}.
\end{aligned} \tag{5.78}$$

Note that the last limit was computed as $\lim_{t\to\infty} e^{(a-s)t} = 0$. *This is only true if* $a - s < 0$, *or* $s > a$. *[Actually, a could be complex. In this case we would only need s to be greater than the real part of a,* $s > Re(a)$.*]*

Example 5.20. *Show that* $\mathcal{L}[\cos at] = \frac{s}{s^2+a^2}$ *and* $\mathcal{L}[\sin at] = \frac{a}{s^2+a^2}$.

For these examples, we could again insert the trigonometric functions directly into the transform and integrate. For example,

$$\mathcal{L}[\cos at] = \int_0^\infty e^{-st}\cos at\,dt.$$

Recall how one evaluates integrals involving the product of a trigonometric function and the exponential function. One integrates by parts two times and then obtains an integral of the original unknown integral. Rearranging the resulting integral expressions, one arrives at the desired result. However, there is a much simpler way to compute these transforms.

Recall that $e^{iat} = \cos at + i\sin at$. *Making use of the linearity of the Laplace transform, we have*

$$\mathcal{L}[e^{iat}] = \mathcal{L}[\cos at] + i\mathcal{L}[\sin at].$$

Thus, transforming this complex exponential will simultaneously provide the Laplace transforms for the sine and cosine functions!

The transform is simply computed as

$$\mathcal{L}[e^{iat}] = \int_0^\infty e^{iat}e^{-st}\,dt = \int_0^\infty e^{-(s-ia)t}\,dt = \frac{1}{s-ia}.$$

Note that we could easily have used the result for the transform of an exponential, which was already proven. In this case, $s > \mathrm{Re}(ia) = 0$.

We now extract the real and imaginary parts of the result using the complex conjugate of the denominator:

$$\frac{1}{s-ia} = \frac{1}{s-ia}\frac{s+ia}{s+ia} = \frac{s+ia}{s^2+a^2}.$$

Reading off the real and imaginary parts, we find the sought-after transforms,

$$\begin{aligned} \mathcal{L}[\cos at] &= \frac{s}{s^2+a^2}, \\ \mathcal{L}[\sin at] &= \frac{a}{s^2+a^2}. \end{aligned} \tag{5.79}$$

Example 5.21. *Show that* $\mathcal{L}[t] = \frac{1}{s^2}$.

For this example we evaluate

$$\mathcal{L}[t] = \int_0^\infty te^{-st}\,dt.$$

This integral can be evaluated using the method of integration by parts:

$$\begin{aligned} \int_0^\infty te^{-st}\,dt &= -t\frac{1}{s}e^{-st}\Big|_0^\infty + \frac{1}{s}\int_0^\infty e^{-st}\,dt \\ &= \frac{1}{s^2}. \end{aligned} \tag{5.80}$$

Example 5.22. *Show that* $\mathcal{L}[t^n] = \frac{n!}{s^{n+1}}$ *for nonnegative integer* n.

We have seen the $n = 0$ *and* $n = 1$ *cases:* $\mathcal{L}[1] = \frac{1}{s}$ *and* $\mathcal{L}[t] = \frac{1}{s^2}$. *We now generalize these results to nonnegative integer powers,* $n > 1$, *of* t. *We consider the integral*

$$\mathcal{L}[t^n] = \int_0^\infty t^n e^{-st}\,dt.$$

Following the previous example, we again integrate by parts:[2]

$$\begin{aligned} \int_0^\infty t^n e^{-st}\,dt &= -t^n\frac{1}{s}e^{-st}\Big|_0^\infty + \frac{n}{s}\int_0^\infty t^{-n}e^{-st}\,dt \\ &= \frac{n}{s}\int_0^\infty t^{-n}e^{-st}\,dt. \end{aligned} \tag{5.81}$$

We could continue to integrate by parts until the final integral is computed. However, look at the integral that resulted after one integration by parts. It is just the Laplace transform of t^{n-1}. *So, we can write the result as*

$$\mathcal{L}[t^n] = \frac{n}{s}\mathcal{L}[t^{n-1}].$$

This is an example of a recursive definition of a sequence. In this case, we have a sequence of integrals. Denoting

$$I_n = \mathcal{L}[t^n] = \int_0^\infty t^n e^{-st}\,dt$$

[2] This integral can just as easily be done using differentiation. We note that

$$\left(-\frac{d}{ds}\right)^n \int_0^\infty e^{-st}\,dt = \int_0^\infty t^n e^{-st}\,dt.$$

Since

$$\int_0^\infty e^{-st}\,dt = \frac{1}{s},$$

$$\int_0^\infty t^n e^{-st}\,dt = \left(-\frac{d}{ds}\right)^n \frac{1}{s} = \frac{n!}{s^{n+1}}.$$

We compute $\int_0^\infty t^n e^{-st}\,dt$ by turning it into an initial value problem for a first-order difference equation and finding the solution using an iterative method.

and noting that $I_0 = \mathcal{L}[1] = \frac{1}{s}$, we have the following:

$$I_n = \frac{n}{s} I_{n-1}, \quad I_0 = \frac{1}{s}. \tag{5.82}$$

This is also what is called a difference equation. It is a first-order difference equation with an "initial condition," I_0. The next step is to solve this difference equation.

Finding the solution of this first-order difference equation is easy to do using simple iteration. Note that replacing n with $n-1$, we have

$$I_{n-1} = \frac{n-1}{s} I_{n-2}.$$

Repeating the process, we find

$$\begin{aligned} I_n &= \frac{n}{s} I_{n-1} \\ &= \frac{n}{s}\left(\frac{n-1}{s} I_{n-2}\right) \\ &= \frac{n(n-1)}{s^2} I_{n-2} \\ &= \frac{n(n-1)(n-2)}{s^3} I_{n-3}. \end{aligned} \tag{5.83}$$

We can repeat this process until we get to I_0, which we know. We have to carefully count the number of iterations. We do this by iterating k times and then figure out how many steps will get us to the known initial value. A list of iterates is easily written out:

$$\begin{aligned} I_n &= \frac{n}{s} I_{n-1} \\ &= \frac{n(n-1)}{s^2} I_{n-2} \\ &= \frac{n(n-1)(n-2)}{s^3} I_{n-3} \\ &= \ldots \\ &= \frac{n(n-1)(n-2)\ldots(n-k+1)}{s^k} I_{n-k}. \end{aligned} \tag{5.84}$$

Since we know $I_0 = \frac{1}{s}$, we choose to stop at $k = n$ obtaining

$$I_n = \frac{n(n-1)(n-2)\ldots(2)(1)}{s^n} I_0 = \frac{n!}{s^{n+1}}.$$

Therefore, we have shown that $\mathcal{L}[t^n] = \frac{n!}{s^{n+1}}$.

Such iterative techniques are useful in obtaining a variety of integrals, such as $I_n = \int_{-\infty}^{\infty} x^{2n} e^{-x^2}\, dx$.

As a final note, one can extend this result to cases when n is not an integer. To do this, we use the Gamma function, which was discussed in Section 3.5. Recall that the Gamma function is the generalization of the factorial function and is defined as

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt. \tag{5.85}$$

Note the similarity to the Laplace transform of t^{x-1} :

$$\mathcal{L}[t^{x-1}] = \int_0^\infty t^{x-1} e^{-st}\, dt.$$

For $x-1$ an integer and $s=1$, we have that

$$\Gamma(x) = (x-1)!.$$

Thus, the Gamma function can be viewed as a generalization of the factorial and we have shown that

$$\mathcal{L}[t^p] = \frac{\Gamma(p+1)}{s^{p+1}}$$

for $p > -1$.

Now we are ready to introduce additional properties of the Laplace transform in Table 5.3. We have already discussed the first property, which is a consequence of the linearity of integral transforms. We will prove the other properties in this and the following sections.

Table 5.3: Table of selected Laplace transform properties.

Laplace Transform Properties
$\mathcal{L}[af(t) + bg(t)] = aF(s) + bG(s)$
$\mathcal{L}[tf(t)] = -\frac{d}{ds}F(s)$
$\mathcal{L}\left[\frac{df}{dt}\right] = sF(s) - f(0)$
$\mathcal{L}\left[\frac{d^2f}{dt^2}\right] = s^2F(s) - sf(0) - f'(0)$
$\mathcal{L}[e^{at}f(t)] = F(s-a)$
$\mathcal{L}[H(t-a)f(t-a)] = e^{-as}F(s)$
$\mathcal{L}[(f * g)(t)] = \mathcal{L}[\int_0^t f(t-u)g(u)\, du] = F(s)G(s)$

Example 5.23. *Show that* $\mathcal{L}\left[\frac{df}{dt}\right] = sF(s) - f(0)$.

We have to compute

$$\mathcal{L}\left[\frac{df}{dt}\right] = \int_0^\infty \frac{df}{dt} e^{-st}\, dt.$$

We can move the derivative off f by integrating by parts. This is similar to what we had done when finding the Fourier transform of the derivative of a function. Letting $u = e^{-st}$ and $v = f(t)$, we have

$$\begin{aligned} \mathcal{L}\left[\frac{df}{dt}\right] &= \int_0^\infty \frac{df}{dt} e^{-st}\, dt \\ &= f(t)e^{-st}\Big|_0^\infty + s\int_0^\infty f(t)e^{-st}\, dt \\ &= -f(0) + sF(s). \end{aligned} \tag{5.86}$$

Here we have assumed that $f(t)e^{-st}$ vanishes for large t.

The final result is that

$$\mathcal{L}\left[\frac{df}{dt}\right] = sF(s) - f(0).$$

Example 6: *Show that* $\mathcal{L}\left[\frac{d^2f}{dt^2}\right] = s^2F(s) - sf(0) - f'(0)$.

We can compute this Laplace transform using two integrations by parts, or we could make use of the last result. Letting $g(t) = \frac{df(t)}{dt}$, *we have*

$$\mathcal{L}\left[\frac{d^2f}{dt^2}\right] = \mathcal{L}\left[\frac{dg}{dt}\right] = sG(s) - g(0) = sG(s) - f'(0).$$

But,

$$G(s) = \mathcal{L}\left[\frac{df}{dt}\right] = sF(s) - f(0).$$

So,

$$\begin{aligned}\mathcal{L}\left[\frac{d^2f}{dt^2}\right] &= sG(s) - f'(0)\\ &= s\left[sF(s) - f(0)\right] - f'(0)\\ &= s^2F(s) - sf(0) - f'(0). \end{aligned} \quad (5.87)$$

We will return to the other properties in Table 5.3 after looking at a few applications.

5.8 Applications of Laplace Transforms

Although the Laplace transform is a very useful transform, it is often encountered only as a method for solving initial value problems in introductory differential equations. In this section we will show how to solve simple differential equations. Along the way we will introduce step and impulse functions and show how the Convolution Theorem for Laplace transforms plays a role in finding solutions. However, we will first explore an unrelated application of Laplace transforms. We will see that the Laplace transform is useful in finding sums of infinite series.

5.8.1 Series Summation Using Laplace Transforms

We saw in Chapter 2 that Fourier series can be used to sum series. For example, in Problem 2.13, one proves that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

In this section we will show how Laplace transforms can be used to sum series.[3] There is an interesting history of using integral transforms to sum series. For example, Richard Feynman[4] (1918 - 1988) described how one can use the Convolution Theorem for Laplace transforms to sum series with denominators that involved products. We will describe this and simpler sums in this section.

We begin by considering the Laplace transform of a known function,

$$F(s) = \int_0^\infty f(t)e^{-st}\,dt.$$

[3] Albert D. Wheelon, *Tables of Summable Series and Integrals Involving Bessel Functions*, Holden-Day, 1968.

[4] R. P. Feynman, 1949, *Phys. Rev.* **76**, p. 769

Inserting this expression into the sum $\sum_n F(n)$ and interchanging the sum and integral, we find

$$\begin{aligned} \sum_{n=0}^{\infty} F(n) &= \sum_{n=0}^{\infty} \int_0^\infty f(t) e^{-nt}\,dt \\ &= \int_0^\infty f(t) \sum_{n=0}^{\infty} \left(e^{-t}\right)^n\,dt \\ &= \int_0^\infty f(t) \frac{1}{1-e^{-t}}\,dt. \end{aligned} \tag{5.88}$$

The last step was obtained using the sum of a geometric series. The key is being able to carry out the final integral as we show in the next example.

Example 5.24. *Evaluate the sum* $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$.

Since, $\mathcal{L}[1] = 1/s$, *we have*

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} &= \sum_{n=1}^{\infty} \int_0^\infty (-1)^{n+1} e^{-nt}\,dt \\ &= \int_0^\infty \frac{e^{-t}}{1+e^{-t}}\,dt \\ &= \int_1^2 \frac{du}{u} = \ln 2. \end{aligned} \tag{5.89}$$

Example 5.25. *Evaluate the sum* $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

This is a special case of the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{5.90}$$

The Riemann zeta function[5] *is important in the study of prime numbers and more recently has seen applications in the study of dynamical systems. The series in this example is* $\zeta(2)$. *We have already seen in Problem 2.13 that*

$$\zeta(2) = \frac{\pi^2}{6}.$$

Using Laplace transforms, we can provide an integral representation of $\zeta(2)$.

The first step is to find the correct Laplace transform pair. The sum involves the function $F(n) = 1/n^2$. *So, we look for a function* $f(t)$ *whose Laplace transform is* $F(s) = 1/s^2$. *We know by now that the inverse Laplace transform of* $F(s) = 1/s^2$ *is* $f(t) = t$. *As before, we replace each term in the series by a Laplace transform, exchange the summation and integration, and sum the resulting geometric series:*

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n^2} &= \sum_{n=1}^{\infty} \int_0^\infty t e^{-nt}\,dt \\ &= \int_0^\infty \frac{t}{e^t - 1}\,dt. \end{aligned} \tag{5.91}$$

So, we have that

$$\int_0^\infty \frac{t}{e^t-1}\,dt = \sum_{n=1}^{\infty} \frac{1}{n^2} = \zeta(2).$$

[5] A translation of Riemann, Bernhard (1859), "Über die Anzahl der Primzahlen unter einer gegebenen Grösse" is in H. M. Edwards (1974). *Riemann's Zeta Function.* Academic Press. Riemann had shown that the Riemann zeta function can be obtained through contour integral representation, $2\sin(\pi s)\Gamma\zeta(s) = i\oint_C \frac{(-x)^{s-1}}{e^x-1}\,dx$, for a specific contour C.

Integrals of this type occur often in statistical mechanics in the form of Bose-Einstein integrals. These are of the form

$$G_n(z) = \int_0^\infty \frac{x^{n-1}}{z^{-1}e^x - 1}\, dx.$$

Note that $G_n(1) = \Gamma(n)\zeta(n)$.

In general, the Riemann zeta function must be tabulated through other means. In some special cases, one can use closed form expressions. For example,

$$\zeta(2n) = \frac{2^{2n-1}\pi^{2n}}{(2n)!} B_n,$$

where the B_n's are the Bernoulli numbers. Bernoulli numbers are defined through the Maclaurin series expansion

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n.$$

The first few Riemann zeta functions are

$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945}.$$

We can extend this method of using Laplace transforms to summing series whose terms take special general forms. For example, from Feynman's 1949 paper, we note that

$$\frac{1}{(a+bn)^2} = -\frac{\partial}{\partial a} \int_0^\infty e^{-s(a+bn)}\, ds.$$

This identity can be shown easily by first noting

$$\int_0^\infty e^{-s(a+bn)}\, ds = \left[\frac{-e^{-s(a+bn)}}{a+bn} \right]_0^\infty = \frac{1}{a+bn}.$$

Now, differentiate the result with respect to a and the result follows.

The latter identity can be generalized further as

$$\frac{1}{(a+bn)^{k+1}} = \frac{(-1)^k}{k!} \frac{\partial^k}{\partial a^k} \int_0^\infty e^{-s(a+bn)}\, ds.$$

In Feynman's 1949 paper, he develops methods for handling several other general sums using the Convolution Theorem. Wheelon gives more examples of these. We will just provide one such result and an example. First, we note that

$$\frac{1}{ab} = \int_0^1 \frac{du}{[a(1-u)+bu]^2}.$$

However,

$$\frac{1}{[a(1-u)+bu]^2} = \int_0^\infty t e^{-t[a(1-u)+bu]}\, dt.$$

So, we have

$$\frac{1}{ab} = \int_0^1 du \int_0^\infty t e^{-t[a(1-u)+bu]}\, dt.$$

We see in the next example how this representation can be useful.

Example 5.26. *Evaluate* $\sum_{n=0}^{\infty} \frac{1}{(2n+1)(2n+2)}$.

We sum this series by first letting $a = 2n+1$ *and* $b = 2n+2$ *in the formula for* $1/ab$. *Collecting the n-dependent terms, we can sum the series leaving a double integral computation in ut-space. The details are as follows:*

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{1}{(2n+1)(2n+2)} &= \sum_{n=0}^{\infty} \int_0^1 \frac{du}{[(2n+1)(1-u)+(2n+2)u]^2} \\
&= \sum_{n=0}^{\infty} \int_0^1 du \int_0^{\infty} te^{-t(2n+1+u)}\, dt \\
&= \int_0^1 du \int_0^{\infty} te^{-t(1+u)} \sum_{n=0}^{\infty} e^{-2nt}\, dt \\
&= \int_0^{\infty} \frac{te^{-t}}{1-e^{-2t}} \int_0^1 e^{-tu}\, du\, dt \\
&= \int_0^{\infty} \frac{te^{-t}}{1-e^{-2t}} \frac{1-e^{-t}}{t}\, dt \\
&= \int_0^{\infty} \frac{e^{-t}}{1+e^{-t}}\, dt \\
&= -\ln(1+e^{-t})\Big|_0^{\infty} = \ln 2. \qquad (5.92)
\end{aligned}$$

5.8.2 Solution of ODEs Using Laplace Transforms

One of the typical applications of Laplace transforms is the solution of nonhomogeneous linear constant coefficient differential equations. In the following examples we will show how this works.

The general idea is that one transforms the equation for an unknown function $y(t)$ into an algebraic equation for its transform, $Y(t)$. Typically, the algebraic equation is easy to solve for $Y(s)$ as a function of s. Then, one transforms back into t-space using Laplace transform tables and the properties of Laplace transforms. The scheme is shown in Figure 5.27.

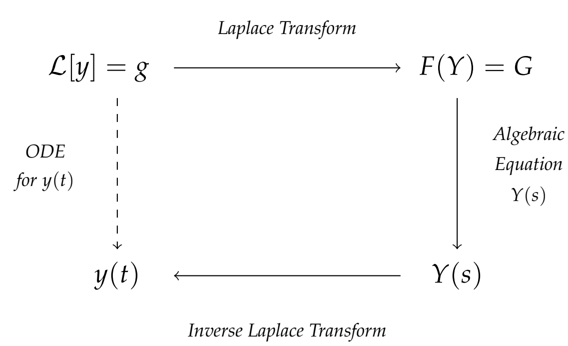


Figure 5.27: The scheme for solving an ordinary differential equation using Laplace transforms. One transforms the initial value problem for $y(t)$ and obtains an algebraic equation for $Y(s)$. Solve for $Y(s)$ and the inverse transform gives the solution to the initial value problem.

Example 5.27. *Solve the initial value problem* $y' + 3y = e^{2t}$, $y(0) = 1$.

The first step is to perform a Laplace transform of the initial value problem. The transform of the left side of the equation is

$$\mathcal{L}[y' + 3y] = sY - y(0) + 3Y = (s+3)Y - 1.$$

Transforming the righthand side, we have

$$\mathcal{L}[e^{2t}] = \frac{1}{s-2}.$$

Combining these two results, we obtain

$$(s+3)Y - 1 = \frac{1}{s-2}.$$

The next step is to solve for $Y(s)$:

$$Y(s) = \frac{1}{s+3} + \frac{1}{(s-2)(s+3)}.$$

Now we need to find the inverse Laplace transform. Namely, we need to figure out what function has a Laplace transform of the above form. We will use the tables of Laplace transform pairs. Later we will show that there are other methods for carrying out the Laplace transform inversion.

The inverse transform of the first term is e^{-3t}. However, we have not seen anything that looks like the second form in the table of transforms that we have compiled, but we can rewrite the second term using a partial fraction decomposition. Let's recall how to do this.

The goal is to find constants A and B such that

$$\frac{1}{(s-2)(s+3)} = \frac{A}{s-2} + \frac{B}{s+3}. \tag{5.93}$$

We picked this form because we know that recombining the two terms into one term will have the same denominator. We just need to make sure the numerators agree afterward. So, adding the two terms, we have

This is an example of carrying out a partial fraction decomposition.

$$\frac{1}{(s-2)(s+3)} = \frac{A(s+3)+B(s-2)}{(s-2)(s+3)}.$$

Equating numerators,

$$1 = A(s+3) + B(s-2).$$

There are several ways to proceed at this point.

a. *Method 1.*

   We can rewrite the equation by gathering terms with common powers of s, we have

$$(A+B)s + 3A - 2B = 1.$$

   The only way that this can be true for all s is that the coefficients of the different powers of s agree on both sides. This leads to two equations for A and B:

$$\begin{aligned} A + B &= 0, \\ 3A - 2B &= 1. \end{aligned} \tag{5.94}$$

   The first equation gives $A = -B$, so the second equation becomes $-5B = 1$. The solution is then $A = -B = \frac{1}{5}$.

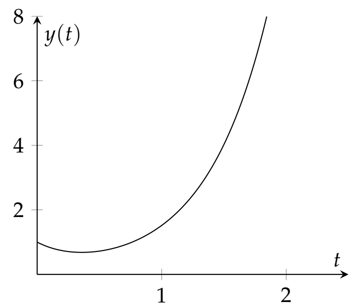


Figure 5.28: A plot of the solution to Example 5.27.

b. *Method 2.*

Since the equation $\frac{1}{(s-2)(s+3)} = \frac{A}{s-2} + \frac{B}{s+3}$ *is true for all* s*, we can pick specific values. For* $s = 2$*, we find* $1 = 5A$*, or* $A = \frac{1}{5}$*. For* $s = -3$*, we find* $1 = -5B$*, or* $B = -\frac{1}{5}$*. Thus, we obtain the same result as Method 1, but much quicker.*

c. *Method 3.*

We could just inspect the original partial fraction problem. Since the numerator has no s *terms, we might guess the form*

$$\frac{1}{(s-2)(s+3)} = \frac{1}{s-2} - \frac{1}{s+3}.$$

But, recombining the terms on the right hand side, we see that

$$\frac{1}{s-2} - \frac{1}{s+3} = \frac{5}{(s-2)(s+3)}.$$

Since we were off by 5, we divide the partial fractions by 5 to obtain

$$\frac{1}{(s-2)(s+3)} = \frac{1}{5}\left[\frac{1}{s-2} - \frac{1}{s+3}\right],$$

which once again gives the desired form.

Returning to the problem, we have found that

$$Y(s) = \frac{1}{s+3} + \frac{1}{5}\left(\frac{1}{s-2} - \frac{1}{s+3}\right).$$

We can now see that the function with this Laplace transform is given by

$$y(t) = \mathcal{L}^{-1}\left[\frac{1}{s+3} + \frac{1}{5}\left(\frac{1}{s-2} - \frac{1}{s+3}\right)\right] = e^{-3t} + \frac{1}{5}\left(e^{2t} - e^{-3t}\right)$$

works. Simplifying, we have the solution of the initial value problem

$$y(t) = \frac{1}{5}e^{2t} + \frac{4}{5}e^{-3t}.$$

We can verify that we have solved the initial value problem.

$$y' + 3y = \frac{2}{5}e^{2t} - \frac{12}{5}e^{-3t} + 3\left(\frac{1}{5}e^{2t} + \frac{4}{5}e^{-3t}\right) = e^{2t}$$

and $y(0) = \frac{1}{5} + \frac{4}{5} = 1$.

Example 5.28. *Solve the initial value problem* $y'' + 4y = 0$, $y(0) = 1$, $y'(0) = 3$.

We can probably solve this without Laplace transforms, but it is a simple exercise. Transforming the equation, we have

$$\begin{aligned} 0 &= s^2Y - sy(0) - y'(0) + 4Y \\ &= (s^2+4)Y - s - 3. \end{aligned} \quad (5.95)$$

Solving for Y*, we have*

$$Y(s) = \frac{s+3}{s^2+4}.$$

We now ask if we recognize the transform pair needed. The denominator looks like the type needed for the transform of a sine or cosine. We just need to play with the numerator. Splitting the expression into two terms, we have

$$Y(s) = \frac{s}{s^2+4} + \frac{3}{s^2+4}.$$

The first term is now recognizable as the transform of $\cos 2t$. *The second term is not the transform of* $\sin 2t$. *It would be if the numerator were a 2. This can be corrected by multiplying and dividing by 2:*

$$\frac{3}{s^2+4} = \frac{3}{2}\left(\frac{2}{s^2+4}\right).$$

The solution is then found as

$$y(t) = \mathcal{L}^{-1}\left[\frac{s}{s^2+4} + \frac{3}{2}\left(\frac{2}{s^2+4}\right)\right] = \cos 2t + \frac{3}{2}\sin 2t.$$

The reader can verify that this is the solution of the initial value problem.

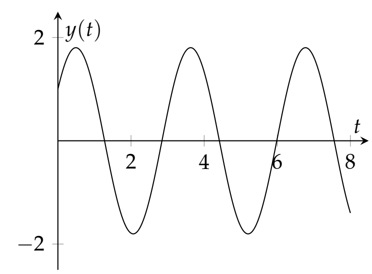


Figure 5.29: A plot of the solution to Example 5.28.

5.8.3 *Step and Impulse Functions*

OFTEN, THE INITIAL VALUE PROBLEMS THAT ONE FACES in differential equations courses can be solved using either the Method of Undetermined Coefficients or the Method of Variation of Parameters. However, using the latter can be messy and involves some skill with integration. Many circuit designs can be modeled with systems of differential equations using Kirchoff's Rules. Such systems can get fairly complicated. However, Laplace transforms can be used to solve such systems, and electrical engineers have long used such methods in circuit analysis.

In this section we add a couple more transform pairs and transform properties that are useful in accounting for things like turning on a driving force, using periodic functions like a square wave, or introducing impulse forces.

We first recall the Heaviside step function, given by

$$H(t) = \begin{cases} 0, & t < 0, \\ 1, & t > 0. \end{cases} \tag{5.96}$$

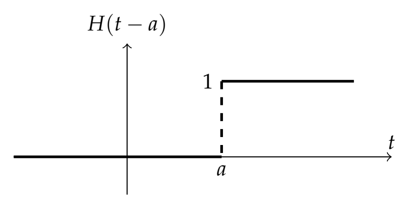


Figure 5.30: A shifted Heaviside function, $H(t-a)$.

A more general version of the step function is the horizontally shifted step function, $H(t-a)$. This function is shown in Figure 5.30. The Laplace transform of this function is found for $a > 0$ as

$$\begin{aligned} \mathcal{L}[H(t-a)] &= \int_0^\infty H(t-a)e^{-st}\,dt \\ &= \int_a^\infty e^{-st}\,dt \\ &= \left.\frac{e^{-st}}{s}\right|_a^\infty = \frac{e^{-as}}{s}. \end{aligned} \tag{5.97}$$

Just like the Fourier transform, the Laplace transform has two Shift Theorems involving the multiplication of the function, $f(t)$, or its transform, $F(s)$, by exponentials. The First and Second Shift Properties/Theorems are given by

The Shift Theorems.

$$\mathcal{L}[e^{at}f(t)] = F(s-a), \quad (5.98)$$
$$\mathcal{L}[f(t-a)H(t-a)] = e^{-as}F(s). \quad (5.99)$$

We prove the First Shift Theorem and leave the other proof as an exercise for the reader. Namely,

$$\begin{aligned} \mathcal{L}[e^{at}f(t)] &= \int_0^\infty e^{at}f(t)e^{-st}\,dt \\ &= \int_0^\infty f(t)e^{-(s-a)t}\,dt = F(s-a). \end{aligned} \quad (5.100)$$

Example 5.29. *Compute the Laplace transform of $e^{-at}\sin\omega t$.*

This function arises as the solution of the underdamped harmonic oscillator. We first note that the exponential multiplies a sine function. The First Shift Theorem tells us that we first need the transform of the sine function. So, for $f(t) = \sin\omega t$, we have

$$F(s) = \frac{\omega}{s^2+\omega^2}.$$

Using this transform, we can obtain the solution to this problem as

$$\mathcal{L}[e^{-at}\sin\omega t] = F(s+a) = \frac{\omega}{(s+a)^2+\omega^2}.$$

More interesting examples can be found using piecewise defined functions. First we consider the function $H(t) - H(t-a)$. For $t < 0$, both terms are zero. In the interval $[0,a]$, the function $H(t) = 1$ and $H(t-a) = 0$. Therefore, $H(t) - H(t-a) = 1$ for $t \in [0,a]$. Finally, for $t > a$, both functions are one and therefore the difference is zero. The graph of $H(t) - H(t-a)$ is shown in Figure 5.31.

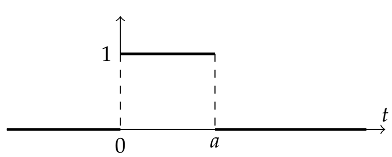


Figure 5.31: The box function, $H(t) - H(t-a)$.

We now consider the piecewise defined function:

$$g(t) = \begin{cases} f(t), & 0 \le t \le a, \\ 0, & t < 0, t > a. \end{cases}$$

This function can be rewritten in terms of step functions. We only need to multiply $f(t)$ by the above box function,

$$g(t) = f(t)[H(t) - H(t-a)].$$

We depict this in Figure 5.32.

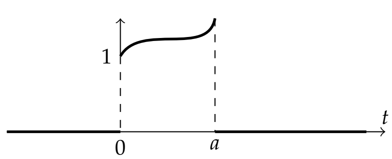


Figure 5.32: Formation of a piecewise function, $f(t)[H(t) - H(t-a)]$.

Even more complicated functions can be written in terms of step functions. We only need to look at sums of functions of the form $f(t)[H(t-a) - H(t-b)]$ for $b > a$. This is similar to a box function. It is nonzero between a and b and has height $f(t)$.

We show as an example the square wave function in Figure 5.33. It can be represented as a sum of an infinite number of boxes,

$$f(t) = \sum_{n=-\infty}^{\infty} [H(t-2na) - H(t-(2n+1)a)],$$

for $a > 0$.

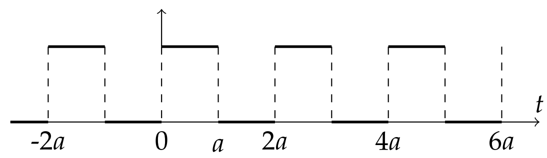


Figure 5.33: A square wave, $f(t) = \sum_{n=-\infty}^{\infty}[H(t-2na) - H(t-(2n+1)a)]$.

Example 5.30. *Find the Laplace Transform of a square wave "turned on" at $t = 0$.*

.

We let

$$f(t) = \sum_{n=0}^{\infty}[H(t-2na) - H(t-(2n+1)a)], \quad a > 0.$$

Using the properties of the Heaviside function, we have

$$\begin{aligned}
\mathcal{L}[f(t)] &= \sum_{n=0}^{\infty}\left[\mathcal{L}[H(t-2na)] - \mathcal{L}[H(t-(2n+1)a)]\right] \\
&= \sum_{n=0}^{\infty}\left[\frac{e^{-2nas}}{s} - \frac{e^{-(2n+1)as}}{s}\right] \\
&= \frac{1-e^{-as}}{s}\sum_{n=0}^{\infty}\left(e^{-2as}\right)^n \\
&= \frac{1-e^{-as}}{s}\left(\frac{1}{1-e^{-2as}}\right) \\
&= \frac{1-e^{-as}}{s(1-e^{-2as})}. \qquad (5.101)
\end{aligned}$$

Note that the third line in the derivation is a geometric series. We summed this series to get the answer in a compact form since $e^{-2as} < 1$.

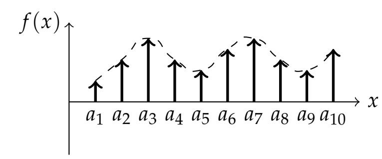


Figure 5.34: Plot representing impulse forces of height $f(a_i)$. The sum $\sum_{i=1}^{n} f(a_i)\delta(x-a_i)$ describes a general impulse function.

Other interesting examples are provided by the delta function. The Dirac delta function can be used to represent a unit impulse. Summing over a number of impulses, or point sources, we can describe a general function as shown in Figure 5.34. The sum of impulses located at points a_i, $i = 1, \ldots, n$, with strengths $f(a_i)$ would be given by

$$f(x) = \sum_{i=1}^{n} f(a_i)\delta(x-a_i).$$

A continuous sum could be written as

$$f(x) = \int_{-\infty}^{\infty} f(\xi)\delta(x-\xi)\, d\xi.$$

This is simply an application of the sifting property of the delta function. We will investigate a case when one would use a single impulse. While a mass on a spring is undergoing simple harmonic motion, we hit it for an instant at time $t = a$. In such a case, we could represent the force as a multiple of $\delta(t-a)$.

$\mathcal{L}[\delta(t-a)] = e^{-as}$.

One would then need the Laplace transform of the delta function to solve the associated initial value problem. Inserting the delta function into the

Laplace transform, we find that for $a > 0$,

$$\begin{aligned}\mathcal{L}[\delta(t-a)] &= \int_0^\infty \delta(t-a)e^{-st}\,dt \\ &= \int_{-\infty}^\infty \delta(t-a)e^{-st}\,dt \\ &= e^{-as}. \end{aligned} \quad (5.102)$$

Example 5.31. *Solve the initial value problem* $y'' + 4\pi^2 y = \delta(t-2)$, $y(0) = y'(0) = 0$.

This initial value problem models a spring oscillation with an impulse force. Without the forcing term, given by the delta function, this spring is initially at rest and not stretched. The delta function models a unit impulse at $t = 2$. *Of course, we anticipate that at this time the spring will begin to oscillate. We will solve this problem using Laplace transforms.*

First, we transform the differential equation:

$$s^2Y - sy(0) - y'(0) + 4\pi^2 Y = e^{-2s}.$$

Inserting the initial conditions, we have

$$(s^2 + 4\pi^2)Y = e^{-2s}.$$

Solving for $Y(s)$, *we obtain*

$$Y(s) = \frac{e^{-2s}}{s^2 + 4\pi^2}.$$

We now seek the function for which this is the Laplace transform. The form of this function is an exponential times some Laplace transform, $F(s)$. *Thus, we need the Second Shift Theorem since the solution is of the form* $Y(s) = e^{-2s}F(s)$ *for*

$$F(s) = \frac{1}{s^2 + 4\pi^2}.$$

We need to find the corresponding $f(t)$ *of the Laplace transform pair. The denominator in* $F(s)$ *suggests a sine or cosine. Since the numerator is constant, we pick sine. From the tables of transforms, we have*

$$\mathcal{L}[\sin 2\pi t] = \frac{2\pi}{s^2 + 4\pi^2}.$$

So, we write

$$F(s) = \frac{1}{2\pi}\frac{2\pi}{s^2 + 4\pi^2}.$$

This gives $f(t) = (2\pi)^{-1}\sin 2\pi t$.

We now apply the Second Shift Theorem, $\mathcal{L}[f(t-a)H(t-a)] = e^{-as}F(s)$, *or*

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}\left[e^{-2s}F(s)\right] \\ &= H(t-2)f(t-2) \\ &= \frac{1}{2\pi}H(t-2)\sin 2\pi(t-2). \end{aligned} \quad (5.103)$$

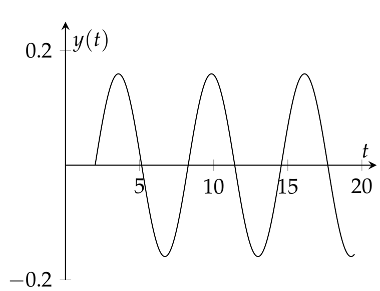


Figure 5.35: A plot of the solution to Example 5.31 in which a spring at rest experiences an impulse force at $t = 2$.

This solution tells us that the mass is at rest until $t = 2$ *and then begins to oscillate at its natural frequency. A plot of this solution is shown in Figure 5.35*

Example 5.32. *Solve the initial value problem*

$$y'' + y = f(t), \quad y(0) = 0, y'(0) = 0,$$

where

$$f(t) = \begin{cases} cos\pi t, & 0 \leq t \leq 2, \\ 0, & otherwise. \end{cases}$$

We need the Laplace transform of $f(t)$. This function can be written in terms of a Heaviside function, $f(t) = \cos \pi t H(t-2)$. In order to apply the Second Shift Theorem, we need a shifted version of the cosine function. We find the shifted version by noting that $\cos \pi(t-2) = \cos \pi t$. Thus, we have

$$\begin{aligned} f(t) &= \cos \pi t\,[H(t) - H(t-2)] \\ &= \cos \pi t - \cos \pi(t-2)H(t-2), \quad t \geq 0. \end{aligned} \tag{5.104}$$

The Laplace transform of this driving term is

$$F(s) = (1 - e^{-2s})\mathcal{L}[\cos \pi t] = (1 - e^{-2s})\frac{s}{s^2 + \pi^2}.$$

Now we can proceed to solve the initial value problem. The Laplace transform of the initial value problem yields

$$(s^2 + 1)Y(s) = (1 - e^{-2s})\frac{s}{s^2 + \pi^2}.$$

Therefore,

$$Y(s) = (1 - e^{-2s})\frac{s}{(s^2 + \pi^2)(s^2 + 1)}.$$

We can retrieve the solution to the initial value problem using the Second Shift Theorem. The solution is of the form $Y(s) = (1 - e^{-2s})G(s)$ for

$$G(s) = \frac{s}{(s^2 + \pi^2)(s^2 + 1)}.$$

Then, the final solution takes the form

$$y(t) = g(t) - g(t-2)H(t-2).$$

We only need to find $g(t)$ in order to finish the problem. This is easily done using the partial fraction decomposition

$$G(s) = \frac{s}{(s^2 + \pi^2)(s^2 + 1)} = \frac{1}{\pi^2 - 1}\left[\frac{s}{s^2 + 1} - \frac{s}{s^2 + \pi^2}\right].$$

Then,

$$g(t) = \mathcal{L}^{-1}\left[\frac{s}{(s^2 + \pi^2)(s^2 + 1)}\right] = \frac{1}{\pi^2 - 1}\left(\cos t - \cos \pi t\right).$$

The final solution is then given by

$$y(t) = \frac{1}{\pi^2 - 1}\left[\cos t - \cos \pi t - H(t-2)(\cos(t-2) - \cos \pi t)\right].$$

A plot of this solution is shown in Figure 5.36

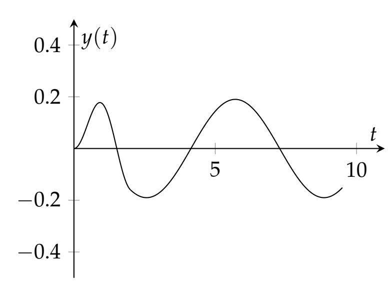


Figure 5.36: A plot of the solution to Example 5.32 in which a spring at rest experiences an piecewise defined force.

5.9 *The Convolution Theorem*

FINALLY, WE CONSIDER THE CONVOLUTION of two functions. Often, we are faced with having the product of two Laplace transforms that we know and we seek the inverse transform of the product. For example, let's say we have obtained $Y(s) = \frac{1}{(s-1)(s-2)}$ while trying to solve an initial value problem. In this case, we could find a partial fraction decomposition. But, there are other ways to find the inverse transform, especially if we cannot perform a partial fraction decomposition. We could use the Convolution Theorem for Laplace transforms or we could compute the inverse transform directly. We will look into these methods in the next two sections.We begin with defining the convolution.

We define the convolution of two functions defined on $[0, \infty)$ much the same way as we had done for the Fourier transform. The convolution $f * g$ is defined as

$$(f * g)(t) = \int_0^t f(u)g(t-u)\,du. \tag{5.105}$$

Note that the convolution integral has finite limits as opposed to the Fourier transform case.

The convolution operation has two important properties:

The convolution is commutative.

1. The convolution is commutative: $f * g = g * f$

   Proof. The key is to make a substitution $y = t - u$ in the integral. This makes f a simple function of the integration variable.

$$\begin{aligned} (g * f)(t) &= \int_0^t g(u)f(t-u)\,du \\ &= -\int_t^0 g(t-y)f(y)\,dy \\ &= \int_0^t f(y)g(t-y)\,dy \\ &= (f * g)(t). \end{aligned} \tag{5.106}$$

□

The Convolution Theorem for Laplace transforms.

2. The Convolution Theorem: The Laplace transform of a convolution is the product of the Laplace transforms of the individual functions:

$$\mathcal{L}[f * g] = F(s)G(s).$$

   Proof. Proving this theorem takes a bit more work. We will make some assumptions that will work in many cases. First, we assume that the functions are causal, $f(t) = 0$ and $g(t) = 0$ for $t < 0$. Second, we will assume that we can interchange integrals, which needs more rigorous attention than will be provided here. The first assumption will allow us to write the finite integral as an infinite integral. Then a change of variables will allow us to split the integral into the product of two integrals that are recognized as a product of two Laplace transforms.

Carrying out the computation, we have

$$\begin{aligned}
\mathcal{L}[f*g] &= \int_0^\infty \left(\int_0^t f(u)g(t-u)\,du \right) e^{-st}\,dt \\
&= \int_0^\infty \left(\int_0^\infty f(u)g(t-u)\,du \right) e^{-st}\,dt \\
&= \int_0^\infty f(u) \left(\int_0^\infty g(t-u)e^{-st}\,dt \right) du \qquad (5.107)
\end{aligned}$$

Now, make the substitution $\tau = t - u$. We note that

$$int_0^\infty f(u) \left(\int_0^\infty g(t-u)e^{-st}\,dt \right) du = \int_0^\infty f(u) \left(\int_{-u}^\infty g(\tau)e^{-s(\tau+u)}\,d\tau \right) du$$

However, since $g(\tau)$ is a causal function, we have that it vanishes for $\tau < 0$ and we can change the integration interval to $[0, \infty)$. So, after a little rearranging, we can proceed to the result.

$$\begin{aligned}
\mathcal{L}[f*g] &= \int_0^\infty f(u) \left(\int_0^\infty g(\tau)e^{-s(\tau+u)}\,d\tau \right) du \\
&= \int_0^\infty f(u)e^{-su} \left(\int_0^\infty g(\tau)e^{-s\tau}\,d\tau \right) du \\
&= \left(\int_0^\infty f(u)e^{-su}\,du \right) \left(\int_0^\infty g(\tau)e^{-s\tau}\,d\tau \right) \\
&= F(s)G(s). \qquad (5.108)
\end{aligned}$$

□

We make use of the Convolution Theorem to do the following examples.

Example 5.33. *Find* $y(t) = \mathcal{L}^{-1}\left[\frac{1}{(s-1)(s-2)}\right]$.

We note that this is a product of two functions:

$$Y(s) = \frac{1}{(s-1)(s-2)} = \frac{1}{s-1}\frac{1}{s-2} = F(s)G(s).$$

We know the inverse transforms of the factors: $f(t) = e^t$ *and* $g(t) = e^{2t}$.

Using the Convolution Theorem, we find $y(t) = (f*g)(t)$. *We compute the convolution:*

$$\begin{aligned}
y(t) &= \int_0^t f(u)g(t-u)\,du \\
&= \int_0^t e^u e^{2(t-u)}\,du \\
&= e^{2t} \int_0^t e^{-u}\,du \\
&= e^{2t}[-e^t + 1] = e^{2t} - e^t. \qquad (5.109)
\end{aligned}$$

One can also confirm this by carrying out a partial fraction decomposition.

Example 5.34. *Consider the initial value problem,* $y'' + 9y = 2\sin 3t$, $y(0) = 1$, $y'(0) = 0$.

The Laplace transform of this problem is given by

$$(s^2+9)Y - s = \frac{6}{s^2+9}.$$

Solving for $Y(s)$*, we obtain*

$$Y(s) = \frac{6}{(s^2+9)^2} + \frac{s}{s^2+9}.$$

The inverse Laplace transform of the second term is easily found as $\cos(3t)$*; however, the first term is more complicated.*

We can use the Convolution Theorem to find the Laplace transform of the first term. We note that

$$\frac{6}{(s^2+9)^2} = \frac{2}{3}\frac{3}{(s^2+9)}\frac{3}{(s^2+9)}$$

is a product of two Laplace transforms (up to the constant factor). Thus,

$$\mathcal{L}^{-1}\left[\frac{6}{(s^2+9)^2}\right] = \frac{2}{3}(f*g)(t),$$

where $f(t) = g(t) = sin3t$*. Evaluating this convolution product, we have*

$$\begin{aligned}
\mathcal{L}^{-1}\left[\frac{6}{(s^2+9)^2}\right] &= \frac{2}{3}(f*g)(t) \\
&= \frac{2}{3}\int_0^t \sin 3u \sin 3(t-u)\, du \\
&= \frac{1}{3}\int_0^t [\cos 3(2u-t) - \cos 3t]\, du \\
&= \frac{1}{3}\left[\frac{1}{6}\sin(6u-3t) - u\cos 3t\right]_0^t \\
&= \frac{1}{9}\sin 3t - \frac{1}{3}t\cos 3t. \qquad (5.110)
\end{aligned}$$

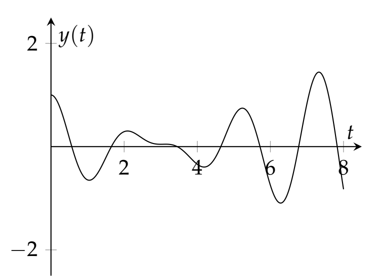


Figure 5.37: Plot of the solution to Example 5.34 showing a resonance.

Combining this with the inverse transform of the second term of $Y(s)$*, the solution to the initial value problem is*

$$y(t) = -\frac{1}{3}t\cos 3t + \frac{1}{9}\sin 3t + \cos 3t.$$

Note that the amplitude of the solution will grow in time from the first term. You can see this in Figure 5.37. This is known as a resonance.

Example 5.35. *Find* $\mathcal{L}^{-1}[\frac{6}{(s^2+9)^2}]$ *using partial fraction decomposition.*

If we look at Table 5.2, we see that the Laplace transform pairs with the denominator $(s^2+\omega^2)^2$ *are*

$$\mathcal{L}[t\sin\omega t] = \frac{2\omega s}{(s^2+\omega^2)^2},$$

and

$$\mathcal{L}[t\cos\omega t] = \frac{s^2-\omega^2}{(s^2+\omega^2)^2}.$$

So, we might consider rewriting a partial fraction decomposition as

$$\frac{6}{(s^2+9)^2} = \frac{A6s}{(s^2+9)^2} + \frac{B(s^2-9)}{(s^2+9)^2} + \frac{Cs+D}{s^2+9}.$$

Combining the terms on the right over a common denominator, we find

$$6 = 6As + B(s^2-9) + (Cs+D)(s^2+9).$$

Collecting like powers of s, we have

$$Cs^3 + (D+B)s^2 + 6As + (D-B) = 6.$$

Therefore, $C = 0$, $A = 0$, $D+B = 0$, and $D-B = \frac{2}{3}$. Solving the last two equations, we find $D = -B = \frac{1}{3}$.

Using these results, we find

$$\frac{6}{(s^2+9)^2} = -\frac{1}{3}\frac{(s^2-9)}{(s^2+9)^2} + \frac{1}{3}\frac{1}{s^2+9}.$$

This is the result we had obtained in the last example using the Convolution Theorem.

5.10 The Inverse Laplace Transform

Up to this point we have seen that the inverse Laplace transform can be found by making use of Laplace transform tables and properties of Laplace transforms. This is typically the way Laplace transforms are taught and used in a differential equations course. One can do the same for Fourier transforms. However, in the case of Fourier transforms, we introduced an inverse transform in the form of an integral. Does such an inverse integral transform exist for the Laplace transform? Yes, it does! In this section we will derive the inverse Laplace transform integral and show how it is used.

We begin by considering a causal function $f(t)$, which vanishes for $t < 0$, and define the function $g(t) = f(t)e^{-ct}$ with $c > 0$. For $g(t)$ absolutely integrable,

$$\int_{-\infty}^{\infty} |g(t)|\, dt = \int_0^{\infty} |f(t)|e^{-ct}\, dt < \infty,$$

A function $f(t)$ is said to be of exponential order if $\int_0^\infty |f(t)|e^{-ct}\, dt < \infty$

we can write the Fourier transform,

$$\hat{g}(\omega) = \int_{-\infty}^{\infty} g(t)e^{i\omega t} dt = \int_0^{\infty} f(t)e^{i\omega t - ct} dt$$

and the inverse Fourier transform,

$$g(t) = f(t)e^{-ct} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{g}(\omega)e^{-i\omega t}\, d\omega.$$

Multiplying by e^{ct} and inserting $\hat{g}(\omega)$ into the integral for $g(t)$, we find

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty} f(\tau)e^{(i\omega - c)\tau} d\tau e^{-(i\omega-c)t}\, d\omega.$$

Letting $s = c - i\omega$ (so $d\omega = ids$), we have

$$f(t) = \frac{i}{2\pi}\int_{c+i\infty}^{c-i\infty}\int_0^\infty f(\tau)e^{-s\tau}d\tau e^{st}\,ds.$$

Note that the inside integral is simply $F(s)$. So, we have

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(s)e^{st}\,ds. \qquad (5.111)$$

The integral in the last equation is the inverse Laplace transform, called the Bromwich Integral and is named after Thomas John I'Anson Bromwich (1875 - 1929) . This inverse transform is not usually covered in differential equations courses because the integration takes place in the complex plane. This integral is evaluated along a path in the complex plane called the Bromwich contour. The typical way to compute this integral is to first choose c so that all poles are to the left of the contour. This guarantees that $f(t)$ is of exponential type. The contour is a closed semicircle enclosing all the poles. One then relies on a generalization of Jordan's Lemma to the second and third quadrants.[6]

[6] Closing the contour to the left of the contour can be reasoned in a manner similar to what we saw in Jordan's Lemma. Write the exponential as $e^{st} = e^{(s_R+is_I)t} = e^{s_Rt}e^{is_It}$. The second factor is an oscillating factor and the growth in the exponential can only come from the first factor. In order for the exponential to decay as the radius of the semicircle grows, $s_Rt < 0$. Since $t > 0$, we need $s < 0$ which is done by closing the contour to the left. If $t < 0$, then the contour to the right would enclose no singularities and preserve the causality of $f(t)$.

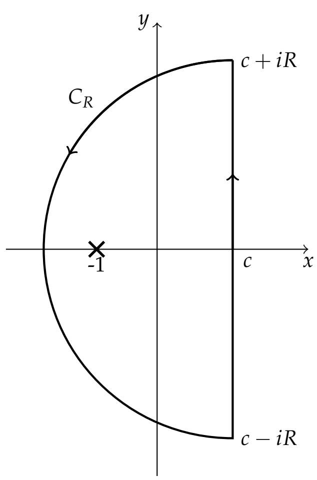


Figure 5.38: The contour used for applying the Bromwich integral to the Laplace transform $F(s) = \frac{1}{s(s+1)}$.

Example 5.36. *Find the inverse Laplace transform of* $F(s) = \frac{1}{s(s+1)}$.

The integral we have to compute is

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{st}}{s(s+1)}\,ds.$$

This integral has poles at $s = 0$ *and* $s = -1$*. The contour we will use is shown in Figure 5.38. We enclose the contour with a semicircle to the left of the path in the complex s-plane. One has to verify that the integral over the semicircle vanishes as the radius goes to infinity. Assuming that we have done this, then the result is simply obtained as* $2\pi i$ *times the sum of the residues. The residues in this case are*

$$Res\left[\frac{e^{zt}}{z(z+1)}; z = 0\right] = \lim_{z\to 0}\frac{e^{zt}}{(z+1)} = 1$$

and

$$Res\left[\frac{e^{zt}}{z(z+1)}; z = -1\right] = \lim_{z\to -1}\frac{e^{zt}}{z} = -e^{-t}.$$

Therefore, we have

$$f(t) = 2\pi i\left[\frac{1}{2\pi i}(1) + \frac{1}{2\pi i}(-e^{-t})\right] = 1 - e^{-t}.$$

We can verify this result using the Convolution Theorem or using a partial fraction decomposition. The latter method is simplest. We note that

$$\frac{1}{s(s+1)} = \frac{1}{s} - \frac{1}{s+1}.$$

The first term leads to an inverse transform of 1 *and the second term gives* e^{-t}*. So,*

$$\mathcal{L}^{-1}\left[\frac{1}{s} - \frac{1}{s+1}\right] = 1 - e^{-t}.$$

Thus, we have verified the result from doing contour integration.

Example 5.37. *Find the inverse Laplace transform of* $F(s) = \frac{1}{s(1+e^s)}$.

In this case, we need to compute

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{st}}{s(1+e^s)}\,ds.$$

This integral has poles at complex values of s such that $1 + e^s = 0$, *or* $e^s = -1$. *Letting* $s = x + iy$, *we see that*

$$e^s = e^{x+iy} = e^x(\cos y + i\sin y) = -1.$$

We see $x = 0$ *and y satisfies* $\cos y = -1$ *and* $\sin y = 0$. *Therefore,* $y = n\pi$ *for n an odd integer. Therefore, the integrand has an infinite number of simple poles at* $s = n\pi i$, $n = \pm 1, \pm 3, \ldots$. *It also has a simple pole at* $s = 0$.

In Figure 5.39, we indicate the poles. We need to compute the residues at each pole. At $s = n\pi i$, *we have*

$$\begin{aligned} Res\left[\frac{e^{st}}{s(1+e^s)}; s = n\pi i\right] &= \lim_{s\to n\pi i}(s - n\pi i)\frac{e^{st}}{s(1+e^s)} \\ &= \lim_{s\to n\pi i}\frac{e^{st}}{se^s} \\ &= -\frac{e^{n\pi i t}}{n\pi i}, \quad n \text{ odd}. \end{aligned} \tag{5.112}$$

At $s = 0$, *the residue is*

$$Res\left[\frac{e^{st}}{s(1+e^s)}; s = 0\right] = \lim_{s\to 0}\frac{e^{st}}{1+e^s} = \frac{1}{2}.$$

Summing the residues and noting the exponentials for $\pm n$ *can be combined to form sine functions, we arrive at the inverse transform.*

$$\begin{aligned} f(t) &= \frac{1}{2} - \sum_{n \text{ odd}} \frac{e^{n\pi i t}}{n\pi i} \\ &= \frac{1}{2} - 2\sum_{k=1}^{\infty}\frac{\sin(2k-1)\pi t}{(2k-1)\pi}. \end{aligned} \tag{5.113}$$

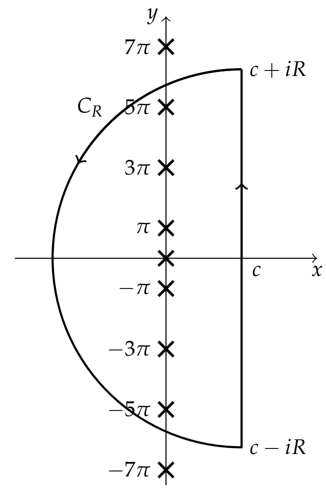


Figure 5.39: The contour used for applying the Bromwich integral to the Laplace transform $F(s) = \frac{1}{1+e^s}$.

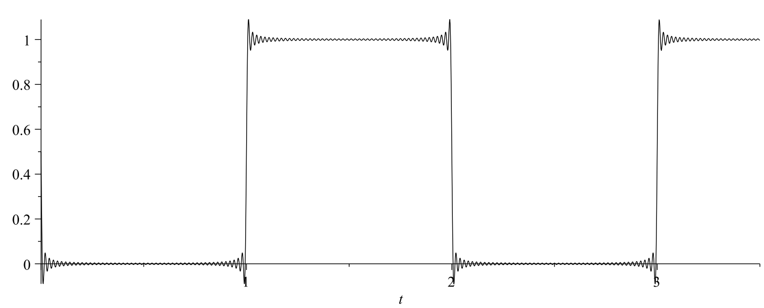


Figure 5.40: Plot of the square wave result as the inverse Laplace transform of $F(s) = \frac{1}{s(1+e^s)}$ with 50 terms.

The series in this example might look familiar. It is a Fourier sine series with odd harmonics whose amplitudes decay like $1/n$. *It is a vertically shifted square* *wave. In fact, we had computed the Laplace transform of a general square wave in Example 5.30.*

In that example, we found

$$\mathcal{L}\left[\sum_{n=0}^{\infty}[H(t-2na)-H(t-(2n+1)a)]\right] = \frac{1-e^{-as}}{s(1-e^{-2as})} = \frac{1}{s(1+e^{-as})}. \quad (5.114)$$

In this example, one can show that

$$f(t)=\sum_{n=0}^{\infty}[H(t-2n+1)-H(t-2n)].$$

The reader should verify that this result is indeed the square wave shown in Figure 5.40.

5.11 *Transforms and Partial Differential Equations*

As another application of the transforms, we will see that we can use transforms to solve some linear partial differential equations. We will first solve the one-dimensional heat equation and the two-dimensional Laplace equations using Fourier transforms. The transforms of the partial differential equations lead to ordinary differential equations which are easier to solve. The final solutions are then obtained using inverse transforms.

We could go further by applying a Fourier transform in space and a Laplace transform in time to convert the heat equation into an algebraic equation. We will also show that we can use a finite sine transform to solve nonhomogeneous problems on finite intervals. Along the way we will identify several Green's functions.

5.11.1 *Fourier Transform and the Heat Equation*

We will first consider the solution of the heat equation on an infinite interval using Fourier transforms. The basic scheme was discussed earlier and is outlined in Figure 5.41.

Consider the heat equation on the infinite line:

$$\begin{aligned} u_t = \alpha u_{xx}, \quad -\infty < x < \infty, \quad t > 0, \\ u(x,0) = f(x), \quad -\infty < x < \infty. \end{aligned} \quad (5.115)$$

We can Fourier transform the heat equation using the Fourier transform of $u(x,t)$,

$$\mathcal{F}[u(x,t)] = \hat{u}(k,t) = \int_{-\infty}^{\infty} u(x,t)e^{ikx}\,dx.$$

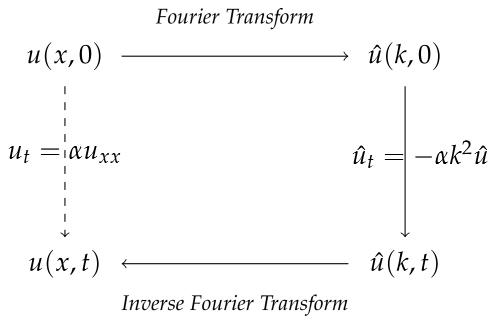


Figure 5.41: Using Fourier transforms to solve a linear partial differential equation.

We need to transform the derivatives in the equation. First we note that

$$\begin{aligned} \mathcal{F}[u_t] &= \int_{-\infty}^{\infty} \frac{\partial u(x,t)}{\partial t} e^{ikx}\, dx \\ &= \frac{\partial}{\partial t} \int_{-\infty}^{\infty} u(x,t) e^{ikx}\, dx \\ &= \frac{\partial \hat{u}(k,t)}{\partial t}. \end{aligned} \tag{5.116}$$

Assuming that $\lim_{|x|\to\infty} u(x,t) = 0$ and $\lim_{|x|\to\infty} u_x(x,t) = 0$, we also have that

$$\begin{aligned} \mathcal{F}[u_{xx}] &= \int_{-\infty}^{\infty} \frac{\partial^2 u(x,t)}{\partial x^2} e^{ikx}\, dx \\ &= -k^2 \hat{u}(k,t). \end{aligned} \tag{5.117}$$

Therefore, the heat equation becomes

The transformed heat equation.

$$\frac{\partial \hat{u}(k,t)}{\partial t} = -\alpha k^2 \hat{u}(k,t).$$

This is a first-order differential equation which is readily solved as

$$\hat{u}(k,t) = A(k) e^{-\alpha k^2 t},$$

where $A(k)$ is an arbitrary function of k. The inverse Fourier transform is

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{u}(k,t) e^{-ikx}\, dk. \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{A}(k) e^{-\alpha k^2 t} e^{-ikx}\, dk. \end{aligned} \tag{5.118}$$

We can determine $A(k)$ using the initial condition. Note that

$$\mathcal{F}[u(x,0)] = \hat{u}(k,0) = \int_{-\infty}^{\infty} f(x) e^{ikx}\, dx.$$

But we also have from the solution that

$$u(x,0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{A}(k) e^{-ikx}\, dk.$$

Comparing these two expressions for $\hat{u}(k,0)$, we see that

$$A(k) = \mathcal{F}[f(x)].$$

We note that $\hat{u}(k,t)$ is given by the product of two Fourier transforms, $\hat{u}(k,t) = A(k)e^{-\alpha k^2 t}$. So, by the Convolution Theorem, we expect that $u(x,t)$ is the convolution of the inverse transforms:

$$u(x,t) = (f * g)(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\xi,t)g(x-\xi,t)\,d\xi,$$

where

$$g(x,t) = \mathcal{F}^{-1}[e^{-\alpha k^2 t}].$$

In order to determine $g(x,t)$, we need only recall Example 5.5. In that example, we saw that the Fourier transform of a Gaussian is a Gaussian. Namely, we found that

$$\mathcal{F}[e^{-ax^2/2}] = \sqrt{\frac{2\pi}{a}}e^{-k^2/2a},$$

or,

$$\mathcal{F}^{-1}[\sqrt{\frac{2\pi}{a}}e^{-k^2/2a}] = e^{-ax^2/2}.$$

Applying this to the current problem, we have

$$g(x) = \mathcal{F}^{-1}[e^{-\alpha k^2 t}] = \sqrt{\frac{\pi}{\alpha t}}e^{-x^2/4t}.$$

Finally, we can write the solution to the problem:

$$u(x,t) = (f * g)(x,t) = \int_{-\infty}^{\infty} f(\xi,t)\frac{e^{-(x-\xi)^2/4t}}{\sqrt{4\pi\alpha t}}\,d\xi,$$

The function in the integrand,

$$K(x,t) = \frac{e^{-x^2/4t}}{\sqrt{4\pi\alpha t}},$$

$K(x,t)$ is called the heat kernel.

is called the heat kernel and acts as an initial value Green's function. The solution takes the form

$$u(x,t) = \int_{-\infty} \infty f(\xi,t)K(x,\xi;t)\,d\xi.$$

5.11.2 *Laplace's Equation on the Half Plane*

We consider a steady-state solution in two dimensions. In particular, we look for the steady-state solution, $u(x,y)$, satisfying the two-dimensional Laplace equation on a semi-infinite slab with given boundary conditions as shown in Figure 5.42. The boundary value problem is given as

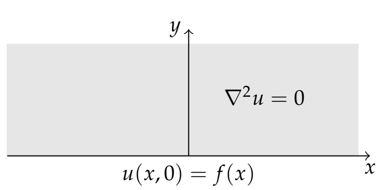


Figure 5.42: This is the domain for a semi-infinite slab with boundary value $u(x,0) = f(x)$ and governed by Laplace's equation.

$$\begin{gathered} u_{xx} + u_{yy} = 0, \quad -\infty < x < \infty, \quad y > 0, \\ u(x,0) = f(x), \quad -\infty < x < \infty, \\ \lim_{y\to\infty} u(x,y) = 0, \quad \lim_{|x|\to\infty} u(x,y) = 0. \end{gathered} \tag{5.119}$$

This problem can be solved using a Fourier transform of $u(x,y)$ with respect to x. The transform scheme for doing this is shown in Figure 5.43. We begin by defining the Fourier transform

$$\hat{u}(k,y) = \mathcal{F}[u] = \int_{-\infty}^{\infty} u(x,y)e^{ikx}\, dx.$$

We can transform Laplace's equation. We first note from the properties of Fourier transforms that

$$\mathcal{F}\left[\frac{\partial^2 u}{\partial x^2}\right] = -k^2 \hat{u}(k,y),$$

if $\lim_{|x|\to\infty} u(x,y) = 0$ and $\lim_{|x|\to\infty} u_x(x,y) = 0$. Also,

$$\mathcal{F}\left[\frac{\partial^2 u}{\partial y^2}\right] = \frac{\partial^2 \hat{u}(k,y)}{\partial y^2}.$$

Thus, the transform of Laplace's equation gives $\hat{u}_{yy} = k^2\hat{u}$.

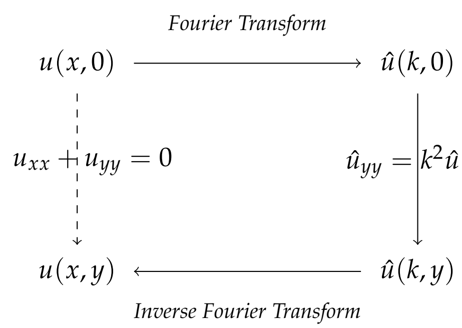


Figure 5.43: The transform scheme used to convert Laplace's equation to an ordinary differential equation, which is easier to solve.

This is a simple ordinary differential equation. We can solve this equation using the boundary conditions. The general solution is

The transformed Laplace equation.

$$\hat{u}(k,y) = a(k)e^{ky} + b(k)e^{-ky}.$$

Since $\lim_{y\to\infty} u(x,y) = 0$ and k can be positive or negative, we have that $\hat{u}(k,y) = a(k)e^{-|k|y}$. The coefficient $a(k)$ can be determined using the remaining boundary condition, $u(x,0) = f(x)$. We find that $a(k) = \hat{f}(k)$ since

$$a(k) = \hat{u}(k,0) = \int_{-\infty}^{\infty} u(x,0)e^{ikx}\, dx = \int_{-\infty}^{\infty} f(x)e^{ikx}\, dx = \hat{f}(k).$$

We have found that $\hat{u}(k,y) = \hat{f}(k)e^{-|k|y}$. We can obtain the solution using the inverse Fourier transform,

$$u(x,t) = \mathcal{F}^{-1}[\hat{f}(k)e^{-|k|y}].$$

We note that this is a product of Fourier transforms and use the Convolution Theorem for Fourier transforms. Namely, we have that $a(k) = \mathcal{F}[f]$ and $e^{-|k|y} = \mathcal{F}[g]$ for $g(x,y) = \frac{2y}{x^2+y^2}$. This last result is essentially proven in Problem 6.

Then, the Convolution Theorem gives the solution

$$\begin{aligned} u(x,y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\xi)g(x-\xi)\,d\xi \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\xi)\frac{2y}{(x-\xi)^2+y^2}\,d\xi. \end{aligned} \tag{5.120}$$

We note that this solution is in the form

$$u(x,y) = \int_{-\infty}^{\infty} f(\xi)G(x,\xi;y)\,d\xi,$$

where

$$G(x,\xi;y) = \frac{2y}{\pi((x-\xi)^2+y^2)}$$

is the Green's function for this problem.

The Green's function for the Laplace equation.

5.11.3 *Heat Equation on Infinite Interval, Revisited*

We next consider the initial value problem for the heat equation on an infinite interval,

$$\begin{aligned} u_t = u_{xx}, &\quad -\infty < x < \infty, \quad t > 0, \\ u(x,0) = f(x), &\quad -\infty < x < \infty. \end{aligned} \tag{5.121}$$

We can apply both a Fourier and a Laplace transform to convert this to an algebraic problem. The general solution will then be written in terms of an initial value Green's function as

$$u(x,t) = \int_{-\infty}^{\infty} G(x,t;\xi)f(\xi)\,d\xi.$$

For the time dependence, we can use the Laplace transform; and, for the spatial dependence, we use the Fourier transform. These combined transforms lead us to define

$$\hat{u}(k,s) = \mathcal{F}[\mathcal{L}[u]] = \int_{-\infty}^{\infty}\int_{0}^{\infty} u(x,t)e^{-st}e^{ikx}\,dtdx.$$

Applying this to the terms in the heat equation, we have

$$\begin{aligned} \mathcal{F}[\mathcal{L}[u_t]] &= s\hat{u}(k,s) - \mathcal{F}[u(x,0)] \\ &= s\hat{u}(k,s) - \hat{f}(k) \\ \mathcal{F}[\mathcal{L}[u_{xx}]] &= -k^2\hat{u}(k,s). \end{aligned} \tag{5.122}$$

Here we have assumed that

$$\lim_{t\to\infty} u(x,t)e^{-st} = 0, \quad \lim_{|x|\to\infty} u(x,t) = 0, \quad \lim_{|x|\to\infty} u_x(x,t) = 0.$$

Therefore, the heat equation can be turned into an algebraic equation for the transformed solution:

$$(s+k^2)\hat{u}(k,s) = \hat{f}(k),$$

tion.

or

$$\hat{u}(k,s) = \frac{\hat{f}(k)}{s+k^2}.$$

The solution to the heat equation is obtained using the inverse transforms for both the Fourier and Laplace transforms. Thus, we have

$$\begin{aligned} u(x,t) &= \mathcal{F}^{-1}[\mathcal{L}^{-1}[\hat{u}]] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{1}{2\pi i}\int_{c-i\infty}^{c+\infty}\frac{\hat{f}(k)}{s+k^2}e^{st}\,ds\right)e^{-ikx}\,dk. \qquad (5.123) \end{aligned}$$

Since the inside integral has a simple pole at $s=-k^2$, we can compute the Bromwich Integral by choosing $c > -k^2$. Thus,

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+\infty}\frac{\hat{f}(k)}{s+k^2}e^{st}\,ds = \text{Res}\left[\frac{\hat{f}(k)}{s+k^2}e^{st}; s=-k^2\right] = e^{-k^2t}\hat{f}(k).$$

Inserting this result into the solution, we have

$$\begin{aligned} u(x,t) &= \mathcal{F}^{-1}[\mathcal{L}^{-1}[\hat{u}]] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(k)e^{-k^2t}e^{-ikx}\,dk. \qquad (5.124) \end{aligned}$$

This solution is of the form

$$u(x,t) = \mathcal{F}^{-1}[\hat{f}\hat{g}]$$

for $\hat{g}(k) = e^{-k^2t}$. So, by the Convolution Theorem for Fourier transforms,the solution is a convolution:

$$u(x,t) = \int_{-\infty}^{\infty} f(\xi)g(x-\xi)\,d\xi.$$

All we need is the inverse transform of $\hat{g}(k)$.

We note that $\hat{g}(k) = e^{-k^2t}$ is a Gaussian. Since the Fourier transform of a Gaussian is a Gaussian, we need only recall Example 5.5:

$$\mathcal{F}[e^{-ax^2/2}] = \sqrt{\frac{2\pi}{a}}e^{-k^2/2a}.$$

Setting $a = 1/2t$, this becomes

$$\mathcal{F}[e^{-x^2/4t}] = \sqrt{4\pi t}e^{-k^2t}.$$

So,

$$g(x) = \mathcal{F}^{-1}[e^{-k^2t}] = \frac{e^{-x^2/4t}}{\sqrt{4\pi t}}.$$

Inserting $g(x)$ into the solution, we have

$$\begin{aligned} u(x,t) &= \frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{\infty} f(\xi)e^{-(x-\xi)^2/4t}\,d\xi \\ &= \int_{-\infty}^{\infty} G(x,t;\xi)f(\xi)\,d\xi. \qquad (5.125) \end{aligned}$$

Here we have identified the initial value Green's function

The initial value Green's function for the heat equation.

$$G(x,t;\xi) = \frac{1}{\sqrt{4\pi t}}e^{-(x-\xi)^2/4t}.$$

5.11.4 Nonhomogeneous Heat Equation

We now consider the nonhomogeneous heat equation with homogeneous boundary conditions defined on a finite interval.

$$\begin{aligned} u_t - ku_{xx} &= h(x,t), \quad 0 \le x \le L, \quad t > 0, \\ u(0,t) &= 0, \quad u(L,t) = 0, \quad t > 0, \\ u(x,0) &= f(x), \quad 0 \le x \le L. \end{aligned} \tag{5.126}$$

When $h(x,t) \equiv 0$, the general solution of the heat equation satisfying the above boundary conditions, $u(0,t) = 0$, $u(L,t) = 0$, for $t > 0$, can be written as a Fourier Sine Series:

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L}.$$

So, when $h(x,t) \neq 0$, we might assume that the solution takes the form

$$u(x,t) = \sum_{n=1}^{\infty} b_n(t) \sin \frac{n\pi x}{L}$$

where the b_n's are the Finite Fourier Sine Transform of the desired solution,

$$b_n(t) = \mathcal{F}_s[u] = \frac{2}{L} \int_0^L u(x,t) \sin \frac{n\pi x}{L}\, dx$$

Note that the Finite Fourier Sine Transform is essentially the Fourier Sine Series which we encountered in Section 2.4.

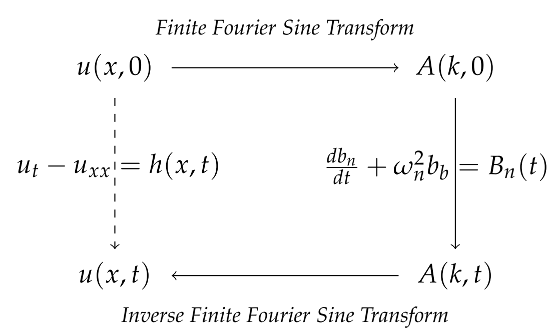


Figure 5.44: Using finite Fourier transforms to solve the heat equation by solving an ODE instead of a PDE.

The idea behind using the Finite Fourier Sine Transform is to solve the given heat equation by transforming the heat equation to a simpler equation for the transform, $b_n(t)$, solve for $bn(t)$, and then do an inverse transform, that is, insert the $b_n(t)$'s back into the series representation. This is depicted in Figure 5.44. Note that we had explored a similar diagram earlier when discussing the use of transforms to solve differential equations.

First, we need to transform the partial differential equation. The finite transform of the derivative terms are given by

$$\mathcal{F}_s[u_t] \quad = \quad \frac{2}{L} \int_0^L \frac{\partial u}{\partial t}(x,t) \sin \frac{n\pi x}{L}\, dx$$

$$\begin{aligned} &= \frac{d}{dt}\left(\frac{2}{L}\int_0^L u(x,t)\sin\frac{n\pi x}{L}\,dx\right) \\ &= \frac{db_n}{dt}. \end{aligned} \tag{5.127}$$

$$\begin{aligned} \mathcal{F}_s[u_{xx}] &= \frac{2}{L}\int_0^L \frac{\partial^2 u}{\partial x^2}(x,t)\sin\frac{n\pi x}{L}\,dx \\ &= \left[u_x \sin\frac{n\pi x}{L}\right]_0^L - \left(\frac{n\pi}{L}\right)\frac{2}{L}\int_0^L \frac{\partial u}{\partial x}(x,t)\cos\frac{n\pi x}{L}\,dx \\ &= -\left[\frac{n\pi}{L}u\cos\frac{n\pi x}{L}\right]_0^L - \left(\frac{n\pi}{L}\right)^2 \frac{2}{L}\int_0^L u(x,t)\sin\frac{n\pi x}{L}\,dx \\ &= \frac{n\pi}{L}[u(0,t) - u(L,0)\cos n\pi] - \left(\frac{n\pi}{L}\right)^2 b_n^2 \\ &= -\omega_n^2 b_n^2, \end{aligned} \tag{5.128}$$

where $\omega_n = \frac{n\pi}{L}$.

Furthermore, we define

$$H_n(t) = \mathcal{F}_s[h] = \frac{2}{L}\int_0^L h(x,t)\sin\frac{n\pi x}{L}\,dx.$$

Then, the heat equation is transformed to

$$\frac{db_n}{dt} + \omega_n^2 b_n = H_n(t), \quad n = 1,2,3,\ldots.$$

This is a simple linear first-order differential equation. We can supplement this equation with the initial condition

$$b_n(0) = \frac{2}{L}\int_0^L u(x,0)\sin\frac{n\pi x}{L}\,dx.$$

The differential equation for b_n is easily solved using the integrating factor, $\mu(t) = e^{\omega_n^2 t}$. Thus,

$$\frac{d}{dt}\left(e^{\omega_n^2 t} b_n(t)\right) = H_n(t)e^{\omega_n^2 t}$$

and the solution is

$$b_n(t) = b_n(0)e^{-\omega_n^2 t} + \int_0^t H_n(\tau)e^{-\omega_n^2(t-\tau)}\,d\tau.$$

The final step is to insert these coefficients (Finite Fourier Sine Transform) into the series expansion (inverse finite Fourier sine transform) for $u(x,t)$. The result is

$$u(x,t) = \sum_{n=1}^{\infty} b_n(0)e^{-\omega_n^2 t}\sin\frac{n\pi x}{L} + \sum_{n=1}^{\infty}\left[\int_0^t H_n(\tau)e^{-\omega_n^2(t-\tau)}\,d\tau\right]\sin\frac{n\pi x}{L}.$$

This solution can be written in a more compact form in order to identify the Green's function. We insert the expressions for $b_n(0)$ and $H_n(t)$ in terms of the initial profile and source term and interchange sums and integrals.

This leads to

$$\begin{aligned}
u(x,t) &= \sum_{n=1}^{\infty}\left(\frac{2}{L}\int_0^L u(\xi,0)\sin\frac{n\pi\xi}{L}\,d\xi\right)e^{-\omega_n^2 t}\sin\frac{n\pi x}{L} \\
&\quad + \sum_{n=1}^{\infty}\left[\int_0^t\left(\frac{2}{L}\int_0^L h(\xi,\tau)\sin\frac{n\pi\xi}{L}\,d\xi\right)e^{-\omega_n^2(t-\tau)}\,d\tau\right]\sin\frac{n\pi x}{L} \\
&= \int_0^L u(\xi,0)\left[\frac{2}{L}\sum_{n=1}^{\infty}\sin\frac{n\pi x}{L}\sin\frac{n\pi\xi}{L}e^{-\omega_n^2 t}\right]d\xi \\
&\quad + \int_0^t\int_0^L h(\xi,\tau)\left[\frac{2}{L}\sum_{n=1}^{\infty}\sin\frac{n\pi x}{L}\sin\frac{n\pi\xi}{L}e^{-\omega_n^2(t-\tau)}\right] \\
&= \int_0^L u(\xi,0)G(x,\xi;t,0)d\xi + \int_0^t\int_0^L h(\xi,\tau)G(x,\xi;t,\tau)\,d\xi d\tau.
\end{aligned} \tag{5.129}$$

Here we have defined the Green's function

$$G(x,\xi;t,\tau) = \frac{2}{L}\sum_{n=1}^{\infty}\sin\frac{n\pi x}{L}\sin\frac{n\pi\xi}{L}e^{-\omega_n^2(t-\tau)}.$$

We note that $G(x,\xi;t,0)$ gives the initial value Green's function.

Evaluating the Green's function at $t=\tau$, we have

$$G(x,\xi;t,t) = \frac{2}{L}\sum_{n=1}^{\infty}\sin\frac{n\pi x}{L}\sin\frac{n\pi\xi}{L}.$$

This is actually a series representation of the Dirac delta function. The Fourier Sine Transform of the delta function is

$$\mathcal{F}_s[\delta(x-\xi)] = \frac{2}{L}\int_0^L \delta(x-\xi)\sin\frac{n\pi x}{L}\,dx = \frac{2}{L}\sin\frac{n\pi\xi}{L}.$$

Then, the representation becomes

$$\delta(x-\xi) = \frac{2}{L}\sum_{n=1}^{\infty}\sin\frac{n\pi x}{L}\sin\frac{n\pi\xi}{L} = G(x,\xi;t,\tau).$$

Also, we note that

$$\begin{aligned}
\frac{\partial G}{\partial t} &= -\omega_n^2 G \\
\frac{\partial^2 G}{\partial x^2} &= -\left(\frac{n\pi}{L}\right)^2 G.
\end{aligned}$$

Therefore, $G_t = G_{xx}$, at least for $\tau \neq t$ and $\xi \neq x$.

We can modify this problem by adding nonhomogeneous boundary conditions.

$$\begin{aligned}
u_t - ku_{xx} &= h(x,t), \quad 0 \le x \le L, \quad t > 0, \\
u(0,t) &= A, \quad u(L,t) = B, \quad t > 0, \\
u(x,0) &= f(x), \quad 0 \le x \le L.
\end{aligned} \tag{5.130}$$

One way to treat these conditions is to assume $u(x,t) = w(x) + v(x,t)$ where $v_t - kv_{xx} = h(x,t)$ and $w_{xx} = 0$. Then, $u(x,t) = w(x) + v(x,t)$ satisfies the original nonhomogeneous heat equation.

If $v(x,t)$ satisfies $v(0,t) = v(L,t) = 0$ and $w(x)$ satisfies $w(0) = A$ and $w(L) = B$, then $u(0,t) = w(0) + v(0,t) = A$ $u(L,t) = w(L) + v(L,t) = B$

Finally, we note that

$$v(x,0) = u(x,0) - w(x) = f(x) - w(x).$$

Therefore, $u(x,t) = w(x) + v(x,t)$ satisfies the original problem if

$$\begin{aligned} v_t - kv_{xx} &= h(x,t), \quad 0 \le x \le L, \quad t > 0, \\ v(0,t) &= 0, \quad v(L,t) = 0, \quad t > 0, \\ v(x,0) &= f(x) - w(x), \quad 0 \le x \le L, \end{aligned} \tag{5.131}$$

and

$$\begin{aligned} w_{xx} &= 0, \quad 0 \le x \le L, \\ w(0) &= A, \quad w(L) = B. \end{aligned} \tag{5.132}$$

We can solve the last problem to obtain $w(x) = A + \frac{B-A}{L}x$. The solution to the problem for $v(x,t)$ is simply the problem we solved already in terms of Green's functions with the new initial condition, $f(x) = A - \frac{B-A}{L}x$.

Problems

1. In this problem you will show that the sequence of functions

$$f_n(x) = \frac{n}{\pi}\left(\frac{1}{1+n^2x^2}\right)$$

approaches $\delta(x)$ as $n \to \infty$. Use the following to support your argument:

a. Show that $\lim_{n\to\infty} f_n(x) = 0$ for $x \neq 0$.

b. Show that the area under each function is one.

2. Verify that the sequence of functions $\{f_n(x)\}_{n=1}^{\infty}$, defined by $f_n(x) = \frac{n}{2}e^{-n|x|}$, approaches a delta function.

3. Evaluate the following integrals:

a. $\int_0^{\pi} \sin x \delta\left(x - \frac{\pi}{2}\right)\, dx$.

b. $\int_{-\infty}^{\infty} \delta\left(\frac{x-5}{3}e^{2x}\right)\left(3x^2 - 7x + 2\right)\, dx$.

c. $\int_0^{\pi} x^2 \delta\left(x + \frac{\pi}{2}\right)\, dx$.

d. $\int_0^{\infty} e^{-2x}\delta(x^2 - 5x + 6)\, dx$. [See Problem 4.]

e. $\int_{-\infty}^{\infty}(x^2 - 2x + 3)\delta(x^2 - 9)\, dx$. [See Problem 4.]

4. For the case that a function has multiple roots, $f(x_i) = 0$, $i = 1, 2, \ldots$, it can be shown that

$$\delta(f(x)) = \sum_{i=1}^{n} \frac{\delta(x - x_i)}{|f'(x_i)|}.$$

Use this result to evaluate $\int_{-\infty}^{\infty} \delta(x^2 - 5x - 6)(3x^2 - 7x + 2)\, dx$.

5. Find a Fourier series representation of the Dirac delta function, $\delta(x)$, on $[-L, L]$.

6. For $a > 0$, find the Fourier transform, $\hat{f}(k)$, of $f(x) = e^{-a|x|}$.

7. Use the result from Problem 6 plus properties of the Fourier transform to find the Fourier transform, of $f(x) = x^2 e^{-a|x|}$ for $a > 0$.

8. Find the Fourier transform, $\hat{f}(k)$, of $f(x) = e^{-2x^2+x}$.

9. Prove the Second Shift Property in the form

$$F\left[e^{i\beta x} f(x)\right] = \hat{f}(k+\beta).$$

10. A damped harmonic oscillator is given by

$$f(t) = \begin{cases} Ae^{-\alpha t} e^{i\omega_0 t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

.

a. Find $\hat{f}(\omega)$ and

b. the frequency distribution $|\hat{f}(\omega)|^2$.

c. Sketch the frequency distribution.

11. Show that the convolution operation is associative: $(f * (g * h))(t) = ((f * g) * h)(t)$.

12. In this problem, you will directly compute the convolution of two Gaussian functions in two steps.

a. Use completing the square to evaluate

$$\int_{-\infty}^{\infty} e^{-\alpha t^2 + \beta t}\, dt.$$

b. Use the result from part a. to directly compute the convolution in Example 5.16:

$$(f * g)(x) = e^{-bx^2} \int_{-\infty}^{\infty} e^{-(a+b)t^2 + 2bxt}\, dt.$$

13. You will compute the (Fourier) convolution of two box functions of the same width. Recall, the box function is given by

$$f_a(x) = \begin{cases} 1, & |x| \leq a \\ 0, & |x| > a. \end{cases}$$

Consider $(f_a * f_a)(x)$ for different intervals of x. A few preliminary sketches will help. In Figure 5.45, the factors in the convolution integrand are show for one value of x. The integrand is the product of the first two functions. The convolution at x is the area of the overlap in the third figure. Think about how these pictures change as you vary x. Plot the resulting areas as a function of x. This is the graph of the desired convolution.

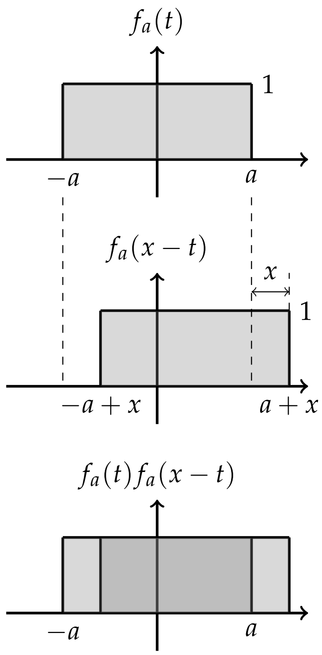


Figure 5.45: Sketch used to compute the convolution of the box function with itself. In the top figure is the box function. The middle figure shows the box shifted by x. The bottom figure indicates the overlap of the functions.

14. Define the integrals $I_n = \int_{-\infty}^{\infty} x^{2n} e^{-x^2}\, dx$. Noting that $I_0 = \sqrt{\pi}$,

a. Find a recursive relation between I_n and I_{n-1}.

b. Use this relation to determine I_1, I_2, and I_3.

c. Find an expression in terms of n for I_n.

15. Find the Laplace transform of the following functions:

a. $f(t) = 9t^2 - 7$.

b. $f(t) = e^{5t-3}$.

c. $f(t) = \cos 7t$.

d. $f(t) = e^{4t} \sin 2t$.

e. $f(t) = e^{2t}(t + \cosh t)$.

f. $f(t) = t^2 H(t-1)$.

g. $f(t) = \begin{cases} \sin t, & t < 4\pi, \\ \sin t + \cos t, & t > 4\pi. \end{cases}$

h. $f(t) = \int_0^t (t-u)^2 \sin u\, du$.

i. $f(t) = (t+5)^2 + te^{2t} \cos 3t$ and write the answer in the simplest form.

16. Find the inverse Laplace transform of the following functions using the properties of Laplace transforms and the table of Laplace transform pairs.

a. $F(s) = \dfrac{18}{s^3} + \dfrac{7}{s}$.

b. $F(s) = \dfrac{1}{s-5} - \dfrac{2}{s^2+4}$.

c. $F(s) = \dfrac{s+1}{s^2+1}$.

d. $F(s) = \dfrac{3}{s^2+2s+2}$.

e. $F(s) = \dfrac{1}{(s-1)^2}$.

f. $F(s) = \dfrac{e^{-3s}}{s^2-1}$.

g. $F(s) = \dfrac{1}{s^2+4s-5}$.

h. $F(s) = \dfrac{s+3}{s^2+8s+17}$.

17. Compute the convolution $(f * g)(t)$ (in the Laplace transform sense) and its corresponding Laplace transform $\mathcal{L}[f * g]$ for the following functions:

a. $f(t) = t^2$, $g(t) = t^3$.

b. $f(t) = t^2$, $g(t) = \cos 2t$.

c. $f(t) = 3t^2 - 2t + 1$, $g(t) = e^{-3t}$.

d. $f(t) = \delta\left(t - \frac{\pi}{4}\right)$, $g(t) = \sin 5t$.

18. For the following problems, draw the given function and find the Laplace transform in closed form.

a. $f(t) = 1 + \sum_{n=1}^{\infty}(-1)^n H(t-n)$.

b. $f(t) = \sum_{n=0}^{\infty}[H(t-2n+1) - H(t-2n)]$.

c. $f(t) = \sum_{n=0}^{\infty}(t-2n)[H(t-2n) - H(t-2n-1)] + (2n+2-t)[H(t-2n-1) - H(t-2n-2)]$.

19. Use the Convolution Theorem to compute the inverse transform of the following:

a. $F(s) = \dfrac{2}{s^2(s^2+1)}$.

b. $F(s) = \dfrac{e^{-3s}}{s^2}$.

c. $F(s) = \dfrac{1}{s(s^2+2s+5)}$.

20. Find the inverse Laplace transform in two different ways: (i) Use tables. (ii) Use the Bromwich Integral.

a. $F(s) = \dfrac{1}{s^3(s+4)^2}$.

b. $F(s) = \dfrac{1}{s^2-4s-5}$.

c. $F(s) = \dfrac{s+3}{s^2+8s+17}$.

d. $F(s) = \dfrac{s+1}{(s-2)^2(s+4)}$.

e. $F(s) = \dfrac{s^2+8s-3}{(s^2+2s+1)(s^2+1)}$.

21. Use Laplace transforms to solve the following initial value problems. Where possible, describe the solution behavior in terms of oscillation and decay.

a. $y'' - 5y' + 6y = 0$, $y(0) = 2$, $y'(0) = 0$.
b. $y'' - y = te^{2t}$, $y(0) = 0$, $y'(0) = 1$.
c. $y'' + 4y = \delta(t-1)$, $y(0) = 3$, $y'(0) = 0$.
d. $y'' + 6y' + 18y = 2H(\pi - t)$, $y(0) = 0$, $y'(0) = 0$.

22. Use Laplace transforms to convert the following system of differential equations into an algebraic system and find the solution of the differential equations.

$$\begin{aligned} x'' &= 3x - 6y, \quad x(0) = 1, \quad x'(0) = 0, \\ y'' &= x + y, \quad y(0) = 0, \quad y'(0) = 0. \end{aligned}$$

23. Use Laplace transforms to convert the following nonhomogeneous systems of differential equations into an algebraic system and find the solutions of the differential equations.

a.

$$\begin{aligned} x' &= 2x + 3y + 2\sin 2t, \quad x(0) = 1, \\ y' &= -3x + 2y, \quad y(0) = 0. \end{aligned}$$

b.

$$\begin{aligned} x' &= -4x - y + e^{-t}, \quad x(0) = 2, \\ y' &= x - 2y + 2e^{-3t}, \quad y(0) = -1. \end{aligned}$$

c.

$$\begin{aligned} x' &= x - y + 2\cos t, \quad x(0) = 3, \\ y' &= x + y - 3\sin t, \quad y(0) = 2. \end{aligned}$$

24. Use Laplace transforms to sum the following series:

a. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{1+2n}$.
b. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+3)}$.
c. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n(n+3)}$.
d. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n^2-a^2}$.

e. $\displaystyle\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2 - a^2}$.

f. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n} e^{-an}$.

25. Use Laplace transforms to prove

$$\sum_{n=1}^{\infty} \frac{1}{(n+a)(n+b)} = \frac{1}{b-a} \int_0^1 \frac{u^a - u^b}{1-u}\, du.$$

Use this result to evaluate the following sums:

a. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$.

b. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(n+2)(n+3)}$.

26. Do the following:

a. Find the first four nonvanishing terms of the Maclaurin series expansion of $f(x) = \dfrac{x}{e^x - 1}$.

b. Use the result in part a. to determine the first four nonvanishing Bernoulli numbers, B_n.

c. Use these results to compute $\zeta(2n)$ for $n = 1, 2, 3, 4$.

27. Given the following Laplace transforms, $F(s)$, find the function $f(t)$. Note that in each case there are an infinite number of poles, resulting in an infinite series representation.

a. $F(s) = \dfrac{1}{s^2(1+e^{-s})}$.

b. $F(s) = \dfrac{1}{s \sinh s}$.

c. $F(s) = \dfrac{\sinh s}{s^2 \cosh s}$.

d. $F(s) = \dfrac{\sinh(\beta\sqrt{s}x)}{s \sinh(\beta\sqrt{s}L)}$.

28. Consider the initial boundary value problem for the heat equation:

$$\begin{aligned}
u_t &= 2u_{xx}, & 0 < t, \quad 0 \le x \le 1, \\
u(x,0) &= x(1-x), & 0 < x < 1, \\
u(0,t) &= 0, & t > 0, \\
u(1,t) &= 0, & t > 0.
\end{aligned}$$

Use the finite transform method to solve this problem. Namely, assume that the solution takes the form $u(x,t) = \sum_{n=1}^{\infty} b_n(t) \sin n\pi x$ and obtain an ordinary differential equation for b_n and solve for the b_n's for each n.

6 From Analog to Discrete Signals

You don't have to be a mathematician to have a feel for numbers. - John Forbes Nash (1928 - 2015)

As you may recall, a goal of this course has been to introduce some of the tools of applied mathematics with an underlying theme of finding the connection between analog and discrete signals. We began our studies with Fourier series, which provided representations of periodic functions. We then moved on to the study of Fourier transforms, which represent functions defined over all space. Such functions can be used to describe analog signals. However, we cannot record and store analog signals. There is an uncountable amount of information in analog signals. We record a signal for a finite amount of time and even then we can only store samples of the signal over that time interval. The resulting signal is discrete. These discrete signals are represented using the Discrete Fourier Transform (DFT). In this chapter we will discuss the general steps of relating analog, periodic and discrete signals. Then we will go further into the properties of the Discrete Fourier Transform (DFT) and in the next chapter we will apply what we have learned to the study of real signals.

6.1 Analog to Periodic Signals

We begin by considering a typical analog signal and its Fourier transform as shown in Figure 6.1. Analog signals can be described as piecewise continuous functions defined over infinite time intervals. The resulting Fourier transforms are also piecewise continuous and defined over infinite intervals of frequencies. We represent an analog signal, $f(t)$, and its transform, $\hat{f}(\omega)$, using the inverse Fourier transform and the Fourier transform, respectively:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{-i\omega t}\,d\omega, \tag{6.1}$$

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}\,dt. \tag{6.2}$$

Note that the figures in this section are drawn as if the transform is real-valued. (See Figure 6.1.) However, in general they are not and we will investigate how this can be handled in the next chapter.

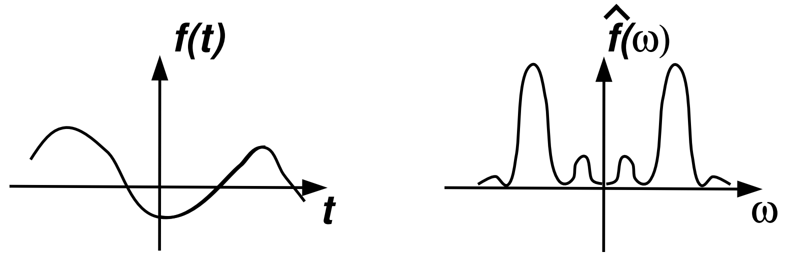


Figure 6.1: Plot of an analog signal $f(t)$ and its Fourier transform $\hat{f}(\omega)$.

Real signals cannot be studied on an infinite interval. Such signals are captured as data over a finite time interval. Let's assume that the recording starts at $t = 0$. Then, the record interval will be written as $[0, T]$, where T is called the *record length*.

The natural representation of a function, $f(t)$, for $t \in [0, T]$, is obtained by extending the signal to a periodic signal knowing that the physical signal is only defined on $[0, T]$. Recall that periodic functions can be modeled by a Fourier exponential series. We will denote the periodic extension of $f(t)$ by $f_p(t)$. The Fourier series representation of f_p and its Fourier coefficients are given by

Recall from Chapter 6 that we defined Fourier exponential series for intervals of length $2L$. It is easy to map that series expansion to one of length T, resulting in the representation used here.

$$\begin{aligned} f_p(t) &= \sum_{n=-\infty}^{\infty} c_n e^{-i\omega_n t}, \\ c_n &= \frac{1}{T}\int_0^T f_p(t) e^{i\omega_n t}\, dt. \end{aligned} \tag{6.3}$$

Here we have defined the discrete angular frequency $\omega_n = \frac{2\pi n}{T}$. The associated frequency is then $\nu_n = \frac{n}{T}$.

Here we begin with a signal $f(t)$ defined on $[0, T]$ and obtain the Fourier series representation (6.3) which gives the periodic extension of this function, $f_p(t)$. The frequencies are discrete as shown in Figure 6.2. We will determine the Fourier transform as

$$\hat{f}_p(\omega) = \frac{2\pi}{T}\hat{f}(\omega)\text{comb}_{\frac{2\pi}{T}}(\omega)$$

and conclude that f_p is the convolution of the signal f with a comb function.

Given that $f_p(t)$ is a periodic function, we would like to relate its Fourier series representation to the Fourier transform, $\hat{f}_p(\omega)$, of the corresponding signal $f_p(t)$. This is done by simply computing its Fourier transform:

$$\begin{aligned} \hat{f}_p(\omega) &= \int_{-\infty}^{\infty} f_p(t) e^{i\omega t}\, dt \\ &= \int_{-\infty}^{\infty} \left(\sum_{n=-\infty}^{\infty} c_n e^{-i\omega_n t} \right) e^{i\omega t}\, dt \\ &= \sum_{n=-\infty}^{\infty} c_n \int_{-\infty}^{\infty} e^{i(\omega-\omega_n)t}\, dt. \end{aligned} \tag{6.4}$$

Recalling from Equation (5.47) that

$$\int_{-\infty}^{\infty} e^{i\omega x}\, dx = 2\pi\delta(\omega),$$

we obtain the Fourier transform of $f_p(t)$:

$$\hat{f}_p(\omega) = \sum_{n=-\infty}^{\infty} 2\pi c_n \delta(\omega - \omega_n). \tag{6.5}$$

Thus, the Fourier transform of a periodic function is a series of spikes at discrete frequencies $\omega_n = \frac{2\pi n}{T}$ of strength $2\pi c_n$. This is represented in Figure 6.2. Note that the spikes are of finite height representing the factor $2\pi c_n$.

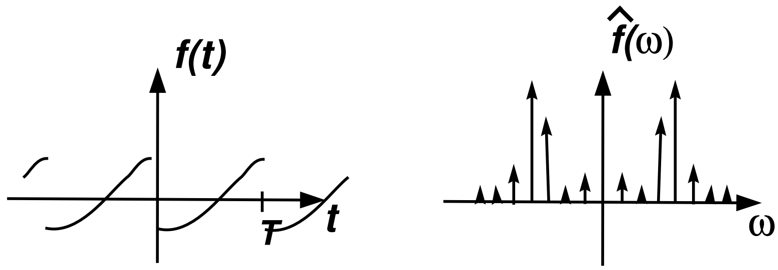


Figure 6.2: A periodic signal contains discrete frequencies $\omega_n = \frac{2\pi n}{T}$.

A simple example displaying this behavior is a signal with a single frequency, $f(t) = \cos\omega_0 t$. Restricting this function to a finite interval $[0, T]$, one obtains a version of the finite wave train, which was first introduced in Example 5.10 of the last chapter.

Example 6.1. *Find the real part of the Fourier transform of the finite wave train $f(t) = \cos\omega_0 t$, $t \in [0, T]$.*

Computing the real part of the Fourier transform, we find

$$\begin{aligned}
Re\hat{f}(\omega) &= Re\int_{-\infty}^{\infty} f(t)e^{i\omega t}\,dt \\
&= Re\int_0^T \cos\omega_0 t e^{i\omega t}\,dt \\
&= \int_0^T \cos\omega_0 t \cos\omega t\,dt \\
&= \frac{1}{2}\int_0^T \left[\cos((\omega+\omega_0)t) + \cos((\omega-\omega_0)t)\right]\,dt \\
&= \frac{1}{2}\left[\frac{\sin((\omega+\omega_0)T)}{\omega+\omega_0} + \frac{\sin((\omega-\omega_0)T)}{\omega-\omega_0}\right] \\
&= \frac{T}{2}\left[\text{sinc}\,((\omega+\omega_0)T) + \text{sinc}\,((\omega-\omega_0)T)\right]. \qquad (6.6)
\end{aligned}$$

Thus, the real part of the Fourier transform of this finite wave train consists of the sum of two sinc functions.

Two examples are provided in Figures 6.3 and 6.4. In both cases we consider $f(t) = \cos 2\pi t$, but with $t \in [0, 5]$ or $t \in [0, 2]$. The corresponding Fourier transforms are also provided. We first see the sum of the two sinc functions in each case. Furthermore, the main peaks, centered at $\omega = \pm\omega_0 = \pm 2\pi$, are better defined for $T = 5$ than for $T = 2$. This indicates that a larger record length will provide a better frequency resolution.

Next, we determine the relationship between the Fourier transform of $f_p(t)$ and the Fourier coefficients. Namely, evaluating $\hat{f}_p(\omega)$ at $\omega = \omega_k$, we have

$$\hat{f}_p(\omega_k) = \sum_{n=-\infty}^{\infty} 2\pi c_n \delta(\omega_k - \omega_n) = 2\pi c_k.$$

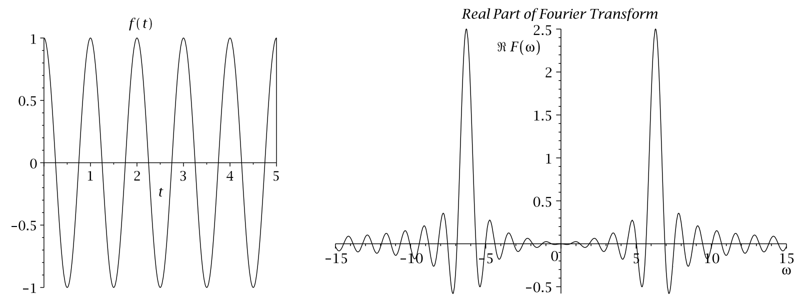


Figure 6.3: Plot of $f(t) = \cos 2\pi t$, $t \in [0,5]$ and the real and imaginary part of the Fourier transform.

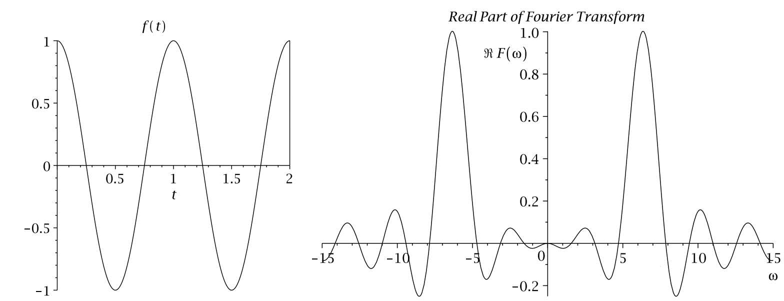


Figure 6.4: Plot of $f(t) = \cos 2\pi t$, $t \in [0,2]$ and the real and imaginary part of the Fourier transform.

Therefore, we can write the Fourier transform as

$$\hat{f}_p(\omega) = \sum_{n=-\infty}^{\infty} \hat{f}_p(\omega_n)\delta(\omega - \omega_n).$$

Further manipulation yields

$$\begin{aligned} \hat{f}_p(\omega) &= \sum_{n=-\infty}^{\infty} \hat{f}_p(\omega_n)\delta(\omega - \omega_n) \\ &= \sum_{n=-\infty}^{\infty} \hat{f}_p(\omega)\delta(\omega - \omega_n) \\ &= \hat{f}_p(\omega) \sum_{n=-\infty}^{\infty} \delta(\omega - \frac{2\pi n}{T}). \end{aligned} \tag{6.7}$$

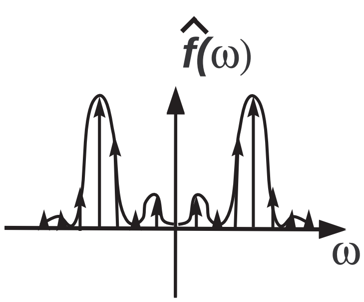


Figure 6.5: The discrete spectrum obtained from the Fourier transform of the periodic extension of $f(t)$, $t \in [0,T]$ is superimposed on the continuous spectrum of the analog signal.

This shows that the Fourier transform of a periodic signal is the product of the continuous function and a sum of δ function spikes of strength $\hat{f}_p(\omega_n)$ at frequencies $\omega_n = \frac{2\pi n}{T}$ as was shown in Figure 6.2. In Figure 6.5 the discrete spectrum is superimposed on the continuous spectrum emphasizing this connection.

The sum in the last line of Equation (6.7) is a special function, the Dirac comb function,

$$\text{comb}_{\frac{2\pi}{T}}(\omega) = \sum_{n=-\infty}^{\infty} \delta(\omega - \frac{2\pi n}{T}),$$

which we discuss further in the next section.

6.2 The Dirac Comb Function

A FUNCTION THAT OFTEN OCCURS in signal analysis is the *Dirac comb function* defined by

$$\text{comb}_a(t) = \sum_{n=-\infty}^{\infty} \delta(t - na). \tag{6.8}$$

This function is simply a set of translated delta function spikes as shown in Figure 6.6. It is also called an *impulse train* or a *sampling function.* It is a periodic distribution and care needs to be taken in studying the properties of this function. We will show how the comb function can be used to relate analog and finite length signals.

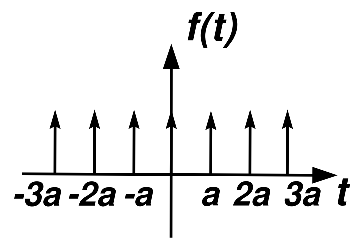


Figure 6.6: The Dirac comb function, $\text{comb}_a(t) = \sum_{n=-\infty}^{\infty} \delta(t - na)$.

In some fields, the comb function is written using the Cyrillic uppercase *Sha* function,

$$Ш(x) = \sum_{n=-\infty}^{\infty} \delta(x - n).$$

This is just a comb function with unit spacing, $a = 1$.

Employing the properties of the Dirac delta function, we can derive several properties of the Shah function. First, we have dilation and translation properties:

$$\begin{aligned} Ш(ax) &= \sum_{n=-\infty}^{\infty} \delta(ax - n) \\ &= \frac{1}{|a|} \sum_{n=-\infty}^{\infty} \delta(x - \frac{n}{a}) \\ &= \frac{1}{|a|} \text{comb}_{\frac{1}{a}}(x). & (6.9) \\ Ш(x + k) &= Ш(x), \quad k \text{ an integer}. & (6.10) \end{aligned}$$

Also, we have the *Sampling Property*,

$$Ш(x)f(x) = \sum_{n=-\infty}^{\infty} f(n)\delta(x - n) \tag{6.11}$$

and the *Replicating Property*

$$(Ш * f)(x) = \sum_{n=-\infty}^{\infty} f(x - n). \tag{6.12}$$

These properties are easily confirmed. The Sampling Property is shown by

$$\begin{aligned} Ш(x)f(x) &= \sum_{n=-\infty}^{\infty} \delta(x - n)f(x) \\ &= \sum_{n=-\infty}^{\infty} f(n)\delta(x - n), & (6.13) \end{aligned}$$

while the Replicating Property is verified by a simple computation:

$$(Ш * f)(x) = \int_{-\infty}^{\infty} Ш(\xi)f(x - \xi)\, d\xi$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left(\sum_{n=-\infty}^{\infty} \delta(\xi - n) \right) f(x-\xi)\, d\xi \\
&= \sum_{n=-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \delta(\xi - n) f(x-\xi)\, d\xi \right) \\
&= \sum_{n=-\infty}^{\infty} f(x-n). \qquad (6.14)
\end{aligned}$$

Thus, the convolution of a function with the Shah function results in a sum of translations of the function.

The comb function inherits these properties, since it can be written using the Shah function,

$$\begin{aligned}
\text{comb}_T(t) &= \sum_{n=-\infty}^{\infty} \delta(t - nT) \\
&= \sum_{n=-\infty}^{\infty} \delta(T(\frac{t}{T} - n)) \\
&= \sum_{n=-\infty}^{\infty} \frac{1}{|T|} \delta(\frac{t}{T} - n) \\
&= \frac{1}{|T|} \text{Ш} \left(\frac{t}{T} \right). \qquad (6.15)
\end{aligned}$$

In the following we will only use the comb function.

We further note that $\text{comb}_T(t)$ has period T :

$$\begin{aligned}
\text{comb}_T(t+T) &= \sum_{n=-\infty}^{\infty} \delta(t + T - nT) \\
&= \sum_{n=-\infty}^{\infty} \delta(t - nT) = \text{comb}_T(t). \qquad (6.16)
\end{aligned}$$

Useful results from this discussion are the exponential Fourier series representation,

$$\text{comb}_T(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{-2\pi int/T},$$

and the sum of exponentials

$$\sum_{n=-\infty}^{\infty} e^{-inat} = \frac{2\pi}{a} \text{comb}_{\frac{2\pi}{a}}(t).$$

Thus, the comb function has a Fourier series representation on $[0, T]$.

Example 6.2. *Find the Fourier exponential series representation of* $\text{comb}_T(t)$.

Since $\text{comb}_T(t)$ *has period* T, *the Fourier series representation takes the form*

$$\text{comb}_T(t) = \sum_{n=-\infty}^{\infty} c_n e^{-i\omega_n t},$$

where $\omega_n = \frac{2\pi n}{T}$.

We can easily compute the Fourier coefficients:

$$\begin{aligned}
c_n &= \frac{1}{T} \int_0^T \text{comb}_T(t) e^{i\omega_n t}\, dt \\
&= \frac{1}{T} \int_0^T \sum_{k=-\infty}^{\infty} \delta(t - kT) e^{i\omega_n t}\, dt \\
&= \frac{1}{T} e^0 = \frac{1}{T}. \qquad (6.17)
\end{aligned}$$

Note, the sum collapses because the only δ *function contribution comes from the* $k = 0$ *term (since* $kT \in [0, T]$ *only for* $k = 0$*). As a result, we have obtained the Fourier series representation of the comb function,*

$$\text{comb}_T(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{-2\pi int/T}.$$

From this result we note that

$$\begin{aligned}\sum_{n=-\infty}^{\infty} e^{inat} &= \sum_{n=-\infty}^{\infty} e^{2\pi int/(2\pi/a)} \\ &= \frac{2\pi}{a}\text{comb}_{\frac{2\pi}{a}}(t). \qquad (6.18)\end{aligned}$$

Example 6.3. *Find the Fourier transform of* $\text{comb}_a(t)$.

We compute the Fourier transform of the comb function directly:

$$\begin{aligned}F[\text{comb}_a(t)] &= \int_{-\infty}^{\infty} \text{comb}_a(t)e^{i\omega t}\,dt \\ &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(t-na)e^{i\omega t}\,dt \\ &= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t-na)e^{i\omega t}\,dt \\ &= \sum_{n=-\infty}^{\infty} e^{i\omega na} \\ &= \frac{2\pi}{a}\text{comb}_{\frac{2\pi}{a}}(\omega). \qquad (6.19)\end{aligned}$$

The Fourier transform of a comb function is a comb function.

$$F[\text{comb}_a(t)] = \frac{2\pi}{a}\text{comb}_{\frac{2\pi}{a}}(\omega).$$

Example 6.4. *Show that the convolution of a function with* $\text{comb}_T(x)$ *is a periodic function with period T. Since* $\text{comb}_T(t)$ *is a periodic function with period* T*, we have*

$$\begin{aligned}(f*\text{comb}_T)(t+T) &= \int_{-\infty}^{\infty} f(\tau)\text{comb}_T(t+T-\tau)\,d\tau \\ &= \int_{-\infty}^{\infty} f(\tau)\text{comb}_T(t-\tau)\,d\tau \\ &= (f*\text{comb}_T)(t). \qquad (6.20)\end{aligned}$$

Next, we will show that the periodic function $f_p(t)$ is nothing but a (Fourier) convolution of the analog function and the comb function. We first show that $(f*\text{comb}_a)(t) = \sum_{n=-\infty}^{\infty} f(t-na)$.

Example 6.5. *Evaluate the convolution* $(f*\text{comb}_a)(t)$ *directly.*

We carry out a direct computation of the convolution integral. We do this by first considering the convolution of a function $f(t)$ *with a shifted Dirac delta function,* $\delta_a(t) = \delta(t-a)$*. This convolution is easily computed as*

$$(f*\delta_a)(t) = \int_{-\infty}^{\infty} f(t-\tau)\delta(\tau-a)\,d\tau = f(t-a).$$

Therefore, the convolution of a function with a shifted delta function is a copy of $f(t)$ *that is shifted by* a*.*

Now convolve $f(t)$ *with the comb function:*

$$\begin{aligned}(f*\text{comb}_a)(t) &= \int_{-\infty}^{\infty} f(t-\tau)\text{comb}_a(\tau)\,d\tau \\ &= \int_{-\infty}^{\infty} f(t-\tau)\sum_{n=-\infty}^{\infty} \delta(\tau-na)\,d\tau\end{aligned}$$

$$\begin{aligned} &= \sum_{n=-\infty}^{\infty} (f * \delta_{na})(t) \\ &= \sum_{n=-\infty}^{\infty} f(t-na). \end{aligned} \tag{6.21}$$

From this result we see that the convolution of a function $f(t)$ with a comb function is then the sum of shifted copies of $f(t)$, as shown in Figure 6.7. If the function has compact support on $[-\frac{a}{2}, \frac{a}{2}]$, i.e., the function is zero for $|t| > 1/a$, then the convolution with the comb function will be periodic.

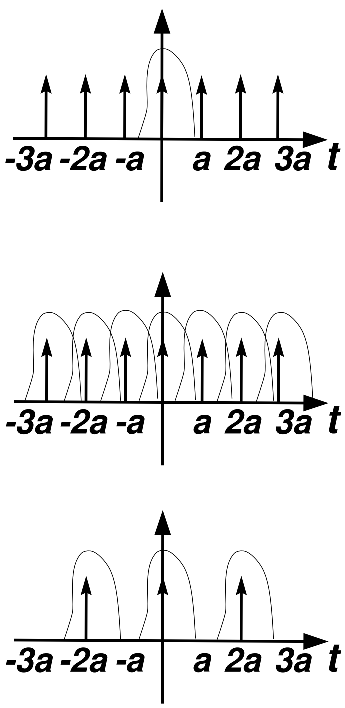


Figure 6.7: The convolution of $f(t)$ with the comb function, $\text{comb}_a(t)$. The first plot shows the function and the comb function. In the second of these plots we add the sum of several translations of $f(t)$. Incorrect sampling will lead to overlap in the translates and cause problems like aliasing. In the last of these plots, one has no overlap and the periodicity is evident.

Example 6.6. *Find the Fourier transform of the convolution* $(f * \text{comb}_a)(t)$.

This is done using the result of the last example and the first shift theorem.

$$\begin{aligned} F[f * comb_a] &= F\left[\sum_{n=-\infty}^{\infty} f(t-na)\right] \\ &= \sum_{n=-\infty}^{\infty} F[f(t-na)] \\ &= \sum_{n=-\infty}^{\infty} \hat{f}(\omega)e^{ina\omega} \\ &= \hat{f}(\omega)\sum_{n=-\infty}^{\infty} e^{ina\omega} \\ &= \frac{2\pi}{a}\hat{f}(\omega)comb_{\frac{2\pi}{a}}(\omega). \end{aligned} \tag{6.22}$$

Another way to compute the Fourier transform of the convolution $(f * comb_a)(t)$ *is to note that the Fourier of this convolution is the product of the transforms of* f *and* $comb_a$ *by the Convolution Theorem. Therefore,*

$$\begin{aligned} F[f * comb_a] &= \hat{f}(\omega)F[comb_a](\omega) \\ &= \frac{2\pi}{a}\hat{f}(\omega)comb_{\frac{2\pi}{a}}(\omega). \end{aligned} \tag{6.23}$$

We have obtained the same result, though in fewer steps.

For a function of period T, $f_p(t) = (f * \text{comb}_T)(t)$, we then have

$$\hat{f}_p(\omega) = \frac{2\pi}{T}\hat{f}(\omega)\text{comb}_{\frac{2\pi}{T}}(\omega).$$

Thus, the resulting spectrum is a series of scaled delta functions separated by $\Delta\omega = \frac{2\pi}{T}$.

The main message of this section has been that the Fourier transform of a periodic function produces a series of delta function spikes. This series of spikes is the transform of the convolution of $f(t)$ and a comb function. This is essentially a sampling of $\hat{f}(\omega)$ in frequency space. A similar result is obtained if we instead had a periodic Fourier transform due to a sampling in t-space. Combining both discrete functions in t and ω spaces, we have discrete signals, as will be described next.

6.3 Discrete Signals

WE WOULD LIKE TO SAMPLE A GIVEN SIGNAL at a discrete set of equally spaced times, $t_n \in [0, T]$. This is how one normally records signals. One samples the signal with a sampling frequency ν_s, such as ten samples per second, or 10 Hz. Therefore, the values are recorded in time steps of $T_s = 1/\nu_s$ For sampling at 10 Hz, this gives a sampling period of 0.1 s.

We can model this sampling at discrete time points by multiplying $f(t)$ by the comb function $\text{comb}_{T_s}(t)$. The Fourier transform will yield a convolution of the Fourier transform of $f(t)$ with the Fourier transform of the comb function. But this is a convolution of $\hat{f}(\omega)$ with another comb function, since the transform of a comb function is a comb function. Therefore, we will obtain a periodic representation in Fourier space.

Example 6.7. *Sample $f(t) = \cos\omega_0 t$ with a sampling frequency of $\nu_s = \frac{1}{T_s}$.*

As noted, the sampling of this function can be represented as

$$f_s(t) = \cos\omega_0 t\, \text{comb}_{T_s}(t).$$

We now evaluate the Fourier transform of $f_s(t)$:

$$\begin{aligned} F[\cos\omega_0 t\, \text{comb}_{T_s}(t)] &= \int_{-\infty}^{\infty} \cos\omega_0 t\, \text{comb}_{T_s}(t) e^{i\omega t}\, dt \\ &= \int_{-\infty}^{\infty} \cos\omega_0 t \left(\sum_{n=-\infty}^{\infty} \delta(t - nT_s) \right) e^{i\omega t}\, dt \\ &= \sum_{n=-\infty}^{\infty} \cos(\omega_0 n T_s) e^{i\omega n T_s} \\ &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \left[e^{i(\omega+\omega_0)nT_s} + e^{i(\omega-\omega_0)nT_s} \right] \\ &= \frac{\pi}{T_s} \left[\text{comb}_{\frac{2\pi}{T_s}}(\omega + \omega_0) + \text{comb}_{\frac{2\pi}{T_s}}(\omega - \omega_0) \right]. \end{aligned} \tag{6.24}$$

Thus, we have shown that sampling $f(t) = \cos\omega_0 t$ with a sampling period of T_s results in a sum of translated comb functions. Each comb function consists of spikes separated by $\omega_s = \frac{2\pi}{T_s}$. Each set of spikes are translated by ω_0 to the left or the right and then added. Since each comb function is periodic with "period" ω_s in Fourier space, then the result turns out to be periodic as noted above.

In collecting data, we not only sample at a discrete set of points, but we also sample for a finite length of time. By sampling like this, we will not gather enough information to obtain the high frequencies in a signal. Thus, there will be a natural cutoff in the spectrum of the signal. This is represented in Figure 6.8. This process will lead to the Discrete Fourier Transform, the topic in the next section.

Here we will show that the sampling of a function defined on $[0, T]$ with sampling period T_s can represented by the convolution of f_p with a comb function. The resulting transform will be a sum of translations of $\hat{f}_p(\omega)$

Once again, we can use the comb function in the analysis of this process. We define a discrete signal as one which is represented by the product of a

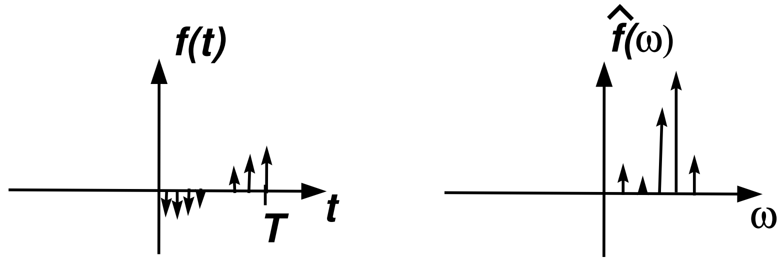


Figure 6.8: Sampling the original signal at a discrete set of times defined on a finite interval leads to a discrete set of frequencies in the transform that are restricted to a finite interval of frequencies.

periodic function sampled at discrete times. This suggests the representation

$$f_d(t) = f_p(t)\text{comb}_{T_s}(t). \tag{6.25}$$

Here T_s denotes the period of the sampling and $f_p(t)$ has the form

$$f_p(t) = \sum_{n=-\infty}^{\infty} c_n e^{-i\omega_n t}.$$

Example 6.8. *Evaluate the Fourier transform of $f_d(t)$.*

The Fourier transform of $f_d(t)$ can be computed:

$$\begin{aligned}
\hat{f}_d(\omega) &= \int_{-\infty}^{\infty} f_p(t) comb_{T_s}(t) e^{i\omega t}\,dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\hat{f}_p(\mu)e^{-i\mu t}\,d\mu\right] comb_{T_s}(t)e^{i\omega t}\,dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}_p(\mu)\underbrace{\int_{-\infty}^{\infty} comb_{T_s}(t)e^{i(\omega-\mu)t}\,dt}_{\text{Transform of } comb_{T_s} \text{ at } \omega-\mu}\,d\mu \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}_p(\mu)\frac{2\pi}{T_s} comb_{\frac{2\pi}{T_s}}(\omega-\mu)\,d\mu \\
&= \frac{1}{T_s}\sum_{n=-\infty}^{\infty}\hat{f}_p(\omega - \frac{2\pi}{T_s}n). \qquad (6.26)
\end{aligned}$$

Note that the Convolution Theorem for the convolution of Fourier transforms needs a factor of 2π.

We note that in the middle of this computation we find a convolution integral:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}_p(\mu)\frac{2\pi}{T_s}\text{comb}_{\frac{2\pi}{T_s}}(\omega-\mu)\,d\mu = \left(\hat{f}_p * \text{comb}_{\frac{2\pi}{T_s}}\right)(\omega)$$

Also, it is easily seen that $\hat{f}_d(\omega)$ is a periodic function with period $\frac{2\pi}{T_s}$:

$$\hat{f}_d(\omega + \frac{2\pi}{T_s}) = \hat{f}_d(\omega).$$

We have shown that sampling a function $f_p(t)$ with sampling frequency $\nu_s = 1/T_s$, one obtains a periodic spectrum, $\hat{f}_d(\omega)$.

6.3.1 Summary

In this chapter we have taken an analog signal defined for $t \in (-\infty, \infty)$ and shown that restricting it to interval $t \in [0, T]$ leads to a periodic function of period T,

$$f_p(t) = (f * \text{comb}_T)(t),$$

whose spectrum is discrete:

$$\hat{f}_p(\omega) = \hat{f}_p(\omega) \sum_{n=-\infty}^{\infty} \delta(\omega - \frac{2\pi n}{T}). \tag{6.27}$$

We then sampled this function at discrete times,

$$f_d(t) = f_p(t)\text{comb}_{T_s}(t).$$

The Fourier transform of this discrete signal was found as

$$\hat{f}_d(\omega) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \hat{f}_p(\omega - \frac{2\pi}{T_s} n). \tag{6.28}$$

This function is periodic with period $\frac{2\pi}{T_s}$.

In Figure 6.9 we summarize the steps for going from analog signals to discrete signals. In the next chapter we will investigate the Discrete Fourier Transform.

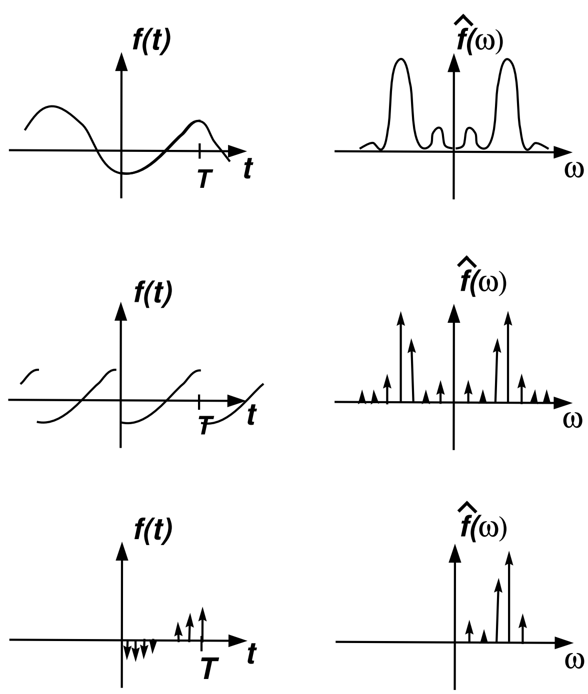


Figure 6.9: Summary of transforming analog to discrete signals. One starts with a continuous signal $f(t)$ defined on $(-\infty, \infty)$ and a continuous spectrum. By only recording the signal over a finite interval $[0, T]$, the recorded signal can be represented by its periodic extension. This in turn forces a discretization of the transform as shown in the second row of figures. Finally, by restricting the range of the sampled, as shown in the last row, the original signal appears as a discrete signal. This is also interpreted as the sampling of an analog signal leads to a restricted set of frequencies in the transform.

6.4 *The Discrete (Trigonometric) Fourier Transform*

Often one is interested in determining the frequency content of signals. Signals are typically represented as time dependent functions. Real signals are continuous, or analog signals. However, through sampling the signal by gathering data, the signal does not contain high frequencies and is finite in duration. The data is then discrete and the corresponding frequencies are discrete and bounded. Thus, in the process of gathering data, one may seriously affect the frequency content of the signal. This is true for a simple superposition of signals with fixed frequencies. The situation becomes more complicated if the data has an overall non-constant (in time) trend or even exists in the presence of noise.

As described earlier in this chapter, we have seen that by restricting the data to a time interval $[0, T]$ for record length T, and periodically extending the data for $t \in (-\infty, \infty)$, one generates a periodic function of infinite duration at the cost of losing data outside the fundamental period. This is not unphysical, as data is typically taken over a finite time period. In addition, if one samples a finite number of values of the data on this finite interval, then the signal will contain frequencies limited to a finite range of values.

This process, as reviewed in the last section, leads us to a study of what is called the *Discrete Fourier Transform,* or *DFT.* We will investigate the discrete Fourier transform in both trigonometric and exponential form. While applications such as MATLAB rely on the exponential version, it is sometimes useful to deal with real functions using the familiar trigonometric functions.

We will change the notation to using $y(t)$ for the signal and f for the frequency.

Recall that in using Fourier series one seeks a representation of the signal, $y(t)$, valid for $t \in [0, T]$, as

$$y(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}[a_n \cos \omega_n t + b_n \sin \omega_n t], \tag{6.29}$$

where the angular frequency is given by $\omega_n = 2\pi f_n = \frac{2\pi n}{T}$. Note: In discussing signals, we will now use $y(t)$ instead of $f(t)$, allowing us to use f to denote the frequency (in Hertz) without confusion.

The frequency content of a signal for a particular frequency, f_n, is contained in both the cosine and sine terms when the corresponding Fourier coefficients, a_n, b_n, are not zero. So, one may desire to combine these terms. This is easily done by using trigonometric identities (as we had seen in Equation(2.1)). Namely, we show that

$$a_n \cos \omega_n t + b_n \sin \omega_n t = C_n \cos(\omega_n t + \psi), \tag{6.30}$$

where ψ is the *phase shift.* Recalling that

$$\cos(\omega_n t + \psi) = \cos \omega_n t \cos \psi - \sin \omega_n t \sin \psi, \tag{6.31}$$

one has

$$a_n \cos \omega_n t + b_n \sin \omega_n t = C_n \cos \omega_n t \cos \psi - C_n \sin \omega_n t \sin \psi.$$

Equating the coefficients of $\sin \omega_n t$ and $\cos \omega_n t$ in this expression, we obtain

$$a_n = C_n \cos \psi, \quad b_n = -C_n \sin \psi.$$

Therefore,

$$C_n = \sqrt{a_n + b_n} \quad \text{and} \quad \tan \psi = -\frac{b_n}{a_n}.$$

Recall that we had used orthogonality arguments in order to determine the Fourier coefficients ($a_n, n = 0, 1, 2, \ldots$ and $b_n, n = 1, 2, 3, \ldots$). In particular, using the orthogonality of the trigonometric basis, we found that

$$\begin{aligned} a_n &= \tfrac{2}{T}\int_0^T y(t)\cos\omega_n t\,dt, \quad n = 0,1,2,\ldots \\ b_n &= \tfrac{2}{T}\int_0^T y(t)\sin\omega_n t\,dt, \quad n = 1,2,\ldots \end{aligned} \tag{6.32}$$

In Chapter 2 we had shown that one can write $a_n \cos\omega_n t + b_n \sin\omega_n t = C_n \sin(\omega_n t + \phi)$. It is easy to shown that *psi* and ϕ are related.

In the next section we will introduce the trigonometric version of the Discrete Fourier Transform. Its appearance is similar to that of the Fourier series representation in Equation (6.29). However, we will need to do a bit of work to obtain the discrete Fourier coefficients using discrete orthogonality.

6.4.1 *Discrete Trigonometric Series*

For the Fourier series analysis of a signal, we had restricted time to the interval $[0, T]$, leading to a Fourier series with discrete frequencies and a periodic function of time. In reality, taking data can only be done at certain frequencies, thus eliminating high frequencies. Such a restriction on the frequency leads to a discretization of the data in time. Another way to view this is that when recording data we sample at a finite number of time steps, limiting our ability to collect data with large oscillations. Thus, we not only have discrete frequencies but we also have discrete times.

We first note that the data is sampled at N equally spaced times

$$t_n = n\Delta t, \quad n = 0, 1, \ldots, N-1,$$

where Δt is the time increment. For a record length of T, we have $\Delta t = T/N$. We will denote the data at these times as $y_n = y(t_n)$.

The DFT representation that we are seeking takes the form:

$$y_n = \frac{1}{2}a_0 + \sum_{p=1}^{M}[a_p \cos\omega_p t_n + b_p \sin\omega_p t_n], \; n = 0, 1, \ldots, N-1. \tag{6.33}$$

Here we note the discretizations used for future reference. Defining $\Delta t = \frac{T}{N}$ and $\Delta f = \frac{1}{T}$, we have

$$t_n = n\Delta t = \frac{nT}{N},$$
$$\omega_p = p\Delta\omega = \frac{2\pi p}{T},$$
$$f_p = p\Delta f = \frac{p}{T},$$
$$\omega_p t_n = \frac{2\pi np}{N},$$

for $n = 0, 1, \ldots, N-1$ and $p = 1, \ldots, M$.

The trigonometric arguments are given by

$$\omega_p t_n = \frac{2\pi p}{T} n\delta t = \frac{2\pi pn}{N}.$$

Note that $p = 1, \ldots, M$, thus allowing only for frequencies $f_p = \frac{\omega_p}{2\pi} = \frac{p}{T}$. Or, we could write

$$f_p = p\Delta f$$

for

$$\Delta f = \frac{1}{T}.$$

We need to determine M and the unknown coefficients. As for the Fourier series, we will need some orthogonality relations, but this time the orthogonality statement will consist of a sum and not an integral.

Since there are N sample values, (6.33) gives us a set of N equations for the unknown coefficients. Therefore, we should have N unknowns. For N

samples, we want to determine N unknown coefficients $a_0, a_1, \ldots, a_{N/2}$ and $b_1, \ldots, b_{N/2-1}$. Thus, we need to fix $M = \frac{N}{2}$. Often the coefficients b_0 and $b_{N/2}$ are included for symmetry. Note that the corresponding sine function factors evaluate to zero at $p = 0, \frac{N}{2}$, leaving these two coefficients arbitrary. Thus, we can take them to be zero when they are included.

We claim that for the Discrete (Trigonometric) Fourier Transform

$$y_n = \frac{1}{2}a_0 + \sum_{p=1}^{M} [a_p \cos\omega_p t_n + b_p \sin\omega_p t_n], \; n = 0, 1, \ldots, N-1. \tag{6.34}$$

the DFT coefficients are given by

$$\begin{aligned}
a_p &= \frac{2}{N}\sum_{n=0}^{N-1} y_n \cos(\frac{2\pi pn}{N}), p = 1, \ldots N/2-1 \\
b_p &= \frac{2}{N}\sum_{n=0}^{N-1} y_n \sin(\frac{2\pi pn}{N}), p = 1, 2, \ldots N/2-1 \\
a_0 &= \frac{1}{N}\sum_{n=0}^{N-1} y_n, \\
a_{N/2} &= \frac{0}{N}\sum_{n=1}^{N-1} y_n \cos n\pi \\
b_0 &= b_{N/2} = 0
\end{aligned} \tag{6.35}$$

6.4.2 *Discrete Orthogonality*

The derivation of the discrete Fourier coefficients can be done using the discrete orthogonality of the discrete trigonometric basis similar to the derivation of the above Fourier coefficients for the Fourier series in Chapter 2. We first prove the following

Theorem 6.1.

$$\begin{aligned}
&\sum_{n=0}^{N-1} \cos\left(\tfrac{2\pi nk}{N}\right) = \begin{cases} 0, & k = 1, \ldots, N-1 \\ N, & k = 0, N \end{cases} \\
&\sum_{n=0}^{N-1} \sin\left(\tfrac{2\pi nk}{N}\right) = 0, \quad k = 0, \ldots, N
\end{aligned} \tag{6.36}$$

Proof. This can be done more easily using the exponential form,

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi nk}{N}\right) + i\sum_{n=0}^{N-1} \sin\left(\frac{2\pi nk}{N}\right) = \sum_{n=0}^{N-1} e^{2\pi ink/N}, \tag{6.37}$$

by using Euler's formula, $e^{i\theta} = \cos\theta + i\sin\theta$ for each term in the sum.

The exponential sum is the sum of a geometric progression. Namely, we note that

$$\sum_{n=0}^{N-1} e^{2\pi ink/N} = \sum_{n=0}^{N-1} \left(e^{2\pi ik/N}\right)^n.$$

Recall from Chapter 2 that a geometric progression is a sum of the form $S_N = \sum_{k=0}^{N-1} ar^k$. This is a sum of N terms in which consecutive terms have a constant ratio, r. The sum is easily computed. One multiplies the sum S_N by r and subtracts the resulting sum from the original sum to obtain

$$S_N - rS_N = (a + ar + \cdots + ar^{N-1}) - (ar + \cdots + ar^N + ar^N) = a - ar^N. \tag{6.38}$$

Factoring on both sides of this chain of equations yields the desired sum,

$$S_N = \frac{a(1-r^N)}{1-r}. \tag{6.39}$$

Thus, we have

$$\begin{aligned} \sum_{n=0}^{N-1} e^{2\pi ink/N} &= \sum_{n=0}^{N-1} \left(e^{2\pi ik/N}\right)^n \\ &= 1 + e^{2\pi ik/N} + \left(e^{2\pi ik/N}\right)^2 + \cdots + \left(e^{2\pi ik/N}\right)^{N-1} \\ &= \frac{1 - \left(e^{2\pi ik/N}\right)^N}{1 - e^{2\pi ik/N}} \\ &= \frac{1 - e^{2\pi ik}}{1 - e^{2\pi ik/N}}. \end{aligned} \tag{6.40}$$

As long as $k \neq 0, N$ the numerator is 0 and $1 - e^{2\pi ik/N}$ is not zero.

In the special cases that $k = 0, N$, we have $e^{2\pi ink/N} = 1$. So,

$$\sum_{n=0}^{N-1} e^{2\pi ink/N} = \sum_{n=0}^{N-1} 1 = N.$$

Therefore,

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi nk}{N}\right) + i \sum_{n=0}^{N-1} \sin\left(\frac{2\pi nk}{N}\right) = \begin{cases} 0, & k = 1, \ldots, N-1 \\ N, & k = 0, N \end{cases} \tag{6.41}$$

and the Theorem is proved. □

We can use this result to establish orthogonality relations.

Example 6.9. *Evaluate the following:*

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right),$$

$$\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right),$$

and

$$\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \sin\left(\frac{2\pi qn}{N}\right).$$

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right) = \frac{1}{2} \sum_{n=0}^{N-1} \left[\cos\left(\frac{2\pi(p-q)n}{N}\right) + \cos\left(\frac{2\pi(p+q)n}{N}\right)\right]. \tag{6.42}$$

Splitting the above sum into two sums and then evaluating the separate sums from earlier in this section,

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi(p-q)n}{N}\right) = \begin{cases} 0, & p \neq q \\ N, & p = q \end{cases},$$

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi(p+q)n}{N}\right) = \begin{cases} 0, & p+q \neq N \\ N, & p+q = N \end{cases}$$

we obtain

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right) = \begin{cases} N/2, & p = q \neq N/2 \\ N, & p = q = N/2 \\ 0, & \textit{otherwise} \end{cases}. \tag{6.43}$$

Similarly, we find

$$\begin{aligned} &\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right) \\ = \;& \frac{1}{2}\sum_{n=0}^{N-1} \left[\sin\left(\frac{2\pi(p-q)n}{N}\right) + \sin\left(\frac{2\pi(p+q)n}{N}\right)\right] \\ = \;& 0. \end{aligned} \tag{6.44}$$

and

$$\begin{aligned} &\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \sin\left(\frac{2\pi qn}{N}\right) \\ = \;& \frac{1}{2}\sum_{n=0}^{N-1} \left[\cos\left(\frac{2\pi(p-q)n}{N}\right) - \cos\left(\frac{2\pi(p+q)n}{N}\right)\right] \\ = \;& \begin{cases} N/2, & p = q \neq N/2 \\ 0, & \textit{otherwise} \end{cases}. \end{aligned} \tag{6.45}$$

We have proven the following orthogonality relations

Theorem 6.2.

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right) = \begin{cases} N/2, & p = q \neq N/2 \\ N, & p = q = N/2 \\ 0, & \textit{otherwise} \end{cases}. \tag{6.46}$$

$$\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \cos\left(\frac{2\pi qn}{N}\right) = 0. \tag{6.47}$$

$$\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \sin\left(\frac{2\pi qn}{N}\right) \begin{cases} N/2, & p = q \neq N/2 \\ 0, & \textit{otherwise} \end{cases}. \tag{6.48}$$

6.4.3 The Discrete Fourier Coefficients

THE DERIVATION OF THE COEFFICIENTS FOR THE DFT is now easily done using the discrete orthogonality from the last section. We start with the expansion

$$y_n = \frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p \cos\left(\frac{2\pi pn}{N}\right) + b_p \sin\left(\frac{2\pi pn}{N}\right)\right],\ n = 0, \ldots, N-1. \tag{6.49}$$

We first sum over n:

$$\begin{aligned}
\sum_{n=0}^{N-1} y_n &= \sum_{n=0}^{N-1}\left(\frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p \cos\left(\frac{2\pi pn}{N}\right) + b_p \sin\left(\frac{2\pi pn}{N}\right)\right]\right) \\
&= \frac{1}{2}a_0 \sum_{n=0}^{N-1} 1 + \sum_{p=1}^{N/2}\left[a_p \sum_{n=0}^{N-1}\cos\left(\frac{2\pi pn}{N}\right) + b_p \sum_{n=0}^{N-1}\sin\left(\frac{2\pi pn}{N}\right)\right] \\
&= \frac{1}{2}a_0 N + \sum_{p=1}^{N/2}\left[a_p \cdot 0 + b_p \cdot 0\right] \\
&= \frac{1}{2}a_0 N.
\end{aligned} \tag{6.50}$$

Therefore, we have $a_0 = \frac{2}{N}\sum\limits_{n=0}^{N-1} y_n$

Now, we can multiply both sides of the expansion (6.77) by $\cos\left(\frac{2\pi qn}{N}\right)$ and sum over n. This gives

$$\begin{aligned}
&\sum_{n=0}^{N-1} y_n \cos\left(\frac{2\pi qn}{N}\right) \\
= &\sum_{n=0}^{N-1}\left(\frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p \cos\left(\frac{2\pi pn}{N}\right) + b_p \sin\left(\frac{2\pi pn}{N}\right)\right]\right)\cos\left(\frac{2\pi qn}{N}\right) \\
= &\frac{1}{2}a_0 \sum_{n=0}^{N-1}\cos\left(\frac{2\pi qn}{N}\right) \\
&+ \sum_{p=1}^{N/2} a_p \sum_{n=0}^{N-1}\cos\left(\frac{2\pi pn}{N}\right)\cos\left(\frac{2\pi qn}{N}\right) \\
&+ \sum_{p=1}^{N/2} b_p \sum_{n=0}^{N-1}\sin\left(\frac{2\pi pn}{N}\right)\cos\left(\frac{2\pi qn}{N}\right) \\
= &\begin{cases} \sum_{p=1}^{N/2}\left[a_p \frac{N}{2}\delta_{p,q} + b_p \cdot 0\right], & q \neq N/2, \\ \sum_{p=1}^{N/2}\left[a_p N \delta_{p,N/2} + b_p \cdot 0\right], & q = N/2, \end{cases} \\
= &\begin{cases} \frac{1}{2}a_q N, & q \neq N/2 \\ a_{N/2}N, & q = N/2 \end{cases}.
\end{aligned} \tag{6.51}$$

So, we have found that

$$a_q = \frac{2}{N}\sum_{n=0}^{N-1} y_n \cos\left(\frac{2\pi qn}{N}\right), \quad q \neq \frac{N}{2}, \tag{6.52}$$

$$\begin{aligned}
a_{N/2} &= \frac{1}{N}\sum_{n=0}^{N-1} y_n \cos\left(\frac{2\pi n(N/2)}{N}\right) \\
&= \frac{1}{N}\sum_{n=0}^{N-1} y_n \cos\left(\pi n\right).
\end{aligned} \tag{6.53}$$

Similarly, $\sum_{n=0}^{N-1} y_n \sin\left(\frac{2\pi qn}{N}\right) =$

$$\begin{aligned}
&= \sum_{n=0}^{N-1}\left(\frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p\cos\left(\frac{2\pi pn}{N}\right) + b_p\sin\left(\frac{2\pi pn}{N}\right)\right]\right)\sin\left(\frac{2\pi qn}{N}\right) \\
&= \frac{1}{2}a_0\sum_{n=0}^{N-1}\sin\left(\frac{2\pi qn}{N}\right) \\
&\quad + \sum_{p=1}^{N/2} a_p \sum_{n=0}^{N-1}\cos\left(\frac{2\pi pn}{N}\right)\sin\left(\frac{2\pi qn}{N}\right) \\
&\quad + \sum_{p=1}^{N/2} b_p \sum_{n=0}^{N-1}\sin\left(\frac{2\pi pn}{N}\right)\sin\left(\frac{2\pi qn}{N}\right) \\
&= \sum_{p=1}^{N/2}\left[a_p\cdot 0 + b_p\frac{N}{2}\delta_{p,q}\right] \\
&= \frac{1}{2}b_q N. \qquad (6.54)
\end{aligned}$$

Finally, we have

$$b_q = \frac{2}{N}\sum_{n=0}^{N-1} y_n \sin\left(\frac{2\pi qn}{N}\right), \quad q = 1, \ldots, \frac{N}{2} - 1. \qquad (6.55)$$

6.5 *The Discrete Exponential Transform*

The derivation of the coefficients for the trigonometric DFT was obtained in the last section using the discrete orthogonality. However, applications like MATLAB do not typically use the trigonometric version of DFT for spectral analysis. MATLAB instead uses a discrete Fourier exponential transform in the form of the Fast Fourier Transform (FFT)[1] . Its description in the help section does not involve sines and cosines. Namely, MATLAB defines the transform and inverse transform (by typing **help fft**) as

[1] The Fast Fourier Transform, or FFT, refers to an efficient algorithm for computing the discrete Fourier transform. It was originally developed by James Cooley and John Tukey in 1965 for computers based upon an algorithm invented by Gauß.

```
For length N input vector x, the DFT is a length N vector X,
with elements
         N
 X(k) = sum  x(n)*exp(-j*2*pi*(k-1)*(n-1)/N), 1 <= k <= N.
        n=1
The inverse DFT (computed by IFFT) is given by
              N
x(n) = (1/N) sum  X(k)*exp( j*2*pi*(k-1)*(n-1)/N), 1 <= n <= N.
             k=1
```

It also provides in the new help system,

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)},$$

$$x(j) = \frac{1}{N}\sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)},$$

where $\omega_N = e^{-2\pi i/N}$ are the Nth roots of unity. You will note a few differences between these representations and the discrete Fourier transform in the last section. First of all, there are no trigonometric functions. Next, the sums do not have a zero index, a feature of indexing in MATLAB. Also, in the older definition, MATLAB uses a "j" and not an "i" for the imaginary unit.

In this section we will derive the discrete Fourier exponential transform in the form

$$F_k = \sum_{j=0}^{N-1} W^{jk} f_j. \tag{6.56}$$

where $W_N = e^{-2\pi i/N}$ and $k = 0, \dots, N-1$. We will find the relationship between what MATLAB is computing and the discrete Fourier trigonometic series. This will also be useful as a preparation for a discussion of the FFT in the next chapter.

We will start with the DFT (Discrete Fourier Transform):

$$y_n = \frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p \cos\left(\frac{2\pi pn}{N}\right) + b_p \sin\left(\frac{2\pi pn}{N}\right)\right] \tag{6.57}$$

for $n = 0, 1, \dots, N-1$.

We use Euler's formula to rewrite the trigonometric functions in terms of exponentials. The DFT formula can be written as

$$\begin{aligned} y_n &= \frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p \frac{e^{2\pi ipn/N} + e^{-2\pi ipn/N}}{2} + b_p \frac{e^{2\pi ipn/N} - e^{-2\pi ipn/N}}{2i}\right] \\ &= \frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[\frac{1}{2}\left(a_p - ib_p\right)e^{2\pi ipn/N} + \frac{1}{2}\left(a_p + ib_p\right)e^{-2\pi ipn/N}\right]. \end{aligned} \tag{6.58}$$

We define $C_p = \frac{1}{2}\left(a_p - ib_p\right)$ and note that the above result can be written as

$$y_n = C_0 + \sum_{p=1}^{N/2}\left[C_p e^{2\pi ipn/N} + \bar{C}_p e^{-2\pi ipn/N}\right], \quad n = 0, 1, \dots, N-1. \tag{6.59}$$

The terms in the sums look similar. We can actually combine them into one form. Note that $e^{2\pi iN} = \cos(2\pi N) + i\sin(2\pi N) = 1$. Thus, we can write

$$e^{-2\pi ipn/N} = e^{-2\pi ipn/N}e^{-2\pi iN} = e^{2\pi i(N-p)n/N}$$

in the second sum. Since $p = 1, \dots, N/2$, we see that $N - p = N-1, N-2, \dots, N/2$. So, we can rewrite the second sum as

$$\sum_{p=1}^{N/2} \bar{C}_p e^{-2\pi ipn/N} = \sum_{p=1}^{N/2} \bar{C}_p e^{2\pi i(N-p)n/N} = \sum_{q=N/2}^{N-1} \bar{C}_{N-q} e^{2\pi iqn/N}.$$

Since q is a *dummy* index (it can be replaced by any letter without changing the value of the sum), we can replace it with a p and combine the terms in both sums to obtain

$$y_n = \sum_{p=0}^{N-1} Y_p e^{2\pi i p n/N}, \quad n = 0, 1, \ldots, N-1, \tag{6.60}$$

where

$$Y_p = \begin{cases} \frac{a_0}{2}, & p = 0, \\ \frac{1}{2}(a_p - ib_p), & 0 < p < N/2, \\ a_{N/2}, & p = N/2, \\ \frac{1}{2}(a_{N-p} + ib_{N-p}), & N/2 < p < N. \end{cases} \tag{6.61}$$

Notice that the real and imaginary parts of the Fourier coefficients obey certain symmetry properties over the full range of the indices since the real and imaginary parts are related between $p \in (0, N/2)$ and $p \in (N/2, N)$. Namely, since

$$Y_{N-j} = \frac{1}{2}(a_j + ib_j) = \overline{Y}_j, \quad j = 1, \ldots, N-1,$$

$Re(Y_{N-j}) = ReY_j$ and $Im(Y_{N-j}) = -ImY_j$ for $j = 1, \ldots, N-1$.

We can now determine the coefficients in terms of the sampled data.

$$\begin{aligned} C_p &= \frac{1}{2}(a_p - ib_p) \\ &= \frac{1}{N} \sum_{n=1}^{N} y_n \left[\cos\left(\frac{2\pi pn}{N}\right) - i \sin\left(\frac{2\pi pn}{N}\right) \right] \\ &= \frac{1}{N} \sum_{n=1}^{N} y_n e^{-2\pi i p n/N}. \end{aligned} \tag{6.62}$$

Thus,

$$Y_p = \frac{1}{N} \sum_{n=1}^{N} y_n e^{-2\pi i p n/N}, \quad 0 < p < \frac{N}{2} \tag{6.63}$$

and

$$\begin{aligned} Y_p &= \bar{C}_{N-p} \\ &= \frac{1}{N} \sum_{n=1}^{N} y_n e^{2\pi i (N-p) n/N}, \quad \frac{N}{2} < p < N \\ &= \frac{1}{N} \sum_{n=1}^{N} y_n e^{-2\pi i p n/N}. \end{aligned} \tag{6.64}$$

We have shown that for all Y's but two, the form is

$$Y_p = \frac{1}{N} \sum_{n=1}^{N} y_n e^{-2\pi i p n/N}. \tag{6.65}$$

However, we can easily show that this is also true when $p = 0$ and $p = \frac{N}{2}$.

$$Y_{N/2} = a_{N/2}$$

$$\begin{aligned} &= \frac{1}{N}\sum_{n=1}^{N} y_n \cos n\pi \\ &= \frac{1}{N}\sum_{n=1}^{N} y_n \left[\cos n\pi - i \sin n\pi\right] \\ &= \frac{1}{N}\sum_{n=1}^{N} y_n e^{-2\pi i n (N/2)/N} \end{aligned} \tag{6.66}$$

and

$$\begin{aligned} Y_0 &= \frac{1}{2} a_0 \\ &= \frac{1}{N}\sum_{n=1}^{N} y_n \\ &= \frac{1}{N}\sum_{n=1}^{N} y_n e^{2\pi i n (0)/N}. \end{aligned} \tag{6.67}$$

Thus, all of the Y_p's are of the same form. This gives us the discrete transform pair

$$y_n = \sum_{p=0}^{N-1} Y_p e^{2\pi i p n/N}, \quad n = 1, \ldots, N, \tag{6.68}$$

$$Y_p = \frac{1}{N}\sum_{n=1}^{N} y_n e^{-2\pi i p n/N}, \quad p = 0, 1, \ldots, N-1. \tag{6.69}$$

Note that this is similar to the definition of the FFT given in MATLAB.

6.6 *FFT: The Fast Fourier Transform*

The usual computation of the discrete Fourier transform (DFT) is done using the Fast Fourier Transform (FFT). There are various implementations of it, but a standard form is the Radix-2 FFT. We describe this FFT in the current section. We begin by writing the DFT compactly using $W = e^{-2\pi i/N}$. Note that $W^{N/2} = -1, \quad W^N = 1$, and $e^{2\pi i jk/N} = W^{jk}$. We can then write

$$F_k = \sum_{j=0}^{N-1} W^{jk} f_j. \tag{6.70}$$

The key to the FFT is that this sum can be written as two similar sums:

$$\begin{aligned} F_k &= \sum_{j=0}^{N-1} W^{jk} f_j \\ &= \sum_{j=0}^{N/2-1} W^{jk} f_j + \sum_{j=N/2}^{N-1} W^{jk} f_j \\ &= \sum_{j=0}^{N/2-1} W^{jk} f_j + \sum_{m=0}^{N/2-1} W^{k(m+N/2)} f_{m+N/2}, \quad \text{for } m = j - \frac{N}{2} \end{aligned}$$

$$= \sum_{j=0}^{N/2-1} \left[W^{jk} f_j + W^{k(j+N/2)} f_{j+N/2} \right]$$
$$= \sum_{j=0}^{N/2-1} W^{jk} \left[f_j + (-1)^k f_{j+N/2} \right] \tag{6.71}$$

since $W^{k(j+N/2)} = W^{kj}(W^{N/2})^k$ and $W^{N/2} = -1$.

Thus, the sum appears to be of the same form as the initial sum, but there are half as many terms with a different coefficient for the W^{jk}'s. In fact, we can separate the terms involving the + or – sign by looking at the even and odd values of k.

For even $k = 2m$, we have

$$F_{2m} = \sum_{j=0}^{N/2-1} \left(W^{2m} \right)^j \left[f_j + f_{j+N/2} \right], \quad m = 0, \dots \frac{N}{2} - 1. \tag{6.72}$$

For odd $k = 2m + 1$, we have

$$F_{2m+1} = \sum_{j=0}^{N/2-1} \left(W^{2m} \right)^j W^j \left[f_j - f_{j+N/2} \right], \quad m = 0, \dots \frac{N}{2} - 1. \tag{6.73}$$

Each of these equations gives the Fourier coefficients in terms of a similar sum using fewer terms and with a different weight, $W^2 = \left(e^{-2\pi i/N} \right)^2 = e^{-2\pi i/(N/2)}$. If N is a power of 2, then this process can be repeated over and over until one ends up with a simple sum.

The process is easily seen when written out for a small number of samples. Let $N = 8$. Then a first pass at the above gives

$$\begin{aligned}
F_0 &= f_0 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7. \\
F_1 &= f_0 + W f_1 + W^2 f_2 + W^3 f_3 + W^4 f_4 + W^5 f_5 + W^6 f_6 + W^7 f_7. \\
F_2 &= f_0 + W^2 f_1 + W^4 f_2 + W^6 f_3 + f_4 + W^2 f_5 + W^4 f_6 + W^6 f_7. \\
F_3 &= f_0 + W^3 f_1 + W^6 f_2 + W f_3 + W^4 f_4 + W^7 f_5 + W^2 f_6 + W^5 f_7. \\
F_4 &= f_0 + W^4 f_1 + f_2 + W^4 f_3 + f_4 + W^4 f_5 + f_6 + W^4 f_7. \\
F_5 &= f_0 + W^5 f_1 + W^2 f_2 + W^7 f_3 + W^4 f_4 + W f_5 + W^6 f_6 + W^3 f_7. \\
F_6 &= f_0 + W^6 f_1 + W^4 f_2 + W^2 f_3 + f_4 + W^6 f_5 + W^4 f_6 + W^2 f_7. \\
F_7 &= f_0 + W^7 f_1 + W^6 f_2 + W^5 f_3 + W^4 f_4 + W^3 f_5 + W^2 f_6 + W f_7.
\end{aligned} \tag{6.74}$$

The point is that the terms in these expressions can be regrouped with $W = e^{-\pi i/8}$ and noting $W^4 = -1$:

$$\begin{aligned}
F_0 &= (f_0 + f_4) + (f_1 + f_5) + (f_2 + f_6) + (f_3 + f_7) \\
&\equiv g_0 + g_1 + g_2 + g_3. \\
F_1 &= (f_0 - f_4) + (f_1 - f_5) W + (f_2 - f_6) W^2 + (f_3 - f_7) W^3 \\
&\equiv g_4 + g_6 + g_5 + g_7. \\
F_2 &= (f_0 + f_4) + (f_1 + f_5) W^2 - (f_2 + f_6) - (f_3 + f_7) W^2
\end{aligned}$$

$$\begin{aligned}
&= g_0 - g_2 + (g_1 - g_3)W^2. \\
F_3 &= (f_0 - f_4) - (f_2 - f_6)W + (f_1 - f_5)WW^2 + (f_3 - f_7)WW^6 \\
&= g_4 - g_6 + g_5W^2 + g_7W^6. \\
F_4 &= (f_0 + f_4) + (f_1 + f_5) - (f_2 + f_6) - (f_3 + f_7) \\
&= g_0 + g_2 - g_1 - g_3. \\
F_5 &= (f_0 - f_4) + (f_2 - f_6)W + (f_1 - f_5)WW^4 + (f_3 - f_7)WW^4 \\
&= g_4 + g_6 + g_5W^4 + g_7W^4. \\
F_6 &= (f_0 + f_4) + (f_1 + f_5)\, W^6 - (f_2 + f_6) - (f_3 + f_7)\, W^6 \\
&= g_0 - g_2 + (g_1 - g_3)W^6. \\
F_7 &= (f_0 - f_4) - (f_2 - f_6)W + (f_1 - f_5)WW^6 + (f_3 - f_7)WW^2 \\
&= g_4 - g_6 + g_5W^6 + g_7W^2. \qquad (6.75)
\end{aligned}$$

However, each of the $g-$series can be rewritten as well, leading to

$$\begin{aligned}
F_0 &= (g_0 + g_2) + (g_1 + g_3) \equiv h_0 + h_1. \\
F_1 &= (g_4 + g_6) + (g_5 + g_7) \equiv h_4 + h_5. \\
F_2 &= (g_0 - g_2) + (g_1 - g_3)W^2 \equiv h_2 + h_3. \\
F_3 &= (g_4 - g_6) + (g_5 - g_7)W^2 \equiv h_6 + h_7. \\
F_4 &= (g_0 + g_2) - (g_1 + g_3) = h_0 - h_1. \\
F_5 &= (g_4 + g_6) - (g_5 + g_7) = h_4 - h_5. \\
F_6 &= g_0 - g_2 - (g_1 - g_3)W^2 = h_2 - h_3. \\
F_7 &= g_4 - g_6 + g_5W^6 + g_7W^2 = h_6 - h_7. \qquad (6.76)
\end{aligned}$$

Thus, the computation of the Fourier coefficients amounts to inputting the f's and computing the g's. This takes 8 additions and 4 multiplications. Then one get the h's, which is another 8 additions and 4 multiplications. There are three stages, amounting to a total of 12 multiplications and 24 additions. Carrying out the process in general, one has $\log_2 N$ steps with $N/2$ multiplications and N additions per step. In the direct computation one has $(N-1)^2$ multiplications and $N(N-1)$ additions. Thus, for $N = 8$, that would be 49 multiplications and 56 additions.

The above process is typically shown schematically in a "butterfly diagram." The basic butterfly transformation is displayed in Figure 6.10. An 8 point FFT is shown in Figure 6.11.

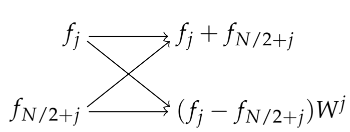


Figure 6.10: This is the basic FFT butterfly.

In the actual implementation, one computes with the h's in the following order:

The binary representation of the index was also listed. Notice that the output is in bit-reversed order as compared to the right side of the table

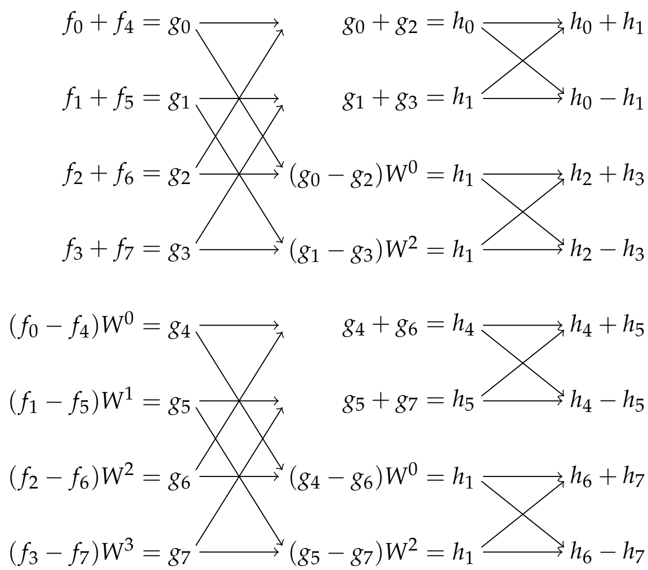


Figure 6.11: This is an 8 point FFT butterfly diagram.

Table 6.1: Output, desired order and binary representation for the Fourier Coefficients.

Output	**Desired Order**
$h_0 + h_1 = F_0$, 000	F_0, 000
$h_0 - h_1 = F_4$, 100	F_1, 001
$h_2 + h_3 = F_2$, 010	F_2, 010
$h_2 - h_3 = F_6$, 110	F_3, 011
$h_4 + h_5 = F_1$, 001	F_4, 100
$h_4 - h_5 = F_5$, 101	F_5, 101
$h_6 + h_7 = F_3$, 011	F_6, 110
$h_6 - h_7 = F_7$, 111	F_7, 111

which shows the coefficients in the correct order. [Just compare the columns in each set of binary representations.] So, typically there is a bit reversal routine needed to unscramble the order of the output coefficients in order to use them.

6.7 *Applications*

IN THE LAST SECTION WE SAW that given a set of data, y_n, $n = 0, 1, \ldots, N-1$, that one can construct the corresponding discrete, finite Fourier series. The series is given by

$$y_n = \frac{1}{2}a_0 + \sum_{p=1}^{N/2}\left[a_p \cos\left(\frac{2\pi pn}{N}\right) + b_p \sin\left(\frac{2\pi pn}{N}\right)\right],\ n = 0, \ldots, N-1. \tag{6.77}$$

and the Fourier coefficients were found as

$$\begin{aligned}
a_p &= \tfrac{2}{N}\sum_{n=0}^{N-1} y_n \cos(\tfrac{2\pi pn}{N}),\ p = 1, \ldots N/2-1\\
b_p &= \tfrac{2}{N}\sum_{n=0}^{N-1} y_n \sin(\tfrac{2\pi pn}{N}),\ p = 1, 2, \ldots N/2-1\\
a_0 &= \tfrac{1}{N}\sum_{n=0}^{N-1} y_n,\\
a_{N/2} &= \tfrac{1}{N}\sum_{n=0}^{N-1} y_n \cos n\pi,\\
b_0 &= b_{N/2} = 0
\end{aligned} \tag{6.78}$$

In this section we show how this is implemented using MATLAB.

Example 6.10. *Analysis of monthly mean surface temperatures.*

Consider the data[2] of monthly mean surface temperatures at Amphitrite Point, Canada shown in Table 6.2. The temperature was recorded in ^{o}C and averaged for each month over a two year period. We would like to look for the frequency content of this time series.

[2] This example is from *Data Analysis Methods in Physical Oceanography*, W. J. Emery and R.E. Thomson, Elsevier, 1997.

Month	1	2	3	4	5	6	7	8	9	10	11	12
1982	7.6	7.4	8.2	9.2	10.2	11.5	12.4	13.4	13.7	11.8	10.1	9.0
1983	8.9	9.5	10.6	11.4	12.9	12.7	13.9	14.2	13.5	11.4	10.9	8.1

Table 6.2: Monthly mean surface temperatures (^{o}C) at Amphitrite Point, Canada for 1982-1983.

In Figure 6.12 we plot the above data as circles. We then use the data to compute the Fourier coefficients. These coefficients are used in the discrete Fourier series an plotted on top of the data in red. We see that the reconstruction fits the data.

The implementations of DFT are done using MATLAB. We provide the code at the end of this section.

Example 6.11. *Determine the frequency content of $y(t) = \sin(10\pi t)$.*

Generally, we are interested in determining the frequency content of a signal. For example, we consider a pure note,

$$y(t) = \sin(10\pi t).$$

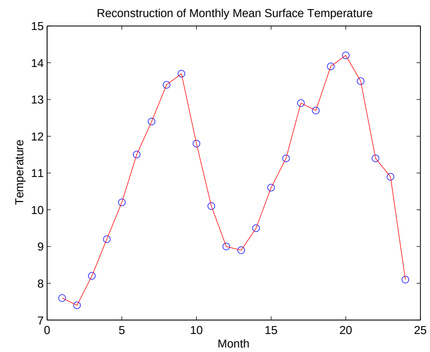


Figure 6.12: Plot and reconstruction of the monthly mean surface temperature data.

Sampling this signal with $N = 128$ points on the interval $[0, 5]$, we find the discrete Fourier coefficients as shown in Figure 6.13. Note the spike at the right place in the B plot. The others spikes are actually very small if you look at the scale on the plot of the A coefficients.

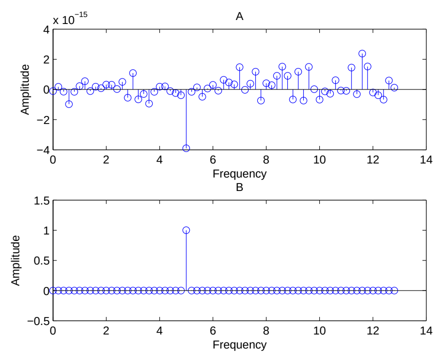


Figure 6.13: Computed discrete Fourier coefficients for $y(t) = \sin(10\pi t)$, with $N = 128$ points on the interval $[0, 5]$.

One can use these coefficients to reconstruct the signal. This is shown in Figure 6.14

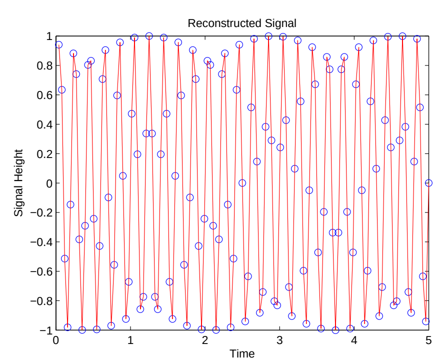


Figure 6.14: Reconstruction of $y(t) = \sin(10\pi t)$ from its Fourier coefficients.

Example 6.12. *Determine the frequency content of $y(t) = \sin(10\pi t) - \frac{1}{2}\cos(6\pi t)$.*

We can look at more interesting functions. For example, what if we add two pure notes together, such as

$$y(t) = \sin(10\pi t) - \frac{1}{2}\cos(6\pi t).$$

We see from Figure 6.15 that the implementation works. The Fourier coefficients for a slightly more complicated signal,

$$y(t) = e^{\alpha t} \sin(10\pi t)$$

for $\alpha = 0.1$, *is shown in Figure 6.16 and the corresponding reconstruction is shown in Figure 6.17. We will look into more interesting features in discrete signals later in the chapter.*

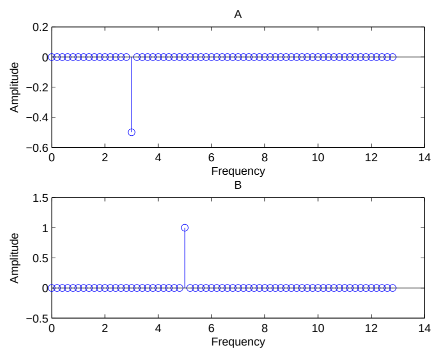


Figure 6.15: Computed discrete Fourier coefficients for $\sin(10\pi t) - \frac{1}{2}\cos(6\pi t)$ with $N = 128$ points on the interval $[0, 5]$.

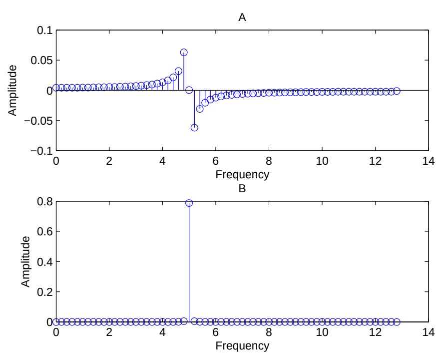


Figure 6.16: Computed discrete Fourier coefficients for $y(t) = e^{\alpha t} \sin(10\pi t)$ with $\alpha = 0.1$ and $N = 128$ points on the interval $[0, 5]$.

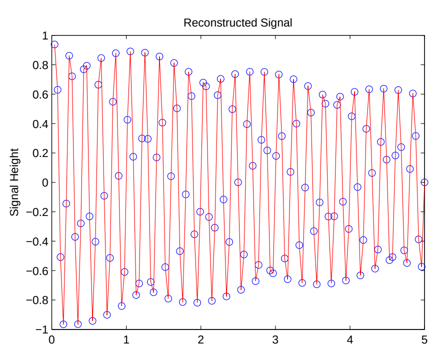


Figure 6.17: Reconstruction of $y(t) = e^{\alpha t} \sin(10\pi t)$ with $\alpha = 0.1$ from its Fourier coefficients.

6.8 MATLAB Implementation

Discrete Fourier Transforms and FFT are easily implemented in computer applications. In this section we describe the MALAB routines used in this course.

6.8.1 MATLAB for the Discrete Fourier Transform

In this section we provide implementations of the discrete trigonometric transform in MATLAB. The first implementation is a straightforward one which can be done in most programming languages. The second implementation makes use of matrix computations that can be performed in MATLAB or similar programs like GNU Octave. Sums can be done with matrix multiplication, as described in the next section. This eliminates the loops in the first program below and speeds up the computation for large data sets.

Direct Implementation

The following code was used to produce Figure 6.12. It shows a direct implementation using loops to compute the trigonometric DFT as developed in this chapter.

```
%
% DFT in a direct implementation
%
% Enter Data in y
y=[7.6 7.4 8.2 9.2 10.2 11.5 12.4 13.4 13.7 11.8 10.1 ...
    9.0 8.9 9.5 10.6 11.4 12.9 12.7 13.9 14.2 13.5 11.4 10.9 8.1];
% Get length of data vector or number of samples
N=length(y);
% Compute Fourier Coefficients
for p=1:N/2+1
    A(p)=0;
    B(p)=0;
    for n=1:N
        A(p)=A(p)+2/N*y(n)*cos(2*pi*(p-1)*n/N)';
        B(p)=B(p)+2/N*y(n)*sin(2*pi*(p-1)*n/N)';
    end
end
A(N/2+1)=A(N/2+1)/2;
% Reconstruct Signal - pmax is number of frequencies used
% in increasing order
pmax=13; for n=1:N
    ynew(n)=A(1)/2;
    for p=2:pmax
        ynew(n)=ynew(n)+A(p)*cos(2*pi*(p-1)*n/N)+B(p) ...
```

```
                *sin(2*pi*(p-1)*n/N);
    end
end
% Plot Data
plot(y,'o')
% Plot reconstruction over data
hold on
plot(ynew,'r')
hold off
title('Reconstruction of Monthly
Mean Surface Temperature')
xlabel('Month')
ylabel('Temperature')
```

The next routine shows how we can determine the spectral content of a signal, given in this case by a function and not a measured time series. The output is the original data and reconstructed Fourier series in Figure 1, the trigonometric DFT coefficients in Figure 2, and the the power spectrum in Figure 3. This code will be referred to as **ftex.m**.

```
% ftex.m
% IMPLEMENTATION OF DFT USING TRIGONOMETRIC FORM
% N = Number of samples
% T = Record length in time
% y = Sampled signal
%
clear
N=128;
T=5;
dt=T/N;
t=(1:N)*dt;
f0=30;
y=sin(2*pi*f0*t);

% Compute arguments of trigonometric functions
for n=1:N
    for p=0:N/2
        Phi(p+1,n)=2*pi*p*n/N;
    end
end

% Compute Fourier Coefficients
for p=1:N/2+1
    A(p)=2/N*y*cos(Phi(p,:))';
    B(p)=2/N*y*sin(Phi(p,:))';
end
A(1)=2/N*sum(y);
```

```
A(N/2+1)=A(N/2+1)/2;
B(N/2+1)=0;

% Reconstruct Signal - pmax is number of frequencies
% used in increasing order
pmax=N/2;
for n=1:N
    ynew(n)=A(1)/2;
    for p=2:pmax
        ynew(n)=ynew(n)+A(p)*cos(Phi(p,n))+B(p)*sin(Phi(p,n));
    end
end

% Plot Data
figure(1)
plot(t,y,'o')

% Plot reconstruction over data
hold on
plot(t,ynew,'r')
xlabel('Time')
ylabel('Signal Height')
title('Reconstructed Signal')
hold off

% Compute Frequencies
n2=N/2;
f=(0:n2)/(n2*2*dt);

% Plot Fourier Coefficients
figure(2)
subplot(2,1,1)
stem(f,A)
xlabel('Frequency')
ylabel('Amplitude')
title('A')
subplot(2,1,2)
stem(f,B)
xlabel('Frequency')
ylabel('Amplitude')
title('B')

% Plot Fourier Spectrum
figure(3)
Power=sqrt(A.^2+B.^2);
stem(f,Power(1:n2+1))
```

```
xlabel('Frequency')
ylabel('Power')
title('Periodogram')

% Show Figure 1
figure(1)
```

Compact Implementation

The next implementation uses matrix products to eliminate the for loops in the previous code. The way this works is described in the next section.

```
%
% DFT in a compact implementation
%
% Enter Data in y
y=[7.6 7.4 8.2 9.2 10.2 11.5 12.4 13.4 13.7 11.8 ...
    10.1 9.0 8.9 9.5 10.6 11.4 12.9 12.7 13.9 14.2 ...
    13.5 11.4 10.9 8.1];
N=length(y);

% Compute the matrices of trigonometric functions
p=1:N/2+1;
n=1:N;
C=cos(2*pi*n'*(p-1)/N);
S=sin(2*pi*n'*(p-1)/N);

% Compute Fourier Coefficients
A=2/N*y*C;
B=2/N*y*S;
A(N/2+1)=A(N/2+1)/2;

% Reconstruct Signal - pmax is number of frequencies used
% in increasing order
pmax=13;
ynew=A(1)/2+C(:,2:pmax)*A(2:pmax)'+S(:,2:pmax)*B(2:pmax)';

% Plot Data
plot(y,'o')

% Plot reconstruction over data
hold on
plot(ynew,'r')
hold off
title('Reconstruction of Monthly
Mean Surface Temperature')
xlabel('Month')
ylabel('Temperature')
```

6.8.2 Matrix Operations for MATLAB

The beauty of using programs like MATLAB or GNU Octave is that many operations can be performed using matrix operations and that one can perform complex arithmetic. This eliminates many loops and make the coding of computations quicker. However, one needs to be able to understand the formalism. In this section we elaborate on these operations so that one can see how the MATLAB implementation of the direct computation of the DFT can be carried out in compact form as shown previously in the MATLAB Implementation section. This is all based upon the structure of MATLAB, which is essentially a MATrix LABoratory.

A key operation between matrices is matrix multiplication. An $n \times m$ matrix is simply a collection of numbers arranged in n rows and m columns. For example, the matrix $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ is a 2×3 matrix. The entries (elements) of a general matrix A can be represented as a_{ij} which represents the ith row and jth column.

Given two matrices, A and B, we can define the multiplication of these matrices when the number of columns of A equals the number of rows of B. The product, which we represent as matrix C, is given by the ijth element of C. In particular, we let A be a $p \times m$ matrix and B an $m \times q$ matrix. The product, C, will be a $p \times q$ matrix with entries

$$\begin{aligned} C_{ij} &= \textstyle\sum_{k=1}^{m} a_{ik} b_{kj}, \quad i = 1, \ldots, p, \quad j = 1, \ldots, q, \\ &= a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots a_{im} b_{mj}. \end{aligned} \tag{6.79}$$

If we wanted to compute the sum $\sum\limits_{n=1}^{N} a_n b_n$, then in a typical programming language we could use a loop, such as

```
Sum = 0 Loop n from 1 to N
    Sum = Sum + a(n)*b(n)
End Loop
```

In MATLAB we could do this with a loop as above, or we could resort to matrix multiplication. We can let a and b be $1 \times n$ and $n \times 1$ matrices, respectively. Then, the product would be a 1×1 matrix; namely, the sum we are seeking. However, these matrices are not always of the suggested size.

A $1 \times n$ matrix is called a row vector and a $1 \times n$ matrix is called a column vector. Often we have that both are of the same type. One can convert a row vector into a column vector, or vice versa, using the matrix operation called a *transpose*. More generally, the transpose of a matrix is defined as follows: A^T has the elements satisfying $\left(A^T\right)_{ij} = a_{ji}$. In MATLAB, the transpose if a matrix A is A'.

Thus, if we want to perform the above sum, we have $\sum\limits_{n=1}^{N} a_n b_n = \sum\limits_{n=1}^{N} a_{1n} b_{n1}$. In particular, if both **a** and **b** are row vectors, the sum in MATLAB is given by **ab**′, and if they are both row vectors, the sum is **a**′**b**. This notation is much easier to type.

In the computation of the DFT, we have many sums. For example, we want to compute the coefficients of the sine functions,

$$b_p = \frac{2}{N}\sum_{n=1}^{N} y_n \sin(\tfrac{2\pi pn}{N}), \quad p = 0, \ldots, N/2 \tag{6.80}$$

The sum can be computed as a matrix product. The function y only has values at times t_n. This is the sampled data. We can represent it as a vector. The sine functions take values at arguments (angles) depending upon p and n. So, we can represent the sines as an $N \times (N/2+1)$ or $(N/2+1) \times N$ matrix. Finding the Fourier coefficients becomes a simple matrix multiplication, ignoring the prefactor $\frac{2}{N}$. Thus, if we put the sampled data in a $1 \times N$ vector **Y** and put the sines in an $N \times (\frac{N}{2}+1)$ vector **S**, the Fourier coefficient will be the product, which has size $1 \times \left(\frac{N}{2}+1\right)$. Thus, in the code we see that these coefficients are computed as **B=2/N*y*S** for the given **y** and **B** matrices. The A coefficients are computed in the same manner. Comparing the two codes in that section, we see how much easier it is to implement. However, the number of multiplications and additions has not decreased. This is why the FFT is generally better. But, seeing the direct implementation helps one to understand what is being computed before seeking a more efficient implementation, such as the FFT.

6.8.3 MATLAB Implementation of FFT

In this section we provide implementations of the Fast Fourier Transform in MATLAB. The MATLAB code provided in Section 6.8.1 can be simplified by using the built-in **fft** function.

Much of the MATLAB code described here can be run directly in GNU Octave or easily mapped to other programming environments.

```
% fanal.m
%
% Analysis Using FFT
%
clear
n=128;
T=5;
dt=T/n;
f0=6.2;
f1=10;
y=sin(2*pi*f0*(1:n)*dt)+2*sin(2*pi*f1*(1:n)*dt);
% y=exp(-0*(1:n)*dt).*sin(2*pi*f0*(1:n)*dt);
Y=fft(y,n);
n2=n/2;
Power=Y.*conj(Y)/n^2;
f=(0:n2)/(n2*2*dt);
```

```
stem(f,2*Power(1:n2+1))
xlabel('Frequency')
ylabel('Power')
title('Periodogram')
figure(2)
subplot(2,1,1)
stem(f,real(Y(1:n2+1)))
xlabel('Frequency')
ylabel('Amplitude')
title('A')
subplot(2,1,2)
stem(f,imag(Y(1:n2+1)))
xlabel('Frequency')
ylabel('Amplitude')
title('B')
figure(1)
```

One can put this into a function which performs the FFT analysis. Below is the function **fanalf**, which can make the main program more compact.

```
function z=fanalf(y,T)
%
% FFT Analysis Function
%
% Enter Data in y and record length T
% Example:
%    n=128;
%    T=5;
%    dt=T/n;
%    f0=6.2;
%    f1=10;
%    fanal(sin(2*pi*f0*(1:n)*dt)+2*sin(2*pi*f1*(1:n)*dt));
% or
%    fanal(sin(2*pi*6.2*(1:128)/128*5),5);
n=length(y);
dt=T/n;
Y=fft(y,n);
n2=floor(n/2);
Power=Y.*conj(Y)/n^2;
f=(0:n2)/(n2*2*dt);
z=Power;
stem(f,2*Power(1:n2+1))
xlabel('Frequency')
ylabel('Power')
title('Periodogram')
```

Examples of the use for doing a spectral analysis of functions, data sets, and sound files are provided below.

This code snippet can be used with **fanalf** to analyze a given function.

```
% fanal2.m

clear
n=128;
T=5;
dt=T/n;
f0=6.2;
f1=10;
t=(1:n)*dt;
y=sin(2*pi*f0*t)+2*sin(2*pi*f1*t);
fanalf(y,T);
```

This code snippet can be used with **fanalf** to analyze a data set inserted in vector **y**.

```
y=[7.6 7.4 8.2 9.2 10.2 11.5 12.4 13.4 13.7 11.8 ...
    10.1 9.0 8.9 9.5 10.6 11.4 12.9 12.7 13.9 14.2 ...
    13.5 11.4 10.9 8.1];
n=length(y);
y=y-mean(y);
T=24;
dt=T/n;
fanalf(y,T);
```

One can also enter data stored in an ASCII file. This code show how this is done.

```
[year,y]=textread('sunspot.txt','%d %f');
n=length(y);
t=year-year(1);
y=y-mean(y);
T=t(n);
dt=T/n;
fanalf(y',T);
```

This code snippet can be used with **fanalf** to analyze a given sound stored in a wav file.

```
[y,NS,NBITS]=wavread('sound1.wav');
n=length(y);
T=n/NS
dt=T/n;
figure(1)
plot(dt*(1:n),y)
figure(2)
fanalf(y,T);
```

Problems

1. Find the imaginary part of the Fourier transform of the finite wave train $f(t) = \cos \omega_0 t$, $t \in [0, T]$ in Example . For $\omega_0 = \pi, 5, frm{-}e\pi, 10$, plot the real, imaginary, and modulus of the Fourier transform.

2. Recall that $\text{Ш}(ax) = \frac{1}{|a|}\text{comb}_{\frac{1}{a}}(x)$. Find a similar relation for $\text{comb}_T(at)$ in terms of the Shah function.

3. Write the following sums in the form $a\text{comb}_b(t)$ for some constants a and b.

a. $\text{comb}_T(\alpha t)$.

b. $\text{comb}_T(t+k)$, k and integer.

4. Evaluate $\mathcal{F}^{-1}[\text{comb}_\Omega(\omega)]$.

5. Evaluate the following sums:

a. $\sum_{n=1}^{8} e^{\pi i(n/4)-2}$

b. $\sum_{n=-7}^{8} \sin\left(\frac{\pi pn}{8}\right) \sin\left(\frac{\pi qn}{8}\right)$

6. Prove

$$\sum_{n=0}^{N-1} \sin\left(\frac{2\pi pn}{N}\right) \sin\left(\frac{2\pi qn}{N}\right) = \begin{cases} N/2, & p = q \neq N/2 \\ 0, & \text{otherwise} \end{cases}.$$

7. Compute a table for the trigonometric discrete Fourier transform for the following and sketch A_k, B_k and $A_k^2 + B_k^2$.

a. $y_n = n^2$, $n \in [0, 15]$.

b. $y_n = \cos n$, $n \in [0, 15]$.

8. Here you will prove a shift theorem for the discrete exponential Fourier transform: The transform of y_k is given as $Y_j = \frac{1}{N}\sum_{k=1}^{N} y_k e^{-2\pi ijk/N}$, $j = 0, 1, \ldots N-1$. Show that for fixed $n \in [0, N-1]$ the discrete transform of y_{k-n} is $Y_j e^{-2\pi ijn/N}$, $j \in [0, N-1]$.

9. Consider the finite wave train

$$f(t) = \begin{cases} 2\sin 4t, & 0 \leq t \leq \pi \\ 0, & \text{otherwise.} \end{cases}$$

a. Plot this function.

b. Find the Fourier transform of $f(t)$.

c. Find the Fourier coefficients in the series expansion

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos 2nt + b_n \sin 2nt.$$

10. A plot of $\hat{f}(\omega)$ from the last problem is shown in Figure 7.72.

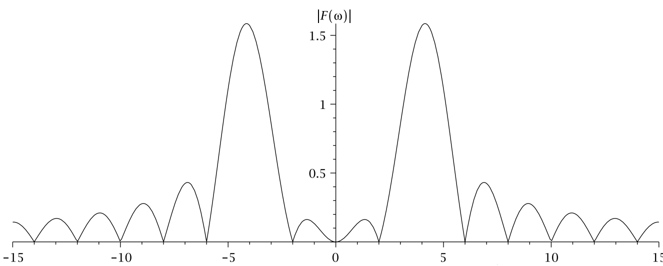


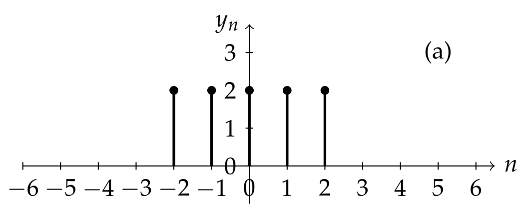
Figure 6.18: Fourier transform of the wave train in Problems 9-10.

a. What do the main peaks tell you about $f(t)$?

b. Use this plot to explain how the Fourier coefficients are related to the Fourier transform.

c. How would this plot appear if $f(t)$ were nonzero on a larger time interval?

11. Consider the sampled function in Figure 6.20. For each case compute the discrete Fourier transform.



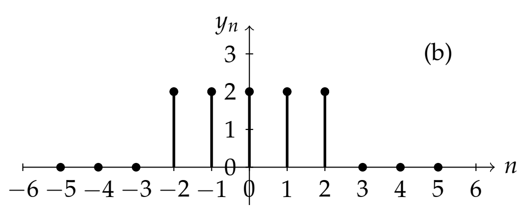
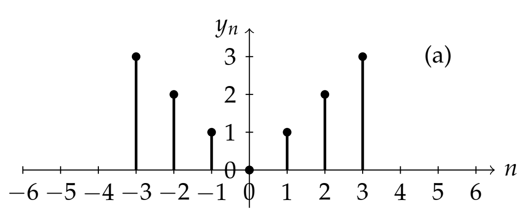


Figure 6.19: Figure for Problem 12.

12. Consider the sampled function in Figure 6.20. For each case compute the discrete Fourier transform.



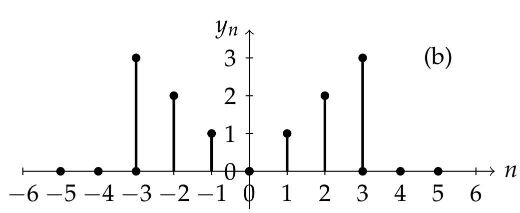


Figure 6.20: Figure for Problem 12.

13. The goal of this problem is to use MATLAB to investigate the use of the Discrete Fourier Transform (DFT) for the spectral analysis of time series. Running the m-files is done by typing the filename without the extension (.m) after the prompt (») in the Command Window, which is the middle panel of the MATLAB program. Note, most of the code provided can be run in GNU Octave.

1. Analysis of simple functions.

   a. File **ftex.m** - See the MATLAB section for the code.
   This file is a MATLAB program implementing the discrete Fourier transform using trigonometric functions like that derived in the text. The input is a function, sometimes with different frequencies. The output is a plot of the data points and the function fit, the Fourier coefficients and the periodogram giving the power spectrum.

      i. Save the file **ftex.m** in your working directory under MATLAB.

      ii. View the file by entering edit **ftex**. Note how the first function is defined by the variable **y**.

      iii. Run the file by typing **ftex** in MATLAB's Command window.

      iv. Change the parameters in **ftex.m**, remembering the original ones. In particular, change the number of points, **N**, (keeping them even), the frequency, **f0**, and the record length, **L**. Note the effects. If you get an error, enter clear and try again. Always save the m-file after making any changes before running ftex.

      v. Reset the parameters to the original values. What happens for frequencies of **f0** $= 5$, 5.5 and 5.6? Is this what you expected?

      vi. Repeat the last set of frequencies for double the record length, **T**. Is there a change?

      vii. Reset the parameters. Put in frequencies of **f0** $= 20, 30, 40$. What frequencies are present in the periodogram? Is this what you expected?

   b. Look at sums of several trigonometric functions.

      i. Reset the parameters.

      ii. Change the function in **y** to `y=sin(2*pi*f0*t)+sin(2*pi*f1*t);` and add a line defining **f1**. Start with **f1** $= 3$; and look at several other values for the two frequencies. Try different amplitudes; for example, `3*sin(2*pi*f0*t)` has an amplitude of 3. Record your observations.

      iii. Change one of the sines to a cosine. What is the effect? What happens when the sine and cosine terms have the same frequency?

c. Investigate non-sinusoidal functions.

  i. Investigate the following functions:

    1. `y=t;`
    2. `y=t.^2;`
    3. `y=sin(2*pi*f0*(t-T/5))./(t-T/5);` What is this function?
    4. `y(1,1:M)=ones(1,M); y(1,M+1:N)=zeros(1,N-M);` Start with $M = N/2$; What is this function? How are the last two problems related? Do they relate to anything from earlier class lectures? What effect results from changing **M**?
    5. Try multiplying the function in 4 by a simple sinusoid; for example, add the line `y=sin(2*pi*f0*t).*y`, for $M = N/2$. How does this affect what you had gotten for the sinusoid without multiplication?

2. Use the FFT function.

   In MATLAB there is a built in set of functions, **fft** and **ifft** for the computation of the Discrete Exponential Transform and its inverse using the Fast Fourier Transform (FFT). The files needed to do this are **fanal.m** or **fanal2.m** and **fanalf.m**. [See Section 6.8.3 for the code.] Put these codes into the MATLAB editor and see what they look like. Note that **fanal.m** was split into the two files **fanal2.m** and **fanalf.m**. This will allow us to be confident when we later create new data files and then using fanalf.m. Test this set of functions for simple sine functions to see that you get results similar to Part 1. First run **fanal** and then **fanal2**. Is there any difference between these last two approaches?

3. Analysis of data sets.

   One often does not have a function to analyze. Some measurements are made over time of a certain quantity. This is called a time series. It could be a set of data describing things like the stock market fluctuations, sunspots activity, ocean wave heights, etc. Large sets of data can be read into y and small sets can be input as vectors. We will look at how this can be done. After the data is entered, one can analyze the time series to look for any periodic behavior, or its frequency content. In the two cases below, make sure you look at the original time series using **plot(y,âĂŹ+âĂŹ)**.

   a. Ocean Waves
      In this example the data consists of monthly mean sea surface temperatures (^{o}C) at one point over a 2 year period. The temperatures are placed in **y** as a row vector. Note how the data is continued to a second line through the use of ellipsis. Also, one typically subtracts the average from the data. What affect should this have on the spectrum? Copy the following code into a new m-file called **fdata.m** and run **fdata**. Determine the dominant periods in the monthly mean sea surface temperature.

```
y=[7.6 7.4 8.2 9.2 10.2 11.5 12.4 13.4 13.7 11.8 ...
    10.1 9.0 8.9 9.5 10.6 11.4 12.9 12.7 13.9 14.2 ...
    13.5 11.4 10.9 8.1];
n=length(y);
y=y-mean(y);
T=24;
dt=T/n;
fanalf(y,T);
```

b. Sunspots

Sunspot data was exported to a text file. Download the file sunspot.txt. You can open the data in Notepad and see the two column format. The times are given as years (like 1850). This example shows how a time series can be read and analyzed. From your spectrum, determine the major period of sunspot activity. Note that we first subtracted the average so as not to have a spike at zero frequency. Copy and paste into the editor and save as **fanaltxt.m**. Note: Copy and Paste of single quotes often does not work correctly. Retype the single quotes after copying.

```
[year,y]=textread('sunspot.txt','%d %f');
n=length(y);
t=year-year(1);
y=y-mean(y);
T=t(n);
dt=T/n;
fanalf(y',T);
```

4. Analysis of sounds. Sounds can be input into MATLAB. You can create your own sounds in MATLAB or sound editing programs like Audacity or Goldwave to create audio files. These files can be input into MATLAB for analysis.

Save the following sample code as **fanalwav.m** and save the first sound file. This code shows how one can read in a WAV file. There will be two plots, the first showing the wave profile and the second giving the spectrum. Try some of the other wav files and report your findings. Note: Copy and Paste of single quotes often does not work correctly. Retype the single quotes after copying.

```
[y,NS,NBITS]=wavread('sound1.wav');
n=length(y);
T=n/NS
dt=T/n;
figure(1)
plot(dt*(1:n),y)
figure(2)
fanalf(y,T);
```

Several wav files are be provided for you to analyze. If you are able to hear the sounds, you can run them from MATLAB by typing `sound(y,NS)`. In fact, you can even create your own sounds based upon simple functions and save them as wav files. For example, try the following code: Note: Copy and Paste of single quotes often does not work correctly. Retype the single quotes after copying.

```
smp=11025;
t=(1:2000)/smp;
y=0.75*sin(2*pi*440*t);
sound(y,smp,8);
wavwrite(y,smp,8, 'myfile.wav');
```

Try some other functions, using several frequencies. If you get a clipping error, then reduce the amplitudes you are using.

7
Signal Analysis

There's no sense in being precise when you don't even know what you're talking about. - John von Neumann (1903 - 1957)

7.1 Introduction

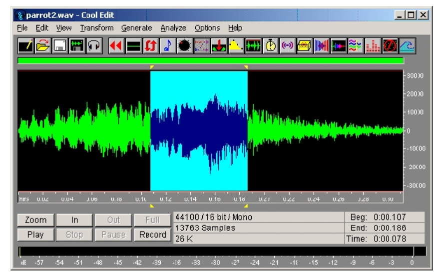


Figure 7.1: Cool Edit displaying a WAV file and its properties.

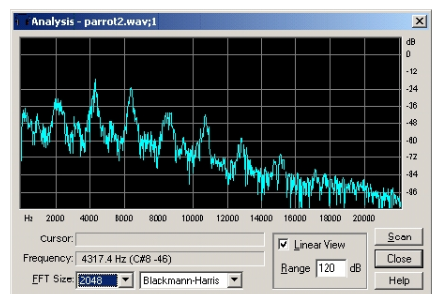


Figure 7.2: Cool Edit displaying the spectrum of a WAV file.

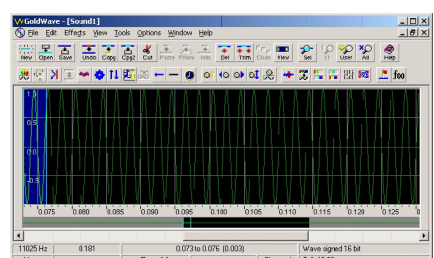


Figure 7.3: Goldwave displaying a simple tone.

It is now time to look back at the Introducton and see what it was that we promised to do in this course. The goal was to develop enough tools to begin to understand what happens to the spectral content of signals when analog signals are discretized. We started with a study of Fourier series and just ended with discrete Fourier transforms. We have seen how Fourier transform pairs, $f(t)$ and $\hat{f}(\omega)$, are affected by recording a signal over a finite time T at a sampling rate of f_s. This in turn lead to the need for discrete Fourier transforms (DFTs). However, we have yet to see some of the effects of this discretization on the information that we obtain from the spectral analysis of signals in practice. In this chapter we will look at results of applying DFTs to a variety of signals.

The simplest application of this analysis is the analysis of sound. Music, which is inherently an analog signal, is recorded over a finite time interval and is sampled at a rate that yields pleasing sounds that can be listened to on the computer, a CD, or in an MP3 player.

You can record and edit sounds yourself. There are many audio editing packages that are available. We have successfully used these packages plus some minimal applets and mathematics packages to introduce high school students and others with a minimal mathematics background to the Fourier analysis of sounds. As we have seen, we need only understand that signals can be represented as sums of sinusoidal functions of different frequencies and amplitudes.

For example, we have had students working with musical instruments, bird sounds, dolphin sounds, ECGs, EEGs, digital images, and other forms of recorded signals or information. One just needs to find a way to determine the frequency content of the signal and then pick out the dominant frequencies to reconstruct the signal.

There are many packages that can be used to display sound and time series. We will see how to use MATLAB, although one could use Maple or Mathematica to import and analyze sounds. There are also stand alone editors like Cool Edit Pro (bought out by Adobe in 2003 and renamed Adobe Audition), Audacity (http://audacity.sourceforge.net/) an open source editor, or Goldwave (http://www.goldwave.com/), which allows one to input a formula and "play" it.

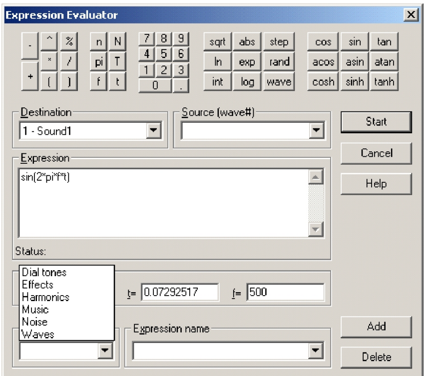


Figure 7.4: Goldwave display of function editor.

A sample Cool Edit session is shown in Figure 7.1. In this figure is displayed the sound print of a parrot. It is obviously a complex signal, made up of many harmonics. Highlighting a part of the signal, one can look at the frequency content of this signal. Such a spectrum is shown in Figure 7.2. Notice the spikes every couple of thousand Hertz.

Not many fancy, but inexpensive, sound editors have a frequency analysis component like Cool Edit had. One has to go out on to the web and search for features that do not just entail editing sounds for MP3 players. Goldwave allows one to enter a formula and then listen to the corresponding sounds. This is also a feature not found in most editors. However, it is a useful tool that takes little "programming" to connect the mathematics to the signal. Cool Edit and others have a feature to generate tones, but this is more exact. The interface for Goldwave is shown in Figure 7.3 and the function editor is in 7.4. However, there are plenty of other editors. In the early 2000's the HASA (Handheld Audio Spectrum Analyzer) application shown in Figure 7.5 was a good tool for pocket PCs. Also, spectrum analyzers are available for mobile devices, such as the iPhone (e.g. Pocket RTA - Spectrum Analyser).

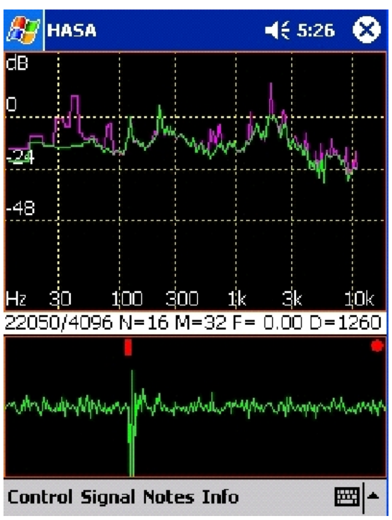


Figure 7.5: HASA for pocket PCs.

7.2 *Periodogram Examples*

THE NEXT STEP IN THE ANALYSIS IS TO UNDERSTAND the output of the discrete transform, or the Fast Fourier Transform (FFT), that is generated by such programs. Often we see spectrograms or periodograms. We should understand what it is that they produce. As an example, lets say we have the sum of a sine and a cosine function with different frequencies and amplitudes. We could represent the discrete Fourier coefficients as a_n's and b_n's, like we have computed many times in the course, in simple plots of the coefficients vs n (or the frequency) such as shown in Figure 7.6. In this case there is a cosine contribution of amplitude two at frequency f_4 and a sine contribution of amplitude one at frequency f_2. It takes two plots to show both the a_n's and b_n's. However, we are often only interested in the energy content at each frequency. For this example, the last plot in Figure 7.7 shows the spectral content in terms of the modulus of the signal.

For example, $\cos 5t$ and $3\sin 5t$ would have spikes in their respective plots at the same frequency, $f = \frac{5}{2\pi}$. As noted earlier in Equation (6.30), we can write the sum $\cos 5t + 3\sin 5t$ as a single sine function with an amplitude $c_n = \sqrt{a_n^2 + b_n^2}$. Thus, a plot of the "modulus" of the signal is used more often. However, in the examples we will display both forms to bring home

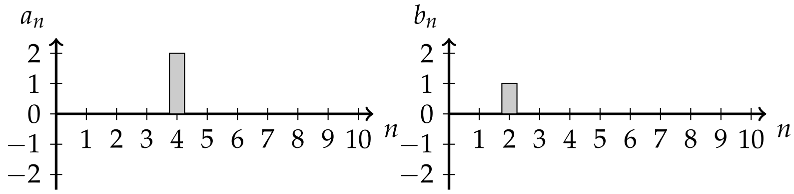


Figure 7.6: This Figure shows the spectral coefficients for a signal of the form $f(t) = 2\cos 4t + \sin 2t$.

the relationship between the trigonometric and exponential forms of the Fourier spectrum of the signal.

Once one has determined the Fourier coefficients, then one can reconstruct the signal. In the case that one has the exact components, then the reconstruction should be perfect as shown for the previous example in Figure 7.6. The reconstruction in this case gave the plot in Figure 7.8. However, for real signals one does not know ahead of time what the actual frequencies are that made up the signal.

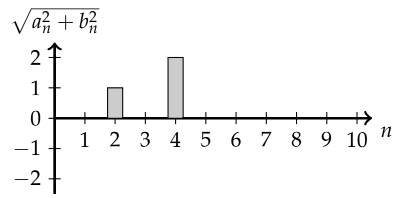


Figure 7.7: This Figure shows the spectrum for a signal of the form $f(t) = 2\cos 4t + \sin 2t$.

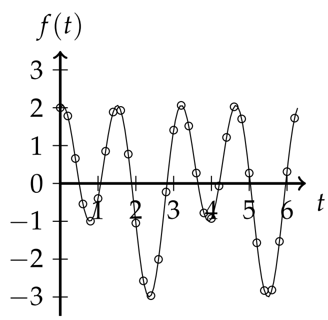


Figure 7.8: This Figure shows the original signal of the form $f(t) = 2\cos 4t + \sin 2t$ and a reconstruction based upon the series expansion.

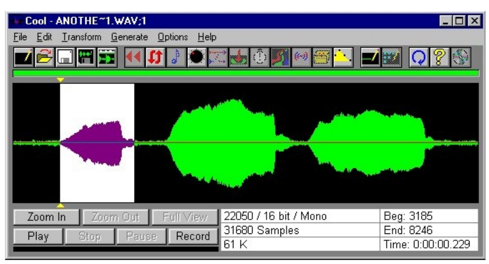


Figure 7.9: A piece of a typical bird sound.

For example, one could analyze a bird sound like the one shown in Figure 7.9. We capture a part of the sound and look at its spectrum. An example is shown in Figure 7.10. Notice that the spectrum is not very clean, although a few peaks stand out. We had a group of high school students carry out this procedure. The students picked out a few of the dominant frequencies and the corresponding amplitudes. Using just a few frequencies, they reconstructed the bird signals. In Figure 7.11 we show the original and reconstructed signals, respectively. While these might not look exactly the same, they do sound very similar.

There are different methods for displaying the Fourier spectrum of signals. Here we define some of these.

Definition 7.1. A *spectrogram* is a three-dimensional plot of the energy of

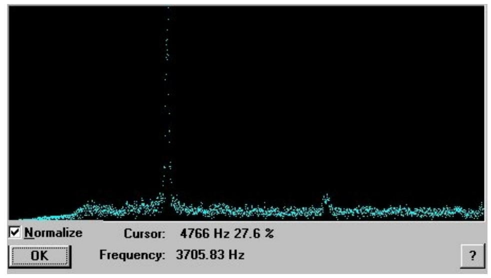


Figure 7.10: A Fourier analysis of the bird sound in Figure 7.9.

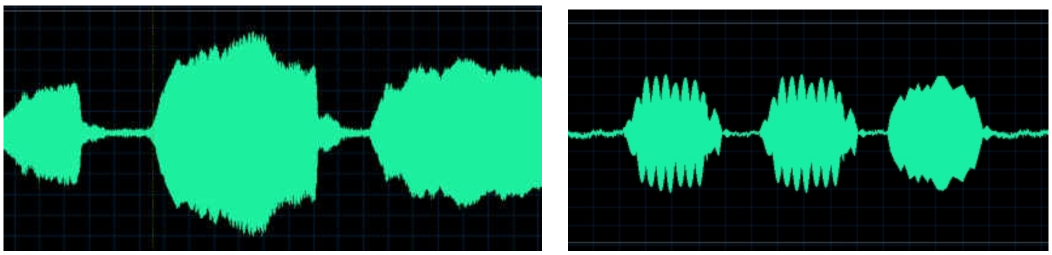


Figure 7.11: Analysis and reconstruction of a bird sound.

the frequency content of a signal as it changes over time.

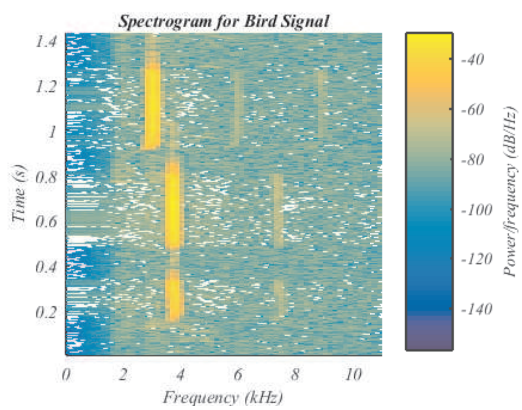


Figure 7.12: Example of spectrogram for the bird sound.

An example of a spectrogram for the bird sound in Figure 7.9 is provided in Figure 7.12. This figure was created using MATLAB's built-in function (in the Signal Processing Toolbox):

```
[y,NS,NBITS]=wavread('firstbird.wav');
spectrogram(y,128,120,128,NS);
title('Spectrogram for Bird Signal')
ylabel('Time (s)')
```

The spectrogram is created using what is called the short-time Fourier transform, or STFT. This function divides a long signal into smaller blocks, or windows, and then computes the Fourier transform on each block. This allows one to track the changes in the spectrum content over time. In Figure 7.12 one can see three different blobs in the 3kHz-4kHz range at different times, indicating how the three chirps of the bird can be picked up. This gives more information than a Fourier analysis over the entire record length.

Definition 7.2. The *power spectrum* is a plot of the portion of a signal's power (energy per unit time) falling within given frequency bins. We can either plot the Fourier coefficients, or the modulus of the Fourier transform.

Definition 7.3. Plots of $c_n = \sqrt{a_n^2 + b_n^2}$ or c_n^2 vs frequency are sometimes called *periodograms*.

An example of a periodogram for the bird sound in Figure 7.9 is provided in Figure 7.13. A periodogram can be created using MATLAB's built-in function (in the Signal Processing Toolbox):

```
[y,NS,NBITS]=wavread('firstbird.wav');
periodogram(y,[],'onesided',1024,NS)
```

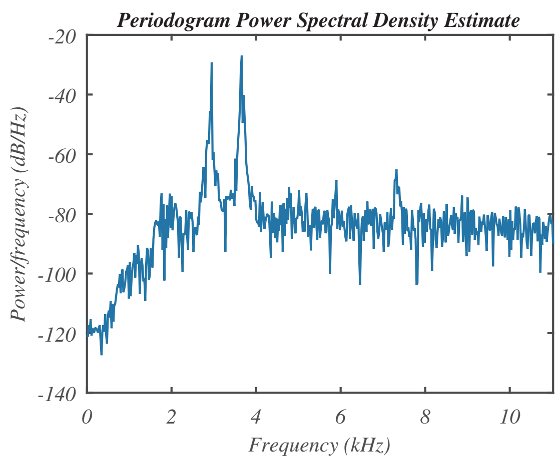


Figure 7.13: Example of periodogram for the bird sound.

There are many other types of applications. We have had students studying the oscillations of mass-spring systems and vibrating beams in differential equations. The setups are shown in Figure 7.14. On the left is a mass-spring system situated above a motion probe. The data is collected using an interface to a handheld computer. (More recently pocket PCs and other mobile devices have been used.) On the left is a similar setup for a clamped two-meter stick, which is clamped at different positions and the motion of end of the stick is monitored.

Of course, a simple mass on a spring exhibits the typical almost pure sinusoidal function as shown in Figure 7.15. The data is then exported to another program for analysis.

Students would learn how to fit their data to sinusoidal functions and then determine the period of oscillation as compared to the theoretical value. They could either do the fits in Maple (Figure 7.16) or Excel (Figure 7.17).

Fitting data to damped oscillations, such as shown in Figure 7.18, is more difficult. This is the type of data one gets when measuring the vertical position of a vibrating beam using the setup shown in Figure 7.14.

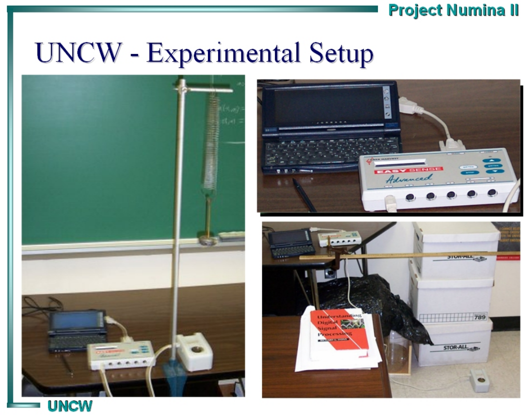


Figure 7.14: Setup for experiments for oscillations of mass-spring systems and vibrating beams. Data is recorded at a 50 Hz sampling rate using handheld devices connected to distance probes.

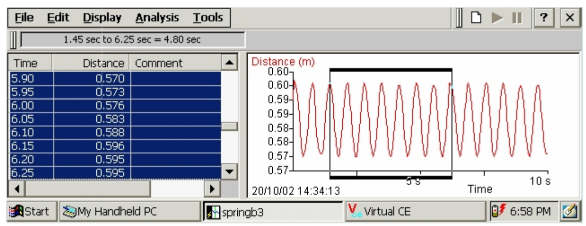


Figure 7.15: Distance vs time plot for a mass undergoing simple harmonic motion.

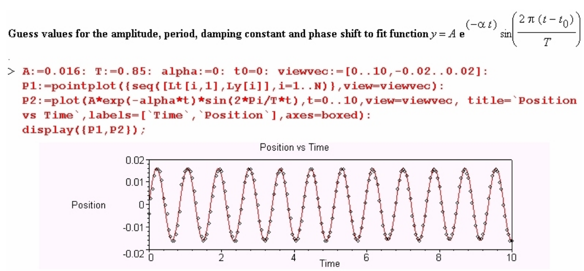


Figure 7.16: Example of fitting data in Maple.

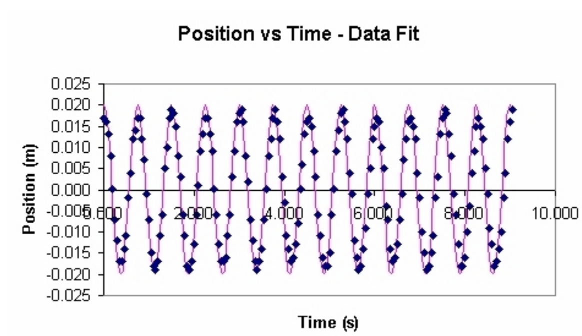


Figure 7.17: Example of fitting data in Excel.

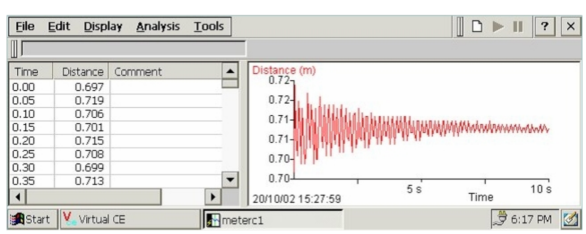


Figure 7.18: Distance vs time plot for a vibrating beam clamped at one end. The motion appears to be damped harmonic motion.

Typically, one has to try to guess several parameters in order to determine the correct period, amplitude and damping. Of course, we know that it is probably better to put such a function into a program like MATLAB and then to perform a Fourier analysis on it to pick out the frequency. We did this for the signal shown in Figure 7.19. The result of the spectral analysis is shown in Figure 7.21. Do you see any predominant frequencies? Is this a better method than trying to fit the data by hand?

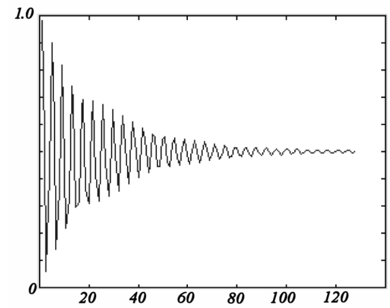


Figure 7.19: Distance vs time plot in MATLAB for a vibrating beam clamped at one end.

7.3 *Effects of Sampling*

We are interested in how well the discrete Fourier transform works with real signals. In the last section we saw a few examples of how signal analysis might be used. We will look into other applications later. For now, we want to examine the effects of discretization on signals so that we can make sense out of the analysis we might do on real signals. We need to begin with the simplest signals and then employ the DFT MATLAB program in Appendix 6.8.1 to show the results of small changes to the data.

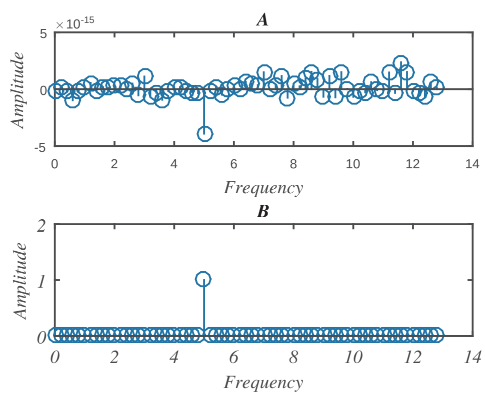


Figure 7.20: Fourier coefficients for the signal $y(t) = \sin(2\pi f_0 t)$ with $f_0 = 5.0$ Hz, $N = 128$, on $[0, 5]$.

We begin by inputting the signal. We consider the function $y(t) = \sin(2\pi f_0 t)$. We sample this function with $f_0 = 5.0$ Hz for $N = 128$ points on the interval $[0, 5]$. The Fourier Trigonometric coefficients are given in Figure 7.20. Note that the A_n's are negligible (on the order of 10^{-15}). There is a spike at the right frequency. We can also plot the periodogram as shown in Figure 7.22. We obtain the expected result of a spike at $f = 5.0$ Hz. We can reconstruct the signal as well. There appears to be agreement between the function $y(t)$ indicated by the line plot and the reconstruction indicated by the circles in 7.23.

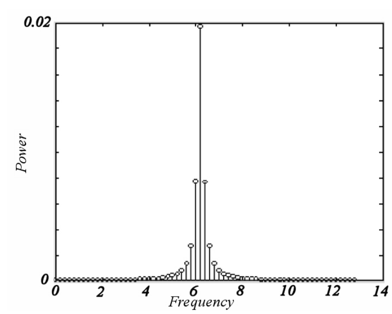


Figure 7.21: Fourier spectrum for the signal shown in Figure 7.19.

In the set of Figures 7.24 and 7.25 we show the Fourier coefficients and periodogram for the function $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ for $f_0 = f_1 = 6$ Hz. We note that the heights in Figure 7.24 are the amplitudes of the sine and cosine functions. The "peaks" are located at the correct frequencies of 6 Hz. However, in the periodogram there is no information regarding the phase shift; i.e., there is no information as to whether the frequency content

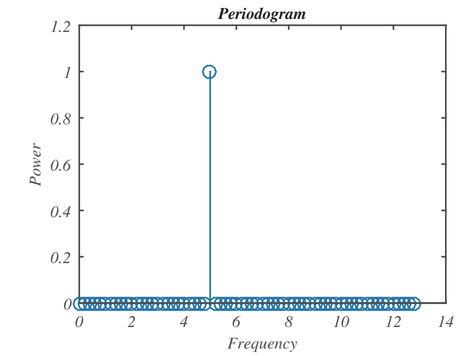


Figure 7.22: Fourier spectrum for the signal $y(t) = \sin(2\pi f_0 t)$ with $f_0 = 5.0$ Hz, $N = 128$, on $[0, 5]$.

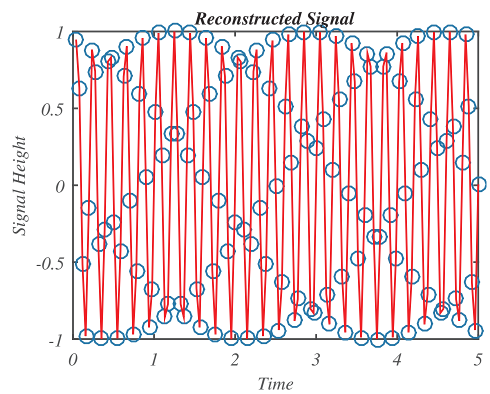


Figure 7.23: The function $y(t)$ is indicated by the line plot and the reconstruction by circles.

arises from a sine or a cosine function. We just know that all of the signal energy is concentrated at one frequency.

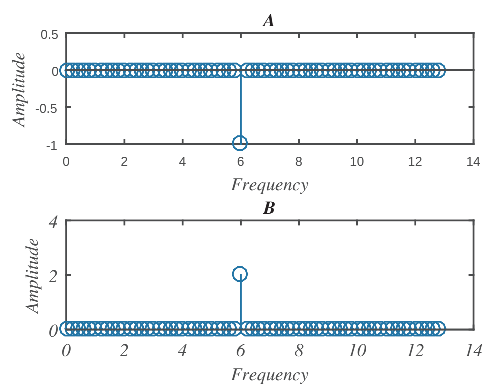


Figure 7.24: Fourier coefficients for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = f_1 = 6.0$ Hz, $N = 128$, on $[0, 5]$.

In the set of Figures 7.26 and 7.27 we show the Fourier coefficients and periodogram for the function $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ for $f_0 = 6$ Hz and $f_1 = 10$ Hz. Once again we see that the amplitudes of the Fourier coefficients are of the right height and in the right location. In the periodogram we see that the energy of the signal is distributed between two frequencies.

In the last several examples we have computed the spectra in using a sampled signal "recorded" over times in the interval $[0, 5]$ and sampled with $N = 128$ points. Sampling at $N = 256$ points leads to the periodogram in Figure 7.28. We note that increasing the number of points leads to a longer

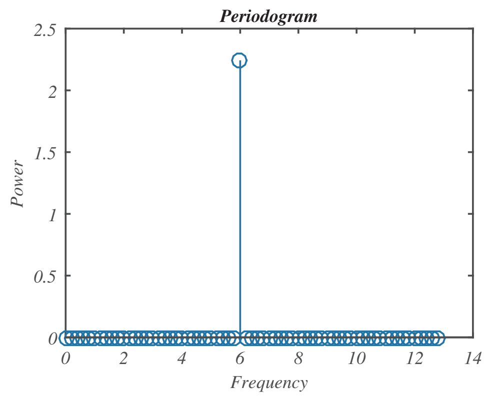


Figure 7.25: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = f_1 = 6.0$ Hz, $N = 128$, on $[0, 5]$.

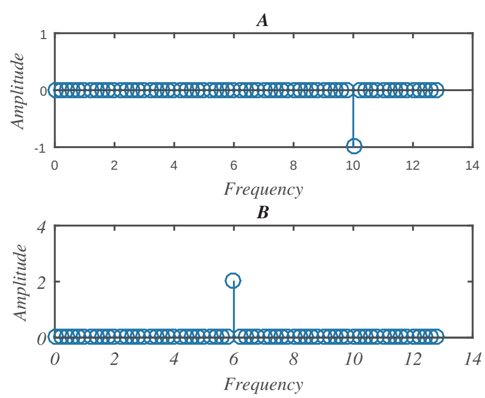


Figure 7.26: Fourier coefficients for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = 6$ Hz and $f_1 = 10$ Hz, $N = 128$, on $[0, 5]$.

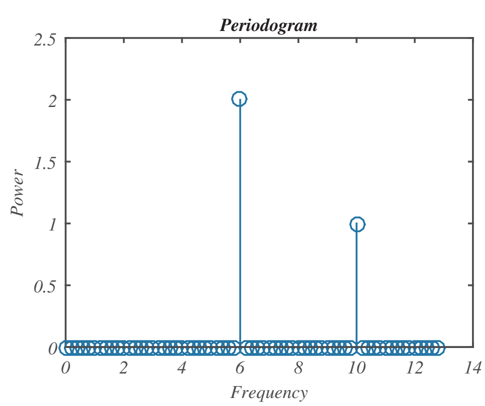


Figure 7.27: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = 6$ Hz and $f_1 = 10$ Hz. The signal was sampled with $N = 128$ points on an interval of $[0, 5]$.

interval in frequency space. In Figure 7.29 we doubled the record length to $T = 10$ and kept the number of points sampled at $N = 128$. In this case the frequency interval has become shorter. We not only lost the 10 Hz frequency, but now we have picked up a 2.8 Hz frequency. We know that the simple signal did not have a frequency term corresponding to 2.8 Hz. So, where did this come from?

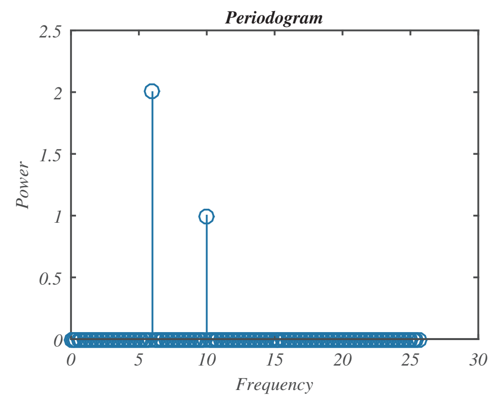


Figure 7.28: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = 6$ Hz and $f_1 = 10$ Hz. The signal was sampled with $N = 256$ points on an interval of $[0, 5]$.

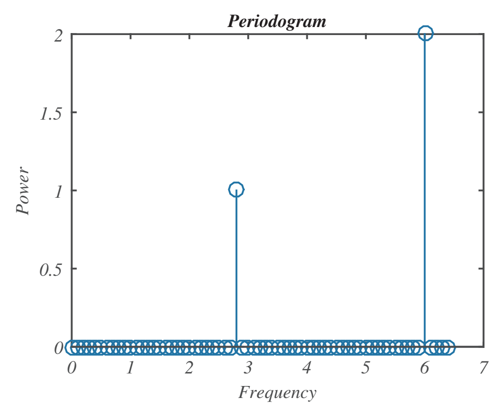


Figure 7.29: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = 6$ Hz and $f_1 = 10$ Hz. The signal was sampled with $N = 128$ points on an interval of $[0, 10]$.

Also, we note that the interval between displayed frequencies has changed. For the cases where $T = 5$ the frequency spacing is 0.2 Hz as seen in Figure 7.29. However, when we increased T to 10 s, we got a frequency spacing of 0.1 Hz. Thus, it appears that $\Delta f = \frac{1}{T}$. This makes sense, because at the beginning of the last chapter we defined

$$\omega_p = 2\pi f_p = \frac{2\pi}{T} p.$$

This gives $f_p = \frac{p}{T}$, or $f_p = 0, \frac{1}{T}, \frac{2}{T}, \frac{3}{T}, \ldots$. Thus, $\Delta f = \frac{1}{T}$. So, for $T = 5$, $\Delta f = 1/5 = 0.2$ and for $T = 10$, $\Delta f = 1/10 = 0.1$.

So, changing the record length will change the frequency spacing. But why does changing T introduce frequencies that are not there? What if we reduced N? We saw that increasing N leads to longer frequency intervals. Will reducing it lead to a problem similar to increasing T? In Figure 7.32 we see the result of using only 64 points. Yes, again we see the occurrence of

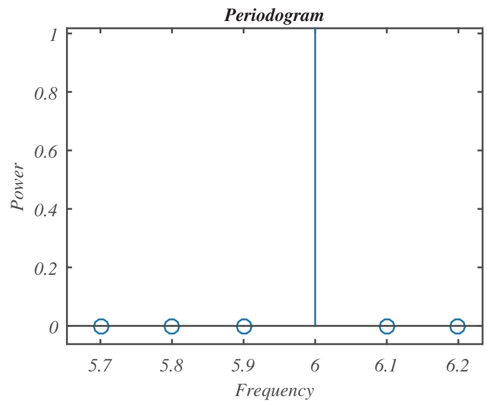


Figure 7.30: Zoomed in view of Figure 7.29.

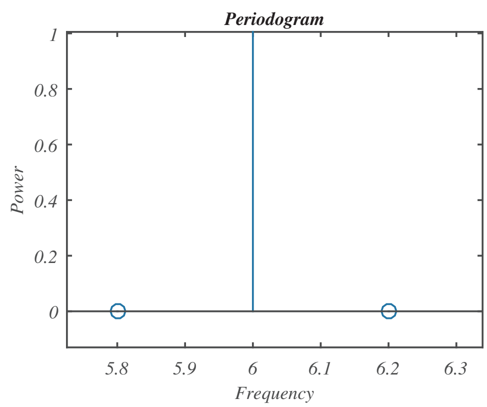


Figure 7.31: Zoomed in view of Figure 7.27. A similar view results for Figure 7.28.

a 2.8 Hz spike. Also, the range of displayed frequencies is shorter. So, the range of displayed frequencies depends upon both the numbers of points at which the signal is sampled and the record length.

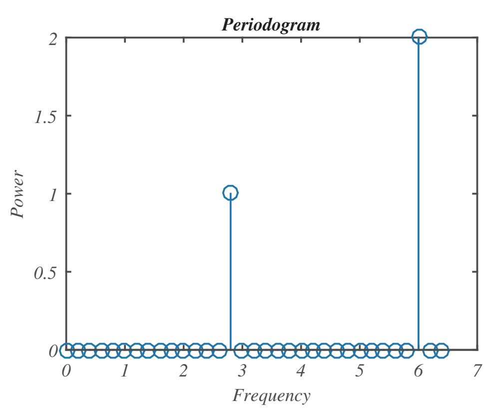


Figure 7.32: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t) - \cos(2\pi f_1 t)$ with $f_0 = 6$ Hz and $f_1 = 10$ Hz. The signal was sampled with $N = 64$ points on an interval of $[0, 5]$.

We will explain this masquerading of frequencies in terms of something called aliasing. However, that is not the whole story. Notice that the frequencies represented in the periodograms is discrete. Even in the case that $T = 5$ and $N = 128$, we only displayed frequencies at intervals of $\frac{1}{T} = 0.2$. What would happen if the signal had a frequency in between these values? For example, what if the 6 Hz frequency was 6.1 Hz? We see the result in Figure 7.33. Since we could not pinpoint the signal's frequencies at one of the allowed discrete frequencies, the periodogram displays a spread in frequencies. This phenomenon is called *ringing* or *spectral leakage*.

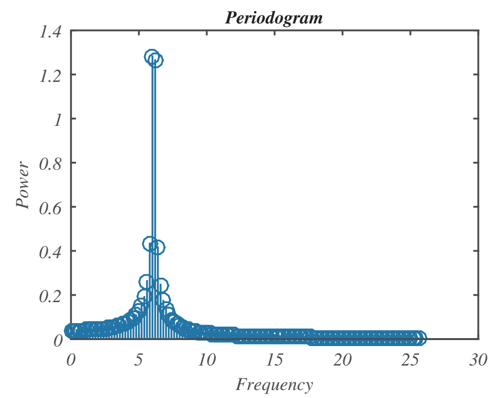


Figure 7.33: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t)$ with $f_0 = 6.1$ Hz. The signal was sampled with $N = 256$ points on an interval of $[0, 5]$.

It is also interesting to see the effects on the individual Fourier coefficients. This is shown in Figure 7.34. While there is some apparent distribution of energy amongst the A_n's, it is still essentially zero. Most of the effects indicate that the energy is distributed amongst sine contributions.

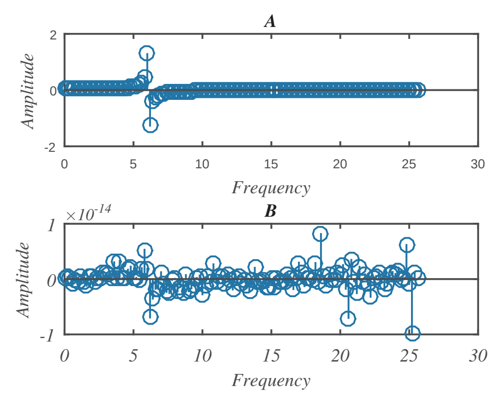


Figure 7.34: Fourier spectrum for the signal $y(t) = 2\sin(2\pi f_0 t)$ with $f_0 = 6.1$ Hz. The signal was sampled with $N = 256$ points on an interval of $[0, 5]$.

What we have learned from these examples is that we need to be careful in picking the record length and number of samples used in analyzing analog signals. Sometimes we have control over these parameters, but other times we are stuck with them depending upon the limitations of the recording devices. Next we will investigate how the effects of ringing and aliasing occur.

7.4 *Effect of Finite Record Length*

In the previous section we saw examples of the effects of finite record length on the spectrum of sampled data. In order to understand these effects for general signals, we will focus on a signal containing only one frequency, such as $y(t) = \sin(2\pi f_0 t)$. We will record this signal over a finite time interval, $t \in [0, T]$. This leads us to studying a finite wave train. (Recall that we has seen examples of finite wave trains earlier in the chapter on Fourier Transforms and in the last chapter. However, in these cases we used a cosine function. Also, in one of these cases we integrated over a symmetric interval.)

We will consider sampling the finite sine wave train given by

$$y(t) = \begin{cases} \sin 2\pi f_0 t, & 0 \le t \le T \\ 0, & \text{otherwise} \end{cases}. \tag{7.1}$$

In order to understand the spectrum of this signal, we will first compute the Fourier transform of this function. Afterwards, we will show how sampling this finite wave train affects the Fourier transform.

Example 7.1. *Compute the Fourier transform of the finite sine wave train.*

We begin by computing the Fourier transform of the finite wave train and write the transform in terms of its real and imaginary parts. The computation is straightforward and we obtain

$$\begin{aligned}
\hat{y}(f) &= \int_{-\infty}^{\infty} y(t) e^{2\pi i f t}\, dt \\
&= \int_0^T \sin(2\pi f_0 t)\cos(2\pi f t)\, dt + i\int_0^T \sin(2\pi f_0 t)\sin(2\pi f t)\, dt \\
&= \frac{1}{2}\int_0^T \left[\sin(2\pi(f+f_0)t) - \sin(2\pi(f-f_0)t)\right]\, dt \\
&\quad + \frac{i}{2}\int_0^T \left[\cos(2\pi(f-f_0)t) - \cos(2\pi(f+f_0)t)\right]\, dt \\
&= \frac{1}{4\pi}\left[\frac{1}{f+f_0} - \frac{1}{f-f_0}\right] \\
&\quad + \frac{1}{4\pi}\left[\frac{\cos 2\pi(f-f_0)T}{f-f_0} - \frac{\cos 2\pi(f+f_0)T}{f+f_0}\right] \\
&\quad - \frac{i}{4\pi}\left[\frac{\sin 2\pi(f-f_0)T}{f-f_0} + \frac{\sin 2\pi(f+f_0)T}{f+f_0}\right]
\end{aligned} \tag{7.2}$$

This, of course, is a complicated result. One might even desire to carry out a further analysis to put this in a more revealing form. However, we could just plot the real and imaginary parts of this result, or we could plot the modulus, $|\hat{y}(f)|$ to get the spectrum. We will consider both types of plots for some special cases.

Let's pick $T = 5$ and $f = 2$. The resulting finite wave train is shown in Figure 7.35.

We now look at the spectral functions. In Figures 7.36 - 7.37 we plot the real and the imaginary parts of the Fourier transform of this finite wave train. In Figure 7.38 we show a plot of the modulus of the Fourier transform. We note in these figures that there are peaks at both $f = 2$ and $f = -2$. Also, there are other smaller peaks, decaying for the frequencies far from the main peaks. It is the appearance of these minor peaks that will contribute to the ringing we had mentioned in the last section. Ringing occurs when we do not sample the data exactly at the frequencies contained in the signal.

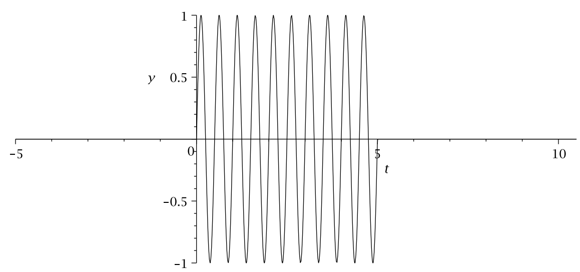


Figure 7.35: The finite sine wave train for $T = 5$ and $f_0 = 2$.

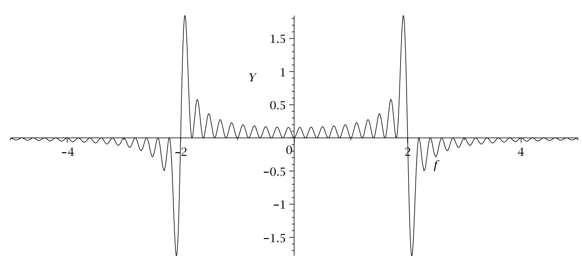


Figure 7.36: The real part of the Fourier transform of the finite wave train for $T = 5$ and $f_0 = 2$.

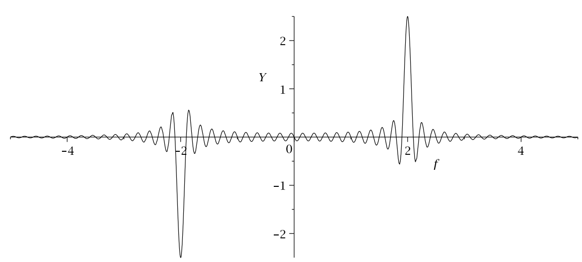


Figure 7.37: The imaginary part of the Fourier transform of the finite wave train for $T = 5$ and $f_0 = 2$.

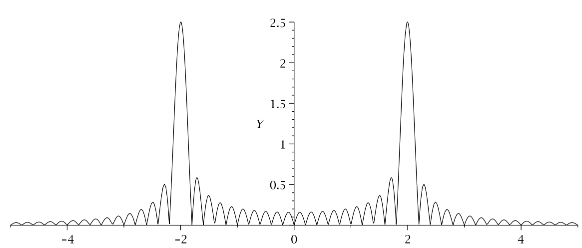


Figure 7.38: The modulus of the Fourier transform, or the spectrum, of the finite wave train for $T = 5$ and $f_0 = 2$.

We now consider what happens when we sample this signal. Let's take $N = 40$ sample points. In Figure 7.39 we show both the original signal and the location of the samples on the signal. (Besides the obvious points at the peaks and valleys, there are actually sampled points located along the time axis.)

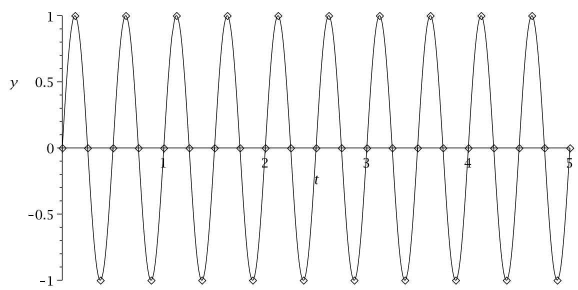


Figure 7.39: The finite wave train for $T = 5$ and $f_0 = 2$ sampled with $N = 40$ points, or $\Delta t = 0.125$.

Using a finite record length and a discrete set of points leads to a sampling of the Fourier transform with $\Delta f = \frac{1}{T} = 0.2$ and an extent of $(N - 1)\Delta f = 7.8$. The sampled Fourier transform superimposed on the original modulus of the Fourier transform is shown in Figure 7.40. (Here we have only shown part of the interval $f \in [0, 7.8]$.) The main peak is captured with the data values at the top of the peak. The other data pints lie at the zeros of the Fourier transform. This is what we would expect for the sampled transform.

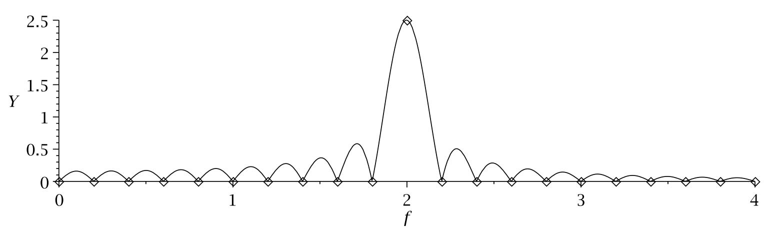


Figure 7.40: The modulus of the Fourier transform and its samples for the finite wave train for $T = 5$ and $f_0 = 2$ sampled with $N = 40$ points and $\Delta f = 0.2$.

Now we ask what happens when we sample a signal with a different frequency. Keeping everything else the same, we consider a finite wave train with frequency $f_0 = 2.1$. We sample this signal with forty points as before. The sampled signal is shown in Figure 7.41 and its transform is displayed in Figure 7.42. Notice that now the main peak of the Fourier transform is at $f_0 = 2.1$, but the sample frequencies do not match this frequency. Instead, there are two nearby frequencies not quite hitting the main peak. In fact, there are other nonzero frequencies in the Fourier transform leading to what is called ringing. It is only a coincidence that the data values have found their way to the peaks of the minor lobes in the transform. If the signal frequency is $f_0 = 2.05$, then we find that even the minor peaks differ from the discrete frequency values as displayed in Figure 7.43.

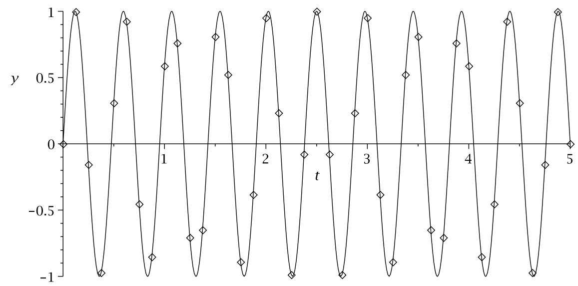


Figure 7.41: The finite wave train for $T = 5$ and $f_0 = 2.1$ sampled with $N = 40$ points, or $\Delta t = 0.125$. Note that the finite wave train does not contain an integer multiple of cycles.

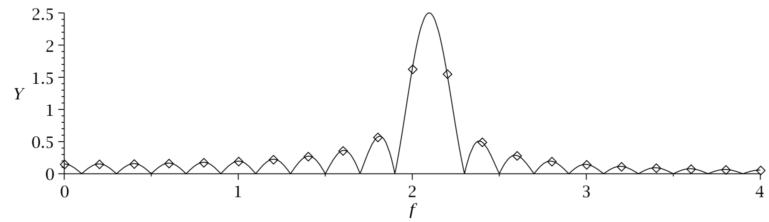


Figure 7.42: The modulus of the Fourier transform and its samples for the finite wave train for $T = 5$ and $f_0 = 2.1$ sampled with $N = 40$ points and $\Delta f = 0.2$.

7.5 *Aliasing*

In the last section we observed that ringing can be described by studying the finite wave train which is due to an improper selection of the record length, T. In this section we will investigate the concept of aliasing, which is due to poor discretization. Aliasing is the masquerading of some frequencies as other ones. Aliasing is a result of a mismatch in sampling rate to the desired frequency analysis. It is due to an inherent ambiguity in the trigonometric functions owing to their periodicity.

Let $y(t) = \sin 2\pi f_0 t$. We will sample the signal at the rate $f_s = \frac{1}{\Delta t}$. Often we have no choice in the sample rate. This is the rate at which we can collect data and may be due to the limitations in the response time of the devices we use. This could be the result of the responses of the recording devices, the data storage devices, or any post processing that is done on the data.

Sampling leads to samples at times $t = nt_s$ for $t_s = \frac{1}{f_s} = \Delta t$. The sampled signal is then

$$y_n = \sin(2\pi f_0 n t_s).$$

Through a little manipulation, we can rewrite this form. Making use of the fact that we can insert any integer multiple of 2π into the argument of the

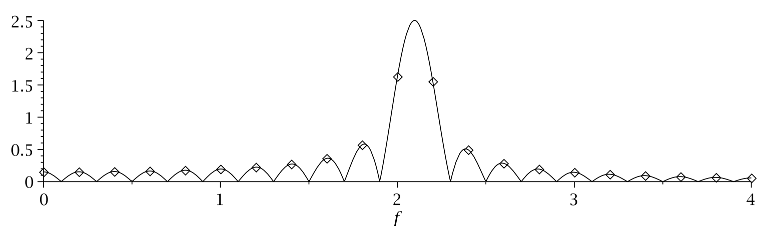


Figure 7.43: The modulus of the Fourier transform and its samples for the finite wave train for $T = 5$ and $f_0 = 2.05$ sampled with $N = 40$ points and $\Delta f = 0.2$.

sine function, we have for integer m that

$$\begin{aligned} y_n &= \sin(2\pi f_0 n t_s) \\ &= \sin(2\pi f_0 n t_s + 2\pi m) \\ &= \sin(2\pi (f_0 + \frac{m}{n t_s}) n t_s) \\ &= \sin(2\pi (f_0 + \frac{k}{t_s}) n t_s), \quad m = kn, \\ &= \sin(2\pi (f_0 + k f_s) n t_s). \end{aligned} \tag{7.3}$$

So, $\sin(2\pi f_0 n t_s) = \sin(2\pi(f_0 + k f_s) n t_s)$. This means that one cannot distinguish signals of frequency f_0 from signals of frequency $f_0 + k f_s$ for k an integer. Thus, if f_0 does not fall into the interval of frequencies, $[0, (N-1)\Delta f]$, then $f = f_0 + k f_s$ might be in this frequency interval for the right value of k. This is why the 2.8 Hz signal showed up in the Figures 7.29 and 7.32 earlier in the chapter.

How do we determine the frequency interval for the allowed discrete frequencies that will show up in the transform? We know that the smallest frequency value that we can resolve is $f_{\text{min}} = \Delta f = \frac{1}{T}$. The largest value of the frequency would be $f_{\text{max}} = \frac{N}{2}\Delta f = \frac{N}{2T} = \frac{1}{2\Delta t}$. Thus, to capture a desired frequency, f_0, we would need the condition that

$$f_{\text{min}} \le f_0 \le f_{\text{max}}.$$

For a signal recorded on $[0, T]$ at N points, we have

$$\Delta t \frac{T}{N}, \Delta f = \frac{1}{T},$$
$$f_s = \frac{1}{\Delta t} = \frac{N}{T}$$
$$f_{\text{max}} = \frac{1}{2} f_s.$$

However, we need to be a bit more careful. Recall that negative frequencies might also contribute. So, we also have that

$$-f_{\text{min}} \ge f_0 \ge -f_{\text{max}}.$$

This leads to the general conclusion that

if $|f_0| \ge f_{\text{max}}$, then the frequency will not be resolved in the analysis; i.e., it will not be captured and there might be an integer k such that

$$f_{\text{min}} \le |f_0 + k f_s| \le f_{\text{max}}, \tag{7.4}$$

in which case we will observe aliasing.

Thus, to capture all frequencies up to f_0, we need to sample at a rate f_s such that $f_0 \le f_{\text{max}} = \frac{1}{2} f_s$. This is what is called the Nyquist criterion.

Nyquist criterion: One needs to sample at a rate at least twice the largest expected frequency.

Example 7.2. *Determine the occurrence of the 2.8 Hz spike in Figure 7.29*

In the earlier sample plots (like Figure 7.29) displaying the results of a DFT analysis of simple sine functions we saw that for certain choices of N and T a 10 Hz signal produced a 2.8 Hz spike. Let's see if we can predict this.

We began by sampling $N = 128$ *points on an interval of* $[0, 5]$. *Thus, the frequency increments were*

$$\Delta f = \frac{1}{T} = 0.2$$

Hz. The sampling rate was

$$f_s = \frac{1}{\Delta t} = \frac{N}{T} = 25.6 \text{ Hz}.$$

Therefore,

$$f_{min} = 0.2\text{Hz and } f_{max} = \frac{1}{2} f_s = 12.8 \text{ Hz}.$$

So, we can sample signals with frequencies up to 12.8 Hz without seeing any aliasing. The frequencies we can resolve without ringing would be of the form $f_n = 0.2n \leq 12.8$.

However, if we change either T or N we may get a smaller interval and the 10 Hz frequency will not be picked up correctly. By picking $N = 64$ *and* $T = 5$, *we have*

$$f_{min} = 0.2 \text{ Hz}, \ f_s = 12.8 \text{ Hz, and } f_{max} = \frac{1}{2} f_s = 6.4 \text{ Hz}.$$

The 10 Hz frequency is larger than f_{max}. *Where will the aliased frequency appear? We just need to determine an integer k in Equation (7.4) such that*

$$0.2 \leq |10 + 12.8k| \leq 6.4. \tag{7.5}$$

The set of values of $10 + 12.8k$ *are* $\{\ldots, -15.6, -2.8, 10, 22.8, \ldots\}$. *We see that for* $k = -1$, $0.2 \leq 2.8 \leq 6.4$. *Thus, the 10 Hz signal will masquerade as a 2.8 Hz signal. This is what we had seen. Similarly, if* $N = 128$ *and* $T = 10$, *we get the same sampling rate* $f_s = 12.8$ *Hz, but* $f_{min} = 0.1$. *The inequality for 10 Hz takes the form*

$$0.1 \leq |10 + 12.8k| \leq 6.4. \tag{7.6}$$

Thus, we once again get a 2.8 Hz frequency.

Example 7.3. *Plotting aliased signals.*

We consider a simple 1.0 Hz signal of the form $y(t) = \sin 2\pi t$ *sampled with 15 points on* $[0, 5]$. *This signal and its samples are shown in figure 7.44. However, we have seen that signals of frequency* $f_0 + kf_s$ *will also pass through the sampled points. In this example the data points were sampled with* $f_s = \frac{N}{T} = \frac{15}{5} = 3.0$ *Hz. For* $f_0 = 1.0$ *and* $k = 1$, *we have* $f_0 + kf_s = 4.0$ *Hz. In Figure 7.45 we plot this signal and the sampled data points of* $y(t) = \sin 2\pi f t$. *Notice that the new signal passes through all of the sampled points.*

We can see this better in Figure 7.46. Here we plot both signals with the sampled points.

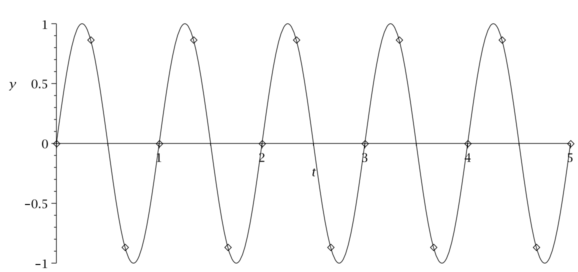


Figure 7.44: A 1.0 Hz signal of the form $y(t) = \sin 2\pi t$ sampled with 15 points on $[0, 5]$.

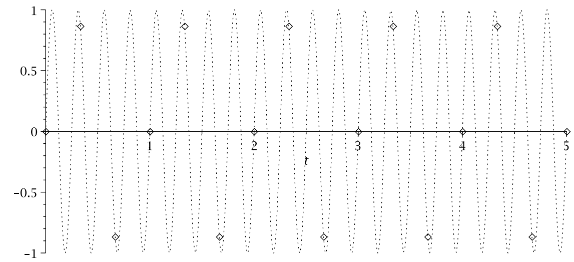


Figure 7.45: A signal of the form $y(t) = \sin 8\pi t$ is plotted with the samples of $y(t) = \sin 2\pi t$ sampled with 15 points on $[0, 5]$.

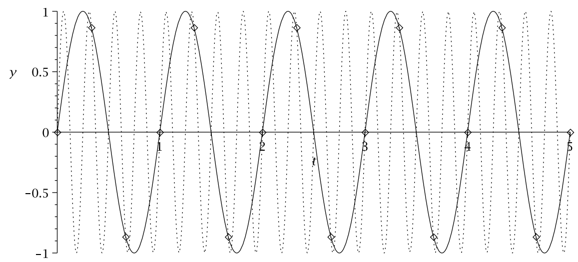


Figure 7.46: Signal $y(t) = \sin 8\pi t$ and $y(t) = \sin 2\pi t$ are plotted with the samples of $y(t) = \sin 2\pi t$ sampled with 15 points on $[0, 5]$.

7.6 The Shannon Sampling Theorem

THE PROBLEM OF RECONSTRUCTING A FUNCTION by interpolating at equidistant samples of the function, using a Cardinal series was presented by Shannon in his well known papers of 1948. He had used the sampling theorem to show that a continuous signal can be completely reconstructed from the proper samples. This well-known sampling theorem, known as the Whittaker-Shannon-Kotel'nikov (WSK) Theorem, can be stated as:

Theorem 7.1. *Let*

$$y(t) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \hat{y}(\omega) e^{i\omega t}\, d\omega, \tag{7.7}$$

where $\hat{y}(\omega) \in L_2[-\Omega, \Omega]$. *Then*

$$y(t) = \sum_{n=-\infty}^{\infty} y\left(\frac{n\pi}{\Omega}\right) \frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}. \tag{7.8}$$

This famous theorem is easily understood in terms of what we have studied in this text. It essentially says that if a signal $y(t)$ is a piecewise smooth, bandlimited ($|\hat{y}(\omega)| \le \Omega$) function, then it can be reconstructed exactly from its samples at $t_n = \frac{n\pi}{\Omega}$. The maximum frequency of the bandlimited signal is $f_{\text{max}} = \frac{\Omega}{2\pi}$. The sampling rate necessary to make this theorem hold is $f_s = \frac{1}{\Delta t} = \frac{\Omega}{\pi}$. Thus,

$$f_s = 2f_{\text{max}}.$$

This is just the Nyquist sampling condition, which we saw previously, so that there is no aliasing of the signal. This sampling rate is called the Nyquist rate. Namely, we need to sample at twice the largest frequency contained in the signal.

Proof. We consider a band-limited function, $y_\Omega(t)$, where $\hat{y}_\Omega(\omega) = 0$ for $|\omega| > \Omega$. This function could have been prepared by multiplying the Fourier transform of an analog signal $\hat{y}(\omega)$ by the gate function[1],

$$G_\Omega(\omega) = \begin{cases} 1, & |\omega| < \Omega, \\ 0, & |\omega| > \Omega. \end{cases}$$

[1] The gate function is also known as a box function or a rectangular function.

Therefore, we consider the band-limited function, show in Figure 7.47

$$\hat{y}_\Omega(\omega) = G_\Omega(\omega)\hat{y}(\omega).$$

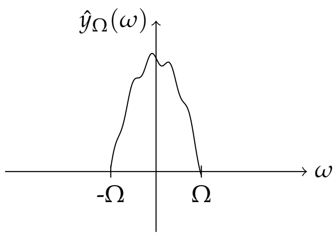


Figure 7.47: Plot of Fourier transform of a bandlimited function, $\hat{y}_\Omega(\omega)$.

We further note that in the language of systems theory, $G_\Omega(\omega)$ is called a low pass filter with a passband of $|\omega| < \Omega$ since only frequencies in this range are allowed through the filter. $G_\Omega(\omega)$ is also referred to as the transfer function of the filter.

We recall that the inverse Fourier transform of the gate function can be easily computed as

$$\begin{aligned} g_\Omega(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} G_\Omega(\omega)e^{-i\omega t}\,d\omega \\ &= \frac{1}{2\pi}\int_{-\Omega}^{\Omega} e^{-i\omega t}\,d\omega \\ &= \frac{\sin \Omega t}{\pi t} \\ &= \frac{\Omega}{\pi}\text{sinc } \Omega t. \end{aligned} \tag{7.9}$$

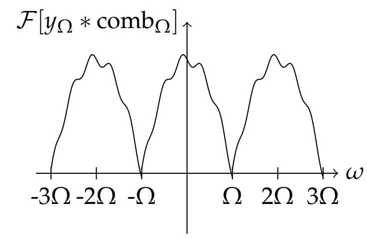


Figure 7.48: The Fourier transform of the product of a bandlimited function and a comb function gives a periodic version of $\hat{y}_\Omega(\omega)$.

The band-limited function, $y_\Omega(t)$, can then be written as

$$y_\Omega(t) = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} \hat{y}_\Omega(\omega)e^{-i\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} G_\Omega(\omega)\hat{y}(\omega)e^{-i\omega t}\,d\omega.$$

We now use the Convolution Theorem, since the Fourier transform of $y_\Omega(t)$ is a product of transforms, $\hat{y}_\Omega(\omega) = G_\Omega(\omega)\hat{y}(\omega)$. Therefore,

$$\begin{aligned} y_\Omega(t) &= \mathcal{F}^{-1}[G_\Omega(\omega)\hat{y}(\omega)] \\ &= (y * g_\Omega)(t) \\ &= \int_{-\infty}^{\infty} y(\tau)\frac{\sin \Omega(\tau - t)}{\pi(\tau - t)}\,d\tau. \end{aligned} \tag{7.10}$$

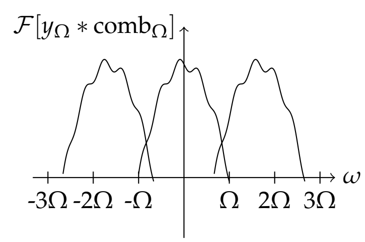


Figure 7.49: The Fourier transform of the the product of a bandlimited function and a comb function gives overlapped translations of $\hat{y}_\Omega(\omega)$ leading to aliasing.

This result is true if we know $y(t)$ for a continuous range of times. What if we were to sample $y(t)$? We will consider sampling a signal in time intervals of Δt. The sampled signal can be written as

$$y_s(t) = y(t)\Delta t\text{comb}_{\Delta t}(t) = \sum_{n=-\infty}^{\infty} y(n\Delta t)\Delta t\delta(t - n\Delta t),$$

where each pulse is weighted by the sampling time, Δt. This weight is needed in order to get the correct Fourier transform. Namely,

$$\mathcal{F}\left[y(t)\text{comb}_{\Delta t}(t)\right] = \frac{1}{2\pi}\left(\hat{y}_\Omega * \frac{2\pi}{\Delta t}\text{comb}_{\frac{2\pi}{\Delta t}}\right)(\omega) = \sum_{n=-\infty}^{\infty} \hat{y}_\omega\left(\omega - \frac{2\pi n}{\Delta t}\right).$$

This gives a sum of translations of $\hat{y}_\Omega(\omega)$ as depicted in Figures 7.48 and 7.49.

Inserting the expression for the sampled signal, $y_s(t)$, into Equation (7.10), we have the approximation

$$\begin{aligned}
\bar{y}_\Omega(t) &= \int_{-\infty}^{\infty} y_s(\tau) \frac{\sin \Omega(\tau - t)}{\pi(\tau - t)}\, d\tau \\
&= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} y(n\Delta t)\delta(\tau - n\Delta t)\Delta t \frac{\sin \Omega(\tau - t)}{\pi(\tau - t)}\, d\tau \\
&= \sum_{n=-\infty}^{\infty} y(n\Delta t)\Delta t \frac{\sin \Omega(t - n\Delta t)}{\pi(t - n\Delta t)} \\
&= \sum_{n=-\infty}^{\infty} y(n\Delta t)\Delta t \frac{\Omega}{\pi} \operatorname{sinc} \Omega(t - n\Delta t). \qquad (7.11)
\end{aligned}$$

If $\Delta t > \frac{\pi}{\Omega}$, then the copies of $\hat{y}_\Omega(\omega)$ will be translated by multiples of

$$\frac{2\pi}{\Delta t} < 2\Omega,$$

which is the bandwidth of the bandlimited function. Therefore, the translations will overlap as in Figure 7.49.

However, if one samples at the Nyquist rate, $f_s = \Omega/\pi$, then the Fourier transform will have no overlap as shown in Figure 7.48. Setting $\Delta t = \frac{1}{f_s} = \frac{\pi}{\Omega}$, we find

$$\begin{aligned}
\bar{y}_\Omega(t) &= \sum_{n=-\infty}^{\infty} y(n\Delta t)\Delta t \frac{\Omega}{\pi} \operatorname{sinc}(\Omega t - n\pi) \\
&= \sum_{n=-\infty}^{\infty} y\left(\frac{n\pi}{\Omega}\right) \frac{\sin \Omega \left(t - \frac{n\pi}{\Omega}\right)}{\Omega \left(t - \frac{n\pi}{\Omega}\right)} \\
&= \sum_{n=-\infty}^{\infty} y\left(\frac{n\pi}{\Omega}\right) \frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}. \qquad (7.12)
\end{aligned}$$

□

Example 7.4. *Multichannel Telephone Conversations*

Consider the amount of information transmitted in telephone channels. Older telephone signals propagated at 56 kbps (kilobytes per second). If each sample is stored as 7 bytes, then we can attain 56 kbps/ 7 bytes = 8000 samples per second = 8 samples per millisecond. Assuming that conversations are transmitted at frequencies of less than 1.0 kHz, then, the Nyquist theorem suggests that we sample conversations at 2 kHz, or every half a millisecond. This means that we can sample up to four conversations simultaneously.

One of the uses of the Shannon Sampling Theorem is in interpolation theory. Given a set of data points as in Figure 7.50, can one find a function, or set of piecewise functions, to approximate the function at points in between the given data points and which pass through the data points? The simplest interpolation is to connect the points with lines, yielding a linear interpolation as shown in Figure 7.51. However, this might not be a close enough match to the actual function as seen in Figure 7.52.

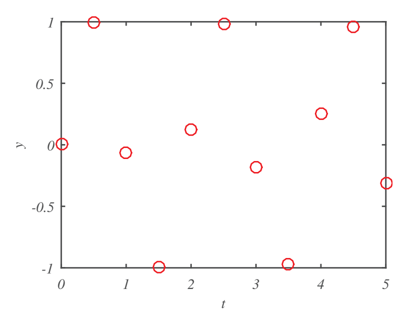


Figure 7.50: Set of data points to be interpolated.

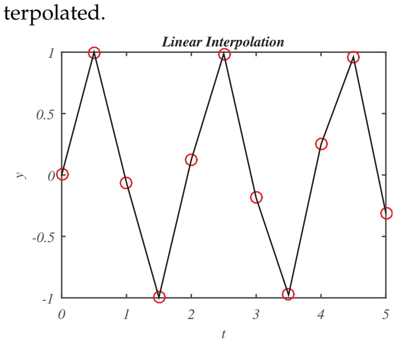


Figure 7.51: Set of data points with linear interpolation.

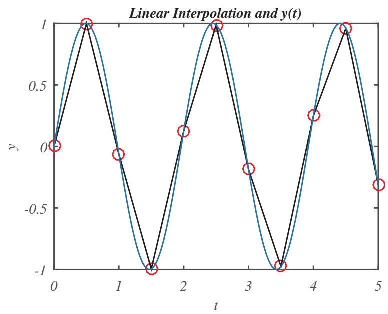


Figure 7.52: Set of data points with lin-

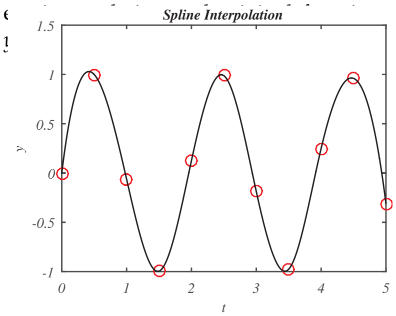


Figure 7.53: Set of data points with spline interpolation.

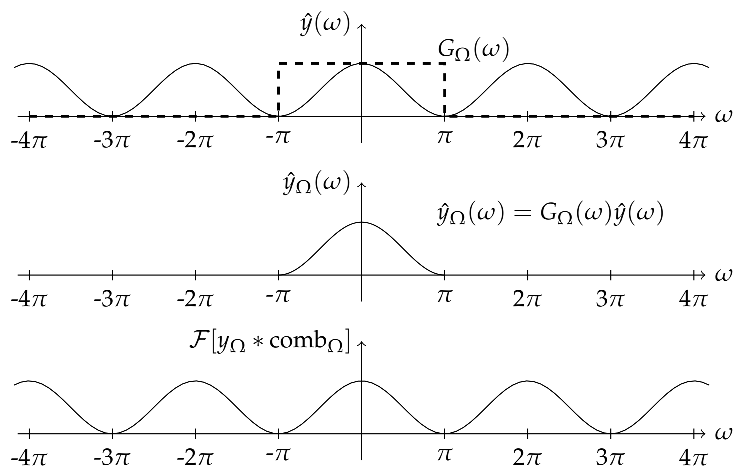
There are other types of interpolation involving higher order polynomial interpolating functions or special functions, such as splines. Such a spline fit is shown in Figure 7.53. This higher order fits have more oscillations. If the data consists of samples of some polynomial, then eventually one can obtain a good fit.

The idea of the sampling series as an interpolating function is a way to introduce high enough oscillations into the picture, but not exceeding a cutoff frequency. If the function is bandlimited, then the series should return exactly the function using the given data points as we have seen.

Example 7.5. *Use the Dirac comb function to derive the Shannon Sampling Theorem.*

Another approach to the proof of the Sampling Theorem is to use the Dirac comb function. As shown in Figure 7.54, the Fourier transform $\hat{y}(\omega)$ is multiplied by the gate function, $G_\Omega(\omega)$ to produce the bandlimited function $\hat{y}_\Omega(\omega)$. One then computes the inverse Fourier transform of this function to obtain $y_\Omega(t)$. Sampling of this function, using the Dirac comb function, gives $(y_\Omega * comb_\Omega)(t)$. The Fourier transform of this function, $\mathcal{F}[y_\Omega * comb_\Omega](\omega)$, is then computed. This results in a periodic function as shown in the last plot of Figure 7.54. One then gets back the transform of the bandlimited function by multiplying by the gate function.

Figure 7.54: The Fourier transform $\hat{y}(\omega)$ is multiplied by the gate function, $G_\Omega(\omega)$ to produce the bandlimited function $\hat{y}_\Omega(\omega)$. One then computes the inverse Fourier transform of this function to obtain $y_\Omega(t)$. Sampling of this function using the comb function gives $(y_\Omega * \text{comb}_\Omega)(t)$. The Fourier transform, $\mathcal{F}[y_\Omega * \text{comb}_\Omega](\omega)$, is then the periodic function shown in the last plot.



This can be captured by the expression

$$\mathcal{F}[y_\Omega] = G_\Omega(\omega)(\hat{y}_\Omega * comb_\Omega)(\omega),$$

or

$$y_\Omega(t) = \mathcal{F}^{-1}\left[G_\Omega(\omega)(\hat{y}_\Omega * comb_\Omega)(\omega)\right].$$

The right hand side is a Fourier transform of a product. The Convolution Theorem gives

$$\begin{aligned} y_\Omega(t) &= g_\Omega * (y_\Omega \mathcal{F}[comb_\Omega]) \\ &= \Omega \, sinc \, \Omega t * \left(y_\Omega \frac{1}{\Omega} \, comb_{\frac{1}{\Omega}} \right) \end{aligned}$$

$$\begin{aligned} &= \Omega\, sinc\, \Omega t * \left(\frac{1}{\Omega} \sum_{k=-\infty}^{\infty} y\left(\frac{k}{\Omega}\right) \delta\left(t - \frac{k}{\Omega}\right)\right) \\ &= \sum_{k=-\infty}^{\infty} y\left(\frac{k}{\Omega}\right) sinc\, \Omega \left(t - \frac{k}{\Omega}\right). \end{aligned} \tag{7.13}$$

This is the Shannon Sampling Theorem.

Example 7.6. *Evaluate the Fourier transform of the sampling series, $\mathcal{F}[y_\Omega(t)]$.*

We seek to find

$$\begin{aligned} \mathcal{F}[y_\Omega(t)] &= \mathcal{F}\left[\sum_{k=-\infty}^{\infty} f\left(\frac{k}{\Omega}\right) sinc\, \Omega \left(t - \frac{k}{\Omega}\right)\right] \\ &= \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\Omega}\right) \mathcal{F}\left[sinc\, \Omega \left(t - \frac{k}{\Omega}\right)\right] \end{aligned} \tag{7.14}$$

We need the Fourier transform of the sinc function.

$$\mathcal{F}\left[sinc\, t\right] = \begin{cases} \pi, & |\omega| < 1, \\ 0, & |\omega| > 1. \end{cases} = \pi G_1(\omega). \tag{7.15}$$

Letting $\tau = \Omega(t - k\Delta t)$, then we have

$$\begin{aligned} \mathcal{F}\left[sinc\, \Omega \left(t - \frac{k}{\Omega}\right)\right] &= \int_{-\infty}^{\infty} sinc\, \Omega \left(t - \frac{k}{\Omega}\right) e^{i\omega t}\, dt \\ &= \frac{1}{\Omega} \int_{-\infty}^{\infty} sinc\, \tau e^{i\omega\tau/\Omega + k\Delta t}\, d\tau \\ &= \frac{e^{i\omega k\Delta t}}{\Omega} \int_{-\infty}^{\infty} sinc\, \tau e^{i\omega\tau/\Omega}\, d\tau \\ &= \frac{e^{i\omega k\Delta t}}{\Omega} \mathcal{F}\left[sinc\right] \left(\frac{\omega}{\Omega}\right) \\ &= \frac{e^{i\omega k\Delta t}}{\Omega} \pi G_1 \left(\frac{\omega}{\Omega}\right) \\ &= \frac{\pi e^{i\omega k\Delta t}}{\Omega} G_\Omega\left(\omega\right). \end{aligned} \tag{7.16}$$

Therefore, we have obtained

$$\mathcal{F}[y_\Omega(t)] = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\Omega}\right) \frac{e^{i\omega k\Delta t}}{\Omega} G_\Omega\left(\omega\right)$$

Noting that

$$\mathcal{F}[y_\Omega] = G_\Omega(\omega)(\hat{y}_\Omega * comb_\Omega)(\omega),$$

we see that

$$(\hat{y}_\Omega * comb_\Omega)(\omega) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\Omega}\right) \frac{e^{i\omega k\Delta t}}{\Omega}.$$

Since

$$(\hat{y}_\Omega * comb_\Omega)(\omega) = \sum_{k=-\infty}^{\infty} \hat{y}_\Omega\left(\omega - k\Omega\right),$$

we have derived the expression

$$\sum_{k=-\infty}^{\infty} \hat{y}_\Omega\left(\omega - k\Omega\right) = \frac{1}{\Omega} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\Omega}\right) e^{i\omega k \Delta t}.$$

This is one form of the Poisson Summation Formula. Taking the Fourier transform, we can show that

The Poisson Summation Formula.

$$\sum_{k=-\infty}^{\infty} y(x+na) = \frac{1}{a} \sum_{k=-\infty}^{\infty} \hat{y}\left(\frac{k\Delta t}{2\pi}\right) e^{ikt/\Delta t}.$$

Example 7.7. *Consider the function*

$$\hat{y}_\Omega(\omega) = \begin{cases} 1-|\omega|, & |\omega| \le 1, \\ 0, & |\omega| > 1. \end{cases}$$

Sketch the function $G_\Omega(\omega)\mathcal{F}[y_\Omega * \Delta t\, comb_{\frac{2\pi}{\Delta t}}]$ *for sampling time* $\Delta t = \frac{4\pi}{3}$.

Correct sampling would be done with $\Delta t = \pi$. *So, we expect overlapping copies in the Fourier transform of the sampled signal. In Figure 7.55 we show the process.*

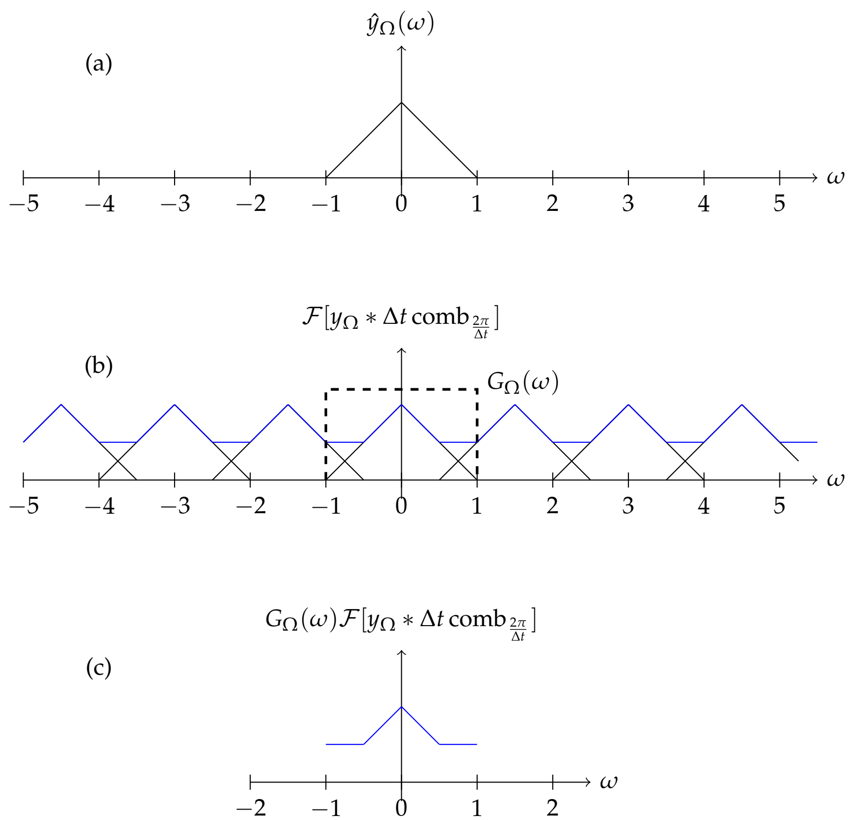


Figure 7.55: Sampling a bandlimited function, $y_\Omega(t)$, where $\Omega = 1$ and $\Delta t = \frac{4\pi}{3}$. (a) $\hat{y}_\Omega(\omega)$ is a triangular function with bandwidth $2\Omega = 2$. (b) A picture of $\mathcal{F}[y_\Omega * \Delta t\,\mathrm{comb}_{\frac{2\pi}{\Delta t}}]$ with the gate function. (c) The spectrum of the sampled signal, $G_\Omega(\omega)\mathcal{F}[y_\Omega * \Delta t\,\mathrm{comb}_{\frac{2\pi}{\Delta t}}]$.

In Figure 7.55(a) we show the sketch of the triangular function, $\hat{y}_\Omega(\omega)$. *The bandwith is* $2\Omega = 2$.

Next, in Figure 7.55(b) we draw the translations of $\hat{y}_\Omega(\omega)$ *in multiples of* $2\pi/\Delta t = 1.5$. *Superimposed on the translations are the sum, represented by* $\mathcal{F}[y_\Omega * \Delta t\, comb_{\frac{2\pi}{\Delta t}}]$.

The Fourier transform of the sampled signal is then obtained by multiplying the sum of the translations in Figure 7.55(b) by $G_\Omega(\omega)$ to obtain the final result, $G_\Omega(\omega)\mathcal{F}[y_\Omega * \Delta t\, comb_{\frac{2\pi}{\Delta t}}]$. This is shown in Figure 7.55(c).

7.7 *Nonstationary Signals*

7.7.1 *Simple examples*

A MAJOR ASSUMPTION MADE IN USING Fourier transforms is that the frequency content of a signal does not change in time. Such signals are called stationary. Consider the following example.

Example 7.8. *Let*

$$f(t) = \begin{cases} 1\sin(2\pi f_0 t), & 0 \le t < 0.25, < \\ 2\sin(2\pi f_1 t), & 0.25 < t < 0.75, \\ 1.5\sin(2\pi f_2 t), & 0.75 < t \le 1, \end{cases}$$

where $f_0 = 20$ Hz, $f_1 = 14$ Hz and $f_3 = 7$ Hz.

A plot of $f(t)$ is shown in Figure 7.56. There are three frequencies present, but occur at during different time intervals. The spectrum of this signal using the discrete Fourier transform over the entire time interval using $N = 256$ gives the plot in Figure 7.56. It indicates many more frequencies are present than just the three we know about. In Figure 7.59 we show a blow-up of the region containing the largest values. While it looked like there might have been ringing, there are peaks at the main frequencies, but somehow we could not capture the fact that the signal is nonstationary.

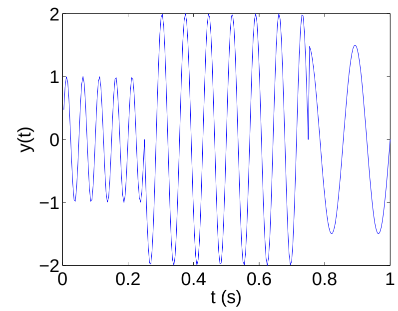


Figure 7.56: A plot of the function $f(t)$ vs t.

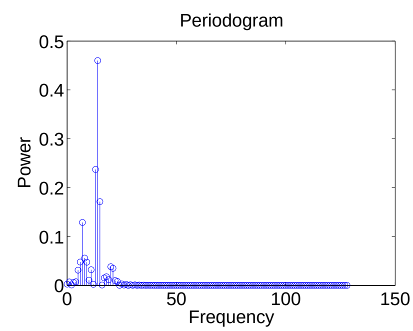


Figure 7.57: The application of the DFT algorithm to $f(t)$.

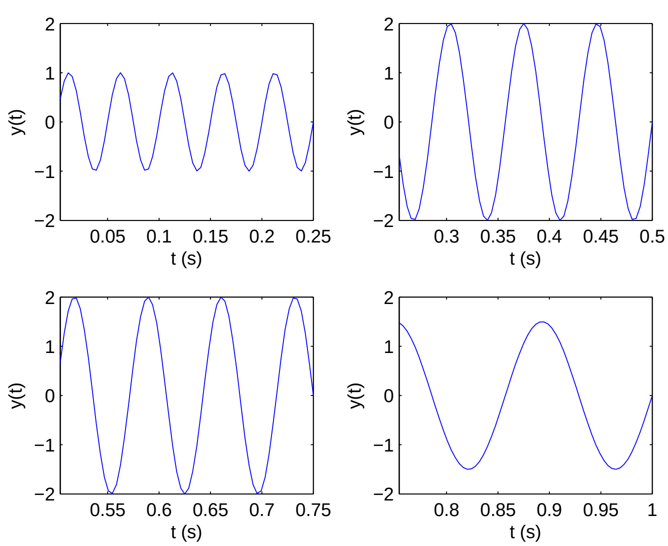


Figure 7.58: A plot of the function $f(t)$ vs t split into four windows.

We can capture the time dependence of the frequency content by splitting the time series into four blocks of width 0.25. This is shown in Figure 7.58. Now we

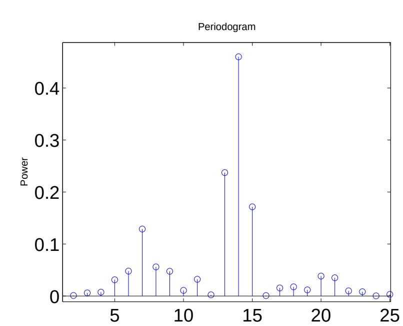


Figure 7.59: A magnified view of the DFT of $f(t)$.

apply the DFT to each block as shown in Figure 7.60. Zooming in further in Figure 7.61, we see that each periodogram displays different frequency content. However, since each block of $f(t)$ is not a perfect sine function, there still is a little inaccuracy in picking out the exact frequency in each block.

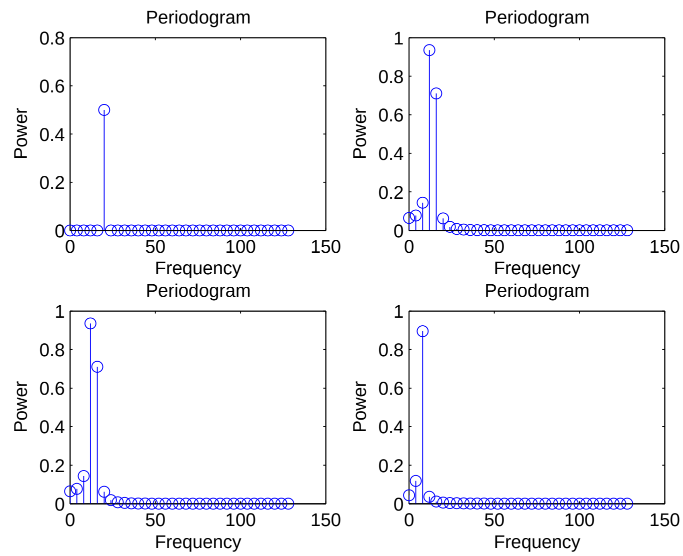


Figure 7.60: The application of the DFT algorithm to each of the four blocks of $f(t)$.

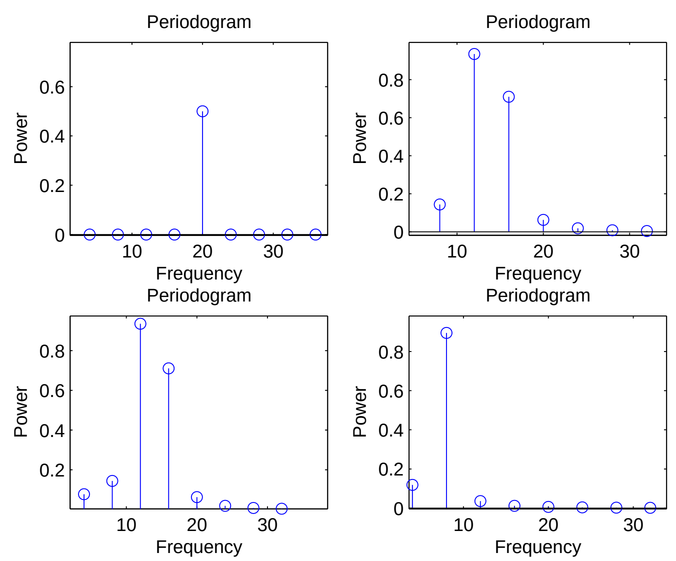


Figure 7.61: A magnified view of the DFT of the four blocks of $f(t)$.

Example 7.9. *Consider changing the time interval to $[0, 4\pi]$ in the previous example. Let*

$$g(t) = \begin{cases} \sin(2\pi f_0 t), & 0 \le t < \pi, \\ 2\sin(2\pi f_1 t), & \pi \le t < 3\pi, \\ 1.5\sin(2\pi f_2 t), & 3\pi \le t \le 4\pi, \end{cases}$$

where $f_0 = 20$ Hz, $f_1 = 14$ Hz and $f_3 = 7$ Hz. The four blocks are shown in

Figure 7.62. The DFT for these blocks with $N = 512$ is shown in Figure 7.63. We see that the frequencies are more defined and correct for the most part.

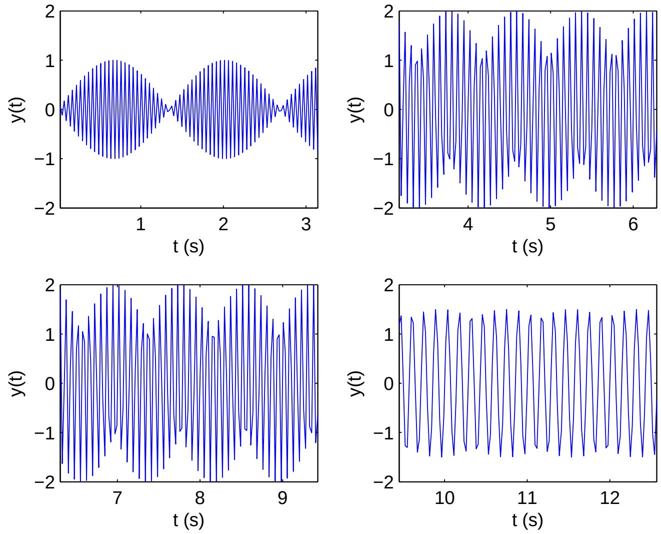


Figure 7.62: The four blocks of $g(t)$.

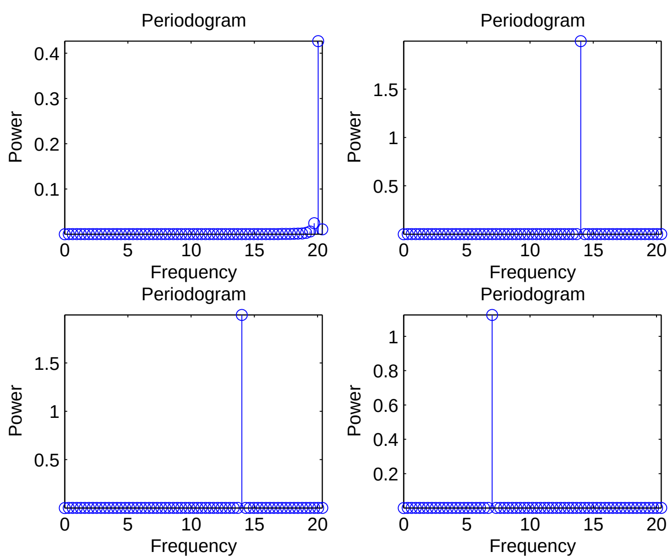


Figure 7.63: A magnified view of the DFT of the four blocks of $g(t)$.

Another example is the "chirp" function. A chirp is a sinusoidal function with a time varying frequency. When turned into a sound of the right length and frequency range, a chirp sounds like the chirp of a bird. a linear chirp is one in which the frequency changes linearly. The next example gives an example of a chirp.

Example 7.10. *Consider the linear chirp signal* $y(t) = \sin(2\pi \left(f_0 + (f_1 - f_0)\frac{t}{T}\right)t)$, $t \in [0, 1]$, *for* $f_0 = 1.0$ *Hz and* $f_1 = 10.0$ *Hz.*

The function takes the form $y(t) = \sin 2\pi f t$, *where the frequency is time-dependent,* $f(t) = f_0 + (f_1 - f_0)\frac{t}{T}$. *In Figure 7.64 we show this linear chirp.*

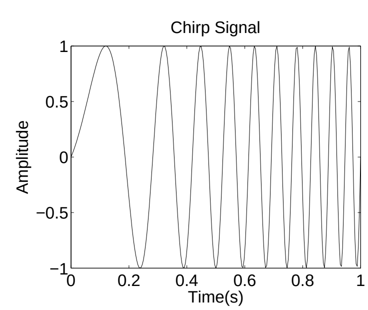


Figure 7.64: A chirp signal.

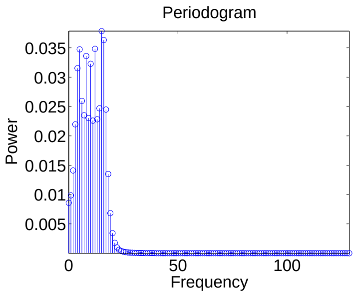


Figure 7.65: DFT of the chirp signal.

The frequency varies from $f_0 = 1.0$ Hz to $f_1 = 10.0$ Hz. When one computes the DTF of this signal, the periodogram in Figure 7.65 results. As one can see, a variety of frequencies appear and there is no indication that the frequency is time varying.

7.7.2 *The Spectrogram*

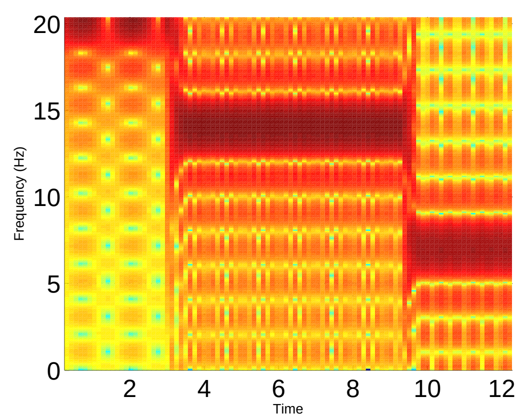
THE EXAMPLES IN THE LAST SECTION POINT TO THE NEED for a modification of the Fourier transform for analog signals and the DFT for discrete signals. In Example 7.9, we saw that computing the DFT over subintervals of the signal, we can attempt to find the time dependence of the frequency spectrum. This idea can be generalized for both continuous and discrete Fourier transforms. In MATLAB the function **spectrogram** produces a plot of the time-dependent frequency of a signal by using similar blocks, but by sliding blocks of a given width in time across the signal and doing a Fourier analysis for each block. The output is a spectrogram.

Example 7.11. *Let*

$$g(t) = \begin{cases} \sin(2\pi f_0 t), & 0 \le t < \pi, \\ 2\sin(2\pi f_1 t), & \pi \le t < 3\pi, \\ 1.5\sin(2\pi f_2 t), & 3\pi \le t \le 4\pi, \end{cases}$$

where $f_0 = 20$ Hz, $f_1 = 14$ Hz and $f_3 = 7$ Hz.

Figure 7.66: The spectrogram plot of $g(t)$.



We sample this signal at 512 points and using the MATLAB command `spectrogram`, *in the form*

```
spectrogram(g,rectwin(20),15,n,1/dt,'yaxis')
```

to generate the spectrogram in Figure 7.66, where **y** *is the signal sampled at* $\mathbf{n} = 512$ *points and* $\mathbf{dt} = 4\pi/n$. *The blocks are 20 pts wide with an overlap of 15 points. Note that the dominant three frequencies appear roughly at the correct locations.*

Example 7.12. *Let* $y(t) = \sin(2\pi\left(f_0 + (f_1 - f_0)\frac{t}{T}\right)t)$, $t \in [0,1]$, *for* $f_0 = 1.0$ *Hz and* $f_1 = 10.0$ *Hz.*

The the spectrogram in Figure 7.67 was produced by sampling this signal at 512 points and using the MATLAB command

```
spectrogram(y,rectwin(20),15,n,1/dt,'yaxis')
```

In the lower part of the figure we see a fuzzy linear region indicating the linear dependence of the frequency on time roughly going from $f = 1$ *Hz to* $f = 10.$ *Hz. The resolution of this frequency content depends partly on the width of the rectangular block used and the overlap of the blocks. In this case the command takes the form* **spectrogram(y,rectwin(w),o)**, *where* **w** *is the width and* **o** *is the size of the overlap.*

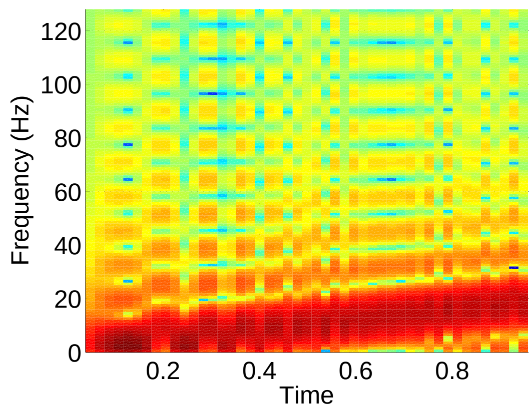


Figure 7.67: The spectrogram plot of the chirp $y(t) = \sin(2\pi\left(f_0 + (f_1 - f_0)\frac{t}{T})t\right)$, $t \in [0, 1]$, for $f_0 = 1.0$ Hz and $f_1 = 10.0$ Hz.

Let's see how to formalize the process described in the last examples. We begin with a sampled signal, y_n, $n = 0, \ldots, N-1$, and a rectangular window (or, block), $w_n = w(t_n)$ of width $M < N$. Then, we compute the transform of the product,

$$Y_{k,\ell} = DFT\left\{[y_\ell w_0, \ldots, y_{\ell+n}w_n, \ldots, y_{\ell+N-1}w_{N-1}]\right\}$$

This gives the spectrum which we can associate with a time associated with a time over which the block is nonzero. Next, the window is translated by a time t_ℓ. The shifted window is given by $w_{n-\ell} = w(t_n - t_\ell)$.

Example 7.13. *A simple example of using overlapping blocks.*

We consider the signal

$$y(t) = \begin{cases} 2\sin 2\pi t, & 0 \le t \le 5, \\ 1.5\sin 3\pi t, & 5 \le t \le 10. \end{cases}$$

This is shown in Figure 7.68(a).

The signal is then sampled with $\Delta t = 0.05$ *as shown in Figure 7.68(b).*

Figure 7.68(c) shows the blocks that can be used. Each block is of width 1.0 *and translated y* 0.75, *leaving an overlap of* 0.25 *between consecutive blocks.*

We can change the values of the width of the rectangular block used and the overlap of the blocks to see the effects on the output. Examples are provided in Figures 7.69-7.70 for Examples 7.9 and 7.10. Each row is a spectrogram of fixed with with a 20%, 40% 60% or 80% blockwidth overlap

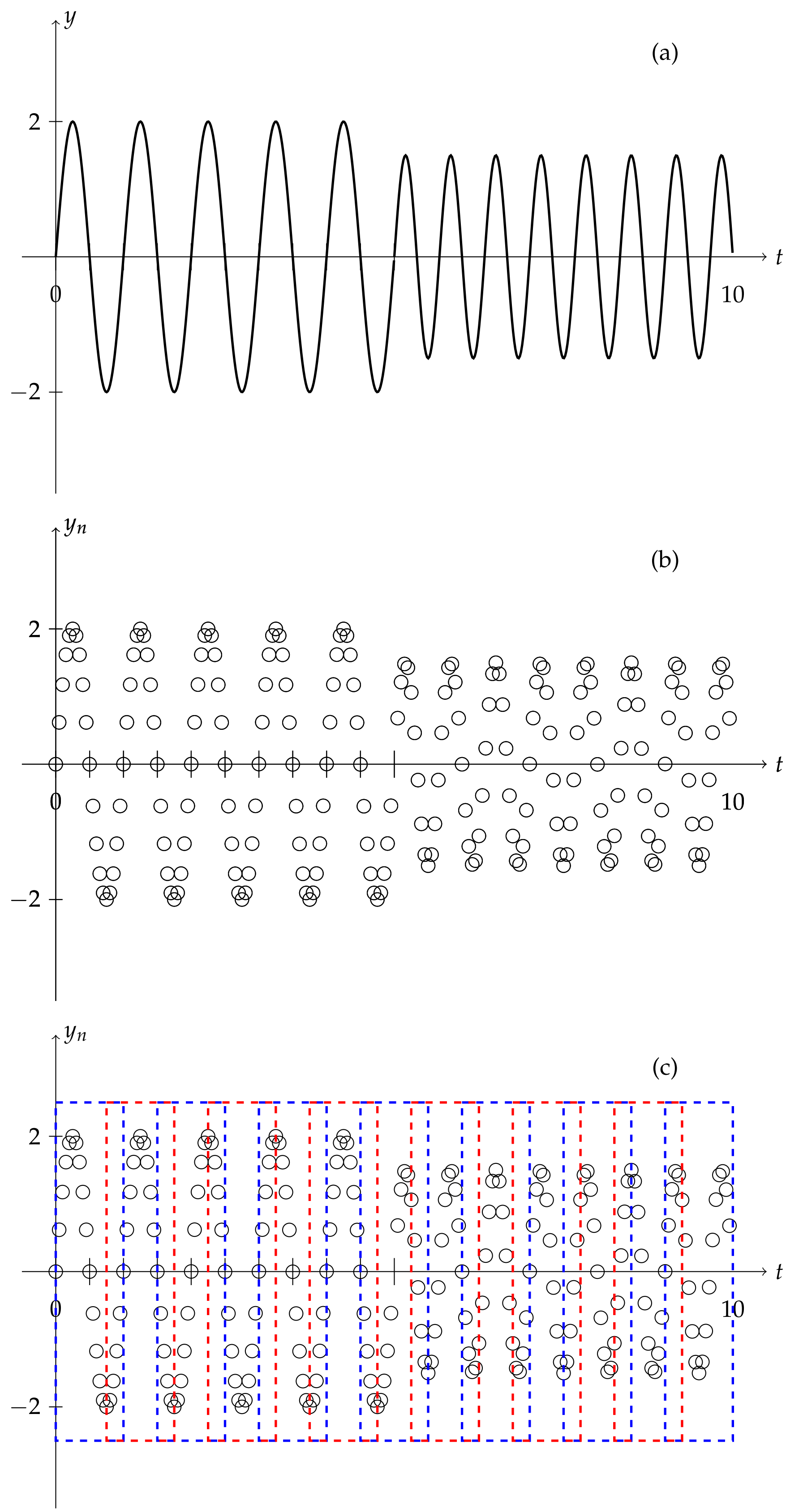


Figure 7.68: (a) Plot of $y(t)$. (b) Sampled signal y_n. (c) Translated windows showing overlapping.

as the blocks are translated across the signal. The block widths down the figure are 10 pts, 20 pts, 30 pts, and 40 pts, respectively.

In Figures 7.69-7.70 we see that there is better frequency resolution for wider blocks, but the time resolution is blurrier. However, increasing the overlap with aid in resolving the time as well. The better frequency resolution is due to using more points in the DFT for that block.

7.7.3 *Short-Time Fourier Transform*

THE KEY TO STUDYING NONSTATIONARY SIGNALS is the Short-Time Fourier Transform (STFT). The continuous version of the discrete Short-Time Fourier Transform is obtained by multiplying the signal by a sliding window function, $w(t)$, which is translated translated along the time axis, and taking the Fourier transform. The idea of Short-Time Fourier Transform is often credited to Dennis Gabor's work in 1946. He used a Gaussian window function.

Formally, we define the window function, $w(t)$, and multiply a shifted window, $w(t-\tau)$, by the signal, $y(t)$, and compute

$$STFT[y](\tau,\omega) \equiv = Y(\tau,\omega) = \int_{-\infty}^{\infty} y(t)w(t-\tau)e^{-i\omega t}\,dt. \tag{7.17}$$

Note that the sign in the exponential is negative, which differs from our earlier convention.

For the purpose of computation we can discretize the time and frequency variables giving the discrete-time Short-Time Fourier Transform and the discrete Short-Time Fourier Transform. The discrete-time Short-Time Fourier Transform is given by

$$Y(n,\omega) = \sum_{m=-\infty}^{\infty} x[m]w[m-n]e^{-i\omega n}. \tag{7.18}$$

The discrete-time Short-Time Fourier Transform is found by restricting ω to a discrete set of frequencies, $\omega = \frac{2\pi k}{N}$. It is given by

$$Y[n,k] = \sum_{m=0}^{N-1} x[m]w[m-n]e^{-2\pi ikn/N}. \tag{7.19}$$

There are a variety of windowing functions used. Different functions aid in time or frequency resolution as compared to the rectangular window used in the earlier sections. Note that the STFT is the Fourier transform of a product of two functions. So, if we fix τ, say $\tau = 0$, then we have

$$Y(0,\omega) = \int_{-\infty}^{\infty} y(t)w(t)e^{-i\omega t}\,dt. \tag{7.20}$$

By the Convolution Theorem, the Fourier transform is the convolution of two Fourier transforms:

$$Y(0,\omega) = (Y * W)(\omega),$$

Thus, an understanding of the Fourier transform of the window helps in understanding the effects of the window on the STFT.

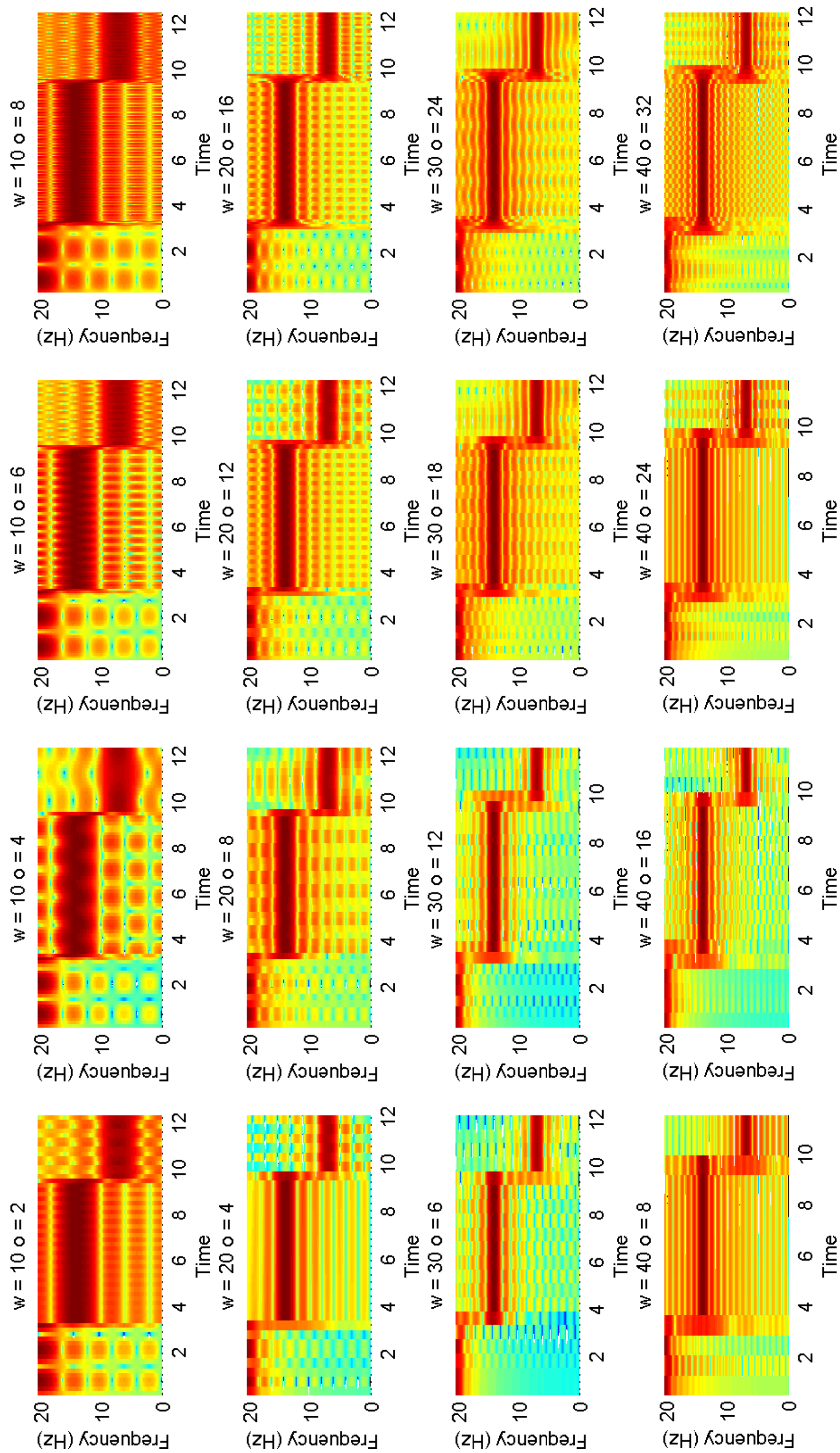


Figure 7.69: The spectrogram plot of $g(t)$ in Example 7.9.

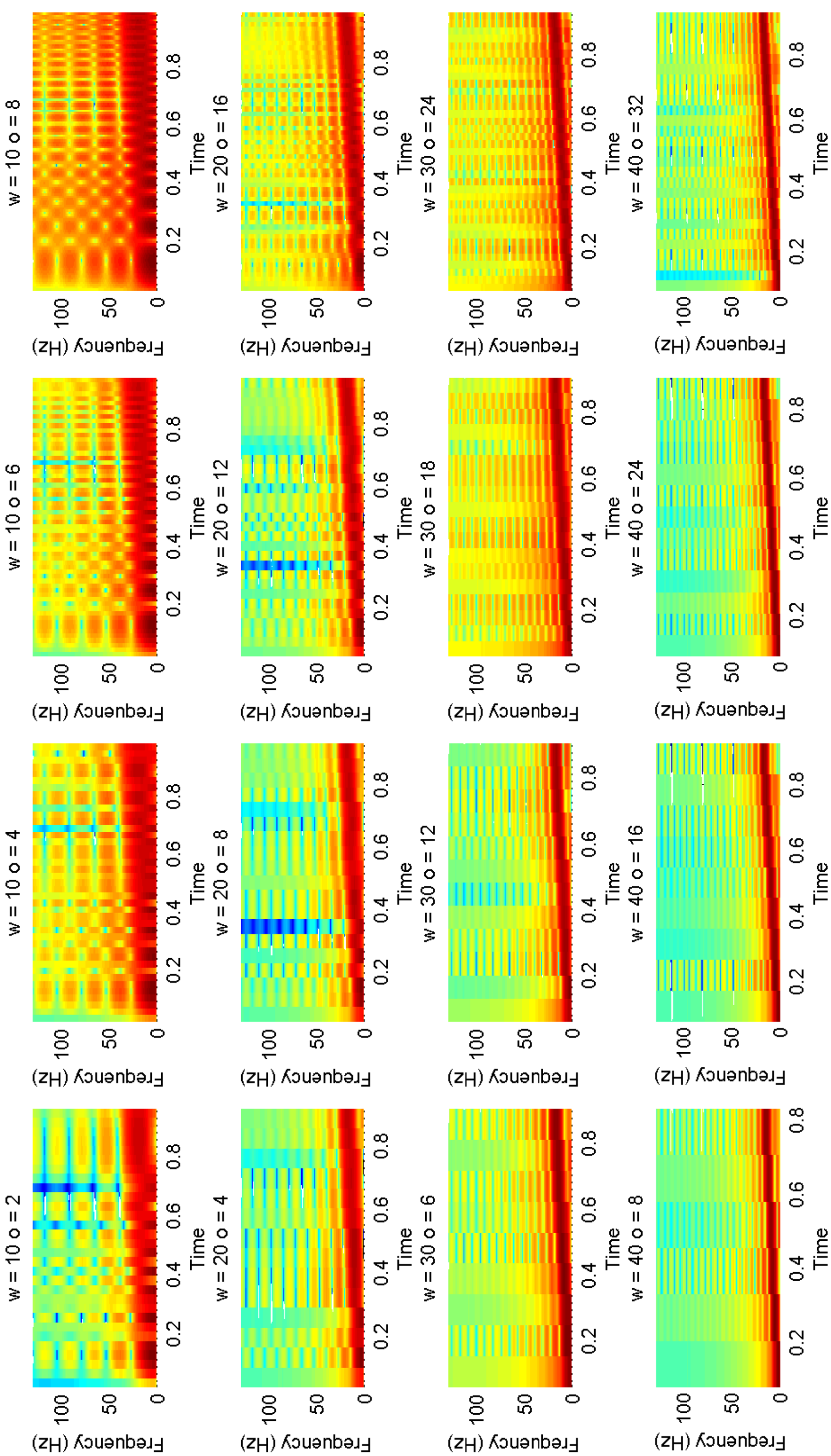


Figure 7.70: The spectrogram plot of the chirp signal, $f(t)$, in Example 7.10.

7.8 Harmonic Analysis

EVEN THOUGH SPECTRAL ANALYSIS, WHICH IS BASED UPON THE DISCRETE FOURIER TRANSFORM, is a readily used technique for analyzing the spectral content of a signal, one might only be interested in the spectral content for particular frequencies that are not included in the discrete set of frequencies provided by the Fourier transform. This can happen when the number of samples is large and the number of frequencies of interest is much smaller. In this section we address the approximation of a given signal, or time series, to a sum over a small set of frequency components. The frequencies of interest are known but the associated amplitude is not known. This will be accomplished using the method of least squares. This is approach is studied in more detail in Emery and Thompson [2].

[2] W. J. Emery and R. E. Thomson. *Data Analysis Methods in Physical Oceanography*. Elsevier, Amsterdam, The Netherlands, second edition, 1991

We begin by reviewing the method of least squares for determining the best fit of a line. The equation of a line only has two unknown parameters, the slope and the intercept. An understanding of the derivation for this simpler approximation should make that for the harmonic analysis more transparent.

We begin with a set of data, presented in the usual pairs (x_i, y_i) for $i = 1, \ldots, N$. We are interested in finding the best approximation, in some sense, of this data by some function. In the case of linear regression, we seek a linear relationship of the form $y = ax + b$. Though this line is not expected to agree with the data, it is expected to be as "close as possible".

What does one mean by "as close as possible"? We could mean that the total distance between the known data points and the line is as small as possible. Though there are many ways we could quantify this, the most natural would be to sum over the standard distance between the points for all data points. Thus, we would look at an expression like

$$F = \sum_{n=1}^{N} \sqrt{(x_i - x_i)^2 + (y_i - (ax_i + b))^2}.$$

However, since the first term under the square root vanishes and the square root only returns a positive number, we could instead just consider the expression

$$E = \sum_{n=1}^{N} [y_i - (ax_i + b)]^2 .$$

Making the later as small as possible only gives the same result as for the first expression, but is easier to compute. It gives the sum of the vertical distances between the data points and the line $y = ax + b$. This is called a "least squares" regression.

We are interested in minimizing E, which could be thought of as a variance about the straight line mean. We minimize this "error" by varying a and b. Thus, we have a two variable minimization problem. In order to minimize a function of one variable, we differentiate the function with respect to the variable and set it equal to zero to determine that value of the independent variable that yields a minimum. In this case, we need to simultaneously set the derivatives with respect to a and b to zero and find the values of a and b that solve the resulting equations.

Differentiating E with respect to a and b separately, gives

$$\begin{aligned} 0 &= 2\sum_{n=1}^{N}[y_i - (ax_i + b)](-x_i) \\ 0 &= 2\sum_{n=1}^{N}[y_i - (ax_i + b)](-1). \end{aligned} \tag{7.21}$$

Regrouping, we find that these are simultaneous equations for the unknowns:

$$\begin{aligned} a\sum_{n=1}^{N}x_i^2 + b\sum_{n=1}^{N}x_i &= \sum_{n=1}^{N}x_i y_i \\ a\sum_{n=1}^{N}x_i + bN &= \sum_{n=1}^{N}y_i. \end{aligned} \tag{7.22}$$

Solving this system of equations gives expressions for a and b in terms of various sums over expressions the involving the data. This is the basis of the so called "best fit line" that is used in many programs, such as those in calculators and in MS Excel. A more detailed discussion is given in [3].

[3] R. L. Herman. *A Course in Mathematical Methods for Physicists*. CRC Press, Taylor & Francis Group, Boca Raton, FL, 2014

However, we are interested in fitting data to a more complicated expression than that of a straight line. In particular, we want to fit the data to a sum over sines and cosines of arguments involving particular frequencies.

We will consider a set of data consisting of N values at equally spaced times, $t_n = n\Delta t,\ n = 1, \ldots, N$. We are interested in finding the best approximation to a function consisting of M particular frequencies, $f_k,\ k = 1, \ldots, M$. Namely, we wish to match the data to the function

$$f(t) = A_0 + \sum_{k=1}^{M}[A_k\cos(2\pi f_k t) + B_k\sin(2\pi f_k t)]. \tag{7.23}$$

The unknown parameters in this case are the A_k's and B_k's. We will not determine a function that exactly fits the data, as in the DFT case, but only seek the best fit curve to the data. [Note: for simplicity, we have redefined the constant term A_0 and have not used the typical form, $a_0/2$.]

We will consider the squared error,

$$E = \sum_{n=1}^{N}[y(t_n) - f(t_n)]^2,$$

summed over times $t_n = n\Delta t = \frac{nT}{N}$. More specifically, we will determine the unknowns, by minimizing the expression

$$E = \sum_{n=1}^{N}[y(t_n) - (A_0 + \sum_{k=1}^{M}[A_k\cos(2\pi f_k t_n) + B_k\sin(2\pi f_k t_n)])]^2. \tag{7.24}$$

We differentiate E with respect to all of the parameters and require these derivative to vanish. Namely, for a particular q, we have for $q = 1, 2, \ldots, M$,

$$\begin{aligned} 0 &= \frac{\partial E}{\partial A_q} \\ &= 2\sum_{n=1}^{N} Y_n \left(-\cos(2\pi f_q t_n)\right) \\ 0 &= \frac{\partial E}{\partial B_q} \\ &= 2\sum_{n=1}^{N} Y_n \left(-\sin(2\pi f_q t_n)\right), \end{aligned} \tag{7.25}$$

where

$$Y_n \equiv \left[y(t_n) - \left(A_0 + \sum_{k=1}^{M} \left[A_k \cos(2\pi f_k t_n) + B_k \sin(2\pi f_k t_n) \right] \right) \right].$$

As with the example of determining a best fit line, we have obtained a system of linear equations for the unknowns, A_k, B_k, $k = 1, 2, \ldots, M$. We can rewrite this system in a more apparent form and then make use of matrix theory to solve for the unknown coefficients.

$$\begin{aligned} & \sum_{n=1}^{N} y(t_n) \cos(2\pi f_q t_n) \\ = \quad & A_0 \sum_{n=1}^{N} \cos(2\pi f_q t_n) \\ & + \sum_{k=1}^{M} \left[A_k \sum_{n=1}^{N} \cos(2\pi f_k t_n) \cos(2\pi f_q t_n) + B_k \sum_{n=1}^{N} \sin(2\pi f_k t_n) \cos(2\pi f_q t_n) \right] \\ & \sum_{n=1}^{N} y(t_n) \sin(2\pi f_q t_n) \\ = \quad & A_0 \sum_{n=1}^{N} \sin(2\pi f_q t_n) \\ & + \sum_{k=1}^{M} \left[A_k \sum_{n=1}^{N} \cos(2\pi f_k t_n) \sin(2\pi f_q t_n) + B_k \sum_{n=1}^{N} \sin(2\pi f_k t_n) \sin(2\pi f_q t_n) \right] \end{aligned} \tag{7.26}$$

for $q = 1, \ldots, M$.

We need to consider the equation for $q = 0$ separately. In this case we have one equation,

$$\begin{aligned} 0 &= \frac{\partial E}{\partial A_0} \\ &= 2\sum_{n=1}^{N} \left[y(t_n) - \left(A_0 + \sum_{k=1}^{M} \left[A_k \cos(2\pi f_k t_n) + B_k \sin(2\pi f_k t_n) \right] \right) \right] (-1)), \end{aligned} \tag{7.27}$$

which reduces to

$$\sum_{n=1}^{N}\left[\left(A_0+\sum_{k=1}^{M}\left[A_k\cos(2\pi f_k t_n)+B_k\sin(2\pi f_k t_n)\right]\right)\right]=\sum_{n=1}^{N}y(t_n). \quad (7.28)$$

We now note that we can write this system of $2M+1$ equations (Equations 7.26 and (7.28)) in matrix form. We first define the $M\times N$ matrices C and S with elements

$$C_{qn}=\cos(2\pi f_k t_n),\quad S_{qn}=\sin(2\pi f_k t_n),$$

where $q=1,\ldots,M$, and $n=1,\ldots,N$.

The sums over n in Equations 7.26 and (7.28) can be written as $M\times M$ matrix products, CC^T, CS^T, and SS^T.[4] The qk-th entries (qth row and kth column) of these matrices are given by

[4] C^T is the transpose of C, satisfying $C^T_{ij}=C_{ji}$ and $(AB)^T=B^TA^T$.

$$\begin{aligned}\left(CC^T\right)_{qk} &= \sum_{n=1}^{N}\cos(2\pi f_k t_n)\cos(2\pi f_q t_n)\\ \left(CS^T\right)_{qk} &= \sum_{n=1}^{N}\sin(2\pi f_k t_n)\cos(2\pi f_q t_n)\\ \left(SS^T\right)_{qk} &= \sum_{n=1}^{N}\sin(2\pi f_k t_n)\sin(2\pi f_q t_n). \quad (7.29)\end{aligned}$$

Inserting these expressions into the system of equations, Equations 7.26 and (7.28), we obtain for $q=1,\ldots,M$,

$$\begin{aligned}A_0\sum_{n=1}^{N}C_{qn}+\sum_{k=1}^{M}\left[\left(CC^T\right)_{qk}A_k+\left(SC^T\right)_{qk}B_k\right] &= \sum_{n=1}^{N}C_{qn}y(t_n),\\ A_0\sum_{n=1}^{N}S_{qn}+\sum_{k=1}^{M}\left[\left(SC^T\right)_{qk}A_k+\left(SS^T\right)_{qk}B_k\right] &= \sum_{n=1}^{N}S_{qn}y(t_n), \quad (7.30)\end{aligned}$$

and

$$A_0N+\sum_{k=1}^{M}A_kc_k+B_ks_k=\sum_{n=1}^{N}y(t_n). \quad (7.31)$$

Finally, these equations can be put into the matrix form $DZ=Y$, where

$$Y=\begin{pmatrix}\bar{y}\\ Cy\\ Sy\end{pmatrix},\quad Z=\begin{pmatrix}A\\ B\end{pmatrix},$$

for $y=[y(t_1),\ dots, y(t_N)]^T$ and

$$D=\begin{pmatrix}N & c^T & s^T\\ c & CC & CS\\ s & CS & SS\end{pmatrix},$$

where

$$c_q=\sum_{n=1}^{N}C_{qn},\quad s_q=\sum_{n=1}^{N}S_{qn}.$$

Note that Y and Z are $2M+1$ dimensional column vectors, c and s are M dimensional column vectors and D is a $(2M+1) \times (2M+1)$ matrix. The system of equations in matrix form, $DZ = Y$, can be solved for the unknowns in the column vector $Z = D^{-1}Y$.

An implementation of this procedure in MATLAB is given below.

```
%
% Harmonic Analysis
%
% Enter Data in y
y=[7.6 7.4 8.2 9.2 10.2 11.5 12.4 13.4 13.7 11.8 10.1 ...
    9.0 8.9 9.5 10.6 11.4 12.9 12.7 13.9 14.2 13.5 11.4 10.9 8.1];
N=length(y);

% Number of Harmonics Desired and frequency dt
M=2;
f=1/12*(1:M);
T=24;
alpha=f*T;

% Compute the matrices of trigonometric functions
n=1:N;
C=cos(2*pi*alpha'*n/N);
S=sin(2*pi*alpha'*n/N);
c_row=ones(1,N)*C';
s_row=ones(1,N)*S';
D(1,1)=N;
D(1,2:M+1)=c_row;
D(1,M+2:2*M+1)=s_row;
D(2:M+1,1)=c_row';
D(M+2:2*M+1,1)=s_row';
D(2:M+1,2:M+1)=C*C';
D(M+2:2*M+1,2:M+1)=S*C';
D(2:M+1,M+2:2*M+1)=C*S';
D(M+2:2*M+1,M+2:2*M+1)=S*S';
yy(1,1)=sum(y);
yy(2:M+1)=y*C';
yy(M+2:2*M+1)=y*S';
z=D^(-1)*yy';
```

Problems

1. Consider the spectra below. In each case the data was sampled at 100 points for 12 seconds.

a. What is the sampling rate?

b. What is Δf?

c. What is the maximum frequency in the plots?

d. Determine the frequencies of the three spikes.

e. In each case write down a possible function, $f(t)$, which would give rise to these spikes.

f. In Figure 7.72 there are two frequencies. The tall spike is an aliased frequency. The data should have been sampled with 150 points. What is the correct frequency that would be displayed using 150 points?

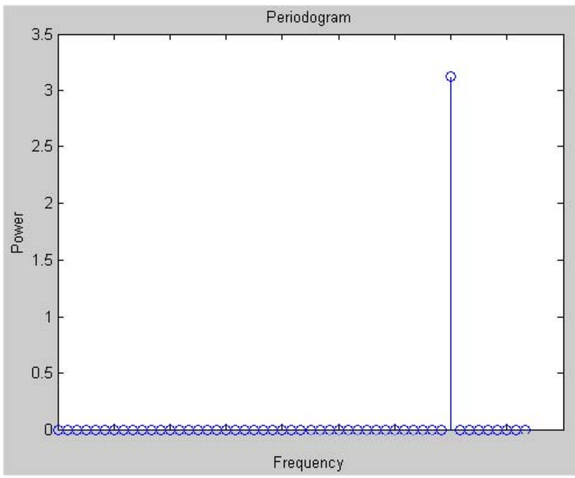


Figure 7.71: Spectrum 1 for Problem 1.

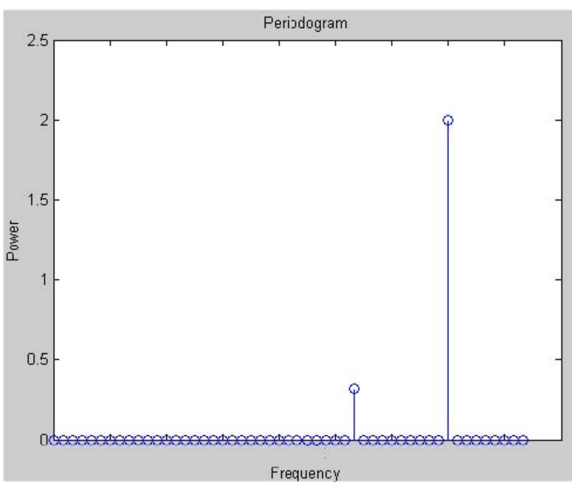


Figure 7.72: Spectrum 2 for Problem 1.

2. Consider the signal $f(t) = 1.5e^{-t}\sin 4t$.

a. Find analytical form for the modulus of the Fourier transform of $f(t)$.

b. Plot $\hat{f}(\omega)$ from part a and $f(t)$.

c. For each of the following, use MATLAB to plot the DFT amplitudes for the given values of T and N.

a. $T = 0.2$ and $N = 32$.

b. $T = 0.1$ and $N = 32$.

c. $T = 0.1$ and $N = 64$.

d. Compare these results with $\hat{f}(\omega)$ and note which are good approximations, which might exhibit aliasing, etc.

3. Consider the signal whose Fourier transform is given by

$$F(\omega) = \begin{cases} 1 - \left|\frac{\omega}{\pi}\right|, & |\omega| < \pi \\ 0 & |\omega| < \pi. \end{cases}$$

Let the signal be sampled at intervals of Δt s.

a. Sketch the discrete Fourier transform that you would expect if $\Delta t = 0.5$ s.

b. Sketch the discrete Fourier transform that you would expect if $\Delta t = 1.0$ s.

c. Sketch the discrete Fourier transform that you would expect if $\Delta t = 2.0$ s.

4. Find the Fourier transform of the sinc function.

$$\mathcal{F}[\operatorname{sinc} t] = \begin{cases} \pi, & |\omega| < 1, \\ 0, & |\omega| > 1. \end{cases} = \pi G_1(\omega).$$

5. Consider the function

$$\hat{y}_\Omega(\omega) = \begin{cases} 4\pi^2 - \omega^2, & |\omega| \leq 4\pi, \\ 0, & |\omega| > 2\pi. \end{cases}$$

a. Sketch $\hat{y}_\Omega(\omega)$.

b. Determine the Nyquist sampling frequency.

c. Sketch the function $G_\Omega(\omega)\mathcal{F}[y_\Omega * \Delta t \operatorname{comb}_{\frac{2\pi}{\Delta t}}]$ for sampling times $\Delta t = 1, 2, 3$.

6. Derive the Poisson Summation Formula,

$$\sum_{k=-\infty}^{\infty} y(x + ka) = \frac{1}{a} \sum_{k=-\infty}^{\infty} \hat{y}\left(\frac{k\Delta t}{2\pi}\right) e^{ikt/\Delta t},$$

by

a. Starting with $\hat{y}_a(\omega) = \mathcal{F}[g_a(t)y(t)]$; and

b. By using the form

$$\sum_{k=-\infty}^{\infty} \hat{y}_\Omega(\omega - k\Omega) = \frac{1}{\Omega} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\Omega}\right) e^{i\omega k\Delta t}.$$

Bibliography

M. J. Ablowitz and A. S. Fokas. *Complex Variables: Introduction and Applications*. Cambridge University Press, Cambridge, U.K. ; New York, 2003.

M. Abromowitz and I. A. Stregum. *Handbook of mathematical functions, with formulas, graphs, and mathematical tables*. Dover Publications, New York, 1965.

G. E. Andrews, R. Askey, and R. Roy. *Special Functions*. Cambridge University Press, Cambridge, U.K. ; New York, 1999.

G. Arfken. *Mathematical Methods for Physicists*. Academic Press, second edition, 1970.

E. T. Bell. *Men of Mathematics*. Fireside Books, 1965.

M. L. Boas. *Mathematical Methods in the Physical Sciences*. John Wiley & Sons, Inc, third edition, 2006.

J. W. Brown and R. V. Churchill. *Complex Variables and Applications*. McGraw-Hill, New York, NY, seventh edition, 2004.

R. V. Churchill. *Introduction to complex variables and applications*. McGraw-Hill, New York, 1948.

C. C. Cowen, K. R. Davidson, and R. P. Kaufman. Rearranging the alternating harmonic series. *The American Mathematical Monthly*, 87(10):817–819, 1980.

J. W. Dettman. *Applied Complex Variables*. Dover Publications, New York, NY, 2010.

W. J. Emery and R. E. Thomson. *Data Analysis Methods in Physical Oceanography*. Elsevier, Amsterdam, The Netherlands, second edition, 1991.

I. M. Gradshtein, I. M. Ryzhik, and A. Jeffrey. *Table of integrals, series, and products*. Academic Press, New York, N. Y., 1980.

R. Haberman. *Applied Partial Differential Equations with Fourier Series and Boundary Value Problems*. Pearson Education, Inc., Upper Saddle River, NJ, fourth edition, 2004.

S. Hassani. *Foundations of Mathematical Physics*. Allyn and Bacon, Needham Heights, MA, 1991.

R. L. Herman. *A Course in Mathematical Methods for Physicists*. CRC Press, Taylor & Francis Group, Boca Raton, FL, 2014.

A. J. Jerri. *Integral and Discrete Transforms with Applications and Error Analysis*. Marcal Dekker, Inc, 1992.

A. J. Jerri. *Introduction to Integral Equatons*. Sampling Publishing, second edition, 1999.

K. Kaplan. *Advanced Calculus*. Addison Wesley Publishing Company, fourth edition, 1991.

J. P. Keener. *Principles of Applied Mathematics, Transformation ad Approximation*. Perseus Books, Cambridge, MA, second edition, 2000.

T. W. Körner. *Fourier Analysis*. Cambridge University Press, Cambridge, U.K. ; New York, 1988.

R. G. Lyons. *Understnading Digital Signal Processing*. Prentice Hall PTR, Upper Saddle River, NJ, 2001.

R. J. Marks II. *Introduction to the Shannon Sampling Theorem and Interpolation Theory*. Springer-Verlag, New Yorkl NY, 1991.

A. I. Markushevich. *Theory of Functions*. Chelsea Publishing Company, New York, NY, second edition, 1977.

D. Richards. *Advanced Mathematical Methods with Maple*. Cambridge University Press, Cambridge, U.K. ; New York, 2002.

S. D. Stearns. *Digital Signal Analysis*. Hayden Book Company, Inc, Rochelle Park, NJ, 1975.

J. Stewart. *Calculus: Early Transcendentals*. Brooks Cole, sixth edition, 2007.

G. B. Thomas and R. L. Finney. *Calculus and Analytic Geometry*. Addison-Wesley Press, Cambridge, MA, ninth edition, 1995.

J. S. Walker. *Fourier Analysis*. Oxford University Press, New York, NY, 1988.

H. F. Weinberger. *A First Course in Partial Differential Equations: with Complex Variables and Transform Methods*. Dover Publications, New York, NY, 1995.

Index