An Introduction to Fourier Analysis
Fourier Series, Partial Differential Equations
and Fourier Transforms

Notes prepared for MA3139

Arthur L. Schoenstadt
Department of Applied Mathematics
Naval Postgraduate School
Code MA/Zh
Monterey, California 93943

August 18, 2005

Contents

List of Figures

MA 3139
Fourier Analysis and Partial Differential Equations

Introduction

These notes are, at least indirectly, about the human eye and the human ear, and about a philosophy of physical phenomena. (Now don't go asking for your money back yet! This really will be a mathematics - not an anatomy or philosophy - text. We shall, however, develop the fundamental ideas of a branch of mathematics that can be used to interpret much of the way these two intriguing human sensory organs function. In terms of actual anatomical descriptions we shall need no more than a simple high school-level concept of these organs.)

The branch of mathematics we will consider is called *Fourier Analysis*, after the French mathematician Jean Baptiste Joseph Fourier[1] (1768-1830), whose treatise on heat flow first introduced most of these concepts. Today, Fourier analysis is, among other things, perhaps the single most important mathematical tool used in what we call *signal processing*. It represents the fundamental procedure by which complex physical "signals" may be decomposed into simpler ones and, conversely, by which complicated signals may be created out of simpler building blocks. Mathematically, Fourier analysis has spawned some of the most fundamental developments in our understanding of infinite series and function approximation - developments which are, unfortunately, much beyond the scope of these notes. Equally important, Fourier analysis is the tool with which many of the everyday phenomena - the perceived differences in sound between violins and drums, sonic booms, and the mixing of colors - can be better understood. As we shall come to see, Fourier analysis does this by establishing a simultaneous dual view of phenomena - what we shall come to call the *frequency domain* and the *time domain* representations.

As we shall also come to argue later, what we shall call the time and frequency domains immediately relate to the ways in which the human ear and eye interpret stimuli. The ear, for example, responds to minute variations in atmospheric pressure. These cause the ear drum to vibrate and, the various nerves in the inner ear then convert these vibrations into what the brain interprets as sounds. In the eye, by contrast, electromagnetic waves fall on the rods and cones in the back of the eyeball, and are converted into what the brain interprets as colors. But there are fundamental differences in the way in which these interpretations occur. Specifically, consider one of the great American pastimes - watching television. The speaker in the television vibrates, producing minute compressions and rarefactions (increases and decreases in air pressure), which propagate across the room to the viewer's ear. These variations impact on the ear drum as a *single* continuously varying pressure. However, by the time the result is interpreted by the brain, it has been *separated* into different actors' voices, the background sounds, etc. That is, the nature of the human ear is to take a single complex signal (the sound pressure wave), *decompose* it into simpler components, and *recognize the simultaneous existence* of those different components. This is the essence of what we shall come to view, in terms of Fourier analysis, as frequency domain analysis of a signal.

---

[1]see: http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Fourier.html

Now contrast the ear's response with the behavior of the eye. The television uses electron "guns" to illuminate groups of three different colored - red, green and blue - phosphor dots on the screen. If illuminated, each dot emits light of the corresponding color. These different colored light beams then propagate across the room, and fall on the eye. (The phosphor dots are so closely grouped on the screen that if more than one dot in a group of three were illuminated the different light beams will seem to have come from the same location.) The eye however interprets the simultaneous reception of these different colors in exactly the reverse manner of the ear. When more than one "pure " color is actually present, the eye perceives a *single, composite* color, e.g. yellow. That is, the nature of the human eye is to take a multiple simple component signals (the pure colors), and *combine or synthesize* them into a single complex signal. In terms of Fourier analysis, this is a time domain interpretation of the signal.

1 Infinite Sequences, Infinite Series and Improper Integrals

1.1 Introduction

The concepts of infinite series and improper integrals, i.e. entities represented by symbols such as

$$\sum_{n=-\infty}^{\infty} a_n \ , \quad \sum_{n=-\infty}^{\infty} f_n(x) \ , \quad \text{and} \quad \int_{-\infty}^{\infty} f(x) \ dx$$

are central to Fourier Analysis. (We assume the reader is already at least somewhat familiar with these. However, for the sake of both completeness and clarity, we shall more fully define all of them later.) These concepts, indeed almost *any* concepts dealing with infinity, can be mathematically extremely subtle.

The ancient Greeks, for example, wrestled, and not totally successfully with such issues. Perhaps the best-known example of the difficulty they had in dealing with these concepts is the famous *Zeno's paradox*. This example concerns a tortoise and a fleet-footed runner, reputed to be Achilles in most versions. The tortoise was assumed to have a certain head start on the runner, but the runner could travel twice as fast as the tortoise. (I know that's either an awfully fast tortoise or a slow runner - but I didn't make up the paradox.) Anyway, as the paradox goes, on a given signal both the runner and the tortoise start to move. After some interval of time, the runner will reach the tortoise's original starting point. The tortoise, having moved, will now be half as far ahead of the runner as he was at the start of the race. But then, by the time the runner reaches the tortoise's new position, the tortoise will have moved ahead, and still lead the runner, albeit by only a quarter of the original distance - and so on!
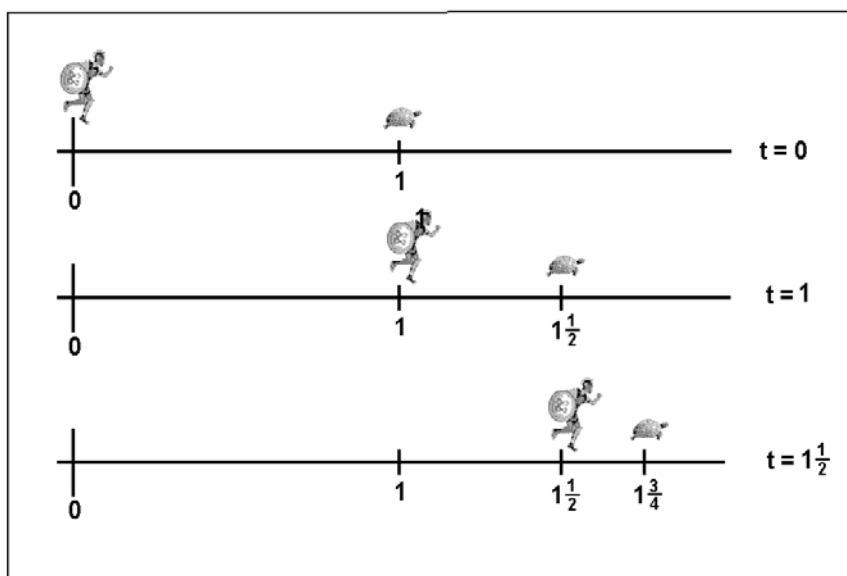
Figure 1: Zeno's Paradox

The paradox is that it appears as if every time the runner reaches the tortoise's old location, the tortoise has continued and is still ahead, even though the distance is closing. So how can the runner catch the tortoise - which experience clearly tells us he will - if the tortoise is always some small distance in the lead? The resolution of this paradox really requires the concept of the infinite series. But first we must establish a little notation.

1.2   Functions and Sequences

Functions are fundamental to mathematics, and to any part of mathematics involving the calculus. There are probably as many different, although basically equivalent, ways to define functions as there are mathematics texts. We shall actually use more than one, although our primary viewpoint will be that of a function as an *input-output* relationship, a "black box, " in which any particular "input" value, x, produces a corresponding "output," $f(x)$.

$$
\begin{array}{ccc}
\text{Input} & & \text{Output} \\
x & \boxed{f(\ \ )} & f(x) \\
\text{(Domain)} & & \text{(Range)}
\end{array}
$$
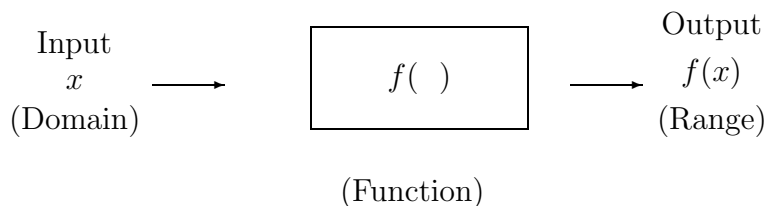
(Function)

Figure 2: The "Black Box" Function

The allowable (or valid) inputs to this black box comprise the *domain* of the function, and the possible outputs its *range*. As an example, consider the natural logarithm function as implemented on most calculators. Somewhere, inside the calculator's circuitry is a "chip," whose function is to take numbers and compute their natural logarithm. The actual algorithm used is totally hidden to the user of the calculator, i.e. it's an opaque box whose inner workings are a mystery. All the user knows is that when they enter a number, for example $x = 1.275$, into the calculator's display, then hit the $ln(x)$ key, the calculator will return with the value $0.2429\ldots\ (=\ ln(1.275))$ in the display. Furthermore, and this is crucial to the concept of a function, the same result will be produced every time for the same input value. (In the "real world," numerical analysts who design computer chips worry a great deal about this very problem, i.e. how to ensure every different manufacturer's chips do in fact produce the same answers.) The domain of the $ln(x)$ function is $0 < x < \infty$, and the range is $-\infty < ln(x) < \infty$. Inputs not in the domain, e.g. $x = -1$, produce, in the calculator, some sort of error message - a flashing display, the word "error," etc.

As we noted above, however, black boxes are not the only way functions may be interpreted. For example, on other occasions, we shall consider real-valued functions to be essentially equivalent to their graphs. Thus, the natural logarithm function could equally well be specified by Figure 3.

From the graphical point of view, the domain of a function consists of all the points on the horizontal axis that correspond to points on the curve, and the range to the equivalent points on the vertical axis. Note that we have deliberately **not** labeled the axes, e.g. called the

2
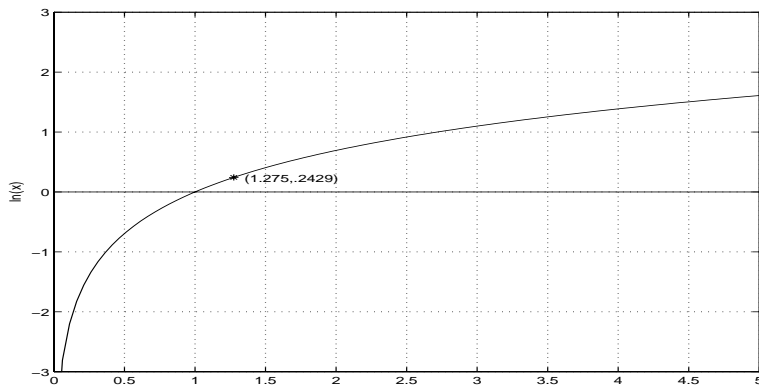
Figure 3: The Natural Logarithm Function (ln( ))

horizontal axis x. We did this to emphasize the essential independence of the function from the particular symbol used to represent the independent (input) variable - i.e. $ln(x)$ as a function of x has the **same** graph as $ln(u)$ as a function of u. In fact, as long as we replace the symbol for the independent variable by the *same* symbol throughout, we will not change the graph. (This is really nothing more than the old Euclidean notion that "equals replaced by equals remain equal.") Therefore, $ln(x^2)$ as a function of x^2 has the same graph as $ln(u)$ as a function of u, even though $ln(x^2)$ as a function of x does not!

The calculus, with which we assume you are already quite familiar, deals with various additional properties and relationships of functions - e.g. limits, the derivative, and the integral. Each of these, in turn, may be viewed in different ways. For example, the derivative,

$$f'(x) = \frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

may be viewed as either the instantaneous rate of change of the function $f(x)$, or as the slope of the tangent line to the curve $y = f(x)$ at x, while the definite integral:

$$\int_a^b f(x)dx$$

may be considered as representing the (net) area under the curve $y = f(x)$ on the interval between the points $x = a$ and $x = b$.

One special class of functions is those whose domain consists only of the integers (the positive and negative whole numbers, plus zero), or some subset of the integers. Such functions are more commonly referred to as **sequences**. Furthermore, in most discussions of sequences, the independent variable is more likely to be represented by one of the letters i, j, k, l, m, or n than by x, y, z, etc. Lastly, although, as functions, they should reasonably be represented by standard functional notation, in practice their special nature is emphasized by writing the independent variable as a **subscript**, i.e. be writing a_n instead of $a(n)$.

Sequences can be specified in several different ways. Among these are:

By listing the terms explicitly, e.g.:
$$a_0 = 1, \quad a_1 = -1, \quad a_2 = \tfrac{1}{2}, \quad a_3 = -\tfrac{1}{6}, \quad a_4 = \tfrac{1}{24}, \quad \text{etc.,}$$

3

By a *recurrence relation*, e.g.:

$$a_0 = 1$$
$$a_n = -\frac{a_{n-1}}{n} \, , \quad n = 1, 2, 3, \dots \quad ,$$

By an explicit formula, e.g.:

$$a_n = \frac{(-1)^n}{n!} \, , \quad n = 0, 1, 2, \dots \, ,$$

Or, by a graph:
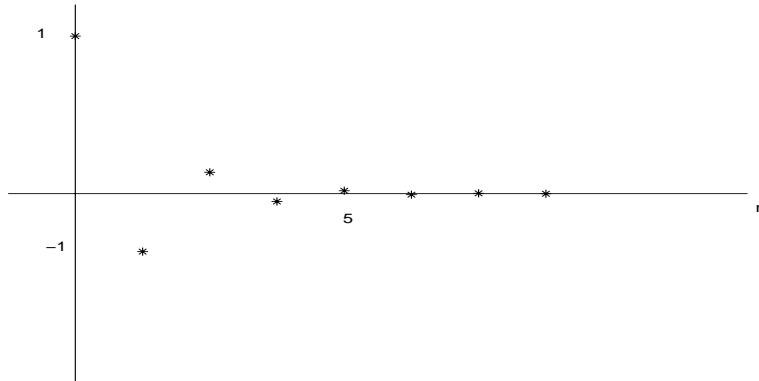
Figure 4: Graph of a Sequence

(Note that since sequences are defined only for integer values of the independent variable, their graphs consist, not of straight lines, but of individual points. In addition, most sequences that we will encounter will be defined only for non-negative integer subscripts.)

Sequences arise in a number of common applications, most of which are outside of the scope of this text. Nevertheless, it's worth mentioning some of them at this point.
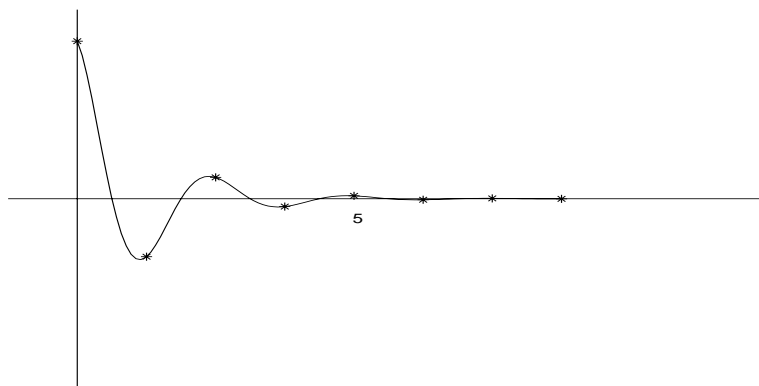
Figure 5: Sampling of a Continuous Signal

Sampling – A common occurrence in signal processing is the conversion of a continuous (analog) signal, say f(t), to a sequence (discrete signal) by simply recording the value of the continuous signal at regular time intervals. Thus,

4

for example, the sampling times would be represented by the sequence t_i, and the resulting sample values by the second sequence $f_i = f(t_i)$.

Probability – Certain random events, for example the arrival of ships in port or of messages at a communications site, occur only in discrete quantities. Thus, for example, the sequence P_n might denote the probability of exactly n arrivals (ships or messages) at a facility during a single unit of time.

Approximation – A significant number of mathematical problems are solved by the process of successive approximations. (So are a number of non-mathematical problems, such as adjusting artillery fire!) In this process, a sequence of (theoretically) better and better approximations to a desired solution are generated. Perhaps the most familiar mathematical instance of this is *Newton's method*, where (most of the time) a sequence of successively better solutions to an equation of the form

$$f(x) = 0$$

can be generated by the algorithm

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad .$$

Because sequences are simply a particular class of functions, they inherit all of the "normal" properties and operations that are valid on functions. This includes the property that the variable (subscript) can be replaced by the same expression, or value, on both sides of any relation, without altering the validity of that relation. Thus, for example, if:

$$a_n = \frac{2^n}{n!} \quad ,$$

then

$$a_k = \frac{2^k}{k!}$$

$$a_{2n-1} = \frac{2^{2n-1}}{(2n-1)!} = \frac{1}{2}\frac{4^n}{(2n-1)!}$$

$$a_4 = \frac{2^4}{4!} = \frac{16}{24} \quad .$$

1.3   Limits

The concepts of limits and limit processes are also central to dealing with sequences (as they are to the calculus). We again assume the reader already has a fairly solid introduction to these concepts, so we shall only briefly review them and present our notation.

When a function of a single real variable has a limit, we shall use the notation:

$$\lim_{x \to x_0} f(x) = A \quad ,$$

or the statement that $f(x)$ **converges** to A as x approaches x_0. We shall treat limits and limiting behavior more as intuitive and pictorial concepts than as mathematical definitions,

although, as you should be aware, these concepts can be defined rigorously (if somewhat abstractly) in the calculus. In our view, a limit exists provided $f(x)$ can be *assured* to be "sufficiently close" to A whenever x is "sufficiently close" (but, strictly speaking, not equal) to x_0. When considered from our black box concept of a function, whether a limit exists or not depends on whether holding the input "within tolerances" is enough *guarantee* the output will be "within specifications." Pictorially, a limit exits when, as the values on the horizontal axis become arbitrarily close to x_0, the corresponding points on the curve become arbitrarily close to A.
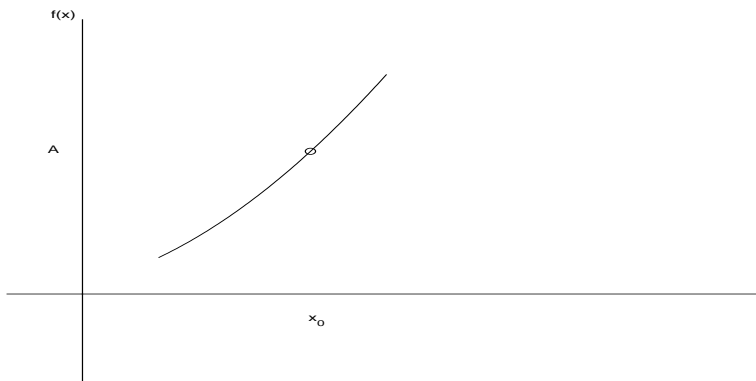
Figure 6: The Pictorial Concept of a Limit

When x_0 is not finite, i.e. when we consider

$$\lim_{x \to \infty} f(x) = A$$

the situation is only slightly changed. Now we consider the limit to exist provided $f(x)$ can be guaranteed to be sufficiently close to A whenever x is "sufficiently large." A look at the corresponding figure (Figure 7) shows that a limit at infinity (or negative infinity) is precisely the condition referred to in the calculus as a *horizontal asymptote*.
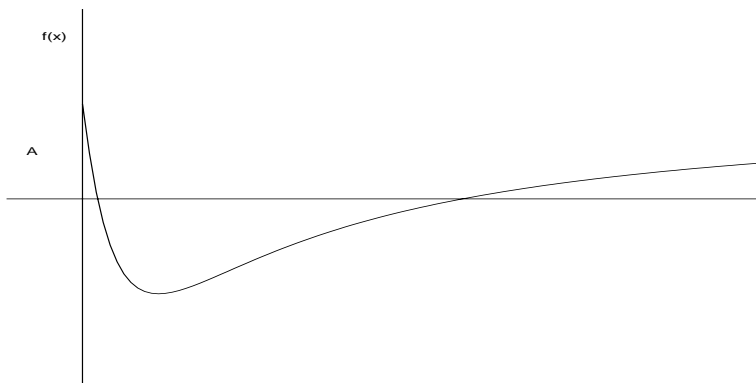
Figure 7: Pictorial Concept of a Limit at Infinity

Limits play a central role in the study of *improper* integrals, one particular class of integral which will be especially important in our studies. Improper integrals are ones whose

values may or may not even exist because either their integrands are discontinuous or the length of the interval of integration is infinite (or both), e.g.

$$\int_0^1 \frac{1}{\sqrt{x}} dx \qquad \text{or} \qquad \int_0^\infty \frac{1}{x^2+1} dx$$

As you should recall, the calculus treats such integrals in terms of limits, and the integrals exist or are said to *converge* whenever the appropriate limits exist. Thus, for example, the second integral above converges (exists), because

$$\lim_{R\to\infty} \int_0^R \frac{1}{x^2+1} dx \;=\; \lim_{R\to\infty} \left. \tan^{-1}(x) \right|_0^R$$

$$= \lim_{R\to\infty} \tan^{-1}(R) = \frac{\pi}{2} \equiv \int_0^\infty \frac{1}{x^2+1} dx$$

If for any given improper integral, the appropriate limit did not exist, then we would say that the improper integral did not converge (*diverged*).

For sequences, however, one important difference emerges. Sequences are defined only for integer values of their argument! Therefore, for a *finite n_0*, the notation

$$\lim_{n\to n_0} a_n$$

is generally really quite silly, since n can **never** be closer to n_0 than one unit away. For sequences, there is no sense in talking about limits *other than at infinity*. The concept, however, of

$$\lim_{n\to\infty} a_n = A$$

is perfectly reasonable! It simply means that a_n can be assured to be arbitrarily close to A if n is sufficiently large. (More colloquially, this means that after some point in the sequence, the values of the terms do not change *effectively* from that point on. For purists, the exact definition is:

$$\lim_{n\to\infty} a_n = A$$

if

> for any arbitrary $\delta > 0$, there exists an N_0 such that
>
> $\mid a_n - A \mid < \delta$ whenever $n > N_0$.   )

Again, as with functions, we shall use the statement that "a_n converges to A" interchangeably with $\lim_{n\to\infty} a_n = A$. We shall also often use the shorthand $a_n \to A$ to denote the existence of the limit. When no limit exists, we use the fairly standard convention of saying the sequence **diverges**.

Several standard techniques may be used to determine the actual existence or non-existence of a limit for a given sequence. The most common perhaps are either inspection or the application of L'Hopital's rule to the sequence (after replacing n by x).

Examples:

$$a_n = \frac{n^3 + 1}{(3n + 1)^3} \quad \text{converges to} \quad \frac{1}{27}$$

$$a_n = \frac{1}{n} \rightarrow 0$$

$$a_n = (-1)^n \quad \text{diverges}$$

$$a_n = n \quad \text{diverges} \quad .$$

(Note that in using L'Hopital's rule, the existence of a L'Hopital's limit as $x \rightarrow \infty$ implies the existence of a limit for the sequence, but not necessarily conversely. One such case where the sequence limit exists but the L'Hopital's limit does not is $a_n = \sin(n\pi)$. You should be able to explain this example both mathematically and graphically!)

The statement that a sequence converges is sufficient to guarantee that *eventually* the terms in the sequence will become arbitrarily close to the limit value. However, knowing **only** that a series converges fails to convey one crucial fact - the *speed* of that convergence, i.e. how many terms must be evaluated before the limit is reasonably well approximated. This can be illustrated by the following table of values of two different sequences:

$$a_n = \frac{n^3 + 1}{3(n + 1)^3} \quad \text{and} \quad b_n = \frac{n^3 - 3}{3n^3} \quad ,$$

both of which converge to the limit $\frac{1}{3}$:

n	a_n	b_n
5	0.1944	0.3253
10	0.2508	0.3323
20	0.2880	0.3332
100	0.3235	0.3333
500	0.3313	0.3333

Note that, at every subscript, the value of b_n contains at least a full significant digit more of accuracy than a_n. As we shall come to see in this course, this is not a trivial consideration. Like any computation, evaluating terms in a sequence is **not** "free," and therefore in most cases it "costs more" to work with slowly converging sequences. In the next section we consider a notation that can significantly simplify the discussion of how fast certain sequences converge.

1.4   The Order Notation

Order notation is used to describe the limiting behavior of a sequence (or function). Specifically, for sequences, order notation relates the behavior of two sequences, one ("α_n") usually

fairly "complicated," the other, ("β_n") fairly "simple," and expresses the concept that, in the limit, either:

(1) The sequences α_n and β_n are essentially identical, or

(2) The sequence β_n "dominates" the sequence α_n in the sense that β_n can serve as a "worst case" approximation to α_n .

Order notation is useful for discussing sequences (and later for discussing infinite series) because the convergence and divergence of these is really determined by their behavior for *large n*. Yet frequently when n is large the terms of a sequence/series can be conveniently approximated by much simpler expressions. For example, when n is large,

$$\alpha_n = \frac{(7n^2 + 1)\cos(n\pi)}{3n^4 + n^3 + 19} \qquad \text{is very close to} \qquad \beta_n = \pm\frac{7}{3n^2} \quad .$$

The first notation we introduce that can be used to convey information on the relative behavior of sequences is:

$$\alpha_n = \mathbf{O}(\beta_n) \qquad \text{if, for some constant } C , \qquad |\alpha_n| \le C|\beta_n| , \quad \text{for all } n .$$

Colloquially, we shall say in the above case either that α_n is *of order* β_n, or, slightly more precisely, that α_n is *"big Oh"* of β_n.

There is one drawback, however, with the order notation as described above. This drawback is that, because the notation involves the *less than or equal to* relation, there are instances where statements made using the notation, while correct, are somewhat unenlightening. For example, since

$$\frac{1}{n^2} \le \frac{1}{n} \qquad \text{then, strictly speaking,} \qquad \frac{1}{n^2} = \mathbf{O}\left(\frac{1}{n}\right) \quad .$$

There are a couple of ways that this can be "cleaned up." One is to introduce a second type of order, sometimes called *"little oh"* or *"small oh."* Specifically, we say

$$\alpha_n = \mathbf{o}(\beta_n) \qquad \text{if} \qquad \lim_{n \to \infty} \frac{|\alpha_n|}{|\beta_n|} = 0 \quad .$$

Note that $\alpha_n = \mathbf{o}(\beta_n)$ implies immediately that $\alpha_n = \mathbf{O}(\beta_n)$, but *not necessarily* the converse. Thus we can now convey that two sequences α_n and β_n involve terms of "about the same size" by saying

$$\alpha_n = \mathbf{O}(\beta_n) , \qquad \text{but} \qquad \alpha_n \neq \mathbf{o}(\beta_n) \quad .$$

A somewhat similar, but slightly stronger statement is conveyed by the notation of asymptotic equivalence. Specifically, we say

$$\alpha_n \text{ is } asymptotically\ equivalent \text{ to } \beta_n \qquad \text{if} \qquad \lim_{n \to \infty} \frac{\alpha_n}{\beta_n} = 1 \quad .$$

Asymptotic equivalence is generally denoted by

$$\alpha_n \sim \beta_n \quad .$$

There are some "standard" rules that simplify determining the order of a sequence. These include:

(1) In any polynomial in n , as $n \to \infty$, the *highest* power will dominate, and, in the limit, the polynomial can be approximated by the highest power term for any computation not involving later subtraction of another polynomial of the same degree. Thus, for example, unless it is to be added or subtracted to another quartic polynomial, in the limit

$$2n^4 + 20n^3 - 500n^2 + n - 1000$$

can simply be replaced by
$$2n^4 \quad .$$

(2) For all values of n and x,

$$|\sin(n\pi x)| \le 1 \quad \text{and} \quad |\cos(n\pi x)| \le 1 \ .$$

(3) For large n ,
$$n! \approx \left[\frac{n}{e}\right]^n \sqrt{2\pi n} \quad .$$

Thus, for example, in the sequence,

$$\alpha_n = \frac{1 + \cos(n\pi)}{(n^2 + 1)\pi^2} \quad ,$$

the following order notation statements are true:

$$\alpha_n \ = \ \mathbf{O}\left(\frac{1}{n^2}\right) \quad , \qquad \alpha_n \ = \ \mathbf{o}\left(\frac{1}{n}\right) \quad ,$$

$$\alpha_n \ \ne \mathbf{o}\left(\frac{1}{n^2}\right) \quad , \qquad \alpha_n \ \sim \frac{2}{\pi^2 n^2} \ , \quad n = 0, 2, 4, \dots \quad .$$

The last important point we would emphasize before leaving this section is that order notation is simply a convenient shorthand for expressing information about the growth or decay of the terms of a sequence. It does not, in and of itself, *prove* that such behavior actually occurs. The information about the actual rates of growth and decay of any sequence must still be obtained by basic calculus techniques. The order notation simply allows to express this information, once found, in a concise form.

1.5 Infinite Series

Again, we shall assume some prior familiarity with this topic on the part of the reader, and cover only some high points. By definition, an infinite series is formed by adding all the terms of a sequence. Note that in general, this theoretically requires adding together an *infinite* number of terms. For example, if

$$a_n = \left(\frac{2}{3}\right)^n , \quad n = 0, 1, 2, \ldots \quad ,$$

then the infinite series formed from a_n is

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + a_3 + \cdots$$

$$= \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n$$

$$= 1 + \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \cdots$$

In the vernacular of FORTRAN, an infinite series is thus equivalent to a **DO** loop whose upper limit is unbounded. Clearly, evaluation of an infinite series by direct addition, i.e. by "brute force," is impractical, even *physically impossible*. This is the heart of Zeno's paradox. As posed by the ancient Greeks, the paradox hinges on adding together an infinite number of time intervals, each half of the previous one. The logical flaw in the Greeks' analysis was the *implicit assumption* that the sum of an infinite number of non-zero terms was *necessarily* infinite. Correctly understanding this paradox requires the recognition that Zeno's paradox, or for that matter any infinite series, involves the sum of an an infinite number of terms, and therefore the only **valid** method of analysis is in terms of limits.

The mathematical analysis of infinite series starts with the recognition that every infinite series in fact involves **two** sequences:

(1) The sequence of **terms** - a_n,

and

(2) The sequence of **partial sums** -

$$S_N = a_0 + a_1 + \cdots + a_N$$

$$= \sum_{n=0}^{N} a_n$$

If the sequence of partial sums itself then has a limit, i.e. if

$$\lim_{N \to \infty} S_N = S$$

exists, then we can meaningfully talk of adding together the infinite number of terms in the series and yet getting a *finite* answer. In such a case, we shall say the infinite series converges. (If the limit doesn't exist, then we say the series diverges.) Fundamental to this definition is the understanding that while ultimately convergence or divergence of the series depends on the terms (a_n), the primary quantity analyzed in deciding whether a series converges or diverges is the sequence of partial sums (S_N). Furthermore, and this frequently causes confusion when students first encounter infinite series, since two sequences are involved, there is always the possibility that one will converge and the other not. To be more precise, while the infinite series (sequence of partial sums) cannot converge if the sequence of terms does not, convergence of the sequence of terms of a series to a limit (even to a limit of zero) **does not** guarantee the infinite series will converge. Many examples exist where the sequence of terms converges, but the series does not. One such series is given by:

$$a_n \quad \equiv 1 \ , \quad n = 0, 1, 2, \ldots$$

$$S_N \quad = \sum_{n=0}^{N} 1 = 1 + 1 + \cdots + 1 = N + 1 \qquad .$$

The above simple example illustrates another fundamental point about infinite series – determining the convergence or divergence of a series *directly* from the definition requires first having (or finding) a *formula* for the partial sums. There are a few, but only a few, classic non-trivial cases in which this is possible. The most well-known example of these is the *geometric series*:

$$S_N = \sum_{n=0}^{N} r^n \quad = 1 + r + r^2 + \cdots + r^N$$

$$= \frac{1 - r^{N+1}}{1 - r} \ , \quad r \neq 1 \quad ,$$

where r is any real number. With this formula, one can then show quite easily that

$$\sum_{n=0}^{\infty} r^n = \begin{cases} \frac{1}{1-r,} & |r| < 1 \\ \\ diverges & otherwise \end{cases}$$

Coincidentally, this formula also resolves Zeno's paradox. As we observed before, each time interval in the paradox is exactly half of the previous interval. Therefore, assuming the first interval, i.e. the time from when the "race" starts until the runner reaches the tortoise's initial position, is one unit of time, the sequence of times is given by:

$$t_n = \left(\frac{1}{2}\right)^n \quad .$$

Hence, the time until the runner catches up with the tortoise is:

$$\sum_{n=0}^{\infty} \left(\tfrac{1}{2}\right)^n = \frac{1}{1 - \frac{1}{2}} = 2 \qquad ,$$

which is *exactly* what common sense tells us it should be.

12

1.6 Convergence Tests

Unfortunately, the majority of infinite series seldom behave as fortuitously as in the last example in the previous section. We almost never, at least in cases of practical interest, have an explicit formula for the partial sums. Yet, because of the importance of infinite series in applications, we cannot, in such cases, simply throw up our hands and say that because we don't have an explicit formula to test for a limit that we then don't know whether the series converges or not. There is a real need for methods which determine whether a series converges or not *without having an explicit formula for the partial sums.* Such methods fall under the general classification of *convergence tests.* These tests are quite valuable, since in general they are applied to the sequence of *terms* of the series, i.e. the a_n. They therefore have the advantage of working with what is almost always a known, analytically expressible sequence. However, one seldom, if ever, gets something for nothing in this world. Convergence tests prove no exception to this rule, and exact two penalties in return for the convenience of allowing us to work only with the a_n. The first of these is that there is no single universally applicable or universally conclusive convergence test. Every such test has at least some series to which it simply cannot be applied, and others for which it will be unable to reach a definite conclusion (i.e. for which it mathematically shrugs it shoulders). The second penalty we must pay for the use of convergence tests is a drawback common to almost all of what are mathematically called *existence* results. This drawback is that, even when convergence tests conclusively state a series will converge, they do not provide *the value* of the sum, i.e. the limit of the partial sums, but only state that the sum (limit) exists. Therefore, in practical applications, convergence tests must be followed by computational algorithms and procedures that allow the sum of a convergent infinite series to be approximated to within acceptable accuracy.

There are a large number of convergence tests. For our purposes, however, only a small subset of these are commonly needed. These are shown in Table 1.

For example, using these tests, we can show that:

$$\sum_{n=0}^{\infty} \frac{3n^2 - n}{n^4 - 3n^3 + 1} \quad \text{converges since} \quad a_n = \mathbf{O}\left(\frac{1}{n^2}\right)$$

$$\sum_{n=0}^{\infty} \frac{n}{n^2 + 1} \quad \text{diverges since} \quad \int_1^{\infty} \frac{x}{x^2 + 1}\, dx \quad \text{diverges}$$

$$\sum_{n=1}^{\infty} \frac{n}{n + 1} \quad \text{diverges since} \quad \lim_{n\to\infty} \frac{n}{n + 1} = 1 \neq 0$$

Note that, as we observed earlier, none of the convergence tests in Table 1 provide formulae to calculate the limit of a convergent infinite series. This limit can, in general, only be *approximated* - by taking a partial sum of "enough" terms in the series so that the partial sum is "sufficiently close" to the final limit. But this procedure implicitly assumes some *a priori* knowledge of how many terms are "enough" to give a good approximation. Such knowledge, however, would be impossible without some information on how "fast" the series is converging. So where would this information come form? After all, convergence

Convergence Tests

1. If $\lim_{n\to\infty} a_n \neq 0$, then $\sum a_n$ *diverges.*
   (This test is inconclusive if $\lim_{n\to\infty} a_n = 0$.)

2. *The Absolute Convergence Test -*
   If $\sum |a_n|$ converges, then $\sum a_n$ converges.
   (This test is inconclusive if $\sum |a_n|$ diverges.)

3. *The Comparison Test -*
   If $\dfrac{a_n}{b_n} \to A$ and $A \neq 0, \pm\infty$, then $\sum a_n$ and $\sum b_n$ either both converge or both diverge.

4. *The Alternating Series Test -*
   If $a_n a_{n+1} < 0$ (i.e. if the terms of the series alternate algebraic signs), if $|a_{n+1}| \leq |a_n|$ and if $\lim_{n\to\infty} a_n = 0$, then $\sum a_n$ *converges.*

5. *The Integral Test -*
   If $a_n > 0$, and $a_n = f(n)$ for a function $f(x)$ which is positive and monotonically decreasing, then $\sum a_n$ and $\int_1^\infty f(x)$ either both converge or both diverge.

6. *The p-Test -*
   $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p > 1$,
              diverges if $p \leq 1$.

7. *The order n^p-Test -*
   If $a_n = \mathbf{O}\left(\frac{1}{n^p}\right)$, $p > 1$, then $\displaystyle\sum_{n=1}^{\infty} a_n$ converges.

8. *The Geometric Series Test -*
   $\displaystyle\sum_{n=0}^{\infty} r^n$ converges if $|r| < 1$.
            diverges if $|r| \geq 1$.

tests generally only provide the information that a given series converges. This only implies that the sequence of partial sums eventually approaches a limit - convergence tests provide no information about *how fast* the limit is being approached. For example, consider the two series:

$$\sum_{n=0}^{\infty} \frac{1}{n!} = e = 2.7183, \quad \text{and}$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} = 1.6449 \quad ,$$

both of which converge according to one or more of the convergence tests, and the following table of their partial sums:

N	$\sum_{n=0}^{N} \frac{1}{n!}$	$\sum_{n=1}^{N} \frac{1}{n^2}$
3	2.6667	1.3611
5	2.7167	1.4636
10	2.7183	1.5498
100	2.7183	1.6350

Clearly, after ten times as many terms, the partial sum of one hundred terms of the second series does not provide as good an approximation to the value of that series than does a ten-term partial sum of the first.

As alluded to above, determination of the rate of convergence of an infinite series is of far more than just academic interest when we try to approximate the value of a convergent infinite series by computing a partial sum. Without some realistic way of determining *how many* terms must be computed in order to obtain a "good" approximation, such a computation would amount to little more than a mathematical shot in the dark. At best, we would take far more terms than really necessary - getting an accurate result, although paying far too much in computational cost for it. At worst, however, we might stop our sum prematurely - and end up with a terribly inaccurate approximation. Fortunately, in the next section we shall see that if we have some idea of the order of the terms in a series, then we can in fact estimate fairly well *a priori* how accurately a given partial sum will approximate the final series sum.

1.7 Error Estimates

Given that in general the only reasonable way to approximate the (unknown) sum of an infinite series is by a partial sum, the natural measure for the accuracy of that approximation would be the difference between the approximation and the actual sum. This measure is "natural" in that since the series converges, we know that

$$S_N = \sum_{n=0}^{N} a_n \to S \quad ,$$

where S is the sum of the series. Therefore, for "large" N , S_N should be "close to" S, and so the *error* in approximating S by S_N is simply:

$$E_N = S - S_N \quad = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{N} a_n$$

$$= \sum_{n=N+1}^{\infty} a_n \quad .$$

On the face of it, this may not seem terribly illuminating, since in general only S_N can be known, and E_N is just another (unknown) convergent series. However, as we shall see, in many cases, E_N can be *estimated*, at least to its order of magnitude, and an order of magnitude estimate to E_N is all that is needed to determine how many digits in S_N are significant. Of course, obtaining such an estimate may not be trivial, since E_N depends on **both** a_n and N. But, for a reasonably large class of series, obtaining such error estimates is also really not all that difficult.

Consider the infinite series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} , \qquad p > 1 \quad .$$

Following our definition above, the error in approximating this series by any partial sum of its terms is thus given by

$$E_N = \sum_{n=N+1}^{\infty} \frac{1}{n^p} = \frac{1}{(N+1)^p} + \frac{1}{(N+2)^p} + \frac{1}{(N+3)^p} + \cdots \quad .$$

But now look at Figure 8. Note that since each rectangle in that figure has a base of length one, then the areas of those rectangles are, respectively,

$$\frac{1}{(N+1)^p} , \quad \frac{1}{(N+2)^p} , \quad \frac{1}{(N+3)^p} , \quad \cdots \quad ,$$

and therefore the total shaded area under the blocks in Figure 8 is precisely equal to the error, E_N.

However, clearly the shaded (rectangular) area in Figure 8 is also strictly smaller than the area under the curve in that figure. The importance of this observation is while we cannot compute the sum of the error series exactly, we can compute the slightly larger integral. Thus we have

$$\sum_{n=N+1}^{\infty} \frac{1}{n^p} \leq \int_{N}^{\infty} \frac{dx}{x^p} = \frac{1}{(p-1)N^{p-1}} \quad ,$$

or, for this series

$$E_N \leq \frac{1}{(p-1)N^{p-1}} \quad .$$

The accuracy of this upper bound to E_N can be checked for the case $p = 2$, whose partial sums were computed in an earlier example, since in that case we know the exact value of the infinite series, i.e.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad .$$

16
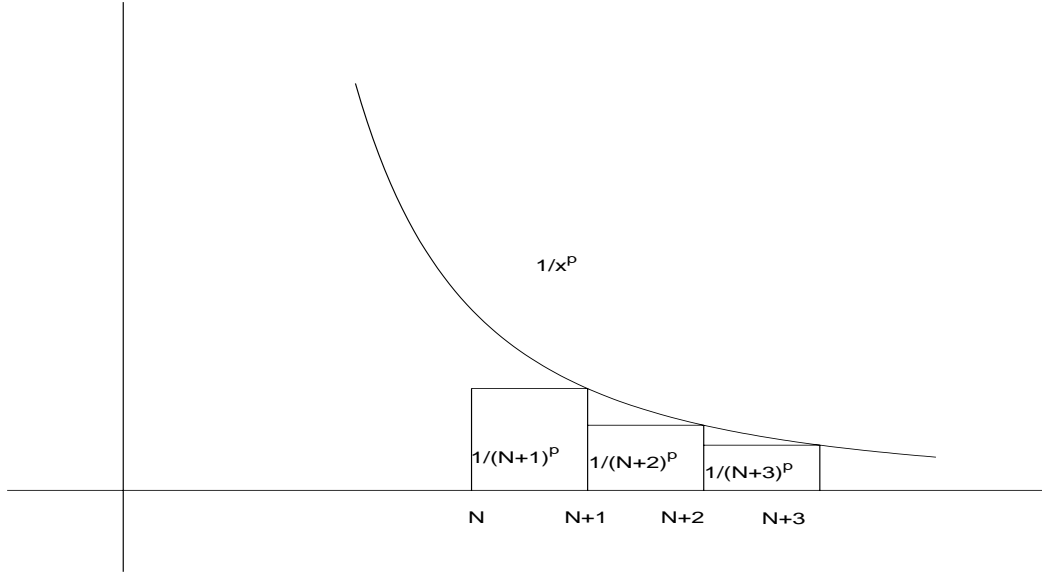
Figure 8: Estimating the Error of a Partial Sum

The results of computing the actual error and comparing it to the upper bound given by the integral are shown in the following table:

N	E_N	$\dfrac{1}{(p-1)N^{p-1}}$
3	0.28382	0.33333
5	0.18132	0.20000
10	0.09517	0.10000
100	0.00995	0.01000

As this table demonstrates, the integral error bound in fact gives an excellent order of magnitude estimate to the error for $N > 5$.

This last result, which was obtained for the very specific series $a_n = \frac{1}{n^p}$ can be fairly straightforwardly extended to any series whose terms are of $\mathbf{O}\left(\frac{1}{n^p}\right)$ for $p > 1$. This happens because when

$$a_n = \mathbf{O}\left(\frac{1}{n^p}\right) \quad,$$

then, for some constant C ,

$$|a_n| \leq \frac{C}{n^p} \quad.$$

But then, because the absolute value of a sum is less than or equal to the sum of the absolute

17

values of the terms in that sum, we have

$$|E_N| = \left| \sum_{n=N+1}^{\infty} a_n \right| \leq \sum_{n=N+1}^{\infty} |a_n|$$

$$\leq \sum_{n=N+1}^{\infty} \frac{C}{n^p}$$

$$\leq \int_{N}^{\infty} \frac{C}{x^p} \, dx = \frac{C}{(p-1)N^{p-1}} \quad ,$$

or

$$|E_N| \leq \frac{C}{(p-1)N^{p-1}} \quad .$$

This last result can be restated as

$$\text{if} \quad a_n = \mathbf{O}\left(\frac{1}{n^p}\right) \quad , \quad \text{then} \quad E_N = \mathbf{O}\left(\frac{1}{n^{p-1}}\right) \quad .$$

Therefore, in such a series, doubling the number of terms used in a partial sum approximation reduces the error by a factor of $\frac{1}{2^{p-1}}$. While this result does not generally provide as tight bounds as in the case when the coefficients are exactly $\frac{1}{n^p}$, it nevertheless works acceptably well in almost all cases.

1.8 Sequences of Functions

In our discussion of sequences and series thus far, we have implicitly been using a more restrictive definition of a sequence than was necessary. Specifically we defined a sequence to be a function defined on the integers, and considered a function to be any unique input/output black box. But in all our examples these were all only *real-valued* functions. There are, however, other sequences than just real-valued ones. For example, there are *sequences of functions* - i.e. unique input/output black boxes, for which the only valid inputs are integers, but for which the outputs are not numbers, but other functions. For example

$$f_n(x) = x^n \ , \quad n \geq 0$$

(with the implicit understanding that $x^0 \equiv 1$) defines a sequence of functions (graphs), the first few members of which are displayed in Figure 9.

Convergence of sequences of functions can be defined in a reasonably similar manner to convergence of sequences of constants, since at a **fixed** x, the sequence of function values is in fact just a sequence of constants. Convergence of sequences of functions is different however in that these values now depend not only on the functions in the sequence itself, but on *the value of x* as well. Thus,

$$f_n(x) \rightarrow f(x)$$

18

Figure 9: A Sequence of Functions
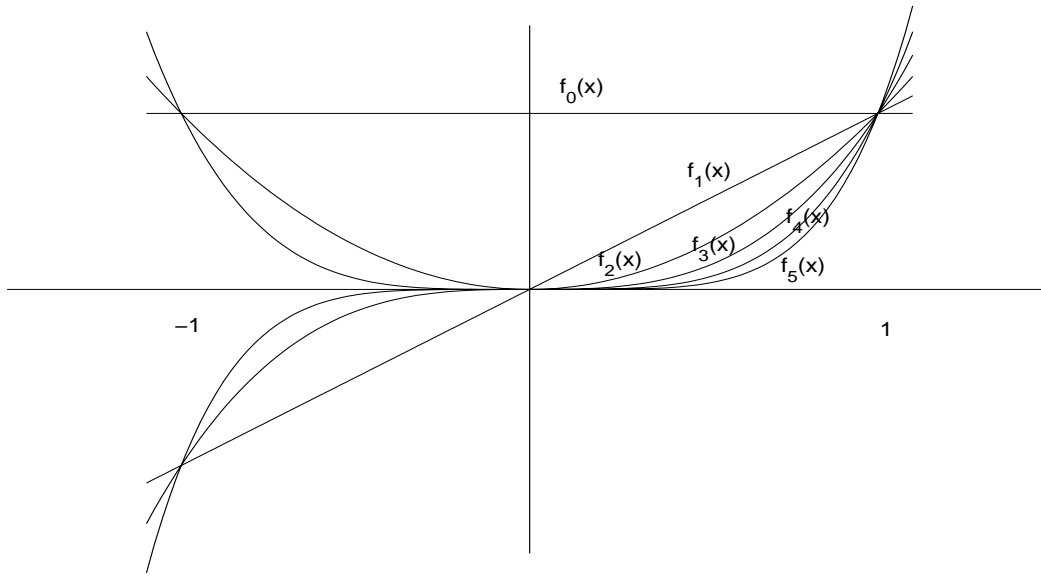
is, in some sense, an incomplete statement, until the values of x at which it holds true are also specified. For example,

$$x^n \to 0$$

but **only** for $-1 < x < 1$. Thus, to completely specify the behavior of this sequence, we would have to state

$$x^n \to 0 , \quad -1 < x < 1 \quad ,$$

while,

$$x^n \to 1 , \quad x = 1 \quad ,$$

and the sequence does not converge if either $x \leq -1$ or $x > 1$. Such behavior is commonly referred to as *pointwise* convergence and divergence.

We would finally note that convergence of sequences of functions has a strong graphical interpretation, just as convergence of a sequence of constants could be viewed in terms of tables where the values do not change after some point. With sequences of functions, the picture is one of a sequence of curves (graphs), which, eventually, begin to lie on top of each other, until they finally become indistinguishable from the limit curve.

Infinite series of functions, e.g.

$$\sum_{n=0}^{\infty} f_n(x)$$

can also be defined similarly to infinite series of constants, since again, at any fixed x, an infinite series of functions reduces to just a series of constants. Again, however, as with the sequence of functions, we must explicitly recognized that the resulting sequence of partial sums

$$S_N(x) = \sum_{n=0}^{N} f_n(x)$$

19

depends on **both** N and x. Therefore, whether a given infinite series of functions converges or diverges is generally affected by the choice of x. For example, for the sequence of functions defined above, the infinite series

$$\sum_{n=0}^{\infty} x^n$$

converges for all $-1 < x < 1$ (by the geometric series test) and therefore

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

(Note that this infinite series does **not** converge for $x = 1$, even though the sequence of terms (functions) does!)

Furthermore, even where a series of functions converges, the *rate* at which it converges, and hence the number of terms necessary for a good approximation, generally depends on the value of x at which it is being evaluated. For example, we have an explicit formula for the partial sum of the geometric series,

$$S_N(x) = \frac{1 - x^{(N+1)}}{1-x} \ ,$$

and therefore it is fairly easily shown that

$$E_N(x) = \frac{x^{(N+1)}}{1-x} \ .$$

Clearly in this latter formula, $E_N(x)$ does depend on both N and x. The degree of this dependence, however is perhaps more strikingly demonstrated by the following table, which displays the error in approximating the sum of the above series for selected values of x and N .

x	$E_5(x)$	$E_{10}(x)$
0.10	1.1×10^{-6}	1.1×10^{-11}
0.25	0.00033	3.1×10^{-7}
0.50	0.03125	0.00098
0.75	0.71191	0.16894
0.95	14.70183	11.37600

As with sequences of functions, this kind of convergence for infinite series of functions - where the convergence properties may vary with the value of x - is commonly referred to as *pointwise convergence*. While better than no convergence at all, pointwise convergence is not particularly desirable, especially from a computational viewpoint, since any program used to compute partial sums in such a case would have to be able to adjust the upper limit of the sum depending on the value of x. If for no other reason than programming convenience, we would far prefer that, given some desired accuracy criterion, there were a *single* number

such that if we took a partial sum with that number of terms, we would be *guaranteed* to be within the desired tolerance of the actual limit (even if taking that number of terms were slightly inefficient at some values of x). That is, we would like to have a convergence behavior which is somewhat *uniform* across **all** the values of x . Such a behavior, when it occurs, is naturally termed *uniform convergence*. Mathematically, it is defined as follows:

The series

$$\sum_{n=0}^{\infty} f_n(x)$$

converges uniformly if, for **every** $\epsilon > 0$, there is an N_ϵ such that

$$|E_N(x)| < \epsilon \quad \text{if} \quad N > N_\epsilon , \quad \text{for all} \quad x \quad .$$

Notice that the English "translation" of this definition is exactly as outlined above. Uniform convergence happens when, given a specified tolerance (ϵ), there is some number of terms (N_ϵ) such that taking a partial sum with at least that many terms will guarantee that the result is within the specified tolerance of the actual value, irrespective of the value of x . Unfortunately, as with most properties of infinite series of functions, whether a series converges uniformly or not depends, more often or not, on the interval of x values of interest. This behavior can again be illustrated by our example series

$$\sum_{n=0}^{\infty} x^n \quad ,$$

for which, as shown above,

$$E_N(x) = \frac{x^{(N+1)}}{1 - x} \quad .$$

Therefore, for $-\frac{1}{2} < x < \frac{1}{2}$,

$$|E_N(x)| = \left| \frac{x^{(N+1)}}{1 - x} \right| \leq \frac{\left(\frac{1}{2}\right)^{N+1}}{1 - \frac{1}{2}} = \frac{1}{2^N} \quad ,$$

and therefore, on this interval, the series converges uniformly. However, no similar relationship can be found on the interval $-1 < x < 1$, and therefore, the series does not converge uniformly on that interval.

The above discussion may seem a bit confusing. If so, don't lose heart completely! Uniform convergence **is** a fairly delicate concept. It is also a central concept in series of functions, and could easily encompass a chapter itself. We shall not devote nearly as much time as is deserved to this topic, but will only scratch the surface in terms of the properties associated with uniform convergence. Fortunately, we shall not need to use most of the results and mathematical theory associated with uniform convergence, because the series that we will consider in this text are in general reasonably straightforward to analyze.

For our purposes, there is one result associated with uniform convergence which transcends all the others. This is covered by the following

Theorem If the series of functions $\sum\limits_{n=0}^{\infty} f_n(x)$ converges uniformly to a limit function $f(x)$, and if each of the $f_n(x)$ is continuous, then $f(x)$ is also continuous.

Note that this theorem does not tell whether any particular given series converges uniformly or not, only what one may conclude after they have determined that series does converge uniformly. Thus, for example, this theorem allows us to conclude that the sequence of functions $f_n(x) = x^n$ will converge to a continuous function for $-1/2 < x < 1/2$, but only after we have determined, as we did above, that the sequence converged uniformly on this interval. The actual determination of whether any particular series converges uniformly, and where, is generally as difficult, or even more so, than determination of convergence of a series of constants. Fortunately however, for our purposes, there is one extremely easy-to-apply test for uniform convergence that will apply to most of the series we must deal with. This test - the so-called *Weierstrass M-test* - is really just another comparison test.

Theorem Given the sequence of functions $f_n(x)$, if there is a sequence of constants, M_n, such that on some interval $a \le x \le b$,

$$|f_n(x)| \le M_n \qquad \text{and} \qquad \sum_{n=0}^{\infty} M_n \ \text{converges,}$$

then

$$\sum_{n=0}^{\infty} f_n(x)$$

converges uniformly on that same interval

(Note that probably the only reason to call this the M-test is that Karl Theodore Wilhelm Weierstrass[2] called his sequence M_n instead of a_n. Had he done the latter, we would probably all call this the Weierstrass a-test today! Furthermore, note that, the Weierstrass test suffers the same limitations as do all convergence tests - there will be series to which it does not apply, and even when it shows that a series does converge uniformly, it does not provide any information on what the limit of that series is.)

We can apply this theorem immediately to a problem very representative of the type of infinite series we shall deal with for the rest of this text -

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \cos(nx) = \cos(x) + \frac{1}{4}\cos(2x) + \frac{1}{9}\cos(3x) + \cdots \quad .$$

For this series, the functions which form the sequence of terms satisfy the conditions of the Wierstauss M-test,

$$|f_n(x)| = \left| \frac{1}{n^2}\cos(nx) \right| \le \frac{1}{n^2}, \quad \text{for all } x .$$

Therefore, since

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \quad \text{converges} \quad ,$$

---

[2] see http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Weierstrass.html

we conclude that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \cos(nx)$$

converges uniformly for *all x*, and therefore the infinite series represents a continuous function.

PROBLEMS

1. For each of the following sequences, determine if the sequence converges or diverges. If the sequence converges, determine the limit

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

   b. $a_n = \dfrac{(n+1)^2}{5n^2 + 2n + 1}$

   c. $a_n = \dfrac{\sin(n)}{n+1}$

   d. $a_n = \cos(n)$

   e. $a_n = \dfrac{2(n+1)^2 + e^{-n}}{3n^2 + 5n + 10}$

   f. $a_n = \dfrac{n\cos(\frac{n\pi}{2})}{n+1}$

   g. $a_n = \dfrac{\cos(n\pi)}{n^2 + 1}$

   h. $a_n = \dfrac{e^n}{n!}$

   i. $a_n = \dfrac{n\sin(n\pi)}{n+1}$

2. Determine the order ("big Oh") of the following sequences

   a. $a_n = \dfrac{n^3 + 2n^2 + 1000}{n^7 + 600N^6 + n}$

   b. $a_n = \dfrac{\cos(n\pi)}{n^2 + 1}$

   c. $a_n = \left[\dfrac{n}{n^2 - 1} - \dfrac{n}{n^2 + 1}\right]\sin((n + \frac{1}{2})\pi)$

   d. $a_n\dfrac{10n^3 e^{-n} + n^2}{(2n+1)^2}\cos(n^2\pi)$

3. Consider the infinite series

$$\sum_{n=0}^{\infty} \frac{(n+1)^2 2^n}{(2n)!}$$

   a. Compute, explicitely, the partial sums S_3 and S_6

   b. Write the equivalent series obtained by replacing n by $k-2$, i.e. by shifting the index.

4. Determine whether each of the following infinite series diverges or converges:

   a. $\displaystyle\sum_{n=0}^{\infty} e^{-n}$

   b. $\displaystyle\sum_{n=0}^{\infty} \frac{n^2 + 1}{(n+1)^3}$

   c. $\displaystyle\sum_{n=0}^{\infty} \frac{n^2\cos(n\pi)}{(n^3 + 1)^2}$

   d. $\displaystyle\sum_{n=0}^{\infty} \frac{n}{n+3}$

   e. $\displaystyle\sum_{n=0}^{\infty} \frac{e^n}{n!}\cos(n\pi)$

   f. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n\ln(n)}$

5. Determine an (approximate) upper bound to the error when each of the following infinite series is approximated by a twenty-term partial sum (S_{20}).

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

   b. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^5}$

   c. $\displaystyle\sum_{n=1}^{\infty} \frac{(2n+1)^2}{n^4}$

6. Consider the series:

$$\sum_{n=0}^{\infty} x^n$$

   a. plot the partial sums $S_1(x)$, $S_5(x)$, $S_{10}(x)$, and $S_{20}(x)$ for $-2 < x < 2$.

   b. What can you conclude about the convergence of the partial sums in this interval?

   c. What, if anything, different can you conclude about the convergence of these partial sums in the interval $-\dfrac{1}{2} < x < \dfrac{1}{2}$.

2 Fourier Series

2.1 Introduction

In the last chapter we reviewed and discussed the concept of infinite series of functions. Such series, the first example of which most students encounter are the Taylor series, are frequently used to approximate "complicated" functions in terms of "simpler" ones. For example, the transcendental function e^x is exactly computable only at $x = 0$. However, the Taylor series,

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

can compute, to any desired degree of accuracy, approximate values for this complicated function in terms of simple powers of x. In this chapter, we introduce what is almost without question the most commonly used infinite series of functions after the Taylor series - the *Fourier series*.

By definition, a Fourier series is an infinite series of the form:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left\{ a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right\}, \qquad (2.1.1)$$

where L is some positive number. (Note that this form is, unfortunately, not quite universal. Some texts write the leading coefficient in the Fourier series as just a_0, rather than as $(a_0/2)$. Still other texts will use the quantity $(T_0/2)$ instead of L. As we shall also try to point out at the appropriate time, while there is nothing wrong with either of these alternative formulations, they simply lead to what we feel are slightly more involved formulas.)

Since the Fourier series is an infinite series, then based on our earlier discussion, several questions should immediately come to mind. Among these, we shall consider the following:

1. For what kind of functions, $f(x)$, can we write a Fourier series?

2. What are the convergence properties of Fourier series? (Equivalently, under what general conditions will a Fourier series converge?)

3. How are the coefficients, a_n and b_n, related to the function $f(x)$?

Our first insight into the general class of functions for which one can write Fourier series comes from the observation that all of the terms in (2.1.1) above are *periodic* of period $2L$, e.g.

$$\cos\left(\frac{n\pi(x + 2L)}{L}\right) = \cos\left(\frac{n\pi x}{L} + 2n\pi\right) = \cos\left(\frac{n\pi x}{L}\right) , \quad \text{etc} .$$

Hence, clearly, the Fourier series itself must also be periodic of period $2L$, i.e.

$$f(x + 2L) = f(x) . \qquad (2.1.2)$$

A function whose general behavior described by (2.1.2) is shown in Figure 10). The key, of course, to periodicity is the repetitive nature of such functions. Furthermore, this repetition should be clearly evident from inspection of the graph. Lastly, observe that once we know that a given function $f(x)$ is periodic, the value of L is very easily related to the graph, since the interval between repetitions is $2L$.
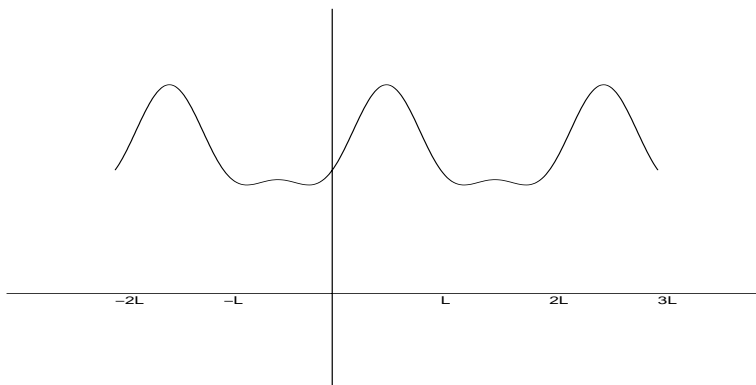
Figure 10: A General Periodic Function

These last few comments also dovetail quite nicely with our earlier ones that infinite series may offer a way of decomposing complicated functions into simple ones. Along these lines, we shall show more fully as this course develops exactly how the Fourier series provides a mechanism by which complicated periodic functions can be broken down into sums of sines and cosines - the simplest of periodic functions.

2.2   Derivation of the Fourier Series Coefficients

We next turn to the question of how the coefficients in the Fourier series are related to the function $f(x)$. There is, unfortunately, no single technique that will produce formulas for the coefficients in any infinite series. You may recall, for example, that for Taylor series,

$$g(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \ ,$$

repeated differentiation, followed by evaluation at $x = x_0$, leads to the conclusion that the coefficients are defined by

$$a_n = \frac{1}{n!} \frac{d^n g}{dx^n}(x_0)$$

(which immediately implies that Taylor series can only be written for functions with an infinite number of derivatives). We now show that a totally different approach determines the Fourier series coefficients (and simultaneously provides important insights into the type of functions for which Fourier series actually can be written.) Our starting point is with a number of integral formulas - the so-called *Orthogonality Integrals*. As we shall see more clearly as our study develops, such integrals are pivotal to Fourier analysis. The precise formulas for these orthogonality integrals are

$$\int_{-L}^{L} \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) dx = \begin{cases} 0 & , m \neq n \\ L & , m = n \end{cases}$$

$$\int_{-L}^{L} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) dx = \begin{cases} 0 & , m \neq n \\ L & , m = n \end{cases} \qquad (2.2.3)$$

$$\int_{-L}^{L} \sin\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) dx = 0, \text{all} \ \ m, n \qquad ,$$

where m and n are any *positive integers*. There is, at least on the surface, nothing particularly extraordinary about these integrals. They can easily be verified by standard calculus techniques. One simply uses standard trigonometric identities, such as

$$\cos(a)\cos(b) = \frac{\cos(a+b) + \cos(a-b)}{2}$$

to reduce the integrand to a sum of elementary functions, then finds the antiderivative, and finally evaluates that antiderivative at the end points, using the fact that

$$\sin(k\pi) = 0 \ \ , \quad \text{and} \quad \cos(k\pi) = \cos(-k\pi)$$

for any integer k.

Given these orthogonality integrals, however, we can proceed to reduce (2.1.1) as follows. Let m denote some fixed, but arbitrary positive integer. (This means that all we may assume about m is that it is positive and an integer. It need not necessarily be even; or odd; large or small; etc. We simply cannot assume that any steps are valid that require more information than the fact that m is a positive integer.) We shall then multiply both sides of (2.1.1) by $\cos\left(\frac{m\pi x}{L}\right)$, then, using the property that the integral of a sum is the sum of the respective integrals, integrate the series formally, term by term, from $-L$ to L. Performing all of these steps leads to the equation

$$\begin{aligned}
\int_{-L}^{L} f(x)\cos\left(\frac{m\pi x}{L}\right) dx = {} & \frac{a_0}{2}\int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right) dx \\
& + \sum_{n=1}^{\infty}\left\{a_n \int_{-L}^{L} \cos\left(\frac{n\pi x}{L}\right)\cos\left(\frac{m\pi x}{L}\right) dx \right. \qquad (2.2.4) \\
& \left. + \ \ b_n \int_{-L}^{L} \sin\left(\frac{n\pi x}{L}\right)\cos\left(\frac{m\pi x}{L}\right) dx \right\}.
\end{aligned}$$

(Mathematically, we say this integration is formal because there is really some analytic uncertainty about the validity of interchanging the operations of integration and *infinite* summation. In fact, in certain other infinite series (none of them Fourier series), interchanging these operations can in fact be shown to produce totally incorrect results. The reader should rest assured however that it can be rigorously proven that, for Fourier series, this interchange can be shown to be valid, although any proof of this claim is far beyond the scope of our discussion here).

Returning to (2.2.4), we can now use elementary integration techniques to show that, since $m \neq 0$, the first integral on the right hand side is identically zero. But now look at the rest of this expression closely. Remember that the series notation implies that the terms inside the brackets must be evaluated for every positive integer n, then these values summed. But wait! If m is a positive integer then, by the orthogonality integrals (2.2.3), every term on the right side will be identically zero, **except one**. The single non-zero term will be the cosine integral that arises the one time inside the summation when the value of n equals m. Somewhat more graphically, this means the entire right-hand of the previous equation obeys

$$\frac{a_0}{2} \underbrace{\int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right) dx}_{=0} + \sum_{n=1}^{\infty} \left\{ a_n \underbrace{\int_{-L}^{L} \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) dx}_{=0 \ , \ m \neq n} \right.$$

$$\left. + \quad b_n \underbrace{\int_{-L}^{L} \sin\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) dx}_{=0 \ , \ \text{all} \ m,n} \right\} \quad .$$

and therefore (2.2.4) simplifies to:

$$\int_{-L}^{L} f(x) \cos\left(\frac{m\pi x}{L}\right) dx = a_m \int_{-L}^{L} \cos^2\left(\frac{m\pi x}{L}\right) dx$$

or, solving for a_m and using the value of the orthogonality integral for cosines with $m = n$,

$$a_m = \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{m\pi x}{L}\right) dx \ , \quad m > 0 \ .$$

But now, note that since m was designated as fixed, but arbitrary, this last equation is in fact simply a *formula* for a_m. Therefore we can equally well now replace m by n on both sides of this expression, and rewrite it as:

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx \ . \tag{2.2.5}$$

We can derive a similar formula for the b_n,

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx \ , \tag{2.2.6}$$

by multiplying both sides of (2.1.1) by a sine and integrating. Finally, a_0 can be determined by observing that the **average** values of both $\cos\left(\frac{n\pi x}{L}\right)$ and $\sin\left(\frac{n\pi x}{L}\right)$ are zero over any whole number of cycles or periods. Thus their average value over the interval from $-L$ to L is zero. Hence the average value of every term under the summation sign is zero, and thus the average value of the entire right-hand side of (2.1.1) is precisely

$$\frac{a_0}{2} \quad .$$

Since the average value of the left-hand side can be computed as

$$\frac{1}{2L}\int_{-L}^{L} f(x)dx \quad,$$

then clearly, after equating these two expressions and canceling the factors of two, we have

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(x)dx \quad. \tag{2.2.7}$$

(Equivalently, if one likes the brute force approach, they may simply integrate both sides of (2.1.1) from $-L$ to L, show the integrals under the summation sign are identically zero, and arrive at exactly the same result. While there would obviously be nothing wrong with that approach, we generally prefer to use physical insight wherever possible.) You might note now that (2.2.7) is precisely what we would obtain by letting $n = 0$ in (2.2.5), even though (2.2.5) was originally derived assuming $n > 0$. This is exactly why we adopted the (seemingly) rather awkward form for the leading coefficient in (2.1.1), since we now see that by writing the constant term in the series as $a_0/2$, we are able to use (2.2.5) to represent all the a_n's in the series, not just those for which n is positive. (Unfortunately, however, we will still generally have to *compute* a_0 separately, in order to avoid an antiderivative with a zero denominator. However our representation still leaves one less formula to remember.)

Example: Consider the function

$$f(x) \;=\; \begin{cases} 0 & ,-2 < x < 0 \\ x & ,0 \le x < 2 \end{cases}$$

$$f(x+4) = f(x)$$

This function is graphed in Figure 11.
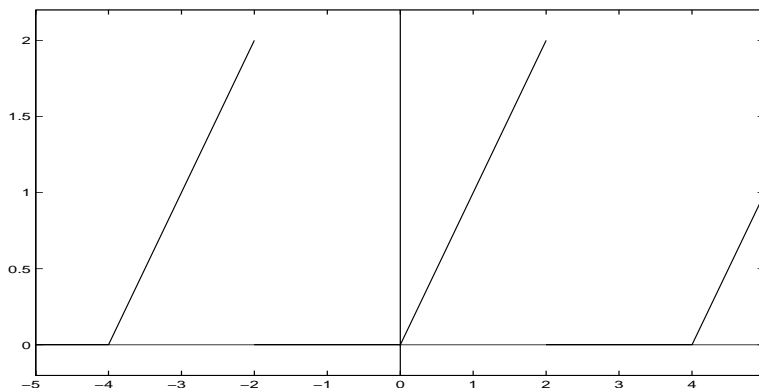
Figure 11: A Piecewise Continuous Function in the Example

Straightforward computation using (2.2.5) and (2.2.6) yields

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(x)dx = \frac{1}{2}\int_0^2 xdx = 1 \Rightarrow \frac{a_0}{2} = \frac{1}{2} \quad,$$

29

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right) dx = \frac{1}{2}\int_{0}^{2} x\cos\left(\frac{n\pi x}{2}\right) dx$$

$$= \frac{1}{2}\left[\frac{2x}{n\pi}\sin\left(\frac{n\pi x}{2}\right) + \left(\frac{2}{n\pi}\right)^2 \cos\left(\frac{n\pi x}{2}\right)\right]\bigg|_{0}^{2}$$

$$= \frac{2}{n^2\pi^2}(\cos(n\pi) - 1) \ ,$$

and

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\left(\frac{n\pi x}{L}\right) dx = \frac{1}{2}\int_{0}^{2} x\sin\left(\frac{n\pi x}{2}\right) dx$$

$$= \frac{1}{2}\left[-\frac{2x}{n\pi}\cos\left(\frac{n\pi x}{2}\right) + \left(\frac{2}{n\pi}\right)^2 \sin\left(\frac{n\pi x}{2}\right)\right]\bigg|_{0}^{2}$$

$$= -\frac{2}{n\pi}\cos(n\pi)) \ .$$

Thus, combining these, we have

$$f(x) = \frac{1}{2} + \sum_{n=1}^{\infty}\left\{\frac{2}{n^2\pi^2}(\cos(n\pi) - 1)\cos\left(\frac{n\pi x}{2}\right) - \frac{2}{n\pi}\cos(n\pi)\sin\left(\frac{n\pi x}{2}\right)\right\}$$

$$= \frac{1}{2} - \frac{4}{\pi^2}\cos\left(\frac{\pi x}{2}\right) + \frac{2}{\pi}\sin\left(\frac{\pi x}{2}\right) - \frac{1}{\pi}\sin(\pi x)$$

$$- \frac{4}{9\pi^2}\cos\left(\frac{3\pi x}{2}\right) + \frac{2}{3\pi}\sin\left(\frac{3\pi x}{2}\right) + \cdots$$

You should realize at this point that the above result may in fact be quite meaningless, for we have not yet proven that this series even converges! In fact, this is an appropriate place to review exactly what the above infinite series notation represents, and what we mean by convergence.

By definition, when we write the Fourier series representation (2.1.1), we mean that, if we were to consider the sequence of partial sums given by:

$$S_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N}\left\{a_n\cos\left(\frac{n\pi x}{L}\right) + b_n\sin\left(\frac{n\pi x}{L}\right)\right\} \ , \qquad (2.2.8)$$

then,

$$\lim_{N\to\infty} S_N(x) = f(x) \ .$$

But, as the notation clearly implies, these partial sums are functions, i.e. they have graphs. Thus, if a Fourier series converges, there should be some sequence of graphs, given by precisely the graphs of these partial sums, that look "more and more like $f(x)$" the "larger" N

is. To demonstrate this in the example above, we have plotted, in Figure 12, several partial sums of the series we just derived. These graphs clearly seem, for larger values of N, to look progressively more like the original $f(x)$, and therefore, in this example, the series found does appear to converge to $f(x)$.
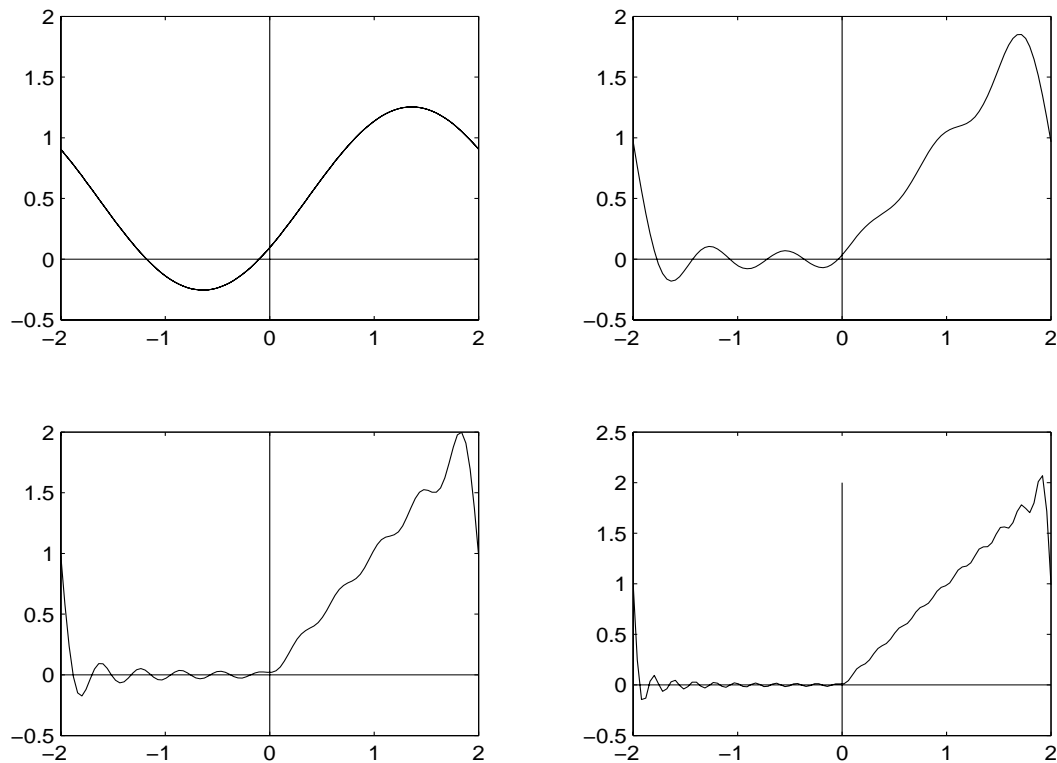
Figure 12: Convergence of the Partial Sums of a Fourier Series

Always having to generate such a set of graphs in order to confirm convergence, however, could clearly become quite cumbersome. Therefore, as part of our study, we will consider the general convergence properties of the Fourier series, and classify those functions for which Fourier series can be *guaranteed* to converge before such graphs are even plotted.

We start by developing a bit more complete characterization of precisely the kind of functions for which it "makes sense" to try to write a Fourier series, i.e. for which we should even try to compute a_n and b_n as given by (2.2.5) and (2.2.6). Observe that computing a_n and b_n requires evaluating definite integrals. Therefore, since $\cos\left(\frac{n\pi x}{L}\right)$ and $\sin\left(\frac{n\pi x}{L}\right)$ are continuous, it seems almost self-evident, from basic calculus considerations, that $f(x)$ cannot be "too far away" from being a continuous function. This initial impression is basically correct, although the precise condition is slightly more involved and its proof far beyond the scope of our discussion here. Therefore, we simply state the most commonly used result on which functions will have convergent Fourier series:

Theorem: The function $f(x)$ will have a convergent Fourier series, with coefficients a_n

and b_n given by:

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx , \quad n \geq 0$$

and

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx , \quad n > 0 \quad.$$

provided $f(x)$ is periodic of period $2L$, and both $f(x)$ and $f'(x)$ are at least piecewise continuous on $-L < x < L$.

A piecewise continuous function is, of course, one composed of a finite number of continuous segments ("pieces") on any finite interval. To be more precise,
<u>Definition</u>. A function $f(x)$ is piecewise continuous in $[a, b]$ if there exists a finite number of points $a = x_1 < x_2 < \ldots < x_n = b$, such that f is continuous in each open interval (x_j, x_{j+1}) and the one sided limits $f(x_{j+})$ and $f(x_{j+1-})$ exist for all $j \leq n - 1$.

The most common examples of such functions are those that are continuous except for "jump" discontinuities. The function we used in our last example is one of these, and satisfies the conditions of this theorem, since it is continuous everywhere except for jumps at $x = \pm 2$, and differentiable everywhere except at the jumps **and** at the sharp point at $x = 0$. (Functions for which $f'(x)$ is piecewise continuous are often called *piecewise smooth*.)

Our second consideration is the question of precisely how a Fourier series converges at a "jump" discontinuity. The proof of the answer, as with the proof of the above existence theorem, is well beyond the scope of our discussion. But it can be shown that, at a "jump," the Fourier series takes essentially the "path of least resistance" by converging to the average value, i.e. at a jump

$$S_N(x) \rightarrow \frac{f(x+) + f(x-)}{2}.$$

where $f(x+)$ and $f(x-)$ refer to the right- and left-hand limits, respectively (e.g. f(x+) means $\lim_{\epsilon \to 0} f(x + \epsilon)$, where ϵ is positive.)

Thus far then, we have introduced the Fourier Series, shown how the Fourier coefficients, a_n and b_n can be computed, and for one example, displayed graphically the convergence of the resulting series.

PROBLEMS

1. Derive the formula for the Fourier sine coefficients, b_n (eqn. (2.2.6)), using a method similar to that used to derive eqn. (2.2.5).

2. For each of the following functions, find the Fourier coefficients, the Fourier series, and sketch the partial sums $S_2(x)$, $S_5(x)$, and $S_{10}(x)$:

a. $f(x) = \begin{cases} 0 & , \quad -1 < x < 0 \\ 1 & , \quad 0 \leq x < 1 \end{cases}$

        $f(x+2) = f(x)$

b. $f(x) = \begin{cases} 3+x & , \quad -3 < x < 0 \\ 3-x & , \quad 0 \leq x < 3 \end{cases}$ $f(x+6) = f(x)$

c. $f(x) = \begin{cases} 0 & , \quad -2 < x < 0 \\ x & , \quad 0 \leq x < 1 \\ 2-x & , \quad 1 \leq x < 2 \end{cases}$

        $f(x+4) = f(x)$

d. $f(x) = 1 - \cos(\pi x), \quad -1 \leq x \leq 1$

3. a. Show that the alternative Fourier Series representation

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

leads to the formulas

$$a_0 = \frac{1}{2L} \int_{-L}^{L} f(x) dx \quad ,$$

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx \quad , \quad n > 0$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx \quad , \quad n > 0$$

where the formula for a_0 is no longer the $n = 0$ special case of the formula for a_n.

b. Show that the alternative Fourier Series representation

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left\{ a_n \cos\left(\frac{2n\pi t}{T_0}\right) + b_n \sin\left(\frac{2n\pi t}{T_0}\right) \right\}$$

(note the independent variable here is t), where L has been replaced by $T_0/2$, leads to

$$a_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(x) \cos\left(\frac{2n\pi t}{T_0}\right) dt \ , \quad n \geq 0$$

$$b_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(x) \sin\left(\frac{2n\pi t}{T_0}\right) dt \ , \quad n > 0 \ .$$

Note that here the formula for a_0 is again the $n = 0$ special case of the formula for a_n.

4. In each of the following, find each point in $-L < x < L$ where $f(x)$ has a discontinuity. Find the left and right-hand limits of $f(x)$ and $f'(x)$ at each point of discontinuity and at the end points of the interval. Without computing the Fourier coefficients, indicate to what values the series should converge at these points.

a. $f(x) = \begin{cases} x^2 & , & 1 \leq x \leq 3 \\ 0 & , & -2 \leq x < 1 \\ 2x & , & -3 \leq x < -2 \end{cases}$

b. $f(x) = \begin{cases} 3 & , & \pi/2 \leq x \leq \pi \\ 2x - 2 & , & -\pi \leq x < \pi/2 \end{cases}$

c. $f(x) = \begin{cases} x^2 & , & -2 \leq x < 0 \\ 0 & , & 0 \leq x < 1 \\ 4(x-1) & , & 1 \leq x \leq 2 \end{cases}$

2.3 Odd and Even Functions

As we saw in the previous section, we can use the Fourier series to represent any appropriate periodic function as a combination of the basic trigonometric functions of sine and cosine. Furthermore, once we have verified that a particular function satisfies the necessary conditions to have a Fourier series representation, determining the coefficients (i.e. a_n and b_n) using (2.2.5) and (2.2.6) is a straightforward, although perhaps quite laborious exercise in integration.

In this section, we show that we can sometimes significantly reduce the amount of computation necessary to construct a particular Fourier series representation by exploiting symmetries in the function $f(x)$. We start by observing that inspection of the graphs of sine and cosine (Figure 13) shows that both of these functions possess certain natural symmetries. Specifically, cosine is a perfect reflection (symmetric) with respect to the y-axis, and sine is an inverted (anti-symmetric) one. Thus, it is not unreasonable to expect a strong connection between the symmetries (if any) in some periodic function, $f(x)$, and the presence (or absence) of sine or cosine terms in its Fourier series. We now show that this is in fact the case.
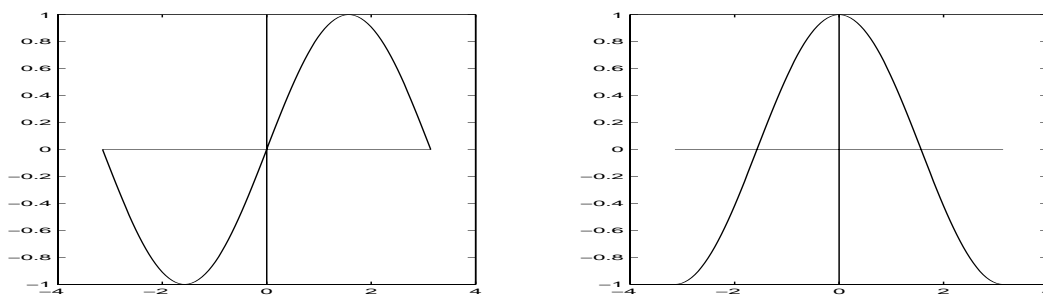
Figure 13: Symmetries in Sine and Cosine

The symmetry and anti-symmetry described above are identical to that of the various powers of x , i.e. x^n, depending on whether the exponent (n) is *even* or *odd*. Therefore, symmetry and anti-symmetry, when they exist, are more commonly referred to, respectively, as even or odd. To be precise,

> DEFINITION:
> A function $f(x)$ is **even** if $f(x) = f(-x)$ for all x.
> A function $f(x)$ is **odd** if $f(x) = -f(-x)$ for all x.

We would hasten to caution that there are functions which have no such symmetries, i.e. which are neither even nor odd. (The first example in the last section is one such function.) Furthermore, no non-trivial function can be both even and odd. But when one is dealing with either an even or odd function, there are several rules that can simplify calculations. Among the most useful of these rules are the following:

1. The product of two even functions is even,

2. The product of two odd functions is even,

3. The product of an even and odd function is odd,

4. The derivative of an even function is odd,

5. The derivative of an odd function is even,

6. If g(x) is even, then

$$\int_{-L}^{L} g(x)dx = 2 \int_{0}^{L} g(x)dx$$

7. If g(x) is odd, then

$$\int_{-L}^{L} g(x)dx = 0 \qquad .$$

All of the above statements are straightforward to prove, and generally follow immediately from the definitions of even and odd. The statements about the integrals reflect the facts, evident from Figure 14, that an even function sweeps out equal areas on both sides of the y axis, and an odd function does almost the same thing, except that with an odd function these areas differ in algebraic sign, and so cancel.
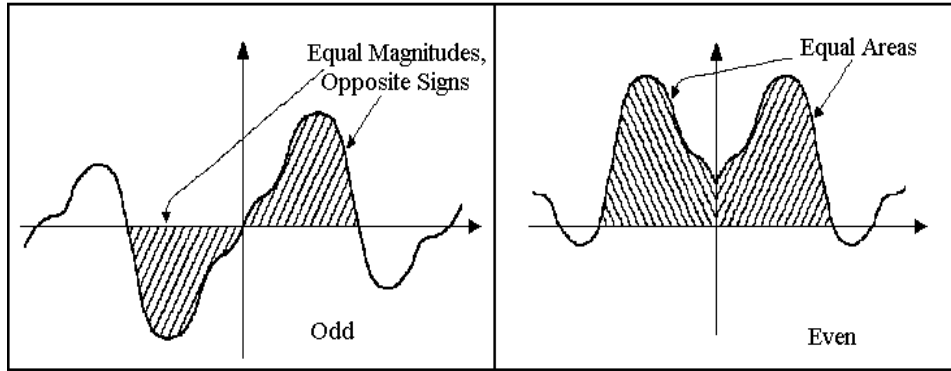
Figure 14: Integrals of Even and Odd Functions

Now, applying these properties to Fourier series leads to the observation:

If $f(x)$ is even, then $\begin{cases} f(x)\cos\left(\frac{n\pi x}{L}\right) & \text{is even, and} \\ f(x)\sin\left(\frac{n\pi x}{L}\right) & \text{is odd} \end{cases}$

However, by the integral properties above, it follows immediately from this last statement that if $f(x)$ is even and has a Fourier series, then by (2.2.5) and (2.2.6) in the previous section,

$$a_n = \frac{2}{L} \int_{0}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right) dx$$

and
$$b_n = 0 \quad .$$

Thus for an even function
$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right)$$

Similarly, we can show that if $f(x)$ is odd, then
$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{L}\right)$$

where
$$b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx \quad .$$

We must emphasize that these results are not really different from, and do not produce any information not already contained in the basic Fourier series coefficient formulas ((2.2.5) and (2.2.6)). They simply are a means to reduce the computational burden in selected cases by indicating, ahead of time, which coefficients should be identically zero. One would obtain identical values the "long way," i.e. using (2.2.5) and (2.2.6) directly, but only after much more effort and an increased chance of algebraic error.

Thus, for example, with the odd function shown below
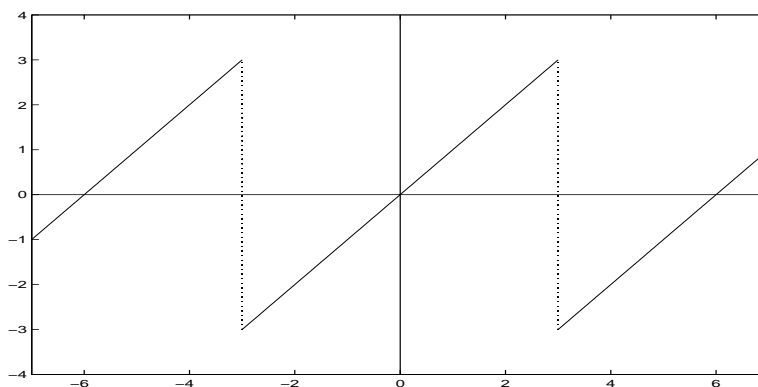
Figure 15: $f(x) = x , \quad -3 < x < 3$

we can immediately claim that
$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{3}\right) \quad ,$$

where
$$b_n = \frac{2}{3} \int_0^3 x \sin\left(\frac{n\pi x}{3}\right) dx = -\frac{6}{n\pi} \cos(n\pi) \quad .$$

and not have to bother computing the a_n, since we know they must be identically zero.

The above results on odd and even functions can also be used to simplify computation in a number of further cases where the addition or subtraction of a constant will convert

an otherwise unsymmetric function to an odd function. (Note addition or subtraction of a constant will not alter the evenness of a function.) For example, the function

$$f(x) \;\; = 4 + x \; , \quad -3 < x < 3,$$
$$f(x+6) = f(x) \;\; ,$$

is neither even nor odd, but $[f(x) - 4]$ is clearly odd. Hence we could write,

$$f(x) = 4 + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{3}\right) \;\; ,$$

where the b_n are as computed above.

PROBLEMS

1. Find the Fourier series for the following functions

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

$f(x + 4) = f(x)$

b. $f(x) = |x| \quad , \qquad f(x + 2) = f(x)$

c. $f(x) = |\sin(x)|$

d. $f(x) = \begin{cases} -2 - x & , & -2 \le x \le -1 \\ x & , & -1 \le x \le 1 \\ 2 - x & , & 1 \le x \le 2 \end{cases}$

$f(x + 4) = f(x)$

2.4 Convergence Properties of Fourier Series

In the previous sections, we have seen that a Fourier series can be written, at least formally, for the periodic (of period $2L$) function $f(x)$, as

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

where

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx , \quad n = 0, 1, 2, \ldots$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx , \quad n = 1, 2, \ldots$$

Obviously, given a particular $f(x)$, there would be little point in bothering to compute these integrals without some assurance the resulting series would converge, and so we cited earlier some theorems involving an extremely straightforward test (i.e. if $f(x)$ is at least piecewise smooth) that, if satisfied, would guarantee that the series constructed using the above coefficients will in fact converge. However, as we shall see later, another, equally important consideration frequently arises in the study of Fourier series. There are often occasions where one must deal with a Fourier series, by itself, with no explicit identification of the function $f(x)$ which it represents. In such cases, one would like to be able to infer, from examining the series itself (i.e. the coefficients), not only whether it converges, but also, if it does, what general characteristics does the function it represents possess? Therefore, motivated by both of these concerns, we shall now investigate in greater detail the exact mechanisms which drive the convergence of Fourier series. During this study, our primary goal will be to more fully understand the interrelationships between the properties of the function $f(x)$ and the general behavior of the coefficients, a_n and b_n.

We start by recalling that the primary vehicle for describing the quantitative convergence properties of **any** infinite series is the error remaining after approximating the series by a partial sum. In the case of the Fourier series, this error is given by the formula:

$$E_N(x) = f(x) - S_N(x) = \sum_{n=N+1}^{\infty} \left\{ a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right\} \qquad (2.4.9)$$

where $S_N(x)$ was given by equation (2.2.8). As we have already discussed at some length during our consideration of general infinite series of functions, and as this formula also emphasizes, both this error and the associated rate of convergence of the series will, in general, depend on the value of x at which we are evaluating the series. Furthermore, the Fourier series will converge pointwise to $f(x)$ if and only if

$$\lim_{N \to \infty} [E_N(x)] = 0 \qquad \text{at } \textbf{all} \text{ values of } x \quad .$$

Lastly, a Fourier series will converge uniformly, and hence the limit function, $f(x)$, be continuous, if and only if

$$|E_N(x)| \leq \epsilon_N \qquad \text{and} \qquad \lim_{N \to \infty} [\epsilon_N] = 0 \quad . \qquad (2.4.10)$$

But now look carefully at (2.4.9). Specifically, note that

$$\left|\cos\left(\frac{n\pi x}{L}\right)\right| \leq 1 \qquad \text{and} \qquad \left|\sin\left(\frac{n\pi x}{L}\right)\right| \leq 1$$

for **all** x. Therefore, if we take absolute values on both sides of (2.4.9), and use the fact that the absolute value of a sum is less than or equal to the sum of the absolute values, we have:

$$|E_N(x)| \leq \sum_{n=N+1}^{\infty} \{|a_n| + |b_n|\} \quad . \tag{2.4.11}$$

There is, of course, no guarantee that the series on the right here will converge, but, if it does, then we have clearly satisfied the condition given by (2.4.10) and hence would have a uniformly convergent series. We formally state these observations in:

THEOREM: The Fourier series for $f(x)$ converges uniformly, and therefore $f(x)$ is continuous, if

$$\sum_{n=1}^{\infty} \{|a_n| + |b_n|\} \qquad \text{converges} \quad .$$

Observe that a major part of the value of this theorem is that it involves only a series of **constants**, whose convergence can therefore be easily tested by the usual tests, e.g. the integral test or the $\mathbf{O}\left(1/n^p\right)$ test. In fact, based on our earlier discussion on series of constants, we can immediately conclude the more practically useful:

COROLLARY: The Fourier series for $f(x)$ will converge uniformly, and therefore $f(x)$ will be continuous, if both a_n and b_n are

$$\mathbf{O}\left(\frac{1}{n^p}\right) \qquad \text{for} \qquad p > 1 \quad .$$

(Before proceeding, we need to emphasize one point about the use of the term "continuous" in discussions of Fourier series. Remember that the Fourier series implicitly considers the function $f(x)$ to be periodic. Thus, when $f(x)$ is given explicitly in some interval, and just assumed to be continued periodically, as in:

$$f(x) = \begin{cases} 0 & , & -2 < x < 0 \\ x & , & 0 \leq x < 2 \end{cases}$$
$$f(x+4) = f(x) \tag{2.4.12}$$

it is continuous in the context used in the above only if $f(x)$ is continuous not only on the given interval, but also at the end points where it begins to periodically repeat itself. Thus the function given above is not continuous in the context of Fourier series theorems, even though it is continuous for $-2 < x < 2$, because it has a jump discontinuity at $x = \pm 2$. Several texts try to avoid this confusion, when discussing continuity, by speaking about the *periodic continuation* of $f(x)$, rather than just referring to $f(x)$.)

The above discussion and theorem fairly well specify the conditions on the coefficients under which a Fourier series will converge to a continuous function. However, we can't really stop yet, since, as we saw when we computed the Fourier series for (2.4.12) in the first section, and as the theorem cited there indicates, formally constructed Fourier series can also converge (at least in some cases) to discontinuous functions. (A discontinuous function, however, clearly can **not** have a *uniformly convergent* Fourier series. Moreover, based on the preceding discussion, it appears virtually certain that when a Fourier series converges to a discontinuous function, then either $\sum_{n=1}^{\infty} |a_n|$ or $\sum_{n=1}^{\infty} |b_n|$ (or both) must diverge - a situation seemingly destined to be the case if either the a_n or the b_n "look like" $1/n$. This seems to be exactly what occurred in the Fourier series for (2.4.12).) But one or two examples of discontinuous functions which have Fourier series do not a general theory make. Thus, we are still left with the unanswered question of whether **all** Fourier series in which at least one of the coefficients is only $\mathbf{O}\,(1/n)$ converge? And, if so, what drives this convergence, since the series on the right in (2.4.11) can not converge in such a case?

In order to understand the answer to these last questions, we first need to introduce the concept of the mean square measure. Specifically, for a given function $f(x)$, the integral

$$\frac{1}{2L} \int_{-L}^{L} [f(x)]^2 dx \tag{2.4.13}$$

is called the mean square value of $f(x)$ on the interval $(-L, L)$. Physically you should recognize that it can be interpreted as essentially a measure of the average energy represented by the function in this interval. If however, we now apply this integral to measure what is in effect the average energy contained in the error in a Fourier series approximation, and expand the squared series, we have

$$\frac{1}{2L} \int_{-L}^{L} [E_N(x)]^2\, dx \;=\; \frac{1}{2L} \int_{-L}^{L} \left[\sum_{n=N+1}^{\infty} \left\{ a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right\} \right]^2 dx$$

$$= \frac{1}{2L} \int_{-L}^{L} \left[\sum_{n=N+1}^{\infty} \left\{ a_n^2 \cos^2\left(\frac{n\pi x}{L}\right) + b_n^2 \sin^2\left(\frac{n\pi x}{L}\right) \right\} \right.$$

$$+ \sum_{\substack{m,n=N+1 \\ m \neq n}}^{\infty} \left\{ a_m a_n \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) \right.$$

$$\left. + b_m b_n \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \right\}$$

$$\left. + \sum_{m,n=N+1}^{\infty} a_m b_n \cos\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) \right]\, dx$$

But now look what happens if we interchange the summations and integration (assuming again that such interchanges are mathematically valid). All of the integrals involved become

42

orthogonality integrals (3), and more precisely, all of the terms in the double series are orthogonality integrals which integrate to zero! Thus, this last equation simply reduces to:

$$\frac{1}{2L} \int_{-L}^{L} [E_N(x)]^2 \, dx = \frac{1}{2} \sum_{n=N+1}^{\infty} \left\{ a_n^2 + b_n^2 \right\} \tag{2.4.14}$$

While this equation may not look especially significant, you should look at it carefully, for it actually embodies one of the basic properties of Fourier series! Because of the importance of this relation to the understanding of these series, we shall investigate it in some detail, starting with the observation that (2.4.14) implies that

$$\lim_{N \to \infty} \left\{ \frac{1}{2L} \int_{-L}^{L} [E_N(x)]^2 \, dx \right\} = 0 \qquad \text{if} \qquad \sum_{n=1}^{\infty} \left\{ a_n^2 + b_n^2 \right\} \qquad \text{converges.}$$

Note the key difference between this last situation (i.e. convergence of the sequence of squares) and our earlier theorem on uniform convergence! Specifically, the series of squares can converge even though

$$\sum_{n=1}^{\infty} \left\{ |a_n| + |b_n| \right\}$$

diverges! (This, again, will almost certainly arise when either a_n and b_n, or both, are only $\mathbf{O}\left(1/n\right)$ - precisely the case we frequently encounter.) Clearly, now,

$$\lim_{N \to \infty} \left\{ \frac{1}{2L} \int_{-L}^{L} [E_N(x)]^2 \, dx \right\} = 0 \tag{2.4.15}$$

represents some type of convergence - but what kind? Is this simply one of the types we have already discussed, in a different form, or is it fundamentally new? By our last comments, this "new" convergence can apparently occur even when the series does not converge uniformly, therefore it must be different from uniform convergence. Even more interestingly, this "new" convergence can occur even when the series does not converge pointwise to $f(x)$, i.e. even when $E_N(x)$ does not approach zero for **all** values of x. The reason for this perhaps surprising statement is that it is possible for

$$\int_{-L}^{L} [g(x)]^2 \, dx = 0$$

even though $g(x) \not\equiv 0$, *provided $g(x)$ is discontinuous, bounded, and nonzero at only a finite number of points.* (This is simply a generalization of the notion that the area under a point is zero.) In other words, (2.4.15) may occur even though $S_N(x)$ does not converge to $f(x)$ everywhere, but only *almost everywhere*, that is when

$$\lim_{N \to \infty} S_N(x) = f(x)$$

except at a finite number of points. These points, of course, must have been points of discontinuity in $f(x)$ to begin with.

Since we have now shown this "new" convergence as defined by (2.4.15) does not coincide with either uniform or pointwise convergence, it needs its own distinctive title. Perhaps not very imaginatively, it is normally called convergence in the mean-square, or simply mean-square convergence. To be precise:

DEFINITION: The sequence of functions $S_N(x)$ will be said to converge to the function $f(x)$ in the mean-square sense, or almost everywhere, if

$$\lim_{N \to \infty} \left\{ \frac{1}{2L} \int_{-L}^{L} [S_N(x) - f(x)]^2 \, dx \right\} = 0 \qquad . \tag{2.4.16}$$

From our point of view, mean-square convergence is important because it not only describes a convergence that appears to be different from both uniform and pointwise convergence, but also a convergence which apparently governs the convergence of Fourier series for discontinuous functions. We can incorporate most of the insights of this recent discussion into the following:

THEOREM: The Fourier series for $f(x)$ converges in the mean-square sense if

$$\sum_{n=1}^{\infty} \left\{ a_n^2 + b_n^2 \right\} \qquad \text{converges.}$$

Moreover, as with the theorem for uniform convergence, we can also immediately conclude a more practically applicable

COROLLARY: The Fourier series for $f(x)$ will converge in the mean-square sense if both a_n and b_n are

$$\mathbf{O}\left(\frac{1}{n}\right) \qquad .$$

We must again emphasize most strongly that mean-square convergence does not necessarily imply that the $S_N(x)$ converge to $f(x)$ at **every** point in $(-L, L)$, although this certainly could be the case. mean-square convergence describes essentially a convergence in terms of energy rather than a pointwise convergence. Therefore, when we say a series converges in the mean-square sense to $f(x)$, we are in effect saying that $f(x)$ and the limit of the series differ, at most, in a function which contains no energy and is therefore physically uninteresting.

To recap our development to this point, we have thus far established a hierarchy of convergences, each increasingly more stringent in its conditions. Specifically:

$$\begin{matrix} Uniform \\ Convergence \end{matrix} \Rightarrow \begin{matrix} Pointwise \\ Convergence \end{matrix} \Rightarrow \begin{matrix} Mean-Square \\ Convergence \end{matrix}$$

where the direction of the arrows emphasizes that the more stringent convergences automatically imply the less stringent ones, but not vice versa. The importance of the difference between these in the context of Fourier series is that:

a. If a Fourier series converges uniformly, then it converges to a continuous function, and both a_n and b_n are $\mathbf{O}\left(1/n^p\right)$ for $p > 1$.

b. If a Fourier series converges, but not uniformly, it will converge to a discontinuous function. Moreover, in this case, we expect at least one of the coefficients, a_n or b_n (or both) will not decay faster than $\mathbf{O}\left(1/n\right)$.

c. The difference between pointwise and mean-square convergence is that a Fourier series that converges in mean square need not converge to the function $f(x)$ used to compute to coefficients, a_n and b_n, but only to some "close" (in the sense of differing only by a function whose energy is zero) function.

We shall close this portion of our discussion with the final observation that, if $f(x)$ has a Fourier series, then we can again (at least formally) write

$$f'(x) = \sum_{n=1}^{\infty} \left\{ -\frac{n\pi}{L} a_n \sin\left(\frac{n\pi x}{L}\right) + \frac{n\pi}{L} b_n \cos\left(\frac{n\pi x}{L}\right) \right\} \qquad .$$

But now, by essentially repeating our earlier discussions, we can easily conclude that this series will represent a continuous function if both a_n and b_n are $\mathbf{O}\left(1/n^3\right)$. A similar argument can be repeated for successively higher derivatives. This leads us to conclude that, simply by looking at the order of the coefficients in a Fourier series, one can determine the highest continuous derivative of the resulting function. This final conclusion of our analysis is incorporated in the following

COROLLARY: The Fourier series for $f(x)$ will converge in at least the mean-square sense if both a_n and b_n are

$$\mathbf{O}\left(\frac{1}{n^p}\right) \qquad \text{for} \qquad p \geq 1 \qquad ,$$

and furthermore, in this case, $f^{(p-2)}(x)$ will be continuous.

This last result completes the first part of our study of the convergence of Fourier series, and describes essentially how much information we can infer about the function represented by a Fourier series by simply observing the behavior of the coefficients. We now shall look at the other side of the question - given the function $f(x)$, what can we infer about the general behavior of the coefficients in its Fourier series, before we even compute them? More specifically, given the continuity properties of the function, can we predict the order of its Fourier series coefficients?

The answer to this question is that a very predictable relationship exists. This relationship is expressed by the following theorem, which is presented without proof:

THEOREM: If k is an integer such that $f^{(k)}(x)$ is the *highest order* derivative of $f(x)$ that is continuous, and $f^{(k+1)}(x)$ is piecewise smooth but **not** continuous, then, both a_n and b_n in the Fourier series for $f(x)$ are $\mathbf{O}\left(1/n^{k+2}\right)$ and at least one (or perhaps both) will **not** be $\mathbf{O}\left(1/n^{k+3}\right)$.

Thus, for example, if $f(x)$ is continuous, and $f'(x)$ is discontinuous but piecewise smooth, then $k = 0$ for this theorem and we should expect both a_n and b_n to be at least $\mathbf{O}\left(1/n^2\right)$ (and hence the series to converge uniformly), but at least one of them will not be $\mathbf{O}\left(1/n^3\right)$. For instance, the function (2.4.12), whose series was computed in the first section, was discontinuous, but piecewise smooth. Therefore, although perhaps not quite so obviously, the appropriate value for k to use in this problem with the theorem should be $k = -1$ (since it is only the *integral* of $f(x)$ which is continuous). With this value the theorem predicts both coefficients will be $\mathbf{O}\left(1/n\right)$ and at least will not be $\mathbf{O}\left(1/n^2\right)$. This was in fact precisely the case, since direct computation in that example showed that

$$a_n = \mathbf{O}\left(\frac{1}{n^2}\right) \qquad \text{and} \qquad b_n = \mathbf{O}\left(\frac{1}{n}\right) \qquad .$$

A principal value of this theorem is that it provides a convenient and quick check on the accuracy of our computations of the Fourier coefficients. For example, it tells us that it would be inconsistent to find a coefficient

$$a_n = \frac{2(1 - cos(n\pi))}{n} = \mathbf{O}\left(\frac{1}{n}\right)$$

arising from a known continuous $f(x)$. (This check, unfortunately, is not totally foolproof in that if had we computed any coefficient a_n here which was $\mathbf{O}\left(1/n^2\right)$ the theorem could not guarantee that our exact algebra was correct, but only that we had arrived at the correct order. Nevertheless, this theorem can frequently, and "cheaply" indicate many cases where the algebra is wrong!)

This last theorem also completes our study of the relationships between the function represented by a Fourier series and the coefficients of that series. Besides introducing the new category of mean-square convergence, these results have two major consequences. First of all, given only the Fourier series for a function, they allow us to infer, by studying only the coefficients in that series, the continuity of both the basic function involved, and of its derivatives. Secondly, given the function, they provide a convenient and quick, though not foolproof, check on the accuracy of our computations of the Fourier coefficients.

PROBLEMS

1. For each of the following Fourier series, determine whether the series will converge uniformly, converge only in mean square, or diverge:

a. $\dfrac{1}{3} + \displaystyle\sum_{n=1}^{\infty} \dfrac{1}{n^3 \pi^3} \sin\left(\dfrac{n\pi x}{3}\right)$

b. $1 + \displaystyle\sum_{n=1}^{\infty} \dfrac{1}{n\pi} \cos(n\pi) \cos\left(\dfrac{n\pi x}{2}\right)$

c. $-2 + \displaystyle\sum_{n=1}^{\infty} \left\{ \dfrac{n \cos(n\pi)}{n^2 + 1} \cos(nx) - \dfrac{1}{n^2 \pi^2} \sin(nx) \right\}$

d. $\displaystyle\sum_{n=1}^{\infty} \dfrac{n}{(n+1)\pi} \sin(n\pi x)$

2. For each convergent Fourier series in problem 1 above, determine the highest derivative of the periodic extension of $f(x)$ that should be continuous.

3. Consider the Fourier series for

$$f(x) = |x|, \qquad -1 \leq x \leq 1$$

(found earlier). Differentiate this series once term by term, and compare your answer to the actual series of $f'(x)$.

2.5 Interpretation of the Fourier Coefficients

Thus far in this chapter, we have introduced the Fourier Series, learned how to calculate its coefficients (a_n and b_n) and seen how these coefficients relate to the convergence properties of the series. In this section we shall turn to a slightly different question - how should we *interpret* the Fourier coefficients, i.e. what do they really represent? To do this, we shall need the following trigonometric identity:

$$a \cos(\omega x) + b \sin(\omega x) = A \cos(\omega x - \phi) = A \cos(\omega(x - \delta))$$

where

$$A = \sqrt{a^2 + b^2}$$

$$\phi = \tan^{-1}\left(\frac{b}{a}\right)$$

$$\delta = \frac{1}{\omega} \tan^{-1}\left(\frac{b}{a}\right) = \frac{1}{\omega}\phi$$

When using this identity, most engineering texts refer to ϕ as the *phase angle*, and δ as the *delay*. (You should also note that, in order to use this identity both the sine and cosine must have the same frequency, and you will need to handle the inverse tangent function "delicately" when $a < 0$.) This identity is easily verified with the aid of the standard identity for the cosine of the difference of two angles:

$$\cos(u - v) = \cos(u)\cos(v) + \sin(u)\sin(v)$$

If we then apply this identity to the nth term in the Fourier series (2.1.1), we have

$$a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) = A_n \cos\left(\frac{n\pi x}{L} - \phi_n\right) = A_n \cos\left(\frac{n\pi}{L}(x - \delta_n)\right)$$

where,

$$A_n = \sqrt{a_n^2 + b_n^2}$$

$$\phi_n = \tan^{-1}\left(\frac{b_n}{a_n}\right)$$

$$\delta_n = \frac{L}{n\pi} \tan^{-1}\left(\frac{b_n}{a_n}\right)$$

and then (using the delay form of the identity) write the Fourier series (2.1.1) in the equivalent representation

$$f(x) = A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi}{L}(x - \delta_n)\right) \quad , \quad \left(A_0 = \frac{a_0}{2}\right) \qquad (2.5.17)$$

Representing the Fourier series this way clearly emphasizes that the nature of the Fourier series involves synthesizing a "complicated" periodic function as a combination of "pure" sinusoidal terms (musically, these sinusoids would correspond to pure tones). Furthermore, A_n now clearly represents the amplitude associated with the nth frequency component. Moreover, since the power or energy in a pure sinusoid is proportional to the square of the amplitude, the A_n also affect the way in which the energy in a complex signal is distributed within the different frequencies. Thus, a plot of A_n (or A_n^2) as a function of n (or of the frequency - $(n/2L)$) represents essentially a "fingerprint" of the signal in terms of amplitude (energy). This amplitude plot is commonly referred to as the spectrum of the signal. (Note that a complete characterization, or fingerprint of the signal would also require displaying the phase, ϕ_n as a function of n (or of frequency). A sample spectrum (for the signal computed as the first example in the first section) is shown below.
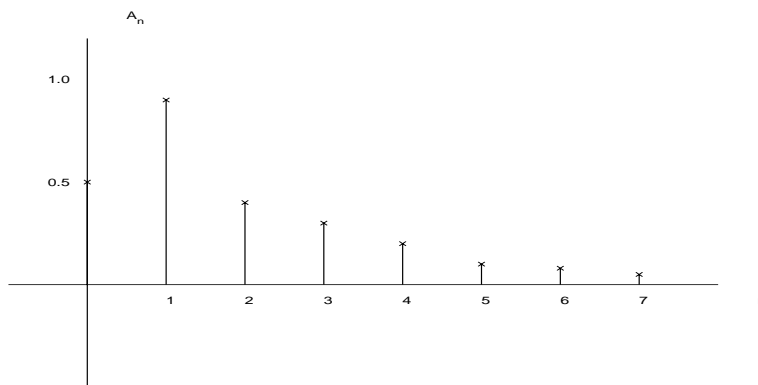
Figure 16: Spectrum of a Signal

The above interpretation can be further expanded by computing the average "energy" contained in a single period of a Fourier series signal, i.e.

$$\frac{1}{2L} \int_{-L}^{L} [f(x)]^2 \, dx$$

This integral is computed similarly to the derivation of (2.4.14), where, after squaring, most of the terms integrate to zero by orthogonality, leaving:

$$\frac{1}{2L} \int_{-L}^{L} [f(x)]^2 \, dx = A_0^2 + \sum_{n=1}^{\infty} \frac{1}{2} A_n^2 \equiv \frac{a_0^2}{4} + \sum_{n=1}^{\infty} \frac{\{a_n^2 + b_n^2\}}{2} \tag{2.5.18}$$

This result, which is frequently referred to as *Parseval's*[3] *Theorem*, has several rather interesting interpretations. First of all, it indicates that the power/energy in the different frequencies add *independently* to give the total energy. Stated another way, the various different frequencies in the series do not "spill over" or interact with each other. Changing the amplitude (or, it can also be shown, the phase) of a single frequency has no effect on the other

---

[3]Marc-Antoine Parseval des Chênes,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Parseval.html

terms in the series. This property has some interesting applications in communications, since it implies that several independent signals could be sent down a single physical "channel," and unambiguously decoded at the far end, provided a properly chosen set of different frequencies are used. In fact, one can show fairly simply that a more general concept of orthogonality than given by (3) is important for sending such signals, since a sufficient condition for separation of the energy in any two signals

$$\int_{-L}^{L} [f(x) + g(x)]^2 \, dx = \int_{-L}^{L} [f(x)]^2 \, dx + \int_{-L}^{L} [g(x)]^2 \, dx$$

is,

$$\int_{-L}^{L} f(x)g(x)dx = 0$$

A second conclusion which follows from (2.5.18) is that a necessary condition for mean-square convergence of a Fourier series is that the function, $f(x)$, must have only finite energy on the interval $(-L, L)$.

The final conclusion which we shall develop from (2.5.18) describes "how good" an approximation to $f(x)$ is obtained by using just a partial Fourier series. Specifically, suppose that we wish to approximate an appropriate $f(x)$ by just a finite trigonometric sum, i.e.

$$f_{\text{approx}}(x) = \frac{\alpha_0}{2} + \sum_{n=1}^{N} \left\{ \alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right) \right\} \qquad (2.5.19)$$

where the α_n and β_n are, for the moment, any set of constants, not necessarily a_n and b_n. Can we find a more accurate approximation than that given by the terms of the Fourier series? To answer this question, we write out the complete Fourier series for $f(x)$, and then subtract the trigonometric approximation from it (term by term), we would find that the error in approximating $f(x)$ by (2.5.19) is:

$$
\begin{aligned}
e(x) \;\; &= f(x) - f_{\text{approx}}(x) \\[2mm]
&= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left\{ a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right\} \\[2mm]
&\quad - \frac{\alpha_0}{2} - \sum_{n=1}^{N} \left\{ \alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right) \right\} \\[2mm]
&= \frac{\gamma_0}{2} + \sum_{n=1}^{\infty} \left\{ \gamma_n \cos\left(\frac{n\pi x}{L}\right) + \eta_n \sin\left(\frac{n\pi x}{L}\right) \right\}
\end{aligned}
$$

where

$$\gamma_n = \begin{cases} a_n - \alpha_n & , \quad 0 \le n \le N \\ a_n & \quad N < n \end{cases}$$

$$\eta_n = \begin{cases} b_n - \beta_n & , \quad 1 \le n \le N \\ b_n & \quad N < n \end{cases}$$

50

But since the error signal, $e(x)$, is itself now just a Fourier series, with coefficients γ_n and η_n, we can apply (2.5.18) and compute the total energy in the error to be

$$\int_{-L}^{L} [e(x)]^2 \, dx = \frac{L}{2} \left\{ \gamma_0^2 + \sum_{n=1}^{\infty} \left\{ \gamma_n^2 + \eta_n^2 \right\} \right\}$$

Now, up until this point, we have not indicated how we should choose the coefficients in the trigonometric approximation. But every term on the right in this last equation is non-negative. Therefore, the mean square error (i.e. the energy in the error signal) can be minimized only by making as many terms as possible in that equation equal to zero. Since we have no flexibility in the choice of a_n and b_n, we cannot affect the values of γ_n and η_n for $n > N$. This implies that we can minimize the error in our approximation (2.5.19) only by choosing

$$\begin{aligned} \gamma_n &= 0 \ , \quad 0 \leq n \leq N \ , \\ \eta_n &= 0 \ , \quad 1 \leq n \ , \end{aligned}$$

or,

$$\begin{aligned} \alpha_n &= a_n \ , \quad 0 \leq n \leq N \ , \\ \beta_n &= b_n \ , \quad 1 \leq n \ , \end{aligned}$$

The importance of this result (which sometimes can get lost in the algebra) is that in order to minimize the energy in the error of a trigonometric approximation, we must use precisely the Fourier coefficients, a_n and b_n, and therefore any purely trigonometric approximation cannot be improved from that given by the partial sum of a Fourier series without increasing the number (N) of frequencies considered. This observation is summarized in the following theorem.

> **Theorem**: The Fourier coefficients, a_n and b_n, are optimum in the sense that they provide the best mean square approximation of a given function by a finite number of trigonometric functions.

This theorem represents the last result we shall derive in this section. We have by no means exhausted all of the interpretations that can be applied to the Fourier coefficients, but we have developed the primary ones that relate these coefficients to the amplitude, phase, and power at the different frequencies in the series. Lastly, we have shown that the Fourier coefficients provide, in the mean square sense, the "best possible" choice for coefficients in a trigonometric expansion of a given function.

PROBLEMS

1. Plot the amplitude and phase as a function of frequency for each of the Fourier series found for the problems in the first section of this chapter.

2. Prove Parseval's Theorem (equation (2.5.18)). (Hint: first show

$$\frac{1}{2L} \int_{-L}^{L} [f(x)]^2 \, dx = \frac{a_0^2}{4} + \sum_{n=1}^{\infty} \frac{\{a_n^2 + b_n^2\}}{2}$$

then use the definition of the A_n.)

2.6 The Complex Form of the Fourier Series

Although the reasons may not be apparent now, it is often advantageous to express the basic concepts of Fourier series in terms of complex variables. Such an alternative formulation is clearly possible since the relations

$$e^{ix} = \cos(x) + i\sin(x) \ ,$$

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}$$

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}$$

allow conversion between the normal, real-valued trigonometric forms and the complex exponential form. (No apology will be made here for the mathematician's use of the symbol i rather than the engineer's j to stand for $\sqrt{-1}$.)

We could now convert (2.1.1) to complex form by replacing the trigonometric functions there with the above equivalent complex exponential forms, and then algebraically regroup the terms. This substitution will clearly produce complex exponential terms of the form

$$e^{in\pi x/L}$$

for all positive *and negative* values of n, i.e. for all values between plus and minus infinity, and therefore reduce (2.1.1) to an expression like:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L} \tag{2.6.20}$$

where the c_n are (probably complex) constants. (Note that the $n = 0$ term of this series must represent the constant (DC) term of the original signal.) As already noted, we could determine the values for c_n by algebraically regrouping the terms, after replacing the trigonometric functions by their equivalent complex exponentials. But there is an easier way, which also happens to be theoretically more pleasant!

Recall how we derived the coefficients in the original, real-valued Fourier series. We multiplied both sides of the series equation by just one of the trigonometric functions, and then integrated and applied the orthogonality integrals. So why not try a similar approach here?

With this general approach in mind, let k stand for some fixed, but arbitrary integer, and multiply both sides of (2.6.20) by $e^{ik\pi x/L}$, and integrate from $-L$ to L. After interchanging the summation and integration, this yields

$$\int_{-L}^{L} f(x)e^{ik\pi x/L}dx = \sum_{n=-\infty}^{\infty} c_n \int_{-L}^{L} e^{i(k+n)\pi x/L}dx$$

But now, direct computation will show that the complex exponentials also obey an orthogonality relation

$$\int_{-L}^{L} e^{i(k+n)\pi x/L}dx = \begin{cases} 0 & , \quad n \neq -k \\ 2L & , \quad n = -k \end{cases} \tag{2.6.21}$$

Therefore, the above equation simplifies to

$$\int_{-L}^{L} f(x)e^{ik\pi x/L}dx = 2L\ c_{-k}$$

or, replacing k by $-n$ on both sides,

$$c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-in\pi x/L}dx \qquad (2.6.22)$$

With this formula for the c_n, we can now express the complex exponent in terms of ordinary trigonometric functions, and use the fact that the integral of a sum in the sum of the integrals, to rewrite (2.6.22) as

$$
\begin{aligned}
c_n &= \frac{1}{2L}\left\{\int_{-L}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right)dx - i\int_{-L}^{L} f(x)\sin\left(\frac{n\pi x}{L}\right)dx\right\} \\
&= \frac{1}{2}\{a_n - ib_n\}\ , \quad n > 0
\end{aligned}
\qquad (2.6.23)
$$

where a_n and b_n are, of course, just the standard Fourier coefficients as given by (2.2.5)-(2.2.6). This shows, as we had expected, the c_n are complex. Furthermore, we can now easily relate them to the A_n and ϕ_n that appeared when we converted the real-valued Fourier series to amplitude-phase form. Specifically, we see from (2.6.23) that:

$$|c_n| = \tfrac{1}{2}A_n$$

$$\arg(c_n) = -\tan^{-1}\left(\frac{b_n}{a_n}\right) = -\phi_n = -\frac{n\pi}{L}\delta_n$$

(Note also that, provided $f(x)$ is real, $c_{-n} = c_n^*$, i.e. c_{-n} is the complex conjugate of c_n. Thus $|c_{-n}| = \tfrac{1}{2}A_n$ and $\arg(c_{-n}) = \phi_n$.) Thus the energy in any single frequency in the real series is split equally between the two (one positive and one negative) respective complex frequencies.

From the above, it is clear that one could plot $|c_n|$ as a function of n (for $-\infty < n < \infty$) and, except for the fact that the amplitude associated with the each real frequency $(n\pi/L)$ is equally divided between the two complex frequencies, we obtain the identical information that would be obtained from plotting the real-valued amplitude spectrum described earlier. In fact, especially in applications in electrical engineering and signal processing, the complex amplitude spectrum is customarily used instead of the real-valued amplitude spectrum. There are many reasons for this, including both the relative ease of working with complex numbers, especially when multiplications are involved, and the often powerful insights that are possible using the complex plane representation.

At this time, however, we will only consider the complex Fourier series as simply an alternative to the real-valued representation, although we shall return to it later when we discuss Fourier transforms.

PROBLEMS

1. Find the complex Fourier series for each of the following functions:
    a. $f(x) = x, -3 \le x \le 3$, $f(x+6) = f(x)$
    b. $f(x) = \begin{cases} 0 & , & -1 < x < 0 \\ 1 & , & 0 < x < 1 \end{cases}$
          $f(x+2) = f(x)$

2. Plot the complex amplitude spectrum for each of the series found in problem 1 above.

3. Show that if we use T_0 for the period of a signal, rather than $2L$, the formula for the complex Fourier series coefficients reduces to

$$c_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(x) e^{-2in\pi x/T_0} dx$$

4. Using the complex form of the Fourier series, prove the following form of Parseval's theorem

$$\frac{1}{2L} \int_{-L}^{L} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |c_n|^2$$

(Hint: Show

$$|f(x)|^2 = f(x)f(x)^* = \left[\sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L} \right] \left[\sum_{m=-\infty}^{\infty} c_m^* e^{-im\pi x/L} \right]$$

$$= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \left[c_n c_m^* e^{i(n-m)\pi x/L} \right] \quad ,$$

then integrate.)

2.7 Fourier Series and Ordinary Differential Equations

Thus far, we have concentrated on developing the basic concepts of Fourier Series themselves. There are many other interesting questions about these series, which we might pursue. The first one we shall study is where Fourier Series "fit" in the scheme of ordinary differential equations.

Consider the ordinary differential equation,

$$py'' + qy' + ry = f(x) \tag{2.7.24}$$

where p, q, and r may be any constants, and $f(x)$ is some general periodic function, i.e. $f(x + 2L) = f(x)$, but $f(x)$ need not be a simple sine or cosine. (For example, $f(x)$ might look like Figure 17.) We now wish to ask the question - are any of the ordinary differential equations methods which you should have seen prior to Fourier Series (variation of parameters, Laplace transforms, or undetermined coefficients) well-suited for this problem?

We know that variation of parameters will produce a solution to this problem. But, you should recall that, besides being usually the least efficient method to produce non-homogenous solutions (when a viable choice exists), variation of parameters produces a solution in terms of *integrals* which involve the forcing function. But, as we have seen in our earlier computations of Fourier coefficients, evaluating an integral involving a piecewise smooth $f(x)$ normally involves splitting the integral into several parts. For example, for the $f(x)$ given in Figure 17, if e^{-x} were a homogeneous solution,
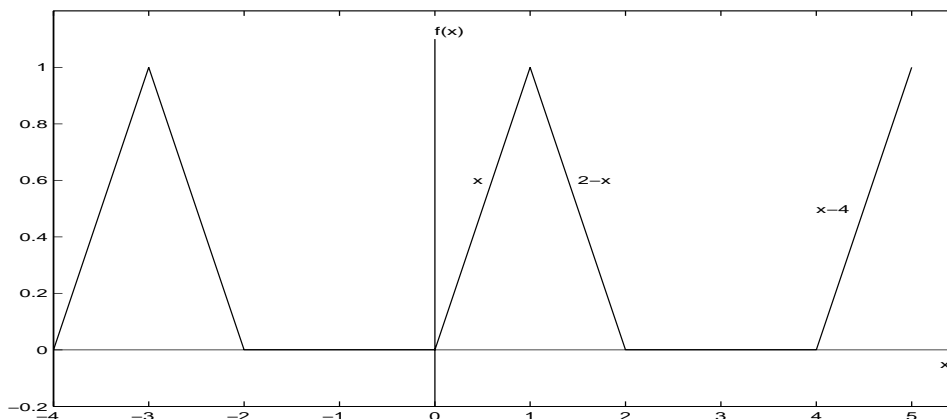
Figure 17: A Typical Periodic Function

variation of parameters would require computing integrals like

$$\int_0^{4.5} e^{-x} f(x) dx$$

which, due to definition of $f(x)$, would have to be expressed as,

$$\int_0^1 x e^{-x} dx + \int_1^2 (2 - x) e^{-x} dx + \int_4^{4.5} (x - 4) e^{-x} dx$$

and would actually require three integrations. Such behavior could, in general, make the evaluation of variation of parameters integrals a computational morass for this class of problems. Thus, if at all possible, we would like to avoid variation of parameters here.

The utility of the Laplace transform in this class of problems is affected by the fact that most "interesting" periodic, non-sinusoidal, functions have jump discontinuities in either the function or one of its derivatives. (The $f(x)$ graphed above has discontinuities in $f'(x)$ at each of the sharp corners.) Furthermore, it is well-known from the study of Laplace transforms that each such jump discontinuity in $f(x)$ will produce an e^{-sc} type term in the transform, where $x = c$ is the location of the discontinuity. But with a periodic function, if there is one discontinuity, there must be an infinite number. Therefore, for example, the transform of a square wave (see Fig. 18)
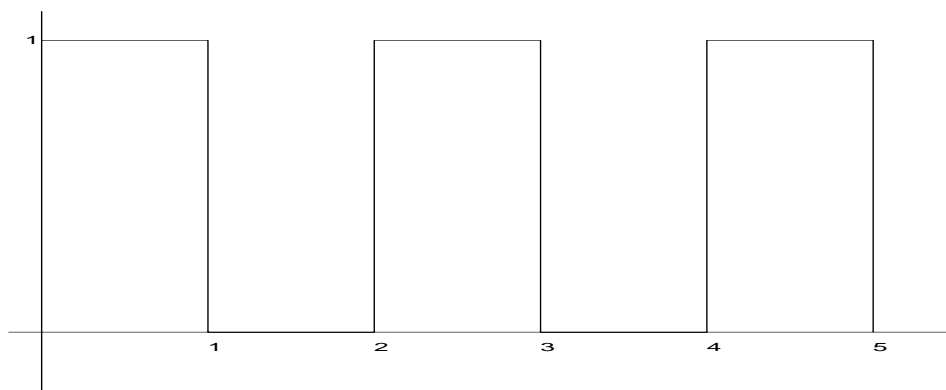
Figure 18: Square Wave

would have an infinite number of different e^{-sc} type terms. (Specifically, you should be able to show for this $f(x)$:

$$F(s) = \frac{1}{s} + \sum_{n=1}^{\infty} (-1)^n \frac{e^{-ns}}{s}$$

But, since using the Laplace transform to solve the ordinary differential equation (2.7.24) involves a term of the form

$$\frac{F(s)}{as^2 + bs + c}$$

it is clear that if $F(s)$ contains an infinite number of different e^{-sc} type terms, you will, at least in principle, have to apply the Laplace inverse (including the Heaviside shift) an infinite number of different times. Like variation of parameters, this is possible, at least in theory, but it normally becomes computationally cumbersome, and leads to results often too complicated to provide any insight into the real properties of the solution. Thus, we conclude that the Laplace transform is also not a particularly suitable method for this class of problems.

With neither variation of parameters nor the Laplace transform seeming to be an especially attractive method for this class of problems, we turn to the last of the standard methods studied with ordinary differential equation methods - undetermined coefficients. Undetermined coefficients, however, can be used only when the forcing function, $f(x)$, meets

fairly stringent criteria. Specifically, the only periodic functions for which the method applies are sines, cosines, and linear combinations of these. Functions such as that shown in Figure 18, at first, do not appear to meet this criterion.

However, we now know that this (and similar) $f(x)$ can be written as a combination of sines and cosines (specifically in the form of a Fourier Series). Furthermore, since the ordinary differential equation is linear, we know the Principle of Superposition holds. Thus, we claim that a particular solution to

$$py'' + qy' + ry = f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left\{ a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right\}$$

is

$$y_p(x) = y_{p_0}(x) + \sum_{n=1}^{\infty} y_{p_n}(x)$$

where

$$py''_{p_0} + qy'_{p_0} + ry_{p_0} = \frac{a_0}{2}$$

$$py''_{p_n} + qy'_{p_n} + ry_{p_n} = a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \quad , \quad n = 1, 2, \ldots$$

It should be obvious now that not only can undetermined coefficients be applied to each of these equations, but that in most cases a general expression for $y_p(x)$ can be obtained by assuming,

$$y_{p_n}(x) = \alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right) \quad , \quad n = 1, 2, \ldots$$

(Note that under some special values of p, q, and r, this form might not be appropriate for one (and only one) value of n.) In fact, if p, q and r are all positive, then we can prove that a particular solution will always have the form

$$y_p(x) = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} \left\{ \alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right) \right\}$$

where,

$$y_{p_0} = \frac{a_0}{2r}$$

provided only that the Fourier Series for f(x) converges.

Thus we conclude that the main value of Fourier Series in ordinary differential equations is that it allows us to decompose complicated periodic forcing functions into a linear combination of sines and cosines, thus permitting the particular solution to a linear, constant coefficient equation to be obtained by using undetermined coefficients and the Principle of Superposition. (As we shall shortly see, the role of Fourier Series in partial differential equations is even more crucial.)

PROBLEM

1. Use Fourier series to construct a non-homogeneous solution to the ordinary differential equations:

$$y'' + 2y' + y = f(x),$$

where:

$$f(x) = \begin{cases} x & , & 0 \le x < 1 \\ 2 - x & , & 1 \le x \le 2 \end{cases}$$
$$f(x + 2) = f(x)$$

2.8   Fourier Series and Digital Data Transmission

The emergence of so-called digital signal processing has had a major impact on electronics, and especially communications. In its simplest form, a digital signal consists of a sequence of ones and zeros, for example,

$$1 \ \ 1 \ \ 0 \ \ 1 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ \ .$$

(You may recognize this as the binary representation of the number 212.) Physically, on transmission lines, a "1" might be represented by a unit voltage being present, and a "0" by no voltage. Thus, the above sample signal might be transmitted, in time sequence, as shown in Figure 19.
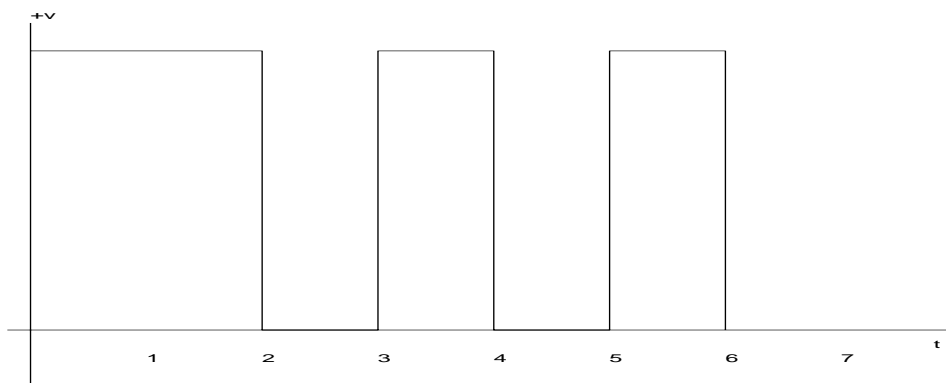
Figure 19: A Transmitted Digital Signal

A major concern of communications engineers is how much will such a signal be altered, distorted or degraded as it passes through both circuits and transmission lines. A very simple model that nevertheless illustrates two major effects can be developed by using some elementary concepts from Fourier series and ordinary differential equations to analyze the response of the circuit shown in Figure 20 to an applied digital test signal.

This circuit consists of a resistor, capacitor, inductor and input voltage source, $E(t)$. The output is assumed to occur by measuring the voltage across the resistor (which is often called the load). If $Q(t)$ and $I(t)$ denote, respectively, the charge on capacitor and the current in the circuit at time t, then $Q(t)$ and $I(t)$ are related by:

$$I(t) = \frac{d}{dt}\left[Q(t)\right] = \frac{dQ}{dt} \quad . \tag{2.8.25}$$

The relations between these variables and the instantaneous voltages across the resistor, capacitor and inductor are given, respectively, by

$$V_R = RI \ , \quad V_C = Q/C \ , \quad \text{and} \quad V_L = L\frac{dI}{dt} \quad .$$
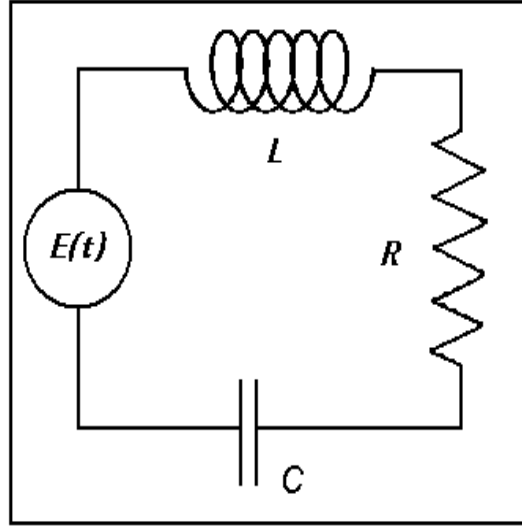
Figure 20: A Simple Circuit

Thus, if we apply Kirchhoff's Law, which states that the sum of the voltage drops around the circuit components must equal the applied (source) voltage, we have

$$L\frac{dI}{dt} + RI + \frac{1}{C}Q = E(t) \quad ,$$

or

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = E(t) \quad . \tag{2.8.26}$$

Suppose now that this circuit is subjected to an $E(t)$ made up of the periodic train of "0"s and "1"s shown in Figure 21.
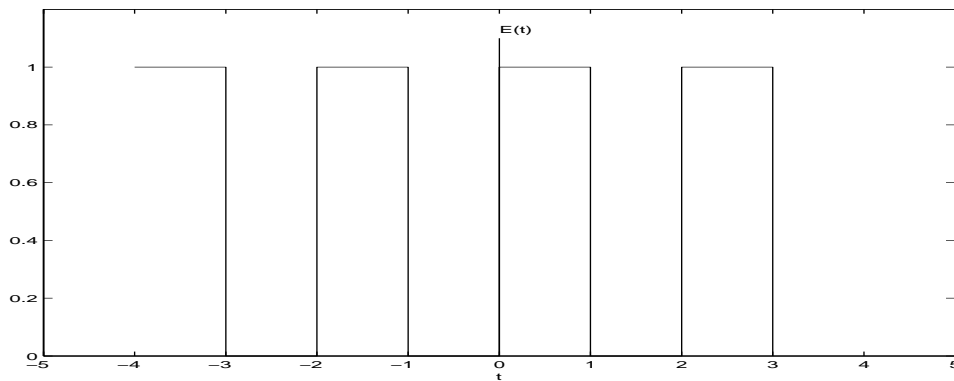
Figure 21: A Periodic Digital Test Signal

Our principal interest is how closely does the voltage across R follows this pattern, or equivalently, could another electrical device attached to R reasonably infer what the original test signal was based on the voltages observed across R? To determine this, we first assume that

61

we look at the circuit after all transients have effectively dropped to zero, i.e. when only the steady-state solution persists. The input signal, E(t), can be represented by the Fourier series:

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

$$= \frac{1}{2} + \sum_{n=1}^{\infty} \frac{(1 - \cos(n\pi))}{n\pi} \sin(n\pi t) \quad , \tag{2.8.27}$$

where the a_n and b_n are determined by the usual formulae,

$$a_n = \int_{-1}^{1} E(t) \cos(n\pi t) \ dt \quad ,$$

$$b_n = \int_{-1}^{1} E(t) \sin(n\pi t) \ dt \quad .$$

Now, since $E(t)$ is thus a sum (albeit an infinite one) of forcing terms, and since the differential equation (2.8.26) is linear, we invoke the Principle of Superposition, and claim the steady-state solution should have the form

$$Q_{ss}(t) = Q_0(t) + \sum_{n=1}^{\infty} Q_n(t) \quad ,$$

where each $Q_n(t)$ is the response to a single frequency in the forcing voltage, $E(t)$. Thus,

$$L\frac{d^2 Q_0}{dt^2} + R\frac{dQ_0}{dt} + \frac{1}{C}Q_0 = \frac{1}{2} \quad ,$$

and $\tag{2.8.28}$

$$L\frac{d^2 Q_n}{dt^2} + R\frac{dQ_n}{dt} + \frac{1}{C}Q_n = \frac{(1 - \cos(n\pi))}{n\pi} \sin(n\pi t) \ , n = 1, 2, 3, \ldots \quad .$$

(In reality, this is a *non-trivial* extension of the usual principle of superposition. The forcing function is **not** a *finite* sum, and the whole structure of infinite series arises precisely because not all properties that hold for finite sums also hold for infinite series. Rigorous justification of this approach at this point would require a detailed analysis of the rates of convergence of the series involved. Fortunately, in the case of Fourier series, such an investigation will almost always show that this method is valid.)

Provided $R > 0$, there will be no undamped resonance, and therefore, by the method of Undetermined Coefficients, we see that the $Q_n(t)$ will have the form

$$Q_0(t) = \frac{\alpha_0}{2} \quad ,$$

$$\tag{2.8.29}$$

$$Q_n(t) = \alpha_n \cos(n\pi t) + \beta_n \sin(n\pi t) \ , \qquad n = 1, 2, 3, \ldots \quad ,$$

62

where α_n and β_n are constants. Substitution of the above expressions into (2.8.28), followed by equating the coefficients of like terms, yields

$$\alpha_0 = C \quad ,$$

and

(2.8.30)

$$- \left[n^2 \pi^2 L - (1/C) \right] \alpha_n + n \pi R \beta_n = 0 \ ,$$

$$-n \pi R \alpha_n - \left[n^2 \pi^2 L - (1/C) \right] \beta_n = \frac{(1 - \cos(n\pi))}{n\pi} \quad .$$

Solution of the system of two equations in two unknowns in (2.8.30) yields

$$\alpha_n = \frac{(1 - \cos(n\pi))}{n\pi} \frac{[-n\pi R]}{\left[\left(n^2 \pi^2 L - \frac{1}{C} \right)^2 + n^2 \pi^2 R^2 \right]} \quad ,$$

(2.8.31)

$$\beta_n = \frac{(1 - \cos(n\pi))}{n\pi} \frac{[-(n^2 \pi^2 L - (1/C))]}{\left[\left(n^2 \pi^2 L - \frac{1}{C} \right)^2 + n^2 \pi^2 R^2 \right]} \quad .$$

Thus

$$Q_{ss}(t) = \frac{C}{2} + \sum_{n=1}^{\infty} \frac{(1 - \cos(n\pi))[-n\pi R \cos(n\pi t) - (n^2 \pi^2 L - (1/C)) \sin(n\pi t)]}{n\pi \left[\left(n^2 \pi^2 L - \frac{1}{C} \right)^2 + n^2 \pi^2 R^2 \right]} \quad . \quad (2.8.32)$$

Recall, however, what we are really interested in is the voltage across the (load) resistor, R. But, according to our basic equations, the steady-state current in the circuit should be just the time derivative of the steady-state charge on the capacitor, i.e.

$$I(t) = \frac{dQ_{ss}}{dt} \quad ,$$

and, therefore, by Ohm's Law, the steady-state voltage across the resistor is then given by

$$V_{ss}(t) = R \frac{dQ_{ss}}{dt} \quad .$$

Differentiating $Q_{ss}(t)$ as given by (2.8.32) term by term (which, if we were mathematically precise, would again require consideration of the rates of convergence of the series involved) and then multiplying by the resistance, yields

$$V_{ss}(t) = R \sum_{n=1}^{\infty} \frac{(1 - \cos(n\pi))[n\pi R \sin(n\pi t) - (n^2 \pi^2 L - (1/C)) \cos(n\pi t)]}{\left[\left(n^2 \pi^2 L - \frac{1}{C} \right)^2 + n^2 \pi^2 R^2 \right]} \quad . \quad (2.8.33)$$

63

Standard trigonometric identities[4] can then be used to reduce this to

$$V_{ss}(t) = \sum_{n=1}^{\infty} \frac{(1 - \cos(n\pi))}{n\pi} A_n \sin(n\pi t - \phi_n) \quad , \tag{2.8.34}$$

where

$$A_n = \frac{n\pi R}{\sqrt{\left[\left(n^2\pi^2 L - \frac{1}{C}\right)^2 + n^2\pi^2 R^2\right]}} \quad ,$$

$$\tag{2.8.35}$$

and

$$\phi_n = \tan^{-1}\left[\frac{(n^2\pi^2 L - (1/C))}{n\pi R}\right] \quad .$$

Careful examination of this representation for $V_{ss}(t)$ yields some very valuable insights into the problems of digital communications, starting with the observation that there are striking similarities between the series for the output, $V_{ss}(t)$, as given by (2.8.34), and the series for the input, $E(t)$, as given by (2.8.27). However, there are also two striking differences:

(1) Each sine term in $V_{ss}(t)$ has an amplitude equal to

$$\frac{(1 - \cos(n\pi))}{n\pi} A_n \quad , \tag{2.8.36}$$

which is precisely A_n times the amplitude of the corresponding sine term (i.e. frequency) in the input.

(2) Each sine term in the output has undergone a phase shift, due to the ϕ_n term, relative to the corresponding input term.

Communications engineers would describe these two effects as *amplitude distortion* and *phase (or delay) distortion*, respectively. We shall see shortly exactly why these effects distort the signal, but first we need to more explicitly define the term "distortion" and illustrate our definition with some simple examples.

By *distortion*, we shall mean any change that serves to alter the fundamental *shape* of a signal, i.e. a change such that the original signal cannot be recovered by any combination of only the following techniques:

---

[4]The identity is:

$$A\sin\Theta - B\cos\Theta \;=\; \sqrt{A^2 + B^2}\left[\frac{A}{\sqrt{A^2 + B^2}}\sin\Theta - \frac{B}{\sqrt{A^2 + B^2}}\cos\Theta\right]$$

$$= R\sin(\Theta - \phi)$$

where

$$R = \sqrt{A^2 + B^2} \;, \quad \sin\phi = \frac{B}{\sqrt{A^2 + B^2}} \;, \quad \text{and} \quad \cos\phi = \frac{A}{\sqrt{A^2 + B^2}} \quad .$$

(1) simple amplification (i.e. change of scale of the vertical axis, or, equivalently, multiplication of the function by a constant)[5],

(2) insertion of a DC bias (i.e. shift up or down of the horizontal (t) axis, or, equivalently, addition of a constant),

(3) a simple time delay (i.e. a shift right or left of the location of the vertical axis, or, equivalently, replacing t by $(t - c)$).[6]

Thus, for example, figures 22(a)-(c) do not represent distortion of the original $E(t)$, since $E(t)$ can be recovered, respectively, by amplification by a factor (gain) of 2, a time shift of $(1/2)$, and a DC bias of $(1/4)$. Figure 22(d), however, does represent a distortion, since there is no way to restore its form to that of $E(t)$ by the above operations.
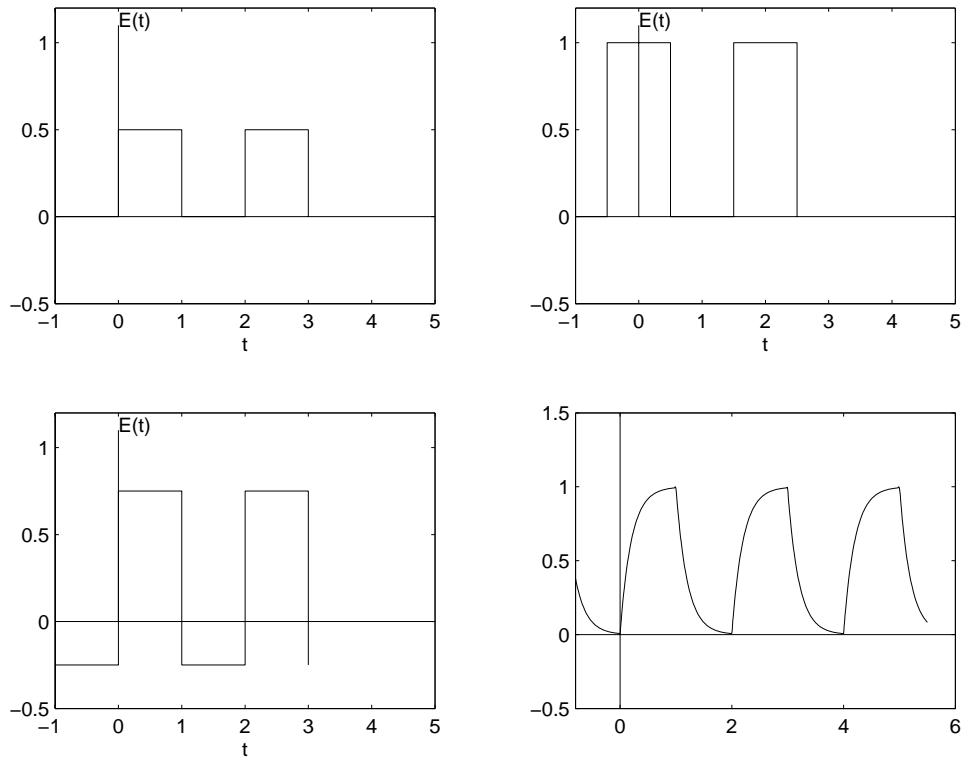
Figure 22: Undistorted and Distorted Signals

It is relatively easy to prove that if a signal is composed of more than one frequency (Fourier series term) and if any one component frequency (term) is multiplied by a constant, then, to remain undistorted, every component must be multiplied by the *same* factor. A similar conclusion can also be shown to be valid for time delays of a signal composed of a

---

[5]The term "amplification" is used here in the general sense which includes cases where the "amplified" signal is actually weaker than the original signal. The latter cases is commonly referred to in engineering as "amplification with a gain less than one."

[6]With periodic functions, one must be slightly careful to ensure that the delays are truly different, since a delay of an integral number of periods is, physically, undetectable.

sum of terms, i.e. any effect which serves to introduce different delays into different terms will distort the resulting signal, just as would "out of step" marchers in a parade.

With these ideas in mind, return to (2.8.36). Note that for each value of n, A_n represents the factor by which the amplitude of that frequency component in the input has been amplified. As is obvious from (2.8.35), A_n is not constant with respect to different values of n. Thus to restore the amplitudes in the output to their original (input) values would require multiplication of each term by a *different* amount, i .e. different amplification at each frequency. Since, as we discussed above, this is not the same as amplifying the output signal by a single factor, we conclude that $V_{ss}(t)$ is a distorted version of $E(t)$. Since this distortion arises due to changes in the relative amplitudes at different frequencies, it is aptly termed amplitude distortion.

However, distortion of the relative amplitudes is not the only source of distortion in $V_{ss}(t)$. Also observe that, since $V_{ss}(t)$ (2.8.34) can be written

$$V_{ss}(t) = \sum_{n=1}^{\infty} \frac{(1 - \cos(n\pi))}{n\pi} A_n \sin\left(n\pi\left(t - \frac{1}{n\pi}\phi_n\right)\right) \quad , \tag{2.8.37}$$

the nth frequency in $V_{ss}(t)$ is shifted, relative to the corresponding frequency in the input. Here, a little careful analysis is required, for if ϕ_n in (2.8.34) is positive, then each term in $V_{ss}(t)$ is *advanced* compared to the same term in $E(t)$. Equivalently, a negative value of ϕ_n corresponds to a *delay* in the response. Therefore, the nth frequency in $V_{ss}(t)$ is delayed, relative to the corresponding frequency in $E(t)$, by a time interval equal to

$$
\begin{aligned}
t_{dn} &= -\frac{1}{n\pi}\phi_n \\
&= -\frac{1}{n\pi}\tan^{-1}\left[\frac{(n^2\pi^2 L - (1/C))}{n\pi R}\right] \quad .
\end{aligned}
\tag{2.8.38}
$$

But this implies that, unless t_{dn} is constant, the different frequencies in the output signal have been delayed selectively, i.e. by different amounts. But, as defined above, t_{dn} is clearly not constant. Thus the output signal can be restored to its original form only by inserting an additional *different* delay at each frequency, not a single delay at all frequencies. As with selective amplification, selective delay must distort the signal, and hence we conclude that any non-constant ϕ_n also distort $V_{ss}(t)$. This type of distortion, as we commented above is commonly referred to as phase or delay distortion. (In transmission through physical media, such as in optical communications, a similar effect is observed. That effect is commonly attributed to a non-constant index of refraction, which causes slightly different transmission velocities at different frequencies, or slightly different transmission path lengths at different frequencies. When encountered in this context, the effect is often referred to as *dispersion*.)

(We close this part of the discussion by noting that a mathematician would more likely conclude that $V_{ss}(t)$ is distorted by noting that for *large n*, the coefficients in (2.8.33) and (2.8.34) obey:

$$\frac{(1 - \cos(n\pi))}{n\pi}A_n = \mathbf{O}\left(\frac{1}{n^2}\right) \quad ,$$

66

and hence the series for $V_{ss}(t)$ must converge *uniformly* (by the Weirstrauss M-test). A standard result from the study of infinite series then guarantees that $V_{ss}(t)$ is a continuous function, and hence cannot have the sharp "jumps" that characterize $E(t)$. Thus, clearly, the "shape" of $V_{ss}(t)$ cannot be the same as that of $E(t)$, and hence $V_{ss}(t)$ is a distortion of $E(t)$.)

Observe, for a moment, what we have just done. We have studied the possible effects of the circuit on the original input signal by analyzing how each of the *individual frequencies* in that signal are affected. This is what we call a frequency domain analysis! This method is almost certainly the best, if not the only way to understand why the output signal may not "look" exactly like the input.

How much or little effect distortion can have on a signal can be graphically demonstrated by two carefully chosen numerical examples. Consider first the case

$$L = 0.02 \quad , \quad C = 5.00 \quad \text{and} \quad R = 0.75 \quad .$$
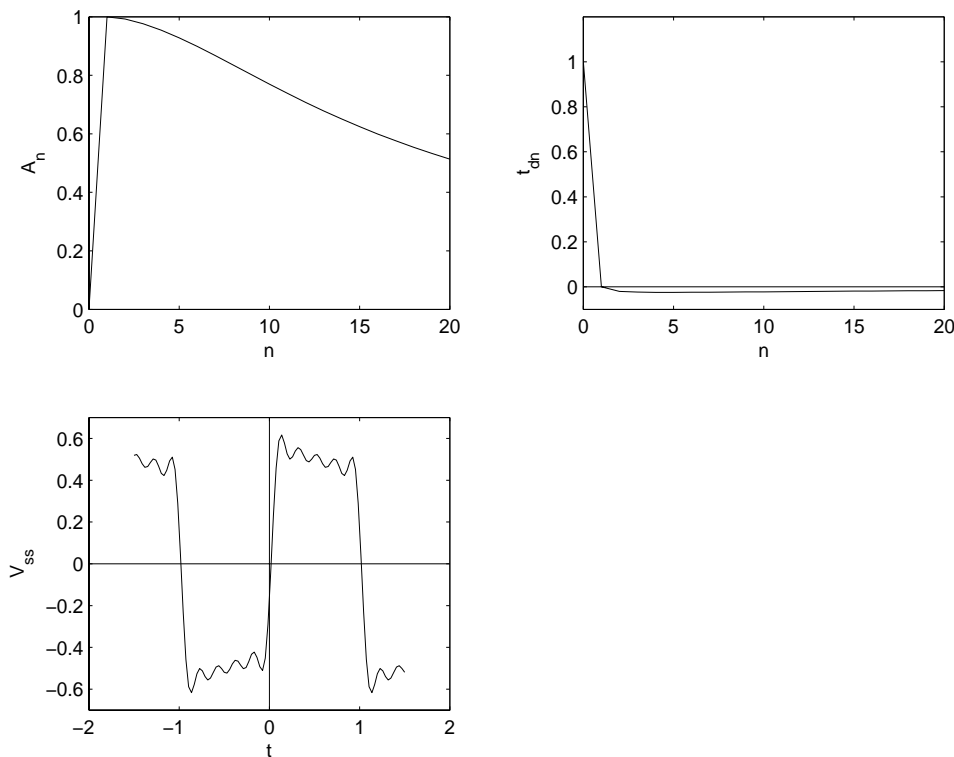
Figure 23: First Sample Output

The graphs of A_n and t_{dn} for these specific values of L, R and C are shown at Figures 23(a)-(b), respectively. (Note that in these figures A_n and t_{dn} are shown for all values, not only the integer (n) values, although when $V_{ss}(t)$ is computed, only the values at the integer points will be used.) The curve for A_n shows noticeable attenuation at higher frequencies, although for $n < 20$ the attenuation never exceeds 50%. (Engineers would refer to this as the 3 dB point.) The curve for t_{dn} appears almost flat for $n > 1$, suggesting almost

67

no noticeable delay distortion should occur in this example. Computation shows that the maximum difference between the delays encountered by different frequencies is:

$$t_{d1} - t_{d5} = 0.0185$$

which is less than 1% of one period. Therefore, with neither the amplitude nor delay distortion seeming especially severe, we might expect the output across R would not be too badly distorted. This is, in fact, the case. Figure 23(c) shows $V_{ss}(t)$, computed for the values of L, R, and C given above by taking a sufficient number of terms in a partial sum of (2.8.33). Note that except for the DC shift of 1/2 unit, it is a fairly recognizable replica of the original signal. Note that Figure 23(c), since it represents the sum of the series, is a time domain view of the phenomenon, while, as discussed above, Figures 23(a) and 23(b) are frequency domain views.)

Lastly, consider the example given by:

$$L = 0.16 \quad , \quad C = 0.01 \quad \text{and} \quad R = 0.75 \quad .$$
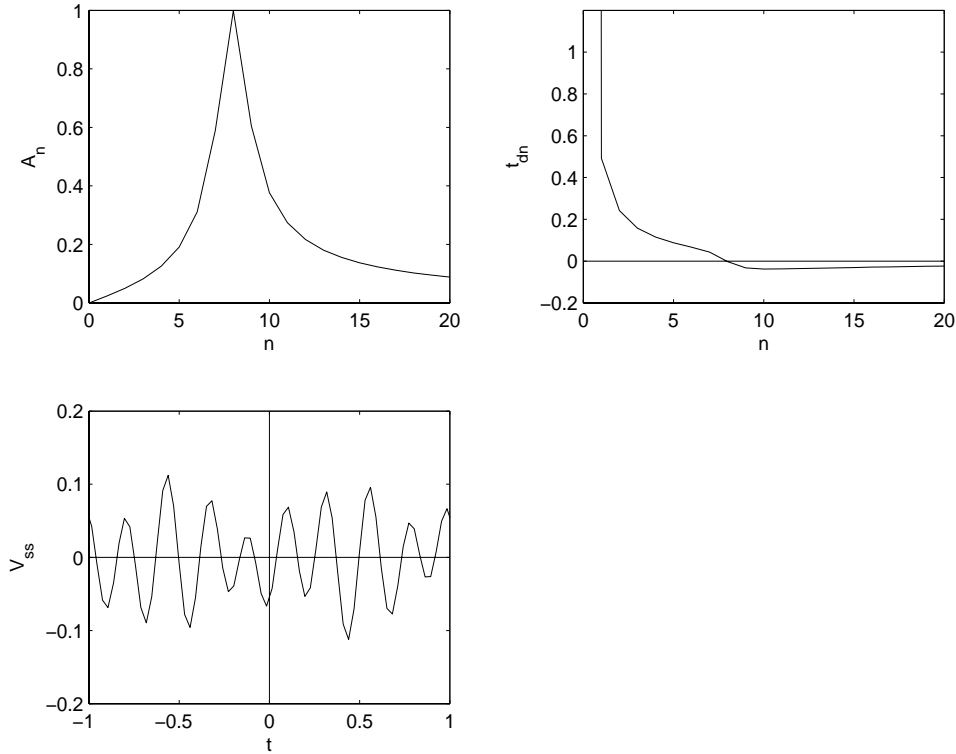
Figure 24: Second Sample Output

The graphs for A_n, and t_{dn} for these values are shown in figures 24(a)-(b), respectively. (Again, the figures plot non-integer values as well as integer values.) Observe that in this case the curves are nowhere near as constant as the corresponding curves in the previous example. The curve for A_n shows that frequencies with $n < 6$ and $n > 12$ will be severely

68

attenuated. Furthermore, the maximum difference between delays at different frequencies in this case is approximately 0.40, or a factor of twenty greater than in the previous example. Thus, we would expect the output from this circuit to be noticeably more distorted than in previous case. As shown by Figure 24(c), it is. In fact, the time domain output represented by Figure 24(c) is so badly distorted that it would be almost impossible to determine what the original (input) signal had been by observing the output. Of course, that would make the second circuit almost useless for communicating the original signal. In contrast, since the output from the first circuit is only slightly distorted from the original input, that circuit would be reasonably appropriate for transmitting this particular signal.

This example demonstrates how the ability of Fourier series to decompose complicated, periodic (time domain) signals into "pure" (frequency domain) sines and cosines allows a communications engineer/designer to analyze the distortion introduced into a digital signal as it passes through a simple circuit, and to compute the (distorted) output signal from that same circuit. This is but one aspect of the power of Fourier series - the ability to provide simple analysis of the response of linear systems.

3 The One-Dimensional Wave Equation

3.1 Introduction

One of the principal applications of Fourier series is to the solution of partial differential equations. Such equations are fundamental to both the physical sciences and to engineering, and describe a tremendous variety of different and important physical processes. In this chapter, we shall study the application of Fourier series to primarily one partial differential equation, the so-called wave equation. Our intent in this study will be two fold - not only to investigate the *mechanics* of solving these problems with Fourier series , but also to *observe and interpret* the physical insights contained in these solutions. As part of these insights, we shall see that the wave equation shares the same dual time-domain, frequency-domain nature that we observed earlier in the Fourier series itself and in the application of Fourier Series to the solution of ordinary differential equations.

The wave equation is really a mathematical model - a set of equations that describe some *idealized*, simplified version of reality. Such mathematical models are common in applications, and generally arise when one tries to include only the *most important* inter-relationships, forces, etc. In modeling most physical processes, we start by identifying the primary forces and relationships. Then, we generally apply various conservation laws to the identified forces to produce the governing equations. Finally, once these equations have been solved, we must show that our solutions accurately reflect the observed actual physical behavior before we can consider the model to be valid. As we proceed with our study of the wave equation, we will endeavor to point out these various aspects of the modeling process when they occur.

3.2 The One-Dimensional Wave Equation

The physical prototype for the wave equation is a thin, tightly stretched elastic string, such as a violin string or a guitar string. These strings are usually wound around some type of peg to produce an extremely strong internal tension. (The magnitude of this tension can be readily appreciated by anyone who has seen a guitar string break while being played.)

The first step in deriving the mathematical model is to identify what we believe are the primary forces and interactions, or, equivalently, to assume which terms and interactions can be neglected. For our model, we assume that the string is extremely thin in comparison to its length, and so tightly stretched that at rest (equilibrium) it does not sag, but lies perfectly horizontal. We further assume that all motion of the string is purely one-dimensional (in the vertical direction), that any displacements from equilibrium of the string and any angles between the string and the horizontal are small. (Any readers who are skeptical about these assumptions should look closely at a plucked guitar string - the displacement is virtually imperceptible. In fact, a reasonable scale for Figure 25 would be a meter along the horizontal axis, and *a millimeter* in the vertical!) To emphasize that the motion can vary with both time and position, we shall denote the vertical displacement from equilibrium by the symbol $u(x,t)$.

We shall also assume that the original cross-sectional area and density of the string
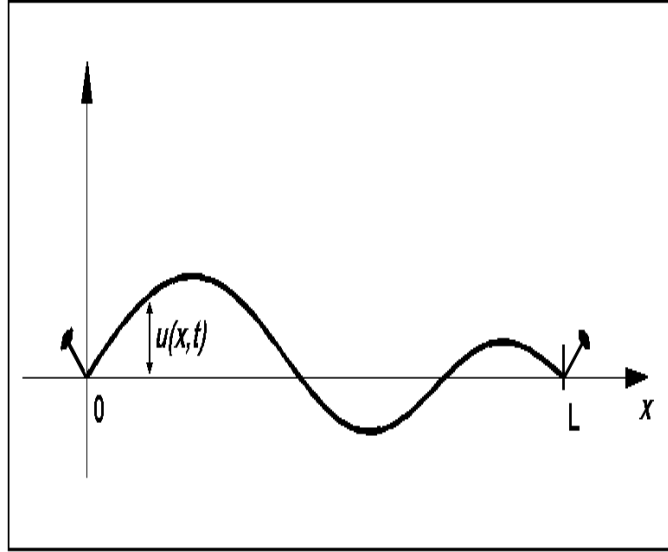
Figure 25: An Elastic String

were uniform (constant), and that any changes in cross-section or density resulting from stretching of the string during its motion are negligible. (The latter part of this assumption ensures consistency with our earlier assumption of extremely small vertical displacements.) Therefore the linear density (mass per unit length) of the string will remain constant. We shall use the symbol ρ to represent this density. Finally, we assume that there is a uniform tension, which we shall denote by τ, throughout the length of the string.

By now, you may feel as if it seems like we are making a lot of assumptions. In a sense, perhaps we are. However, as we observed above, there is effectively one "acid test" for whether these assumptions are in fact reasonable. This test is whether or not the values of the mathematical solutions to the equations that result from the model assumptions agree with the *actually observed and measured* physical motion of such strings. If these values agree well with "reality," then we accept the model as being a relatively accurate representation of reality. If, however, the model leads to solutions that do not agree with observed phenomena, then we would have to go back to our fundamental assumptions and ask - "What have we missed? What have we ignored? What have we neglected?" - because apparently something important is going on which we have not captured.

Once we have identified the basic assumptions in this, or any similar problem, the next step is to set up the complete description of all the interior forces at work. To do this, we consider an idealized segment located at some arbitrary position in the interior of the string (Figure 26). The segment is assumed to be of a "small" length, denoted by Δx, and located between x and $x + \Delta x$.

Assuming some basic background in physics on the reader's part, we now proceed to identify the possible forces that could act on this small segment.

Since we have assumed there is no motion except in the vertical direction, we will not analyze any forces in the x or z direction, and concentrate on only those in the vertical. Figure 26 clearly seems to indicate that there are at most three possible forces acting on

71
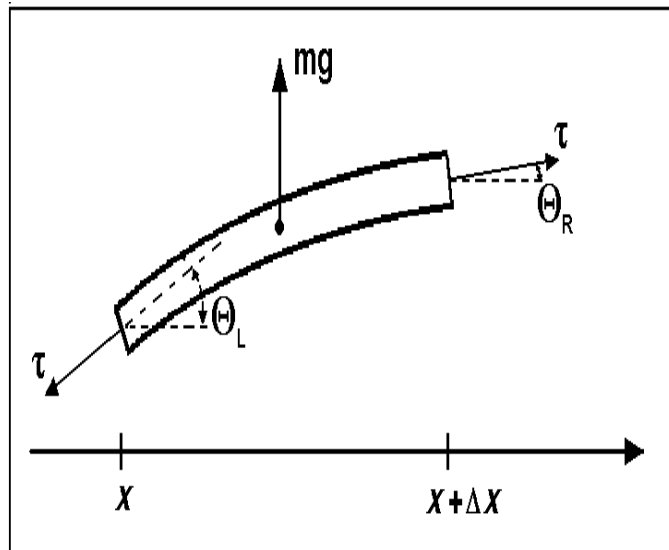
Figure 26: A Small Segment of the String

this segment - a tension force exerted on the left end by the neighboring segment of string there, another similar tension force exerted on the right end, and then, possibly, some kind of external (commonly called body) force. As the figure implies, this body force, if any, effectively acts through the center of mass of the segment, which, since we have assumed the string has a uniform density, coincides with its geometrical center. (Typical body forces arise from gravity (which we have already effectively neglected here by assuming the string lies flat at equilibrium), magnetic fields, damping due to air friction, etc.) In our model, we shall simply represent acceleration due to this body force by $g(x + \frac{1}{2}\Delta x, t)$. Note that this representation emphasized not only the fact that the force acts through the center of mass, but also that such forces generally vary with both time and position. (Some body forces in fact, may also vary with the displacement or vertical velocity of the string, i.e. have the form $g(x + \frac{1}{2}\Delta x, t, u, \frac{\partial u}{\partial t})$.) However, since we have also assumed a constant linear density and no effective lengthening of the string, the mass of our small segment is $\rho \Delta x$, and so the resulting vertical force on the segment due to body forces is represented by

$$\rho g(x + \tfrac{1}{2}\Delta x, t)\Delta x \quad .$$

We turn next to the tension forces exerted on the segment by its neighboring segments. We represent these forces, which must lie along the tangents to the string, by two vectors of equal magnitude, tangent at the left and right ends of the segment. (The fact that the magnitudes are equal is not accidental, but is a direct consequence of our earlier model assumptions that the internal tension is uniform, that there is no net horizontal motion and that the angles between the string and the horizontal are small.) However, if the string has any curvature, the angles at the right and left ends, while small, will generally not be equal. Therefore, even though the magnitudes of the two tension forces are equal, their directions will generally not be identical, and this will result in a *net* vertical force component due to tension. The normal rules for decomposing vectors into their components give the vertical

72

components of the tension at the left and right hand ends, respectively, as

$$-\tau \sin(\theta_L) \quad \text{and} \quad \tau \sin(\theta_R) \qquad ,$$

and therefore their sum,

$$\tau \sin(\theta_R) - \tau \sin(\theta_L) \quad ,$$

represents the net vertical tension on this segment. (You should understand that since these are components of *vectors*, the negative sign associated with the left hand component is mandatory to ensure consistency with our convention that the upward vertical direction is positive.)

But now we return to our assumptions that the displacements and angles (θ_L and θ_R) are small. Therefore, mathematically,

$$\theta_i \approx \sin(\theta_i) \approx \tan(\theta_i) , \quad i = L, R \quad ,$$

and we can now express the net vertical force due to tension as essentially equal to

$$\tau \tan(\theta_R) - \tau \tan(\theta_L) \quad .$$

However, a fundamental result from the calculus is that the slope of the tangent line to a curve is also precisely the tangent of the angle between the curve and the horizontal axis, and is given by the value of the first derivative of the equation for the curve at that point. Therefore,

$$\tan(\theta_R) = \frac{\partial u}{\partial x}(x + \Delta x, t) \quad \text{and} \quad \tan(\theta_L) = \frac{\partial u}{\partial x}(x, t) \quad .$$

(Note we must use partial derivatives here, since $u(x, t)$ is a function of two variables. Further note that the *slope* is given by the partial with respect to x, while the vertical *velocity* would be given by the partial with respect to t.)

Combining all of these components yields that the *net* vertical force acting on the small segment of the string is

$$\tau \left\{ \frac{\partial u}{\partial x}(x + \Delta x, t) - \frac{\partial u}{\partial x}(x, t) \right\} + \rho g(x + \tfrac{1}{2}\Delta x, t) \, \Delta x$$

But, if this net force is not zero, then, according to Newton's second law of motion, the segment will experience a proportional vertical acceleration, i.e. $F = ma$. **This** is our conservation law! (You may recall from your basic college physics courses that Newton's second law is equivalent to conservation of momentum.) Since we already know the mass of this segment ($\rho \Delta x$), and since the vertical acceleration (computed at the center of mass of the segment) is nothing more that the second derivative of position with respect to time, i.e.,

$$\frac{\partial^2 u}{\partial t^2}(x + \tfrac{1}{2}\Delta x, t) \quad ,$$

then Newton's second law reduces to

$$\rho \frac{\partial^2 u}{\partial t^2}(x + \tfrac{1}{2}\Delta x, t)\Delta x = \tau \left\{ \frac{\partial u}{\partial x}(x + \Delta x, t) - \frac{\partial u}{\partial x}(x, t) \right\} + \rho g(x + \tfrac{1}{2}\Delta x, t) \, \Delta x \quad .$$

This, of course, being a mathematics text and written by a mathematician, there is no way we can not succumb to the temptation to divide by Δx, yielding

$$\rho \frac{\partial^2 u}{\partial t^2}(x + \tfrac{1}{2}\Delta x, t) = \tau \left\{ \frac{\frac{\partial u}{\partial x}(x + \Delta x, t) - \frac{\partial u}{\partial x}(x, t)}{\Delta x} \right\} + \rho g(x + \tfrac{1}{2}\Delta x, t) \quad,$$

and then let Δx approach zero. (After all, isn't this what every mathematics book does?) Thus we have

$$\rho \frac{\partial^2 u}{\partial t^2}(x, t) = \tau \lim_{\Delta x \to 0} \left[\frac{\frac{\partial u}{\partial x}(x + \Delta x, t) - \frac{\partial u}{\partial x}(x, t)}{\Delta x} \right] + \rho g(x, t) \quad.$$

But now, wait! That limit should look familiar. It's nothing more, basically, than the definition of a partial derivative - specifically the partial derivative of $\partial u/\partial x$ with respect to x, i.e. exactly what we mean by

$$\frac{\partial^2 u}{\partial x^2}(x, t) \quad.$$

Putting all of these pieces together, we have,

$$\rho \frac{\partial^2 u}{\partial t^2}(x, t) = \tau \frac{\partial^2 u}{\partial x^2}(x, t) + \rho g(x, t) \quad,$$

the writing of which we shall now simplify by dividing through by ρ and dropping the explicit reference to x and t in u, to yield

$$\frac{\partial^2 u}{\partial t^2} = \frac{\tau}{\rho} \frac{\partial^2 u}{\partial x^2} + g(x, t) \quad. \tag{3.2.1}$$

There is one final simplification we shall use. Since τ and ρ are positive constants, then so is their ratio, τ/ρ, and we therefore shall replace the ratio by the single value c^2. The square is deliberately used to unequivocally emphasize the positive character of the ratio. (It will also make the algebra a lot simpler later on!) Thus, we finally have the *one-dimensional wave equation*:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + g(x, t) \quad. \tag{3.2.2}$$

There is one further, minor simplification that we shall mention. Partial derivatives are frequently represented by subscripts, instead of with the ratio notation, except in cases where subscripts would be ambiguous. Thus, for example, many texts will write

$$u_x \quad \text{instead of} \quad \frac{\partial u}{\partial x} \,, \quad \text{etc.}$$

In this notation, which we shall frequently use, the one-dimensional wave equation would then be written

$$u_{tt} = c^2 u_{xx} + g(x, t) \quad. \tag{3.2.3}$$

PROBLEMS

1. Show that if a uniform, thin, tightly stretched elastic string is acted upon by no forces other than internal tension and an external *air resistance* proportional to the vertical velocity, then Newton's second law leads to a partial differential equation of the form:

$$\frac{\partial^2 u}{\partial t^2} + \kappa_d \frac{\partial u}{\partial t} = \frac{\tau}{\rho} \frac{\partial^2 u}{\partial x^2} \quad ,$$

where κ_d is some positive constant of proportionality.

2. Show that if a uniform, thin, tightly stretched elastic string is acted upon by no forces other than internal tension and an external *spring-like restoring force* proportional to the vertical displacement, then Newton's second law leads to a partial differential equation of the form:

$$\frac{\partial^2 u}{\partial t^2} = \frac{\tau}{\rho} \frac{\partial^2 u}{\partial x^2} - \kappa_s u \quad ,$$

where κ_s is some positive constant of proportionality.

3.3　Boundary Conditions

Our just-completed derivation of the one-dimensional wave equation includes, however, one crucial restriction. It arose from considering an internal segment of the string, and really expresses only the physics that are occurring *inside* the string. But guitar strings don't stretch on forever - they're usually only two or three feet long at best! They have ends! Physically, these ends, which we have denoted as $x = 0$ and $x = L$ in our model, represent locations where the physics of the process change, i.e. *locations where the internal physics no longer validly describe what is occurring.* Mathematically, at the ends and beyond, the basic governing partial differential equation ceases to hold true. (In other words, to be precisely correct, we really should have written (3.2.3) as

$$u_{tt} = c^2 u_{xx} + g(x,t) , \qquad 0 < x < L \quad .$$

Therefore, in order to completely specify our model, we must also derive mathematical equations which correctly express the physics that are occurring at these ends.

The derivation of the mathematical conditions at the ends turns out to be a bit more involved than was the derivation of the equation for the interior physics, primarily because, instead of the single basic internal process just studied, there are a fairly large number of *different* physical processes which may occur at the ends. One in particular, however, almost always occurs at the ends of a real guitar string - the ends are rigidly attached to some immovable object, e.g. the bridge on the body of a guitar or a turnable peg on the neck (Figure 25). The key feature here is that where the string meets the fixed object, the string doesn't move, i.e. there is no displacement at the ends. Since we've already decided to represent the displacement at *any* generic point by $u(x,t)$, the only sensible expressions for the displacements at the left and right-hand ends of the strings are $u(0,t)$ and $u(L,t)$, respectively. But if there is no displacement at these ends, then both of these quantities must, of course, equal zero. Thus, the mathematical formulation of the physics at fixed ends must be

$$\begin{aligned} u(0,t) &= 0 \quad , \\ u(L,t) &= 0 \quad . \end{aligned} \qquad (3.3.4)$$

Mathematically, equations of the type of (3.3.4) are called *boundary conditions*. They differ from the more familiar *initial conditions* you should have encountered in ordinary differential equations, e.g.

$$\begin{aligned} y'' + 2y' &+ y = 0 \\ y(0) &= 0 \\ y'(0) &= 1 \end{aligned}$$

in that the initial conditions of an ordinary differential equation, are **all** specified at the *same* value of the independent variable (in this case $x = 0$). The conditions in (3.3.4), in contrast, are specified at two *different* values - $x = 0$ and $x = L$ - values which physically correspond to the boundaries of the region of interest. As we shall see shortly, a number of possible different mathematical general boundary conditions exist for the wave equation. Those with the form of (3.3.4) are commonly referred to as *fixed end* conditions, *Dirichlet*[7]

---

[7]Johann Peter Gustav Lejeune Dirichlet,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Dirichlet.html

conditions, or boundary conditions of the *first kind.*

As alluded to above, we could specify other physics at the ends of the string than simply making them immobile. As a practical matter, such other conditions are never seen at the ends of real guitars strings. However, there are a number of other intriguing physical models whose internal physics are also described by the wave equation that do lead to the same mathematical boundary conditions that arise from what would be unusual boundary physics for a guitar string. We shall examine some of these other conditions now, and determine the kind of mathematical boundary conditions they produce.

One such alternative set of physics is the so-called *free end* condition. A model for this end condition assumes that the string terminates in a small, effectively weightless ring that can slide up and down, without any friction (drag), on a vertical pole which is also located at the end of the string. We shall further assume that for the moment this condition occurs at the right-hand end of the string. In this situation, clearly, the right-hand end of the string can move and therefore a boundary condition such as $u(L, t) = 0$ would be silly. (And furthermore wrong!) To determine the correct boundary condition for these physics, we will consider a very enlarged view of this right-hand end, and adopt a common approach in physics - the two observer approach. In this method, we assume that one observer stands just outside the string, observing the ring and pole and measuring the forces there, while the second measures the forces just inside the end of string. The key to this approach is the realization that since these two observers are in fact observing the *same phenomenon,* their measurements must therefore be identical. This realization becomes the effective conservation principle from which we can derive our boundary condition equation.
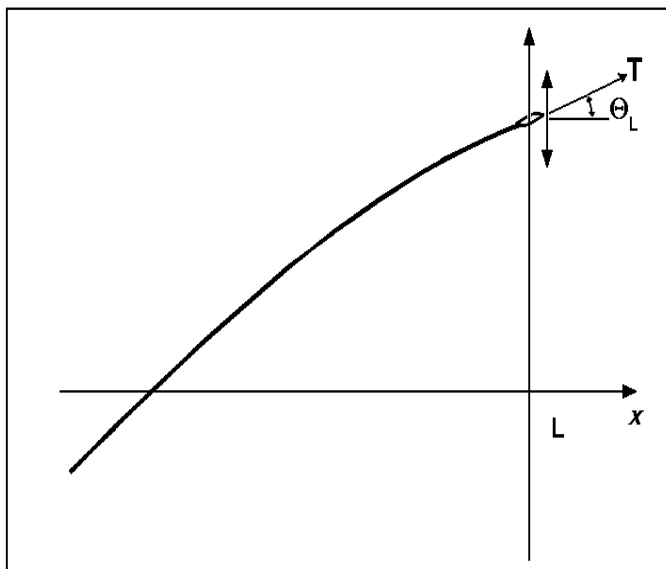
Figure 27: Free End Conditions

Now consider the outside observer in this case. He sees the ring sliding, without friction, up and down along the pole. But without friction, the pole cannot exert any drag on the ring. Therefore the pole exerts *no net vertical force* on the ring, and since any drag on the

ring would be instantly transferred to the end of the string, the outside observer concludes that there is zero vertical force on the string. (There is, of course, a significant *horizontal* force. But that is beside the point!)

The inside observer, by contrast, has no knowledge of exactly what is going on outside the end of the string, but can only measure the tension force and the net angle between the right-hand end of the string and the horizontal. He measures the net vertical force exerted on this end, exactly as in our earlier derivation (Figure 26), as equal to the magnitude of the vertical component of the tension there, i.e. as

$$\tau \sin(\theta_L) \quad .$$

(In Figure 26, we denoted the angle at the right-hand end of the interior segment θ_2. Since the angle here is at the end of the string ($x = L$), we choose to call it θ_L instead.) But if we are to reconcile the observations of these two observers, then the net vertical force measured by the inside observer must equal that measured by the outside observer, i.e. zero. Therefore

$$\tau \sin(\theta_L) = 0 \quad \Longrightarrow \quad \sin(\theta_L) = 0 \quad \Longrightarrow \quad \theta_L = 0 \quad .$$

In other words, the string must, at the exact location of the right-hand end, be *horizontal* - i.e. parallel to the x-axis. However, if this is the case, the slope there must also be zero. Hence, we are led to the equation

$$\frac{\partial u}{\partial x}(L, t) = 0 \quad , \tag{3.3.5}$$

or, in subscript notation,

$$u_x(L, t) = 0 \quad ,$$

and we have our boundary condition at this end for this model. A similar argument shows that a similar frictionless ring and pole placed at the left-hand end of the string result in the boundary condition

$$\frac{\partial u}{\partial x}(0, t) = 0 \quad , \tag{3.3.6}$$

or

$$u_x(0, t) = 0 \quad .$$

Boundary conditions with this form are commonly referred to as *free end* conditions, *Neumann* conditions[8], or boundary conditions of the *second kind*. (We cannot emphasize too strongly at this point that there is no requirement to place the same kind of pole and ring apparatus at the left end. We could equally well have left the left end fixed. The physics at the left and right ends are **totally independent**.)

A third type of boundary physics, which also is essentially never encountered in real guitar strings, but which leads to a mathematical condition not uncommon in other models, occurs if we modify the slip ring arrangement in the previous example by attaching a small vertical spring which *will* exert a force at the end of the string whenever the spring is either compressed or stretched (Figure 27). In this case, we assume that the position of the spring
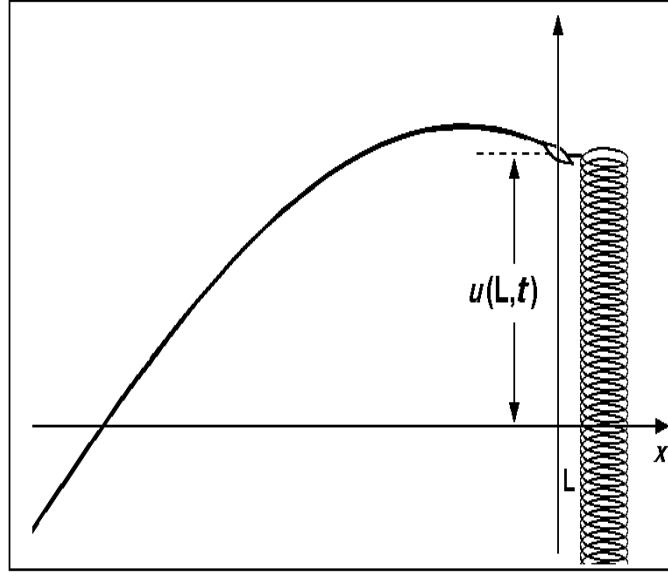
Figure 28: Mixed End Conditions

has already been adjusted so that at equilibrium (when the string is horizontal) the spring is neither compressed nor stretched.

But now consider how our inside and outside observers view this newest situation! The outside observer will observe that the spring has been stretched, and wants to return toward its equilibrium position. Therefore the spring will exert a force on the ring, and through the ring onto the string, proportional to the amount by which it has been stretched. The outside observer will then measure this net force exerted on the end of the string by the spring as equal to

$$-k_s u(L, t) \quad ,$$

where k_s is referred to as the spring constant. (Note that because the quantities involved are forces, the negative sign here explicitly ensures the correct orientation of the spring force, i.e. that the spring exerts a force opposite to the direction in which it has been displaced.) The inside observer, again, has no knowledge of the outside mechanism - he can only measure the net force exerted on the end of the string in terms of the vertical component of the tension he sees at that end. He will therefore (again using the small-angle assumption) compute the net vertical force being exerted at the end as:

$$\tau \sin(\theta_L) \approx \tau \tan(\theta_L) = \tau \frac{\partial u}{\partial x}(L, t) \quad .$$

(Here the *positive* sign is necessary to retain the proper sense of the vector components involved!) However, since these observers are in fact observing the same force, their calculated forces must yield the same result, i.e.

$$-k_s u(L, t) = \tau \frac{\partial u}{\partial x}(L, t) \quad ,$$

---

[8]Carl Gottfried Neumann,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Neumann_Carl.html

or, after collecting both terms on the same side of the equality

$$k_s u(L, t) + \tau \frac{\partial u}{\partial x}(L, t) = 0 \quad , \tag{3.3.7}$$

or

$$k_s u(L, t) + \tau u_x(L, t) = 0 \quad .$$

This type of boundary condition is commonly referred to as the *mixed end*, *Robin* condition, or boundary condition of the *third kind*. If we were to conduct a similar analysis at the left-hand end of the string, we would find that this spring and slip-ring arrangement at that end, would lead to the mathematical boundary condition

$$k_s u(0, t) - \tau \frac{\partial u}{\partial x}(0, t) = 0 \quad , \tag{3.3.8}$$

or

$$k_s u(0, t) - \tau u_x(0, t) = 0 \quad .$$

It is important that you realize that, for any particular string, or for any other model which may be described by the wave equation, at most only one of the above boundary conditions can hold at any one boundary - i.e. they are mutually exclusive. (This should be clear from physical principles!) However, as we noted before, the conditions at different ends are independent. That is, it would be totally permissible to have a string fixed at the left and free at the right, i.e. with boundary conditions

$$\begin{aligned} u(0, t) &= 0 \quad , \\ u_x(L, t) &= 0 \quad . \end{aligned}$$

Equally well, we might attach a spring and slip ring at the left, and fix the right-hand end, yielding

$$\begin{aligned} k_s u(0, t) - \tau u_x(0, t) &= 0 \quad , \\ u(L, t) &= 0 \quad , \end{aligned}$$

or, apply any other appropriate combination. There is, however, one fundamental unifying element to the above three types of conditions - they are all *linear* and *homogeneous*. That is, they are all just special cases of the general form

$$\begin{aligned} \alpha_1 u(0, t) - \beta_1 u_x(0, t) &= 0 \quad , \\ \alpha_2 u(L, t) + \beta_2 u_x(L, t) &= 0 \quad , \end{aligned}$$

for specific values of the α_i and β_i.

PROBLEMS

1. Having physically correct algebraic signs in boundary conditions can be critical. Show, both mathematically any physically, that the following boundary conditions:
a. $u(L, t) - u_x(L, t) = 0$
b. $u(0, t) + u_x(0, t) = 0$
are **not** physically realistic.

2. Show that if a uniform, thin, tightly stretched elastic string is attached at its right-hand boundary to a slip-ring on a pole that is **not** frictionless, but in which the friction is proportional to the vertical velocity along the pole, then the boundary condition at that point becomes

$$\tau \frac{\partial u}{\partial x} + \kappa_d \frac{\partial u}{\partial t} = 0 \quad,$$

where κ_d is some positive constant of proportionality.

3.4 Initial Conditions

Thus far, we have as the most general formulation of the one-dimensional wave equation

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} + g(x,t) \quad , \\
\alpha_1 u(0,t) \ - \ \beta_1 u_x(0,t) &= 0 \ , \\
\alpha_2 u(L,t) + \beta_2 u_x(L,t) &= 0 \quad .
\end{aligned}
$$

These equations, however, are still not sufficient to *uniquely* specify, either mathematically or physically, the motion of any elastic string. The reason is really quite simple - there is still some important information missing. What is missing is the information about how the string "starts out," i.e. what were its initial conditions. (After all, this is a differential equation with two time derivatives.) Based on experience with ordinary differential equations, we then expect that this model will require *two* initial conditions - one on u and another on u_t. Physically, this is equivalent to saying that, in order to have a unique solution, we need to know both the initial *position* and *velocity*. Mathematically, these two conditions would be written

$$
\begin{aligned}
u(x,0) &= f(x) \ , \\
u_t(x,0) &= g(x) \ ,
\end{aligned}
\tag{3.4.9}
$$

where $f(x)$ and $g(x)$ may be virtually any two functions. Thus, finally, the complete specification of a one-dimensional wave equation becomes

$$
\begin{aligned}
u_{tt} = c^2 u_{xx} \ &+ \ g(x,t) \ , \\
\alpha_1 u(0,t) - \beta_1 u_x(0,t) &= \ 0 \\
\alpha_2 u(L,t) + \beta_2 u_x(L,t) &= \ 0 \\
u(x,0) &= \ f(x) \\
u_t(x,0) &= \ g(x)
\end{aligned}
\tag{3.4.10}
$$

This complete problem - the partial differential equation plus appropriate boundary and initial conditions - is commonly referred to as a *boundary value* problem.

3.5 Introduction to the Solution of the Wave Equation

We start our investigation into the application of Fourier series techniques in the solution of partial differential equations with the one-dimensional wave equation derived in the previous section. We chose to begin with this equation for two reasons:

(1) It obviously represents a "real world" model.

(2) Partial differential equations, by definition, involve more than one independent variable. Clearly, the simplest ones should involve *exactly two*. Of the partial differential equations with exactly two independent variables, the wave equation is the easiest in which to observe and interpret our dual time domain/frequency domain approach to problems.

But how should we go about solving a partial differential equation, if we've never solved one before? The answer includes a fundamental insight about mathematics, for one of the great values of mathematics is its power to generalize, i.e. to solve new problems by extending ideas and methods that have already been shown to work in other, similar situations.

Therefore, we shall try, as much as possible, to apply to the study of partial differential equations lessons learned in the study of ordinary differential equations, starting with the observation that, as written, the one dimensional wave equation (3.4.10) conforms to our idea of a *linear, non-homogeneous* problem, since the term $g(x,t)$ do not contain the unknown u or any of its derivatives. To further motivate the general direction we shall take, recall that when first confronting the general second order linear, non-homogeneous ordinary differential equation:

$$P(x)y'' + Q(x)y' + R(x)y = G(x) \quad ,$$

one does not "jump in" and try to immediately develop the complete solution to this problem. Instead, they first study the nature of solutions to a simpler case - the homogeneous, constant coefficient equation:

$$py'' + qy' + ry = 0 \quad .$$

A similar approach then seems logical when confronting our first partial differential equation. Therefore, we shall start by adding to the numerous assumptions that we have already made the assumption that any body forces in (3.4.10) are negligible. (Again, the reader should consider whether such an assumption is reasonable for "real" guitar strings.)

Thus, to summarize, the first case we will solve is:

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \ , & 0 < x < L \ , \quad 0 < t \ , \\
\alpha_1 u(0,t) &- \beta_1 u_x(0,t) = 0 \ , & 0 < t \\
\alpha_2 u(L,t) &+ \beta_2 u_x(L,t) = 0 \ , & 0 < t \\
u(x,0) &= f(x) \ , & 0 < x < L \\
u_t(x,0) &= g(x) \ , & 0 < x < L
\end{aligned}
\tag{3.5.11}
$$

(Observe that this is about the simplest, yet reasonably general case we could deal with. For example, if we had also had homogeneous initial conditions, i.e. if $u(x,0) = 0$ and $u_t(x,0) = 0$, then $u(x,t) \equiv 0$ would satisfy the problem. In fact, we could show this would be the *only* solution! We call this the trivial case, and it is clearly of no practical interest. We have also, for emphasis, specifically identified the domains where the various equations are valid. Henceforth, we shall simply take these domains as implicitly given.)

The homogenous problem (3.5.11) is, however, still somewhat general in that we have made no particular restriction on the α_i and β_i. We could, conceivably, proceed with a general approach at this point, keeping α_i and β_i arbitrary. But we shall not do this. Instead, we shall first consider several special cases, in which we prescribe specific values for these constants and try to arrive at the complete solution for that particular problem. Finally, once we have developed a sufficient "feel" from these examples, we shall compare the different problems to try to identify the common features that arise irrespective of the choice of α_i and β_i, and then try to deduce the general theory. (In a sense, what we will be trying to do is analogous to separating the "rules of the game" (theory) in a sport from the particular teams and players (functions and equations) that may arise in a specific case.)

PROBLEMS

1. Briefly describe, in a few sentences, a physical model for each of the following boundary value problems:

   a.
   $$u_{tt} = 4u_{xx}$$
   $$u(0,t) = u_x(3,t) = 0$$
   $$u(x,0) = \begin{cases} 2x & ,0 < x < 1 \\ 0 & ,1 \le x < 3 \end{cases}$$
   $$u_t(x,0) = 0$$

   b.
   $$u_{tt} = u_{xx}$$
   $$u(0,t) = u(1,t) = 0$$
   $$u(x,0) = 0$$
   $$u_t(x,0) = 1$$

   c.
   $$u_{tt} = 9u_{xx}$$
   $$u_x(0,t) = u_x(2,t) = 0$$
   $$u(x,0) = x$$
   $$u_t(x,0) = 0$$

   d.
   $$u_{tt} = u_{xx}$$
   $$u(0,t) = 0$$
   $$u(3,t) + 2u_x(3,t) = 0$$
   $$u(x,0) = \begin{cases} 2x & ,0 < x < 1 \\ 0 & ,1 \le x < 3 \end{cases}$$
   $$u_t(x,0) = 10$$

3.6   The Fixed End Condition String

As our first special case, we shall start with the one which is physically the most common. This, as we discussed earlier, occurs when the ends of the string are rigidly attached to some immovable object - i.e. the fixed end, or Dirichlet boundary conditions. In terms of the homogeneous problem (3.5.11) this coincides to $\alpha_1 = \alpha_2 = 1$, and $\beta_1 = \beta_2 = 0$, or $u(x,t)$ satisfies:

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u(0,t) &= 0 \\
u(L,t) &= 0 \\
u(x,0) &= f(x) \\
u_t(x,0) &= g(x)
\end{aligned}
\tag{3.6.12}
$$

Now, let's start by reflecting again on our experience with second order, linear ordinary differential equations. The first step in these problems was to identify all of the linearly independent homogeneous solutions. In the case of constant coefficient equations, this identification was accomplished primarily by assuming a certain solution form ($y = e^{rx}$), then substituting that form into the equation, and finally determining the specific values of r for which non-trivial solutions exist. Once the necessary number of linearly independent solutions corresponding to these values of r had been constructed, then the general solution was formed by taking a linear combination of all the linearly independent solutions i.e. by multiplying each linearly independent solution by an arbitrary constant and then summing them together. At this point the initial conditions were considered, in order to determine the values of the constants in the solution. We will try a somewhat analogous approach here.

If we are then to try a similar procedure here, our first step should be to identify what constitutes the *homogenous problem* in the case of partial differential equations. Considering the fixed-end problem (3.5.11), and recalling that in solving for the linearly independent homogenous solutions in ordinary differential equations, the initial conditions do not enter until the very end of the problem, then, according to our normal concept of homogenous, the homogeneous problem without initial conditions for (3.6.12) appears to be

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u(0,t) &= 0, \\
u(L,t) &= 0
\end{aligned}
\tag{3.6.13}
$$

This problem clearly has at least the trivial solution, $u(x,t) = 0$. (As does any linear homogeneous problem! After all, $y \equiv 0$ solved $ay'' + by' + cy = 0$.) But trivial solutions are inherently uninteresting. The *important* question is does (3.6.13) have any *non-trivial* solutions? And, if so, how do we find them? And how many of them are linearly independent? (After all, any non-trivial solutions can only be determined up to an *arbitrary multiplicative constant* anyway.)

But wait! Observe that the homogeneous partial differential equation in (3.6.13) possesses a natural "splitting," in that the left-hand side involves only an operation on the time behavior of $u(x,t)$ and the right-hand side involves only an operation on the spatial behavior of $u(x,t)$. This suggests its solutions might also have a splitting of the time behavior from

the spatial behavior. (Although this is by no means guaranteed!) So we shall first look for solutions which have the form:

$$u(x,t) = X(x)T(t) \quad . \tag{3.6.14}$$

(This is commonly called the *Separation of Variables* assumption.) It's important at this point to emphasize that this form is only an *assumed* solution, just as $y = e^{rx}$ was an assumed solution to $py'' + qy' + ry = 0$. There is of course no guarantee that *any* of the homogeneous solutions to the partial differential equation need to be this product form, and certainly no *a priori* assurance that *all* of them will be. (For example, not all of the solutions to $y'' + 2y' + y = 0$ are "pure" exponentials.) There is only one way to tell whether or not (3.6.14) will really yield valid non-trivial solutions, and that is to substitute it into (3.6.13) and see what follows. That is precisely what we shall do next.

In order to substitute (3.6.14) into (3.6.13), we first need to compute the various indicated derivatives. But, by the basic rules for partial differentiation:

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2}{\partial t^2}[X(x)T(t)] = X(x)T''(t) \quad ,$$

and,

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2}{\partial x^2}[X(x)T(t)] = X''(x)T(t) \quad .$$

Thus, upon substitution, the partial differential equation in (3.6.13) becomes:

$$X(x)T''(t) = c^2 X''(x)T(t) \quad ,$$

or, dividing both sides by the product $c^2 XT$,

$$\frac{T''(t)}{c^2 T(t)} = \frac{X''(x)}{X(x)} \quad . \tag{3.6.15}$$

But consider the implication of this last equation *carefully*. The left-hand side contains only terms which depend on time, hence it is completely independent of x. The right-hand side contains only terms in x, and so is independent of time. But both sides are *equal*. Thus, each side of this equation is independent of both time and x, and therefore each side must just be **constant**.

We now denote this constant by $-\lambda$. (We shall see in a moment there was an ulterior motive, involving essentially facts which we have not yet developed here, for including the explicit minus sign. There is really no basis in the information we have developed up to this time to expect one algebraic sign over of another.) Thus, the equation becomes:

$$\frac{T''(t)}{c^2 T(t)} = \frac{X''(x)}{X(x)} = -\lambda \quad ,$$

or

$$\begin{aligned} T''(t) + \lambda c^2 T(t) &= 0 \ , \\ X''(x) + \lambda X(x) &= 0 \ . \end{aligned} \tag{3.6.16}$$

You should appreciate that at this point there will not necessarily be only a single value of λ for which this approach works. (After all, there are generally two values of r in the second-order ordinary differential equation.) For now, any value for λ should seem as good as any other. (Actually, this is a slight overstatement. From physical arguments, it seems unlikely that λ in (3.6.16) could be negative, since then the general solution to the equation for T(t) would include a growing exponential in time, and such a solution is not physically reasonable for normal guitar strings. But mathematically, we cannot yet discard any possible values for λ.)

Actually, the fact that we can't say anything yet about the value(s) of λ should not be really surprising, since we have not utilized *all* the information contained in the homogeneous problem (3.6.13). Specifically, we have not yet used the boundary conditions. But if our product solution (3.6.14) is to satisfy the entire homogenous problem, it must, of necessity, satisfy not only the partial differential equation, but the boundary conditions in (3.6.13) as well. So, we substitute the product (3.6.14) into the first (left-hand end) condition, yielding

$$u(0,t) = X(0)T(t) = 0 \quad . \tag{3.6.17}$$

But observe that $X(0)$, being the value of a function at one point, is just a number, i.e. a constant. But the product in (3.6.17) is identically zero. Therefore, unless $X(0) = 0$, then $T(t)$ must be identically zero. But if $T(t)$ were identically zero, then $u(x,t) = X(x)T(t)$ would also be identically zero (trivial), and therefore of no interest to us - we already know the trivial solution satisfies the homogeneous problem. What we care about is *non-trivial* solutions! Thus, to retain any hope of finding non-trivial solutions, we must require that

$$X(0) = 0 \ .$$

Similarly, at $x = L$, by an essentially identical argument, we must also have:

$$u(L,t) = X(L)T(t) = 0 \Rightarrow X(L) = 0 \ .$$

Now, if we collect all the information we have so far deduced about $X(x)$ and $T(t)$ as defined by (3.6.14) , we have:

$$T'' + \lambda c^2 T = 0 \qquad \begin{aligned} X'' &+ \lambda X = 0 \\ X(0) &= 0 \\ X(L) &= 0 \quad . \end{aligned} \tag{3.6.18}$$

Observe the equation for $T(t)$ still has a solution for any value of λ, since there are no conditions on $T(t)$ other than the differential equation. But the same is clearly no longer true for $X(x)$. After all, the equation for $X(x)$

$$\begin{aligned} X'' &+ \lambda X = 0 \\ X(0) &= 0 \\ X(L) &= 0 \quad . \end{aligned} \tag{3.6.19}$$

is homogeneous and now has two homogeneous boundary conditions. If these were initial conditions, they would be enough to guarantee an unique solution for $X(x)$, which clearly

would have to be $X(x) \equiv 0$. So can we find any λ for which there are non-trivial solutions as well?

Well, suppose for the moment we can and consider the situation if λ is negative. We can conveniently represent this by writing

$$\lambda = -\zeta^2 \ ,$$

where ζ is some *non-zero* real number. With this substitution, $X(x)$ satisfies:

$$X'' - \zeta^2 X = 0 \ ,$$

the general solution to which is

$$X(x) = c_1 e^{\zeta x} + c_2 e^{-\zeta x} \ .$$

But now we have to determine the constants c_1 and c_2 so that the boundary conditions are satisfied. When we do this at $x = 0$, we have

$$X(0) = c_1 + c_2 = 0 \qquad \Longrightarrow \qquad c_2 = -c_1 \qquad \Longrightarrow \qquad X(x) = c_1 \left(e^{\zeta x} - e^{-\zeta x} \right) \ .$$

Then the condition at $x = L$ becomes,

$$X(L) = c_1 \left(e^{\zeta L} - e^{-\zeta L} \right) = 0 \ .$$

However, it is easily shown that $\left(e^{\zeta L} - e^{-\zeta L} \right)$ cannot be zero if $L > 0$ and $\zeta \neq 0$. Therefore we must have $c_1 = 0$, which means

$$X(x) \equiv 0 \ ,$$

or we have *only* the trivial solution to

$$\begin{aligned} X'' + \lambda X &= 0 \\ X(0) = X(L) &= 0 \end{aligned}$$

when $\lambda < 0$. (In view of our earlier comment about not wanting a growing exponential in the solutions for $T(t)$, this is somewhat of a relief, in that the mathematics confirms our physical intuition - a key attribute if we are to consider the model to be valid.) Since we are not interested in trivial solutions, we drop $\lambda < 0$ from any further mention for this problem.

Well, if λ is not negative, then maybe λ is zero. But, in this case, the differential equation for $X(x)$ becomes

$$X'' = 0 \ ,$$

and the general solution to this is

$$X(x) = d_1 + d_2 x \ .$$

Now, we again try to satisfy the boundary conditions with this solution. When $x = 0$, we have

$$X(0) = d_1 = 0 \qquad \Longrightarrow \qquad X(x) = d_2 x \ ,$$

and then when $x = L$,

$$X(L) = d_2 L = 0 .$$

Now it is easily seen that the only possible way this can happen is if $d_2 = 0$ and again we have only the trivial solution. So $\lambda = 0$ must also be discarded for this problem. (We would note that this should have been clear geometrically, since the differential equation and boundary conditions for $\lambda = 0$ imply that $X(x)$ is a straight line passing through the x-axis at $x = 0$ and $x = L$, and $X(x) \equiv 0$ is obviously the only solution to these conditions!)

We are now about at the end of our rope. Our only remaining hope of finding non-trivial solutions in the form (3.6.14) rests with the possibility that λ is positive. In this case, we could write,

$$\lambda = \xi^2$$

where ξ is some non-zero real number. Then $X(x)$ would satisfy the differential equation

$$X'' + \xi^2 X = 0 ,$$

whose general solution is

$$X(x) = a_1 \sin(\xi x) + a_2 \cos(\xi x) \ .$$

Again we must try to pick a_1 and a_2 so that the boundary conditions are satisfied. This yields first, at $x = 0$

$$X(0) = a_2 = 0 \qquad \Longrightarrow \qquad X(x) = a_1 \sin(\xi x) \ ,$$

and then, at $x = L$,

$$X(L) = a_1 \sin(\xi L) = 0 \ . \tag{3.6.20}$$

But now there is a major difference between this equation and the cases when $\lambda < 0$ or $\lambda = 0$! Neither of the two terms in (3.6.20) are *automatically* non-zero. We could in fact still set $a_1 = 0$, and get the trivial solution, but this is *no longer necessary* in order to satisfy the boundary condition! We could equally well arrange for $\sin(\xi L)$ to be zero, simply by picking ξL to be an integer multiple of π. In this latter case, the a_1 could remain non-zero!

But which integer should we use? $\xi L = \pi$ works, but so do $\xi L = 2\pi$, and $\xi L = 3\pi$, and $\xi L = 4\pi$, and so forth. In fact any time we pick

$$\xi L = n\pi , \quad n = 1, 2, 3, \ldots \quad ,$$

we get a non-zero value for a_1 and hence a non-trivial solution. For the sake of convenience, we shall denote these different values with a subscript, as

$$\xi_n = \frac{n\pi}{L} , \quad n = 1, 2, 3, \ldots \quad . \tag{3.6.21}$$

Thus, for each different value of ξ_n, we get another non-trivial solution to:

$$\begin{aligned} X'' + \xi^2 X &= 0 , \\ X(0) = X(L) &= 0 , \end{aligned}$$

and we shall denote that solution by:

$$X_n(x) = \sin(\xi_n x) = \sin\left(\frac{n\pi x}{L}\right) , \quad n = 1, 2, 3, \ldots \quad . \tag{3.6.22}$$

Note we have dropped the arbitrary constant in the solution $X(x)$. We did this since we have already observed that, as solutions to a linear, homogeneous problem, the $X(x)$ can only be determined up to an arbitrary constant anyway. So we are only interested in the linearly independent solutions to this problem! Interestingly, we have passed, in one step, from having doubts about whether or not the splitting we assumed in (3.6.14) would work for any λ, to the determination that it works for an *infinite number of different values* of λ, given by

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2 , \quad n = 1, 2, 3, \ldots \quad . \tag{3.6.23}$$

Returning to the equation for $T(t)$, we see that for each value of λ_n given by (3.6.23), there are two linearly independent solutions, which we can represent as

$$T_n(t) = A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right) , \quad n = 1, 2, 3, \ldots \quad .$$

Finally, combining the various pieces, we see that for each value of λ given by (3.6.23),the separation of variables form (3.6.14) produces two non-trivial, linearly independent solutions to (3.6.13) given by:

$$
\begin{aligned}
u_n(x, t) \;&=\; X_n(x) T_n(t) \\
&=\; \left[A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right)\right] \sin\left(\frac{n\pi x}{L}\right) , \quad n = 1, 2, 3, \ldots \quad .
\end{aligned}
\tag{3.6.24}
$$

(At this point, we shall pause momentarily to reflect on what has just happened. We had a linear, homogeneous problem, given by (3.6.13), and containing a parameter, λ. This type of problem normally would have only the trivial solution. Upon substitution, however, we found that non-trivial solutions did in fact exist, but only for certain specified values of the parameter. This behavior is qualitatively similar to that encountered in the study of systems of linear algebraic equations with the matrix-vector equation

$$Ax = \lambda x ,$$

where x is an n-dimensional vector, and A is a constant matrix. This equation has non-trivial solutions if and only if λ satisfies

$$\det(A - \lambda I) = 0 ,$$

that is, if λ is an *eigenvalue* of A and x the corresponding *eigenvector*. Because of these similarities, we call the λ as given by (3.6.23) the eigenvalues of the problem (3.6.19). In a similar vein the non-trivial solutions $X_n(x)$ given by (3.6.22) are often referred to as the *eigenfunctions*.)

We now proceed. With (3.6.24) we seem to have identified all the linearly independent solutions to (3.6.13). (Although we have not proven this, their number would indicate it.) Therefore, since the basic problem was linear, we should be able to form the general solution by taking an arbitrary linear combination of the linearly independent homogeneous solutions, just as the general solution to

$$y'' + 2y' + y = 0$$

is given as a linear combination of e^{-x} and xe^{-x}, i.e. as $y(x) = c_1 e^{-x} + c_2 x e^{-x}$.

Following this logic, the general solution to (3.6.13) should be given by

$$u(x,t) = \left[A_1 \cos\left(\frac{\pi ct}{L}\right) + B_1 \sin\left(\frac{\pi ct}{L}\right) \right] \sin\left(\frac{\pi x}{L}\right) +$$

$$\left[A_2 \cos\left(\frac{2\pi ct}{L}\right) + B_2 \sin\left(\frac{2\pi ct}{L}\right) \right] \sin\left(\frac{2\pi x}{L}\right) + \tag{3.6.25}$$

$$\left[A_3 \cos\left(\frac{3\pi ct}{L}\right) + B_3 \sin\left(\frac{3\pi ct}{L}\right) \right] \sin\left(\frac{3\pi x}{L}\right) + \cdots ,$$

where the A_i and B_i, $i = 1, 2, 3, \ldots$ are arbitrary constants.

Note that we can save a great deal of writing by using instead the shorthand

$$u(x,t) = \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right) \right] \sin\left(\frac{n\pi x}{L}\right) . \tag{3.6.26}$$

But, if this really represents the general solution to (3.6.13) , then it must be able, for a proper choice of the constants A_i and B_i, to satisfy any initial conditions we would wish to apply, and specifically to satisfy the initial conditions given in (3.6.13),

$$\begin{aligned} u(x,0) &= f(x) , \\ u_t(x,0) &= g(x) . \end{aligned}$$

Can this, in fact satisfy the conditions? There is really but one way to find out, and that is to take our "general solution" and substitute it into these equations. Therefore we let $t = 0$ in (3.6.26), which is supposed to represent $u(x,t)$ at *all* times and substitute that expression for $u(x,0)$ in the initial conditions. This produces

$$f(x) = u(x,0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right) . \tag{3.6.27}$$

Now can we find the A_n so that this is satisfied? The answer, of course, is yes, provided $f(x)$ is at least *piecewise smooth* on $0 < x < L$. Physically, this must almost certainly be the case, since it is difficult to conceive of any guitar string in the "real world" whose displacements are not at least piecewise smooth. (Really, it's quite difficult to think of any that are not continuous!) So all we need do is extend $f(x)$ oddly into the region $-L < x < 0$. Since (3.6.27) is nothing more than the Fourier sine series for this extension, its coefficients (the A_n) are given by:

$$A_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx . \tag{3.6.28}$$

For the second initial condition, we must first differentiate (3.6.26) to obtain an expression for $u_t(x, t)$,

$$u_t(x, t) = \sum_{n=1}^{\infty} \left[-\frac{n\pi c}{L} A_n \sin\left(\frac{n\pi ct}{L}\right) + \frac{n\pi c}{L} B_n \cos\left(\frac{n\pi ct}{L}\right) \right] \sin\left(\frac{n\pi x}{L}\right) ,$$

and then let $t = 0$ and substitute into the second initial condition, to get

$$u_t(x, 0) = \sum_{n=1}^{\infty} \frac{n\pi c}{L} B_n \sin\left(\frac{n\pi x}{L}\right) . \tag{3.6.29}$$

But this is simply another Fourier sine series, except with an effective coefficient of $\frac{n\pi c}{L} B_n$ instead of b_n. Therefore, by the usual rule for computing these coefficients, we simply imagine $g(x)$ to be oddly extended and then

$$\frac{n\pi c}{L} B_n = \frac{2}{L} \int_0^L g(x) \sin\left(\frac{n\pi x}{L}\right) dx ,$$

or

$$B_n = \frac{2}{n\pi c} \int_0^L g(x) \sin\left(\frac{n\pi x}{L}\right) dx . \tag{3.6.30}$$

This clearly indicates that the B_n can also be computed provided the initial velocity, $g(x)$ is at least piecewise smooth - which from physical considerations it really needs to be.

Therefore, we see that we can satisfy any initial condition for the homogeneous partial differential equation (3.6.13) with the solution form (3.6.26), and hence we seem to have in fact found the general solution.

But wait a minute! The solution as defined by (3.6.26) involves an infinite series, and some infinite series are known not to converge. Is it possible that we might find in some problems that the solution did not converge? Fortunately, here, the answer is no! This is because it is easily shown that, since $\sin(n\pi ct/L)$ and $\cos(n\pi ct/L)$ are always less than or equal to one in magnitude, then (3.6.26) will always converge provided the original Fourier sine series (3.6.27) did. But, the original Fourier sine is known to converge from the basic principles of Fourier series (provided the initial displacement and velocity are piecewise smooth), so there is no problem.

As an example, consider the problem

$$
\begin{aligned}
u_{tt} &= u_{xx} , \quad 0 < x < 2 , \quad 0 < t , \\
u(0, t) &= 0 , \\
u(2, t) &= 0 , \\
u(x, 0) &= \begin{cases} x & , 0 < x \leq 1 \\ 2 - x & , 1 < x < 2 \end{cases} \\
u_t(x, 0) &= 0
\end{aligned}
$$

This problem describes the an elastic string of length two, fixed at both ends, which was displaced by lifting up at the center to the shape shown in Figure 29, held there, and then at $t = 0$ released. (At the instant of release, its velocity is thus zero, even though its acceleration is not.)

92
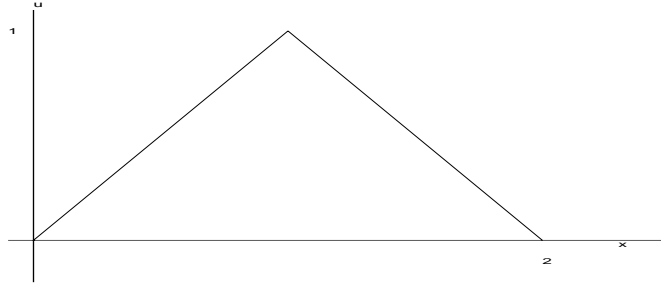
Figure 29: The Initial Displacement - $u(x, 0)$

If we retrace the steps from (3.6.13) on for this problem, we see that the eigenfunctions satisfy

$$X'' + \lambda X = 0 \ ,$$
$$X(0) = X(2) = 0 \ ,$$

and are therefore

$$X_n(x) = \sin\left(\frac{n\pi x}{2}\right) \ , \quad n = 1, 2, 3, \ldots \quad .$$

The eigenvalues are

$$\lambda_n = \xi_n^2 = \left(\frac{n\pi}{2}\right)^2 \ , n = 1, 2, 3, \ldots$$

Thus, the general solution to this problem looks like

$$u(x, t) = \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi t}{2}\right) + B_n \sin\left(\frac{n\pi t}{2}\right)\right] \sin\left(\frac{n\pi x}{2}\right) \quad .$$

The initial displacement becomes,

$$u(x, 0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{2}\right) = \begin{cases} x & , \ 0 \le x \le 1 \\ 2 - x & , \ 1 < x \le 2 \end{cases} \ ,$$

and, therefore, by the usual Fourier formulas

$$A_n = \frac{2}{2} \int_0^2 f(x) \sin\left(\frac{n\pi x}{2}\right) dx$$

$$= \int_0^1 x \sin\left(\frac{n\pi x}{2}\right) dx + \int_1^2 (2 - x) \sin\left(\frac{n\pi x}{2}\right) dx$$

$$= \left\{-\frac{2}{n\pi} x \cos\left(\frac{n\pi x}{2}\right) + \left[\frac{2}{n\pi}\right]^2 \sin\left(\frac{n\pi x}{2}\right)\right\}\Big|_0^1 +$$

$$\left\{-\frac{2}{n\pi}(2 - x)\cos\left(\frac{n\pi x}{2}\right) - \left[\frac{2}{n\pi}\right]^2 \sin\left(\frac{n\pi x}{2}\right)\right\}\Big|_1^2$$

$$= 2\left[\frac{2}{n\pi}\right]^2 \sin\left(\frac{n\pi}{2}\right) \quad .$$

93

Similarly, the initial condition on the velocity, $u_t(x,0)$, yields

$$B_n = 0 \quad ,$$

and therefore,

$$u(x,t) = \sum_{n=1}^{\infty} 2 \left[\frac{2}{n\pi} \right]^2 \sin\left(\frac{n\pi}{2} \right) \cos\left(\frac{n\pi t}{2} \right) \sin\left(\frac{n\pi x}{2} \right) \quad .$$

PROBLEMS

1. Solve:

$$u_{tt} = u_{xx}$$
$$u(0,t) = u(3,t) = 0$$
$$u(x,0) = \begin{cases} 2x & ,0 < x < 1/2 \\ 2 - 2x & ,1/2 \le x < 1 \\ 0 & ,1 \le x < 3 \end{cases}$$
$$u_t(x,0) = 0$$

Sketch the ten-term partial sum of your computed solution at

$$t = 0, 1, 2, 4 .$$

2. Solve:

$$u_{tt} = u_{xx}$$
$$u(0,t) = u(\pi,t) = 0$$
$$u(x,0) = \begin{cases} 0 & , & 0 < x < \pi/4 \\ (4x - \pi)/\pi & , & \pi/4 \le x \le \pi/2 \\ (3\pi - 4x)/\pi & , & \pi/2 \le x \le 3\pi/4 \\ 0 & , & 3\pi/4 < x < \pi \end{cases}$$
$$u_t(x,0) = 0$$

3. Solve:

$$u_{tt} = u_{xx}$$
$$u(0,t) = u(\pi,t) = 0$$
$$u(x,0) = x(\pi - x), 0 < x < \pi$$
$$u_t(x,0) = 0$$

4. Solve:

$$u_{tt} = u_{xx}$$
$$u(0,t) = u(3,t) = 0$$
$$u(x,0) = 0$$
$$u_t(x,0) = x$$

5. Solve:

$$u_{tt} = 9u_{xx}$$
$$u(0,t) = u(\pi,t) = 0$$
$$u(x,0) = sin(x)$$
$$u_t(x,0) = 1$$

6. Solve:

$$u_{tt} = 4u_{xx}$$
$$u(0,t) = u(5,t) = 0$$
$$u_t(x,0) = \begin{cases} x & , & 0 < x < 5/2 \\ 5 - x & , & 5/2 \le x < 5 \end{cases}$$
$$u(x,0) = 0$$

95

7. The dissipation of heat in a "very large" solid slab of thickness L whose faces are held at a fixed reference temperature of 0° is described by the partial differential equation:

$$u_t = ku_{xx}$$
$$u(0,t) = u(L,t) = 0$$
$$u(x,0) = f(x)$$

where $u(x,t)$ denotes the temperature at location x and time t.

    a. Why is only one initial condition required in this problem?

    b. Show that the method of Separation of Variables also "works" in this problem, and leads formally to the general solution:

$$u(x,t) = \sum_{n=1}^{\infty} b_n e^{-k(n\pi/L)^2 t} \sin\left(\frac{n\pi x}{L}\right) \quad ,$$

where:

$$b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx \quad .$$

3.7 The Free End Conditions Problem

As we have already indicated, our general approach to the study of the one-dimensional wave equation (3.5.11) was to first solve several special cases, and then try to deduce the main aspects of the general theory from these specific cases. Each case would correspond to different specific boundary conditions (i.e. values of α and β in (3.5.11)). In the previous section, our first case, we developed the solutions to the fixed end condition (Dirichlet) problem. There we saw that the separation of variables assumption, applied to the homogeneous problem without initial conditions, leads to an infinite number of linearly independent solutions determined by the eigenvalues of the problem. Furthermore, after we formed the general solution as a linear combination of these solutions, we found that satisfying the initial conditions reduced to nothing more than an ordinary Fourier sine series. The obvious question which we must now pursue is how much of this behavior will any one-dimensional wave equation share, and which aspects were purely example (Dirichlet boundary conditions) specific? To answer this, we must of course study more cases.

In this section, we consider a second special case, specifically that of $\alpha_1 = \alpha_2 = 0$, and $\beta_1 = \beta_2 = 1$, or $u(x,t)$ satisfying:

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u_x(0,t) &= 0 \\
u_x(L,t) &= 0 \\
u(x,0) &= f(x) \\
u_t(x,0) &= g(x)
\end{aligned}
\qquad (3.7.31)
$$

Physically, this problem could represent a string with frictionless slip rings located at *both* ends (Figure 30). Mathematically, as mentioned during the derivation of the various boundary conditions, we call this problem the free end condition problem, or, alternatively, say it has Neumann conditions at both ends. (Before reading further, you should pause to reflect, physically, on what qualitative properties you might expect a solutions to this problem would have.)

We will, of course, want to apply to this problem, as much as possible, our experience with the Dirichlet conditions. (After all, there is very little profit in "reinventing the wheel," or in repeating mathematical steps when you know (or should know) ahead of time what the result will be.) Since, in the Dirichlet case, we started with the homogeneous problem without initial conditions, if we were to proceed similarly with (3.7.31), then we should start with

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u_x(0,t) &= 0, \\
u_x(L,t) &= 0
\end{aligned}
\qquad (3.7.32)
$$

Clearly, as also occurred with the Dirichlet case, this problem always has at least the trivial solution, and again the important question becomes what are its linearly independent non-trivial solutions, and how do we find them?

Well, observe that the partial differential equation in (3.7.32) has not changed from the Dirichlet problem, and so still possesses the same natural splitting of the time and spatial operations. Therefore it seems quite reasonable to use the same separation of variables
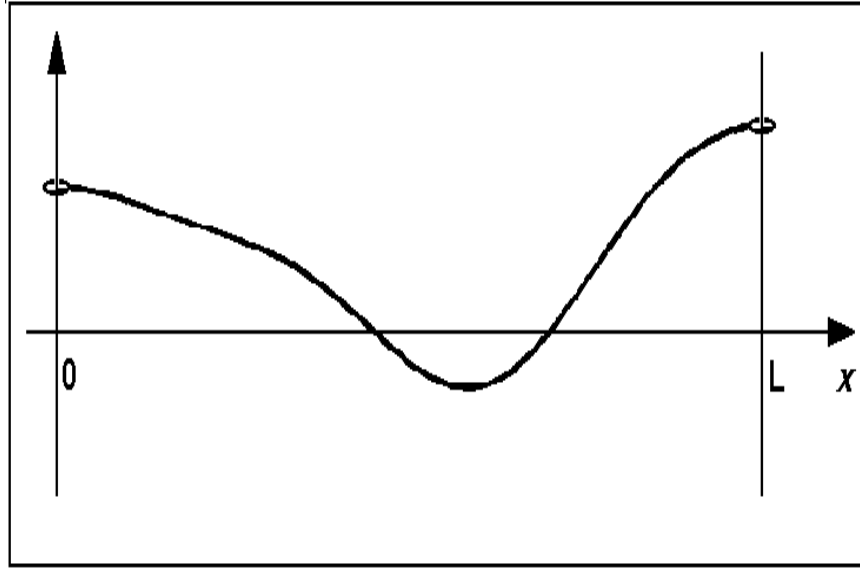
Figure 30: The Free End Conditions Problem

form (3.6.14) here. Moreover, since the partial differential equation is unchanged (only the *boundary conditions* have changed), substituting the product form into the partial differential equation must produce exactly the same separated ordinary differential equations, i.e.

$$T''(t) + \lambda c^2 T(t) = 0 \qquad X''(x) + \lambda X(x) = 0 \ .$$

Keep in mind that at this point λ does not necessarily represent any single value. In fact, based on our last example we should almost certainly expect an infinite number of different values will "work" here also. But what are they? After all, the boundary conditions aren't the same as in the last case, so there's no reason to expect the same solutions. It still does seem unlikely from physical arguments that λ could be negative, since exponentially growing solutions $(T(t))$ seem no more physically reasonable here than they were before. But, mathematically, we can't assume anything yet about possible values for λ for this problem.

This last point shouldn't really be surprising, since, of course, we haven't yet considered the boundary conditions for this problem, and the boundary conditions were crucial in determining the eigenvalues of the fixed end problem. Furthermore, since the boundary conditions for this problem (3.7.31) are not the same as in the Dirichlet case, we can't say $X(0) = X(L) = 0$, but must derive a new set of conditions, appropriate for this case, by substituting the separation of variables form (3.6.14) directly into boundary conditions in (3.7.32). That substitution yields

$$u_x(0, t) = X'(0)T(t) = 0 \quad ,$$

which, since $X'(0)$ must just be a constant, requires,

$$X'(0) = 0 \ .$$

98

if we are to have non-trivial solutions. An essentially identical argument at $x = L$ leads to

$$u_x(L, t) = X'(L)T(t) = 0 \Rightarrow X'(L) = 0 .$$

Thus, with our new set of boundary conditions, we now have the complete set of separated equations

$$
\begin{array}{cc}
T'' + \lambda c^2 T = 0 & \begin{aligned} X'' &+ \lambda X = 0 \\ X'(0) &= 0 \\ X'(L) &= 0 \end{aligned}
\end{array}
\qquad (3.7.33)
$$

Observe that in these equations, as in the fixed end case, the equation for $T(t)$ is solvable for any value of λ. This still seems reasonable since we have not yet applied any initial conditions. Furthermore, the equation for $X(x)$

$$
\begin{aligned}
X'' &+ \lambda X = 0 \\
X'(0) &= 0 \\
X'(L) &= 0
\end{aligned}
\qquad (3.7.34)
$$

is again homogeneous with homogenous boundary conditions, and so clearly still has the trivial solution. Therefore, the fundamental question we must answer has not changed from the Dirichlet problem - for what values of λ are there non-trivial solutions to (3.7.34)? Well, unless we're clairvoyant, there's no way to answer this other than to basically repeat the same procedure as before, i.e. assuming, in sequence, negative, zero and positive values for λ, and seeing if these produce any solutions.

As before, we start by assuming λ could be negative, and again represent this by

$$\lambda = -\zeta^2 , \quad \zeta \neq 0 .$$

This yields

$$X(x) = c_1 e^{\zeta x} + c_2 e^{-\zeta x} ,$$

and therefore the boundary conditions reduce to

$$
\begin{aligned}
X'(0) &= \zeta c_1 - \zeta c_2 = 0 \\
X'(L) &= \zeta c_1 e^{\zeta L} - \zeta c_2 e^{-\zeta L} = 0 ,
\end{aligned}
$$

which since ζ and L are non-zero, can be shown to be satisfied only by $c_1 = c_2 = 0$. Thus, as before, we have *only* the trivial solution when $\lambda < 0$ (which again coincides with our earlier expectation that growing exponentials still should not occur in the solutions for $T(t)$.) So $\lambda < 0$ may be dropped from further consideration for this problem.

Next, we must check whether λ can be zero. Well, if this is the case, then

$$X(x) = d_1 + d_2 x ,$$

and when we substitute this solution into the free boundary conditions, we find

$$
\begin{aligned}
X'(0) &= d_2 = 0 \\
X'(L) &= d_2 = 0 .
\end{aligned}
$$

99

But wait! This requires that $d_2 = 0$, but it doesn't say anything about d_1, which apparently can have any value. Therefore here, unlike the fixed end case, $\lambda = 0$ does not produce only the trivial solution, and so cannot be dropped in this problem. In fact, according to our definitions, zero is clearly an eigenvalue here. The corresponding eigenfunction is a constant, which, since we are only interested in linearly independent solutions, we can assign to be any convenient value. For purposes which should become clear later, we now chose this value to be $\frac{1}{2}$. Thus, in summary, we have

$$\begin{aligned} \lambda_0 &= \quad 0 \\ X_0(x) &= \tfrac{1}{2} \end{aligned}$$

(You might note again that this conclusion should have been obvious from a geometrical interpretation of the $\lambda = 0$ condition, which only requires a straight line parallel to the x-axis at $x = 0$ and $x = L$.)

But, of course, so far we really have only half the answer at $\lambda = 0$! Our real goal is to identify the linearly independent solutions $u(x, t)$. Therefore each $X(x)$ must have a corresponding $T(t)$. When $\lambda_0 = 0$, $T(t)$ must clearly satisfy the differential equation

$$T_0'' = 0 \Rightarrow T_0(t) = A_0 + B_0 t \quad,$$

and therefore the corresponding product solution becomes

$$u_0(x, t) = X_0(x) T_0(t) = \tfrac{1}{2}[A_0 + B_0 t] \quad.$$

But we can't stop yet! We have to identify *all* of the linearly independent homogeneous solutions, and the fixed end problem had an infinite number of different eigenvalues. So it doesn't seem reasonable that this problem would have only one! Therefore, we must also investigate whether any non-trivial solutions exist for λ positive. Thus we write,

$$\lambda = \xi^2$$

where ξ is some non-zero real number. This gives the as general solution for (3.7.34)

$$X(x) = a_1 \sin(\xi x) + a_2 \cos(\xi x) \quad,$$

and again we must try to pick a_1 and a_2, not both zero, so that the boundary conditions are satisfied. For the free boundary conditions, this leads, at $x = 0$, to

$$X'(0) = a_1 \xi = 0 \Rightarrow X(x) = a_2 \cos(\xi x)$$

(since ξ was assumed non-zero). Then, at $x = L$, the boundary condition reduces to

$$X'(L) = -a_2 \xi \sin(\xi L) = 0 \quad, \quad \xi \neq 0 \quad.$$

Here, as in the fixed end case, and unlike the situation when λ is negative, it's not *necessary* to set the a_2 equal to zero (and so have only the trivial solution) in order to satisfy the

boundary condition! We could again keep a_2 non-zero by making $\sin(\xi L)$ zero, using the same eigenvalues as worked before, i.e.

$$\xi_n = \frac{n\pi}{L} \, ,$$

$$n = 1, 2, 3, \ldots \, . \qquad (3.7.35)$$

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2 \, ,$$

But there is an important difference here compared to the fixed-end case - even though the **eigenvalues** have not changed, a review of these last steps shows that **the eigenfunctions have**. In fact, the eigenfunctions here are

$$X_n(x) = \cos(\xi_n x) = \cos\left(\frac{n\pi x}{L}\right) \, , \quad n = 1, 2, 3, \ldots \qquad (3.7.36)$$

(where we have again dropped the arbitrary constant since we only care about linearly independent solutions). Geometrically these eigenfunctions reflect the fact that, for $\lambda > 0$, equation (3.7.34) may be viewed as simply specifying those sinusoidal functions with zero slope at both ends of the interval!

As before, each eigenvalue given by (3.7.35) also determines a corresponding differential equation for the $T(t)$, whose solution involves two linearly independent solutions which can be represented as

$$T_n(t) = A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right) \, , \quad n = 1, 2, 3, \ldots \, .$$

Combining all of this information, we see that each positive eigenvalue produces two non-trivial, linearly independent solutions to (3.7.32) given by:

$$u_n(x,t) = X_n(x)T_n(t)$$

$$= \left[A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right)\right] \cos\left(\frac{n\pi x}{L}\right) \, , \quad n = 1, 2, 3, \ldots \, . \qquad (3.7.37)$$

By this point, we would again seem to have identified all the linearly independent solutions to (3.7.32). (Although again we have not proven this.) Therefore, we now form the general solution by taking an arbitrary linear combination of all the linearly independent homogeneous solutions (including $\lambda = 0$), to yield

$$u(x,t) = \tfrac{1}{2}[A_0 + B_0 t] + \left[A_1 \cos\left(\frac{\pi ct}{L}\right) + B_1 \sin\left(\frac{\pi ct}{L}\right)\right] \cos\left(\frac{\pi x}{L}\right) +$$

$$\left[A_2 \cos\left(\frac{2\pi ct}{L}\right) + B_2 \sin\left(\frac{2\pi ct}{L}\right)\right] \cos\left(\frac{2\pi x}{L}\right) +$$

$$\left[A_3 \cos\left(\frac{3\pi ct}{L}\right) + B_3 \sin\left(\frac{3\pi ct}{L}\right)\right] \cos\left(\frac{3\pi x}{L}\right) + \cdots \, , \qquad (3.7.38)$$

$$= \tfrac{1}{2}[A_0 + B_0 t] + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right)\right] \cos\left(\frac{n\pi x}{L}\right) \, ,$$

101

where the A_i and B_i, $i = 0, 1, 2, \ldots$ still have to be determined. Moreover, as before, if this really represents the general solution to (3.7.32), then we have to be able to determine the constants A_i and B_i so as to satisfy any set of initial conditions we would wish to apply, and specifically for the initial conditions

$$
\begin{aligned}
u(x, 0) &= f(x) \ , \\
u_t(x, 0) &= g(x) \ .
\end{aligned}
$$

To see whether we can accomplish this here, we must, as we did before, substitute (3.7.38), which is supposed to represent $u(x, t)$ at *all* times, directly into the expressions for $u(x, 0)$ and $u_t(x, 0)$. If we do this, we have,

$$
u(x, 0) = f(x) = \tfrac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right) \ , \quad 0 < x < L \ ,
$$

and

$$
u_t(x, 0) = g(x) = \tfrac{1}{2}B_0 + \sum_{n=1}^{\infty} \frac{n\pi c}{L} B_n \cos\left(\frac{n\pi x}{L}\right) \ , \quad 0 < x < L \ .
$$

(3.7.39)

Now can we find the A_n and B_n that satisfy these? The answer, as in the fixed end case, is yes, provided $f(x)$ and $g(x)$ are at least *piecewise smooth* on $0 < x < L$ (Which, from physical considerations, should clearly be the case). There is, however, one change. In this problem we must use *even* extensions, rather than odd ones, since only even function have Fourier cosine series (which are all (3.7.39) represent). Computing these coefficients in the usual way (recognizing that the effective coefficient in the second series is $\frac{n\pi c}{L} B_n$ instead of a_n) yields

$$
A_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx \ , \quad n = 0, 1, 2, \ldots
$$

$$
B_0 = \frac{2}{L} \int_0^L g(x) dx \ ,
$$

(3.7.40)

$$
B_n = \frac{2}{n\pi c} \int_0^L g(x) \cos\left(\frac{n\pi x}{L}\right) dx \ , \quad n = 1, 2, 3, \ldots .
$$

Therefore, we have found that solution form (3.7.38) can satisfy any initial condition for the homogeneous partial differential equation (3.7.31), and hence does in fact represent the general solution. (Mathematically, as in the case of the series solution to the fixed end case, (3.7.38) really only represents only a *formal solution*. That is, it's not valid unless the infinite series involved converge. Fortunately though, here as before, we could show that this series will always converge provided the initial condition Fourier cosine series (3.7.39) did.)

We cannot leave this problem, however, without pointing out one additional intriguing feature of (3.7.38) that we did not see in the Dirichlet solution - the presence of the $\tfrac{1}{2}[A_0 + B_0 t]$ term. What does this represent? What would be the meaning of

$$
\tfrac{1}{2}A_0 = \frac{1}{L} \int_0^L f(x) dx \neq 0 \ ?
$$

And how should we interpret

$$\tfrac{1}{2}B_0 = \frac{1}{L}\int_0^L g(x)dx \neq 0 \quad ?$$

(We will leave the reader to reflect on these questions!)

To illustrate this general class of problems, consider the example

$$
\begin{aligned}
u_{tt} &= u_{xx} , \quad 0 < x < 4 , \quad 0 < t , \\
u_x(0,t) &= 0 , \\
u_x(4,t) &= 0 , \\
u(x,0) &= \begin{cases} 0 & , \; 0 < x \le 1 \\ x-1 & , \; 1 < x < 2 \\ 3-x & , \; 2 \le x \le 3 \\ 0 & , \; 3 < x < 4 \end{cases} \\
u_t(x,0) &= 0
\end{aligned}
$$

This problem describes an elastic string of length four, free at both ends, whose middle half is displaced into a triangular shape, held there, and then released at $t = 0$. (See Figure 31.) Note that $c = 1$ for this example.
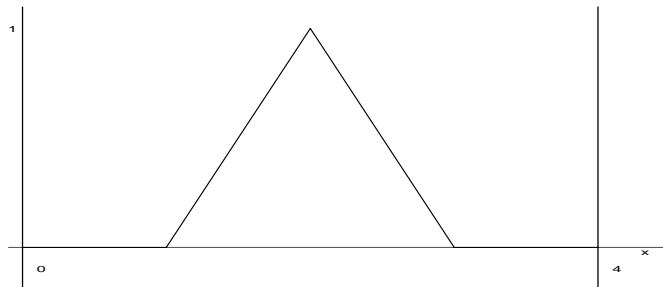
Figure 31: The Initial Displacement $f(x)$

If we retrace the steps from (3.7.32) on for this problem, we see that the eigenfunctions satisfy

$$
\begin{aligned}
X'' + \lambda X &= 0 , \\
X'(0) = X'(4) &= 0 ,
\end{aligned}
$$

and therefore the eigenvalues are

$$\lambda_0 = 0 \quad \text{and} \quad \lambda_n = \frac{n\pi x}{4} , n = 1, 2, 3, \ldots ,$$

with corresponding eigenfunctions

$$X_0(x) = \tfrac{1}{2} \quad \text{and} \quad X_n(x) = \cos\left(\frac{n\pi x}{4}\right) , \quad n = 1, 2, 3, \ldots ,$$

respectively. Thus, according to the steps outlined in this chapter, the general solution to this problem is

$$u(x,t) = \tfrac{1}{2}[A_0 + B_0 t] + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi t}{4}\right) + B_n \sin\left(\frac{n\pi t}{4}\right)\right] \cos\left(\frac{n\pi x}{4}\right) \quad .$$

103

The initial conditions become,

$$u(x,0) = \tfrac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi x}{4}\right) = \begin{cases} 0 & , \ 0 < x \leq 1 \\ x-1 & , \ 1 < x < 2 \\ 3-x & , \ 2 \leq x \leq 3 \\ 0 & , \ 3 < x < 4 \end{cases}$$

and,

$$u_t(x,0) = g(x) = \tfrac{1}{2}B_0 + \sum_{n=1}^{\infty} \frac{n\pi}{L} B_n \cos\left(\frac{n\pi x}{4}\right) = 0 \ .$$

Therefore, by the usual Fourier formulas

$$\tfrac{1}{2}A_0 = \tfrac{1}{4}$$

$$A_n = \tfrac{2}{4} \int_0^4 f(x) \cos\left(\frac{n\pi x}{4}\right) dx$$

$$= \tfrac{1}{2} \int_1^2 (x-1) \cos\left(\frac{n\pi x}{4}\right) dx + \tfrac{1}{2} \int_2^3 (3-x) \cos\left(\frac{n\pi x}{4}\right) dx$$

$$= \frac{1}{2}\left[\frac{4}{n\pi}\right]^2 \left[2\cos\left(\frac{n\pi}{2}\right) - \cos\left(\frac{n\pi}{4}\right) - \cos\left(\frac{3n\pi}{4}\right)\right]$$

$$B_n \equiv 0 \ , \quad n = 0, 1, 2, \ldots \quad .$$

Note the value of $\dfrac{1}{2}A_0$ here - it's simply the average height, along the string, of the initial displacement, i.e. the center of mass in the vertical direction of the string. Furthermore, there are no external forces acting on this string, and the average value of $\cos\left(\frac{n\pi t}{4}\right)$ is zero! This implies that since $B_n \equiv 0$, we could view this solution as a vibration about a fixed, offset center of mass. (If you're still not sure how to interpret the B_0 in the general case, try applying conservation of momentum to the initial conditions in (3.7.31).)

PROBLEMS

1. Solve:
$$u_{tt} = 25u_{xx}$$
$$u_x(0, t) = u_x(1, t) = 0$$
$$u(x, 0) = \begin{cases} 0 & , \quad 0 < x < 1/4 \\ x - 1/4 & , 1/4 < x < 3/4 \\ 1/2 & , \quad 3/4 < x < 1 \end{cases}$$
$$u_t(x, 0) = 0$$

Interpret the solution physically.

2. Solve:
$$u_{tt} = u_{xx}$$
$$u_x(0, t) = u_x(2, t) = 0$$
$$u(x, 0) = \begin{cases} 2x & , \quad 0 < x < 1/2 \\ 2 - 2x & , \quad 1/2 \le x \le 3/2 \\ 2x - 4 & , \quad 3/2 < x < 2 \end{cases}$$
$$u_t(x, 0) = 1$$

Interpret the solution physically.

3.8 The Mixed End Conditions Problem

Thus far, we have solved two special cases of the one-dimensional wave equation in our effort to develop a "feel" for the general behavior of the solutions to this equation. In both of these cases, the separation of variables assumption, applied to the homogeneous parts of the problem, produced an infinite number of linearly independent solutions, as determined by the eigenvalues. These eigenvalues, however, and their associated eigenfunctions, were problem dependent, i.e. they changed with the problem, although in neither case were the eigenvalues negative, and in only one of the two cases zero was an eigenvalue. Furthermore, in both cases, the general solution we formed as a linear combination of these solutions reduced to an ordinary Fourier series when applied to the initial conditions. The fundamental question still in front of us is how many of these properties will remain unchanged as we study more cases.

Therefore, we continue by investigating a third case,

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u(0,t) &= 0 \\
u_x(L,t) &= 0 \\
u(x,0) &= f(x) \\
u_t(x,0) &= g(x) \quad .
\end{aligned}
\tag{3.8.41}
$$

(In terms of our standard form (3.5.11), this newest case corresponds to

$$
\alpha_1 = \beta_2 = 1 \qquad \text{and} \qquad \beta_1 = \alpha_2 = 0 \ . \ )
$$

Physically, this equation could model a string fixed at the left end, with a frictionless slip ring located at the right end (Figure 32). Mathematically, we may either refer to this as the mixed end condition problem, or say the problem has Dirichlet conditions at the left and Neumann conditions at the right. In this point, as in any problem, you should pause, before continuing, and consider what qualitative physical and mathematical properties you might expect to find in the solution.

We will, of course, want to apply here as much as possible of our experience with the two earlier problems. In both of these examples, we started by considering just the homogeneous parts of the problem (without the initial conditions). So it seems reasonable to start here with

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u(0,t) &= 0 \\
u_x(L,t) &= 0
\end{aligned}
\tag{3.8.42}
$$

Given (3.8.42), the important question, as it was in the two previous cases, is what are the linearly independent, non-trivial solutions, and how do we find them? The answer in this new case is not obvious *a priori*, but having already solved the earlier cases, we should at least be able to skip many of the intermediate details.

For example, the partial differential equation has still not changed, and therefore the separation of variables form (3.6.14) should (and does) produce exactly the same two separated ordinary differential equations. Furthermore, we have already separated the left-hand
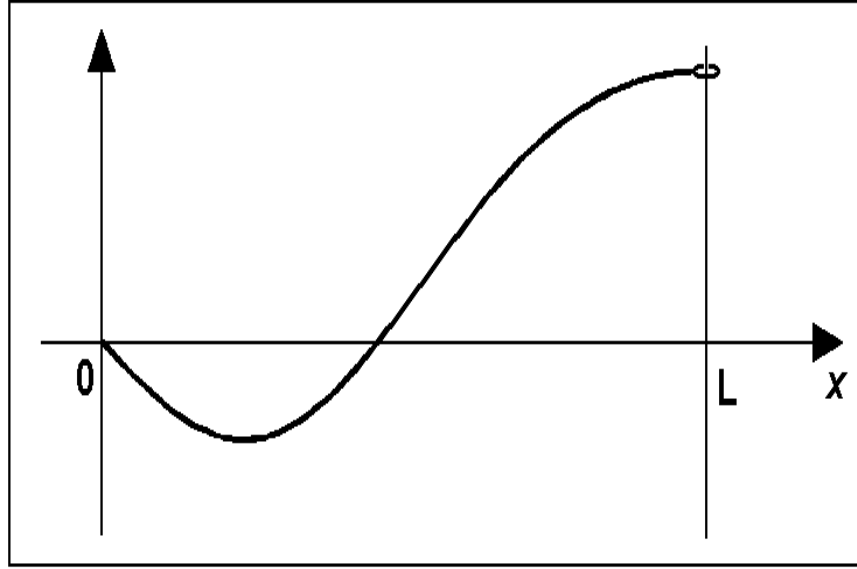
Figure 32: The Mixed End Condition Problem

boundary condition in our first example, and the right-hand condition in our second one. Thus, we proceed to immediately write the separated problems

$$T'' + \lambda c^2 T = 0 \qquad \begin{aligned} X'' &+ \lambda X = 0 \\ X(0) &= 0 \\ X'(L) &= 0 \quad . \end{aligned} \qquad (3.8.43)$$

As before, the equation for $T(t)$ will produce two linearly independent solutions for *any* value of λ. (By now, as long as we don't apply initial conditions, we should expect this behavior.) Therefore, as in the earlier cases, the fundamental question remains the same - for what values of λ does

$$\begin{aligned} X'' + \lambda X &= 0 \\ X(0) &= 0 \\ X'(L) &= 0 \end{aligned} \qquad (3.8.44)$$

have non-trivial solutions? The procedure for answering this question also remains the same, i.e. considering, in sequence, negative, zero and positive values for λ, and seeing which of these do produce solutions.

In the current example, when λ is negative, and represented by $\lambda = -\zeta^2$ ($\zeta \neq 0$), we have as the general solution for $X(x)$

$$X(x) = c_1 e^{\zeta x} + c_2 e^{-\zeta x} \ ,$$

and therefore the boundary conditions reduce to

$$X(0) \quad = c_1 + c_2 = 0$$

$$X'(L) \quad = \zeta c_1 e^{\zeta L} - \zeta c_2 e^{-\zeta L} = 0 \ ,$$

107

which again can be shown to be satisfied only by $c_1 = c_2 = 0$. So negative values of λ still produce only trivial solutions, and therefore growing exponentials for $T(t)$ are still mathematically (and physically) impossible.

When λ is zero, the general solution to the ordinary differential equation is still

$$X(x) = d_1 + d_2 x \ ,$$

and, when substituted into the boundary conditions, this yields

$$\begin{array}{ll} X(0) & = d_1 = 0 \\ X'(L) & = d_2 = 0 \end{array} \ .$$

So we're back, in this problem, to only trivial solutions when $\lambda = 0$, and therefore zero is not an eigenvalue here. (Geometrically, this conclusion should be intuitive, since the problem for $\lambda = 0$ can be viewed as defining a straight line which passes through the x-axis at $x = 0$, and is horizontal at $x = L$.)

Lastly, we must consider whether positive values of λ generate non-trivial solutions. (This should by now seem highly likely, since they have in both previous cases). Following our usual convention of letting $\lambda = \xi^2$, where $\xi \neq 0$, again gives

$$X(x) = a_1 \sin(\xi x) + a_2 \cos(\xi x) \ .$$

The boundary conditions in this case, however, produce a slightly different result than we have seen previously when we try to determine the a_1 and a_2. Specifically,

$$X(0) = a_2 = 0 \Rightarrow X(x) = a_1 \sin(\xi x)$$

and then

$$X'(L) = a_1 \xi \cos(\xi L) = 0 \ , \quad \xi \neq 0 \ .$$

Here, as in the fixed and free end examples, the important point is that we don't *have to* choose $a_1 = 0$ and get only the trivial solution. However, in order to avoid the trivial solution, we do have to change our condition for selecting ξ from $\quad \sin(\xi L) = 0 \quad$ to

$$\cos(\xi L) = 0 \ .$$

This new condition, of course, will be true only for

$$\xi L = \frac{\pi}{2}, \ \frac{3\pi}{2}, \ \frac{5\pi}{2}, \ldots \ ,$$

or, equivalently

$$\xi_n \ = \ \frac{(2n-1)\pi}{2L} \ , \quad n = 1, 2, 3, \ldots \ ,$$

$$\lambda_n \ = \ \left[\frac{(2n-1)\pi}{2L} \right]^2 , \quad n = 1, 2, 3, \ldots \ .$$

(3.8.45)

108

This is a deviation, albeit minor, from the earlier cases we studied - the positive **eigenvalues** are not the same as in the "pure" fixed end or free end cases. Neither are the eigenfunctions

$$X_n(x) = \sin(\xi_n x) = \sin\left(\frac{(2n-1)\pi x}{2L}\right) , \quad n = 1, 2, 3, \ldots \tag{3.8.46}$$

(Note we have continued our practice of dropping the arbitrary constant when only one linearly independent solution exists. Also we use $(2n-1)$ simply as a convenient way of generating the *odd* integers. We could have equally correctly written

$$X_n(x) = \sin\left(\frac{n\pi x}{2L}\right) , \quad n = 1, 3, 5, \ldots ,$$

etc. This latter representation, however, is not generally used since it becomes more cumbersome.)

Now, as before, corresponding to each eigenvalue given by (3.8.45) there are two linearly independent solutions for $T(t)$, which can be represented in this case as

$$T_n(t) = A_n \cos\left(\frac{(2n-1)\pi ct}{2L}\right) + B_n \sin\left(\frac{(2n-1)\pi ct}{2L}\right) , \quad n = 1, 2, 3, \ldots ,$$

and therefore there are also two linearly independent solutions to (3.8.42) given by:

$$\begin{aligned} u_n(x,t) \quad &= X_n(x) T_n(t) \\ &= \left[A_n \cos\left(\frac{(2n-1)\pi ct}{2L}\right) + B_n \sin\left(\frac{(2n-1)\pi ct}{2L}\right)\right] \sin\left(\frac{(2n-1)\pi x}{2L}\right) , \\ &\qquad\qquad n = 1, 2, 3, \ldots . \end{aligned}$$

Again, since we should by now have identified all the linearly independent solutions, we form the general solution in the usual way, as

$$\begin{aligned} u(x,t) \quad &= \left[A_1 \cos\left(\frac{\pi ct}{2L}\right) + B_1 \sin\left(\frac{\pi ct}{2L}\right)\right] \sin\left(\frac{\pi x}{2L}\right) + \\ &\qquad \left[A_2 \cos\left(\frac{3\pi ct}{2L}\right) + B_2 \sin\left(\frac{3\pi ct}{2L}\right)\right] \sin\left(\frac{3\pi x}{2L}\right) + \\ &\qquad \left[A_3 \cos\left(\frac{5\pi ct}{2L}\right) + B_3 \sin\left(\frac{5\pi ct}{2L}\right)\right] \sin\left(\frac{5\pi x}{2L}\right) + \cdots , \\ &= \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{(2n-1)\pi ct}{2L}\right) + B_n \sin\left(\frac{(2n-1)\pi ct}{2L}\right)\right] \sin\left(\frac{(2n-1)\pi x}{2L}\right) . \end{aligned}$$

$$\tag{3.8.47}$$

Continuing as before, we attempt to determine the constants A_i and B_i, $i = 1, 2, 3, \ldots$ by

substituting (3.8.47) into the initial conditions. This yields,

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{(2n-1)\pi x}{2L}\right) \ , \quad 0 < x < L \ ,$$

and

(3.8.48)

$$u_t(x,0) = g(x) = \sum_{n=1}^{\infty} \frac{(2n-1)\pi c}{2L} B_n \sin\left(\frac{(2n-1)\pi x}{2L}\right) \ , \quad 0 < x < L \quad .$$

Again, (3.8.48) poses an identical question to that we faced in earlier examples - how do we find the A_n and B_n that satisfy these? Unlike the earlier problems we studied, however, the answer here is **not** as clear. One difficulty is that the series in (3.8.48), while they are Fourier sine series, have a period of $4L$, **not** $2L$. (Look at the denominators carefully.) Therefore, even if $f(x)$ and $g(x)$ are at least *piecewise smooth* on $0 < x < L$, and we extend them oddly into $-L < x < 0$, we don't have enough values to use the Fourier sine series formulas directly. We're missing the values of $f(x)$ for into $L < x < 2L$ and $-2L < x < -L$! (Actually, since we're using sine series, only the values from $L < x < 2L$ really matter.) So what do we do now?

Well, one choice is to still try to use odd and even Fourier series concepts. After all, the general Fourier sine series of period $2L$ has the form

$$\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{2L}\right) = b_1 \sin\left(\frac{\pi x}{2L}\right) \quad + b_2 \sin\left(\frac{2\pi x}{2L}\right) + b_3 \sin\left(\frac{3\pi x}{2L}\right)$$

$$+ b_4 \sin\left(\frac{4\pi x}{2L}\right) + b_5 \sin\left(\frac{5\pi x}{2L}\right) + \cdots \quad .$$

So then, all (?) we seem need to do in order to satisfy the first of (3.8.48) is figure out how to extend $f(x)$ into the interval $L < x < 2L$ in such a way that

$$b_2 = b_4 = \cdots = 0 \quad \text{and} \quad A_1 = b_1 \ , \quad A_2 = b_3 \ , \quad A_3 = b_5 \ , \quad \ldots \quad .$$

We could actually do this, but, as things turn out, such an approach simply delays the inevitable!

The better way to find the coefficients in (3.8.48) is to return to fundamental principles, and ask why was it so important in our earlier examples that the initial conditions reduce to elementary Fourier series? The straightforward answer would seem to be because we could then immediately compute the coefficients. But that's too superficial. **Why** could we compute the coefficients? The answer is because the **orthogonality** of the sines and cosines, i.e. the integrals like

$$\int_{-2L}^{2L} \sin\left(\frac{n\pi x}{2L}\right) \sin\left(\frac{m\pi x}{2L}\right) dx = 0 \ , m \neq n \ ,$$

allowed use to derive the appropriate formulas. Therefore, before we try various, perhaps artificial, extensions in (3.8.48), we should ask instead if perhaps the functions in those series,

110

i.e. the eigenfunctions, are themselves also orthogonal on the interval $0 < x < L$. This means considering the integral

$$\int_0^L \sin\left(\frac{(2n-1)\pi x}{2L}\right) \sin\left(\frac{(2m-1)\pi x}{2L}\right) dx \quad .$$

Using standard trigonometric identities we have

$$\int_0^L \sin\left(\frac{(2n-1)\pi x}{2L}\right) \sin\left(\frac{(2m-1)\pi x}{2L}\right) dx$$

$$= \frac{1}{2} \left\{ \int_0^L \cos\left(\frac{((2n-1)-(2m-1))\pi x}{2L}\right) dx \right.$$

$$\left. + \int_0^L \cos\left(\frac{((2n-1)+(2m-1))\pi x}{2L}\right) dx \right\}$$

$$= \frac{1}{2} \left\{ \int_0^L \cos\left(\frac{(n-m)\pi x}{L}\right) dx + \int_0^L \cos\left(\frac{(m+n-1)\pi x}{L}\right) dx \right\}$$

$$= \begin{cases} \frac{1}{2}\left\{ \frac{L}{(n-m)\pi} \sin\left(\frac{(n-m)\pi x}{L}\right)\Big|_0^L + \frac{L}{(m+n-1)\pi} \sin\left(\frac{(m+n-1)\pi x}{L}\right)\Big|_0^L \right\} & , \quad m \neq n \quad , \\[2ex] \frac{1}{2}\left\{ x\Big|_0^L + \frac{L}{(m+n-1)\pi} \sin\left(\frac{(m+n-1)\pi x}{L}\right)\Big|_0^L \right\} & , \quad m = n \quad , \end{cases}$$

$$= \begin{cases} 0 & , \quad m \neq n \quad , \\ \frac{1}{2}L & , \quad m = n \end{cases}$$

since m and n are integers! (Note we must assume $m, n \geq 1$ in order to compute the second antiderivative as we did.) Thus, the eigenfunctions **are** orthogonal. (Later we shall show this result was *predictable* based solely on the type of differential equation and boundary conditions satisfied by the $X_n(x)$! But for now we view it as just a fortuitous occurrence.)

Now that we know the eigenfunctions are orthogonal, we should be able to proceed just as in our derivation of the original Fourier series coefficient formulas. That is, to multiply both sides of (3.8.48) by $\sin\left(\frac{(2m-1)\pi x}{2L}\right)$, where m denotes some fixed integer, and integrate from 0 to L. If we do this, taking the $\sin\left(\frac{(2m-1)\pi x}{2L}\right)$ inside the summation sign in the series and interchanging the order of integration and summation, we would have

$$\int_0^L f(x) \ \sin\left(\frac{(2m-1)\pi x}{2L}\right) dx =$$

$$\sum_{n=1}^{\infty} A_n \int_0^L \sin\left(\frac{(2n-1)\pi x}{2L}\right) \sin\left(\frac{(2m-1)\pi x}{2L}\right) dx$$

But, of course, the summation notation indicates that the term on the right should be evaluated for every integer value of n starting with 1, and then summed. Yet, because of orthogonality, this evaluation will result in zero for every value of n except the *one* value $m = n$. Thus the sum reduces to only a single term, specifically the term obtained when $m = n$, and the equation becomes

$$\int_0^L f(x) \sin\left(\frac{(2m-1)\pi x}{2L}\right) dx = \frac{L}{2} A_m \quad,$$

or,

$$A_m = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{(2m-1)\pi x}{2L}\right) dx \quad,$$

Using the A_n computed using this formula, we expect our solution (3.8.47) will satisfy the first initial condition in (3.8.48). By similar reasoning, the second condition in (3.8.48) leads to

$$A_n \quad = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{(2n-1)\pi x}{2L}\right) dx$$

(3.8.49)

$$B_n \quad = \frac{4}{(2n-1)\pi c} \int_0^L g(x) \sin\left(\frac{(2n-1)\pi x}{2L}\right) dx \quad,$$

and so we have solved this problem.

Actually, this last statement is a bit premature. What we have by using (3.8.49) to compute the coefficients is really only a so-called formal solution. That is it is a solution computed formally, based on assumptions that certain mathematical steps were really valid. To be sure that we really have the solution, there are two "minor" points that should be addressed. (These "minor" points worried mathematiciann for many years). First of all, just above, we blandly interchanged summation and integration in an infinite series. Yet many examples exist that show this is not valid for *any arbitrary series.* In fact, generally speaking, a series needs to be uniformly convergent before such interchange is permitted, although it does work for **some** series that do not converge uniformly. Specifically, the rather elegant theory of what are called generalized functions has shown that, for any Fourier or "Fourier-like" series (a term we leave purposely vague), such an interchange is valid. (The proof of this result is far beyond the scope of this course.) So, for the rest of this course, we need not worry about this concern.

The second problem deals with whether we can actually represent an arbitrary $f(x)$ and $g(x)$ (i.e., arbitrary initial conditions) using only the functions $\sin\left(\frac{(2n-1)\pi x}{2L}\right)$. In other words, do we have any guarantees that either equation in (3.8.48) really represents a valid equality! The answer, mathematically, is not obvious! The basic problem is that, unless we are dealing with a "standard" Fourier series, we have, at this point, no theory to assure us that $f(x)$ and $g(x)$ can acually be expanded in such series. But, as we noted above, if we try to look at (3.8.48) as an elementary Fourier series in terms of sines of $\frac{\pi x}{2L}$, we have only odd numbered terms - so it seems as though we have just half the functions we need. Is

this really enough? If not, then of course any conclusions we might draw, including (3.8.49), would be meaningless.

Perhaps an analogy may help to clarify our concern. Go back and look at (3.8.48) again. What we are trying to do there may be viewed as trying to construct ("build up") $f(x)$ and $g(x)$ as linear combinations, using the functions $\sin\left(\frac{(2m-1)\pi x}{2L}\right)$ as building blocks, or *components*. But isn't this exactly what one tries to do in linear algebra, when they write an arbitrary vector $\mathbf{v}$ in terms of some other set of vectors, i.e. when we try to write

$$\mathbf{v} = a_1\mathbf{v}^{(1)} + a_2\mathbf{v}^{(2)} + \cdots + a_n\mathbf{v}^{(n)} \qquad ?$$

And we should recall from linear algebra that there are some conditions under which such a representation may not be possible. For example, consider, in "ordinary" three-dimensional space, the vectors

$$\mathbf{v}^{(1)} = (1, 1, 0) \quad \text{and} \quad \mathbf{v}^{(2)} = (1, -1, 1) \ .$$

They are easily shown to be orthogonal under the usual dot product for R^3. So suppose we try to write the vector $(3,1,2)$ in terms of them. We would have

$$(3, 1, 2) = a_1(1, 1, 0) + a_2(1, -1, 1).$$

Taking dot products on both sides and using orthogonality leads to

$$a_1 = \frac{(3, 1, 2) \cdot (1, 1, 0)}{(1, 1, 0) \cdot (1, 1, 0)} = 2$$

and

$$a_2 = \frac{(1, -1, 1) \cdot (1, -1, 1)}{(1, -1, 1) \cdot (1, -1, 1)} = \frac{4}{3}$$

Then, if our analysis were correct, we should have:

$$(3, 1, 2) = 2\,(1, 1, 0) + \frac{4}{3}\,(1, -1, 1) = (\frac{10}{3}, \frac{2}{3}, \frac{4}{3}) \quad ??????$$

What happened? We computed the coefficients correctly, didn't we. The answer is simple enough. Our vectors were orthogonal, but we used them to try to write a vector in R^3, and R^3 is a vector space of dimension three. Thus, a basis for R^3 requires three vectors, and we has but two. (In linear algebra terminology, our basis was not *complete*.) Thus, our original equality was invalid, even though we were able to compute coefficients from it. But, of course, in R^3 this is obvious, so why all the fuss?

Well, go back again and look at (3.8.48) , which says

$$f(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{(2n-1)\pi x}{2L}\right) \ .$$

Here we are trying to write a function ("vector"), $f(x)$, in terms of the orthogonal functions ("vectors"), $\sin\left(\frac{(2n-1)\pi x}{L}\right)$. But apparently, if our analogy is valid, the dimension of this "vector space" is *infinite*! Hence, it's no longer *obvious* that we have enough functions

("vectors") in our "basis," is it? In fact, we've just argued that it looks like we have left out half of them already.

Well, it turns out that we do have enough - our "basis" is complete. The reason, interestingly enough, lies in the type of differential equation and boundary conditions that generated the eigenfunctions. We can show that, as long as we include all the eigenfunctions that are appropriate to the "right kind" of differential equation and boundary conditions, then we are assured of complete "basis." (We shall comment more on this when we discuss Sturm[9]-Liouville[10] problems. At this point, we will simply proceed on the assumption our basis is complete. One could argue that this is not really right. We should come to an immediate halt until we have rigorously proved all these wild assertions. Or should we? Commonly, in applied mathematics, one does not at first attempt to justify rigorously every intermediate step. Instead, one works problems through under the assumption that those operations that appear reasonable can be later justified. This is done to see whether or not in the "best" case one's method will work. For, if one can't find at least a formal solution in the "best" case, why worry whether all the steps are valid?

We shall close this case with an example. Consider

$$
\begin{aligned}
u_{tt} &= 4 \quad u_{xx} \quad, \quad 0 < x < 3 \,, \quad 0 < t \\
u(0, t) &= 0 \\
u_x(3, t) &= 0 \\
u(x, 0) &= 0 \\
u_t(x, 0) &= 1
\end{aligned}
$$

(Note $c^2 = 4$.) The general solution to this problem, according to our earlier development is

$$
u(x, t) = \sum_{n=1}^{\infty} \left[A_n \cos \left(\frac{(2n-1)\pi t}{3} \right) + B_n \sin \left(\frac{(2n-1)\pi t}{3} \right) \right] \sin \left(\frac{(2n-1)\pi x}{6} \right) \quad,
$$

and the initial condition equations reduce to

$$
u(x, 0) = \sum_{n=1}^{\infty} A_n \sin \left(\frac{(2n-1)\pi x}{6} \right) = 0
$$

$$
u_t(x, 0) = \sum_{n=1}^{\infty} \frac{(2n-1)\pi}{3} B_n \sin \left(\frac{(2n-1)\pi x}{6} \right) = 1 \quad.
$$

Therefore

$$
A_n \equiv 0 \quad,
$$

and

$$
B_n = \frac{3}{(2n-1)\pi} \frac{2}{3} \int_0^3 \sin \left(\frac{(2n-1)\pi x}{6} \right) dx = \frac{12}{(2n-1)^2 \pi^2} \quad,
$$

[9] Jacques Charles François Sturm,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Sturm.html
[10] Joseph Liouville,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Liouville.html

and so
$$u(x,t) = \sum_{n=1}^{\infty} \frac{12}{(2n-1)^2\pi^2} \sin\left(\frac{(2n-1)\pi t}{3}\right) \sin\left(\frac{(2n-1)\pi x}{6}\right) \quad.$$

In summary then, in this example we also found that separation of variables led to a solution, although the eigenvalues and eigenfunctions again differed from the earlier examples. Furthermore, the initial condition here reduced to a series which, while not an elementary Fourier series, was closely akin to one in that the (eigen)functions involved were orthogonal. In fact, it seems that all we may be doing in these problems is just "changing the basis", depending on the particular problem. Perhaps, in a loose way of speaking, separation of variables may be nothing more than picking that "orthogonal basis" of functions which is most natural to the particular boundary value problem at hand.

PROBLEMS

1. Solve:

$$u_{tt} = u_{xx}$$
$$u(0, t) = 0$$
$$u_x(2, t) = 0$$
$$u(x, 0) = \begin{cases} x & , \quad 0 < x \leq 1 \\ 1 & , \quad 1 < x < 2 \end{cases}$$
$$u_t(x, 0) = 0$$

2. Solve:

$$u_{tt} = 4u_{xx}$$
$$u_x(0, t) = 0$$
$$u(1, t) = 0$$
$$u(x, 0) = \begin{cases} 1 & , 0 < x \leq 1/2 \\ 2 - 2x & , 1/2 < x < 1 \end{cases}$$
$$u_t(x, 0) = 0$$

3. Solve:

$$u_{tt} = 9u_{xx}$$
$$u_x(0, t) = u(2, t) = 0$$
$$u(x, 0) = 0$$
$$u_t(x, 0) = (2 - x), 0 < x < 2$$

4. Show that the "normal" Fourier series reduces to

$$f(x) = \sum_{n=1}^{\infty} b_{2n-1} \sin\left(\frac{(2n-1)\pi x}{2L}\right)$$
$$= b_1 \sin\left(\frac{\pi x}{2L}\right) + b_3 \sin\left(\frac{3\pi x}{2L}\right) + b_5 \sin\left(\frac{5\pi x}{2L}\right) + \cdots$$

provided:

    a. $f(x)$ is odd,

    b. $f(x)$ is periodic of period $4L$, and

    c. $f(x + L) = f(L - x)$, $0 < x < L$

3.9   Generalizations on the Method of Separation of Variables

Recall that our goal in this chapter has been to develop a reasonable "feel" for the basic principles ("rules of the game") that apply to partial differential equations by studying special cases of the one-dimensional wave equation. At this point (as the reader is hopefully beginning to suspect), we may have reached the point of diminishing returns in terms of additional general information we can extract from the further study of special cases. In fact, we already appear to have "reinvented of the wheel" a number of times. Specifically:

(i) In cases characterized by "usual" physics, physically unrealistic, negative eigenvalues ($\lambda < 0$) have not occurred, although we have had to prove this in each case by completely solving the appropriate ordinary differential equation, with boundary conditions.

(ii) In every case, there have been an infinite number of eigenvalues and associated eigenfunctions.

(iii) In every case, the eigenfunctions have been orthogonal on the interval of interest, i.e. they have satisfied

$$\int_0^L X_n(x)X_m(x)dx = 0 \quad ,$$

whenever $X_n(x)$ and $X_m(x)$ were eigenfunctions for two different eigenvalues, although in the mixed end conditions case we had to prove this fact by "brute force" integration

(iv) In every example we have worked, the eigenfunctions appeared to be complete in the sense that, at least numerically, the formal solutions seemed able to satisfy whatever initial conditions we had (although we did not mathematically prove this in the mixed end conditions case).

With at least some aspects of these problems now seeming fairly predictable, we will stop solving further special cases of (3.5.11) at this time. Instead, we'll turn to investigating the degree to which the above patterns in fact reflect some *general theory* that underlies all of these cases, i.e. what aspects of the above behavior should we have been able to *predict*, and what aspects will we have to determine individually in any case. (For example, the cases we have studied so far indicate that we should not expect any general theory will predict the *specific* eigenvalues and eigenfunctions in advance, since these changed with each different problem. This shouldn't be completely surprising either. The same situation occurs in ordinary differential equations, where we know that

$$ay'' + by' + cy = 0 \quad ,$$

will have at least some solutions of the form $y = e^{rx}$, but we can only determine the actual values of r on a problem-by-problem basis.)

We start with the observation that the cases and exercises we've already studied suggest (at least for uniform media) that separation of variables "works" whenever both the basic

partial differential equation and boundary conditions are homogeneous. To substantiate this apparent "rule," we return to the most general form of the homogeneous one-dimensional wave equation without initial conditions

$$\rho(x)u_{tt} = \frac{\partial}{\partial x}\left[\tau(x)\frac{\partial u}{\partial x}\right] - q(x)u \quad , \quad 0 < x < L , \quad 0 < t$$

$$\alpha_1 u(0,t) - \beta_1 u_x(0,t) \qquad\qquad = 0 \qquad\qquad (3.9.50)$$

$$\alpha_2 u(L,t) + \beta_2 u_x(L,t) \qquad\qquad = 0 \quad .$$

(This form is more general than the cases we've studied, but still is linear and homogeneous. The added $-q(x)u$ term in this partial differential equation was not addressed in our earlier derivation of the wave equation, but could physically represent a series of springs distributed down the length of the original string. The other coefficients, of course, connote variable density and/or tension.) The fundamental consideration here, as it was in all of our special cases, will be to identify as much as possible about the linearly independent solutions of this problem.

Observe that the partial differential equation in (3.9.50), while significantly more involved than the ones we have solved thus far, still possesses the same natural splitting that we observed in the simplest case

$$u_{tt} = c^2 u_{xx} \quad ,$$

in that the left-hand side differentiates only the time behavior, while the right-hand side differentiates only the spatial behavior. So, just as we have all along, we now look for linearly independent solutions in the form:

$$u(x,t) = X(x)T(t) \quad .$$

Standard differentiation rules, applied when the product form is substituted into the differential equation, yield

$$\rho(x)XT'' = T\frac{d}{dx}\left[\tau(x)\frac{dX}{dx}\right] - q(x)XT \quad ,$$

or, upon simplification

$$\frac{T''}{T} = \frac{[\tau(x)X']' - q(x)X}{\rho(x)X} = -\lambda \quad ,$$

Here again, we may introduce the separation constant since the left-hand side is clearly independent of x and the right-hand side independent of t. (Note there is one slight difference between this formulation and the one we have used in our examples. Specifically, the c^2 term which appears in the constant coefficient case has been absorbed into λ here.)

A brief calculation shows that we can always separate the boundary conditions in (3.9.50), in the same manner as the examples, e.g.

$$\begin{aligned}
\alpha_1 u(0,t) - \beta_1 u_x(0,t) &= \alpha_1 X(0)T(t) - \beta_1 X'(0)T(t) \\
&= T(t)\left[\alpha_1 X(0) - \beta_1 X'(0)\right] = 0 , \\
&\Rightarrow \alpha_1 X(0) - \beta_1 X'(0) = 0 ,
\end{aligned}$$

and similarly at the right-hand end. Combining all of these results shows that the original problem separates into

$$T'' + \lambda T = 0 \qquad \begin{aligned} [\tau(x)X']' + [\lambda\rho(x) - q(x)]X &= 0 \\ \alpha_1 X(0) - \beta_1 X'(0) &= 0 \\ \alpha_2 X(L) + \beta_2 X'(L) &= 0 \ . \end{aligned} \qquad (3.9.51)$$

Here, as was true in all of our special cases, the problem for $X(x)$ is a second-order ordinary differential equation, with two auxiliary conditions and involving an unspecified constant (λ). *Normally*, therefore, we should expect only a unique (and hence trivial) solution for $X(x)$. The problem here is thus, as it was before, to determine the values of λ for which nontrivial solutions for $X(x)$ also exist. (Of course, $T(t)$ is also the solution of an ordinary differential equation. But its equation is still constant coefficient, without boundary or initial conditions, and so we can immediately write its general solution, once we know the value or values of λ for which non-trivial solutions for $X(x)$ can be found.)

For reasons that will hopefully become clearer later, we will choose to view this problem for $X(x)$ as a special case of another general differential equation

$$\begin{aligned} [p(x)y'(x)]' + [\lambda w(x) - q(x)]y(x) &= 0 \ , \quad 0 < x < L \\ \alpha_1 y(0) - \beta_1 y'(0) &= 0 \\ \alpha_2 y(L) + \beta_2 y'(L) &= 0 \end{aligned} \qquad .$$

(For the wave equation, the appropriate correspondence is $p(x) = \tau(x)$, $w(x) = \rho(x)$, and $y(x) = X(x)$. Physical considerations related to such interpretations will motivate additional restrictions on the coefficients in this problem. For example, because of their interpretation in terms of the one-dimensional wave equation, we expect $p(x)$ and $w(x)$ will "normally" be positive, and $q(x)$ nonnegative.) This differential equation, with its attendant boundary conditions and certain other restrictions, constitutes what is known as a *Sturm-Liouville problem*. Since it's linear and homogeneous, we know the trivial solution, $y(x) = 0$, always satisfies the problem, and non-trivial solutions (if any) are determined only up to arbitrary multiplicative constants. So, again, our focus will be what are its non-trivial solutions; when do they arise; and what are their general properties?

Before proceeding further, we must emphasize that there is nothing magical about the Sturm-Liouville problem. It's really just a template against which we match candidate problems. For example, every special case we have solved thus far in this chapter involved the ordinary differential equation

$$X'' + \lambda X = 0 \ ,$$

which is in fact a special case of the Sturm-Liouville differential equation corresponding to

$$p(x) = 1 \ , \quad w(x) = 1 \ , \quad \text{and} \quad q(x) = 0 \ .$$

The importance of Sturm-Liouville theory is that once we can identify, for any particular problem we are solving, a choice of $p(x)$, $q(x)$ and $w(x)$, plus α_i and β_i that reduces the Sturm-Liouville form to that specific problem, then we shall see that there are a number

of conclusions we can immediately draw. (Analogously, in ordinary differential equations, when we recognize that

$$y'' + 3y' + 2y = 0$$

is a special case of the constant coefficient equation,

$$ay'' + by' + cy = 0 \ ,$$

we can immediately conclude that two linearly independent solutions exist, and that $y = e^{rx}$ is a solution for those values of r which satisfy $r^2 + 3r + 2 = 0$.) Having these conclusions should save us work, since in the future we will know more completely what behavior to expect.

There is, unfortunately, one result which the study of the Sturm-Liouville will not provide. There simply is no general formula that expresses the solution for the Sturm-Liouville ordinary differential equation, even when explicit formulas for $p(x)$, $q(x)$ and $w(x)$ are given. (Your earlier study of ordinary differential equations should already have made this fairly clear.) We have really been quite lucky (so far) to have encountered only constant coefficient equations, one of the few instances for which closed form solution is possible. Nevertheless, it is truly amazing how much information about the general character of the solutions of this problem we will be able to extract without having a formula for the solutions and without knowing the specific coefficient functions, the specific values of the α_i and β_i, and without knowing the actual formula for the solution. We shall develop these results in the next section. At this point, we should simply recognize that the study of the Sturm-Liouville problem is important for us in that some special case of it will arise every time we try separation of variables on the homogeneous wave equation (3.9.50), or the heat equation or Laplace's equation (which we mention later). Therefore, any results we can derive about solutions of the Sturm-Liouville problem in general hold the prospect of greatly reducing our workload when we apply the method of separation of variables to specific cases.

3.10 Sturm-Liouville Theory

By definition, a Regular Sturm-Liouville problem consists of describing the existence and behavior of non-trivial solutions to

$$
\begin{aligned}
[p(x)y'(x)]' + [\lambda w(x) - q(x)]y(x) \ &= 0 \ , \ \ 0 < x < L \\
\alpha_1 y(0) - \ \beta_1 y'(0) &= 0 \\
\alpha_2 y(L) + \beta_2 y'(L) &= 0 \ \ \ ,
\end{aligned}
\tag{3.10.52}
$$

where,

$p'(x)$, $w(x)$ and $q(x)$ are continuous for $0 < x < L$, and

$p(x) > 0$ for $0 \le x \le L$.

$w(x) > 0$ for $0 < x < L$.

$\alpha_i^2 + \beta_i^2 > 0$ for $i = 1, 2$.

Our development of the wave equation and earlier discussions have already indicated why differential equations and boundary conditions of this general type are of interest. However a few more words on the origin of the restrictions on the coefficients are in order.

As we have already noted, in the actual wave equation, $p(x)$, $w(x)$ and $q(x)$ represent such physical quantities as tension, density, etc. Since discontinuous changes in tension, density, etc., inside a string are not likely, then the continuity restrictions certainly seem physically reasonable. Moreover, mathematically, because the Sturm-Liouville differential equation can be written

$$p(x)y''(x) + p'(x)y'(x) + [\lambda w(x) - q(x)]y(x) = 0 \quad,$$

the assumptions on continuity can be seen to be important to the existence of well-behaved solutions. However, continuity of the coefficients alone will not ensure well-behaved solutions. The additional requirement that $p(x) \neq 0$ on $0 < x < L$ precludes the existence of singular points in this interval and so avoid some nasty mathematical problems.

Furthermore, physically, for the wave equation, $p(x)$ represents the tension. But if the tension were to vanish somewhere inside the string, i. e., if $p(x) = 0$ at some point in $0 < x < L$, the physical effect would be the same as if the string were cut in two at that point. Lastly, negative tension at a point would correspond to an inward push on the string, collapsing it, and destroying our originally assumed model. Therefore, on physical grounds, the restriction that $p(x) > 0$ inside the string seems eminently reasonable. (For *strings*, $p(x) = 0$ at either end seems equally disturbing, although for other media and problems this situation can in fact occur. The situation $p(0) = 0$, as well as several other cases, lead to what are called *singular* Sturm-Liouville problems. Many of the qualitative properties of the solutions in this singular case are however similar to the those we shall derive for solutions to the regular Sturm-Liouville problem. Singular Sturm-Liouville problems also require some modifications in the types of boundary conditions which are acceptable at the singular point. For example, if $p(0) = 0$, then mathematically and physically, boundary conditions, such as

$$\alpha_1 y(0) - \beta_1 y'(0) = 0$$

are no longer really appropriate.)

We can also interpret the restrictions on $w(x)$ physically. For example, in the wave equation, $w(x)$ represents the density. Thus, requiring $w(x) > 0$ inside the string is clearly physically realistic. (In idealized situations, or other models, however, conditions such as $w(0) = 0$ or $w(L) = 0$ could be possible. For example, imagine a string which gets progressively thinner near an end and finally vanishes precisely at the end!)

Our first major result on the regular Sturm-Liouville problem will be to prove our earlier conjecture that $\lambda < 0$ almost never arises. In order to do this, however, we need some restrictions on the algebraic signs beyond those prescribed above. Specifically we need that

$$q(x) \geq 0 \ \text{ in } \ 0 < x < L, \qquad \text{and} \qquad \alpha_i \beta_i \geq 0 \ , i = 1, 2 \quad . \tag{3.10.53}$$

(As we have already noted that, in the wave equation, the $q(x)$ term can be interpreted as a spring distributed down the length of the string. Therefore, in order to satisfy the action

of normal springs, $q(x)$ must also be positive in the interval (or equal to zero if there is no such mechanism). The restriction on the signs of the α_i and β_i agrees with the signs that arise in the wave equation under normal physics.)

Assuming now that all of the necessary restrictions are satisfied, we let $y(x)$ denote any non-trivial solution to (3.10.52), without regard to how it might have been constructed. If we then multiply both sides of the differential equation by $y(x)$ and integrate from 0 to L. (noting that the right hand side remains identically zero), this yields, in order,

$$[p(x)y'(x)]'y(x) + [\lambda w(x) - q(x)]y^2(x) = 0 \ ,$$

and then

$$\int_0^L [p(x)y'(x)]'y(x)dx + \int_0^L [\lambda w(x) - q(x)]y^2(x)dx = 0 \ .$$

We can simplify the first integral here by integrating the $[p(x)y'(x)]'y(x)$ term by parts, and then moving all the terms not containing λ across the equal sign. Performing these calculations, and dropping the explicit reference to x inside the integrals, simplifies this last equation to:

$$\lambda \int_0^L wy^2 dx = \int_0^L p[y']^2 dx + \int_0^L qy^2 dx - p(x)y(x)y'(x)\Big|_0^L \qquad (3.10.54)$$

We now analyze the algebraic signs of the terms in this equation. Since $p(x)$ and $q(x)$ have both been assumed positive, then clearly the first two integrals on the right are non-negative. The third term, written out fully, is

$$-p(L)y(L)y'(L) + p(0)y(0)y'(0) \ .$$

However, according to the first boundary condition in (3.10.52):

$$\alpha_1 y(0) - \beta_1 y'(0) = 0, \quad \text{or} \quad \alpha_1 y(0) = \beta_1 y'(0) \ .$$

But if $\alpha_1\beta_1 \geq 0$, see (3.10.53), then clearly α_1 and β_1 cannot have *different* algebraic signs. (By convention, zero may have either sign.) But, if this is the case, then to satisfy the above equality, $y(0)$ and $y'(0)$ cannot have different signs either. Furthermore, by assumption, $p(x)$ is positive on the interval, including the end points. Therefore

$$p(0)y(0)y'(0) \geq 0 \ .$$

Similarly, we can show the boundary condition at x = L implies,

$$-p(L)y(L)y'(L) \geq 0 \ ,$$

and so

$$-p(x)y(x)y'(x)\Big|_0^L \geq 0 \ .$$

Thus *every term* on the right of (3.10.54) is non-negative. But, with $w(x) > 0$ and $y(x) \not\equiv 0$, we must also have

$$\int_0^L wy^2 dx > 0 \ ,$$

and therefore we have now shown that, for the terms contained in (3.10.54)

$$\lambda \underbrace{\int_0^L wy^2 dx}_{>\,0} = \underbrace{\int_0^L p[y']^2 dx}_{\geq\,0} + \underbrace{\int_0^L qy^2 dx}_{\geq\,0} \quad \underbrace{-p(x)y(x)y'(x)\Big|_0^L}_{\geq\,0} \quad,$$

and so, clearly, λ cannot be negative, i.e. $\lambda \geq 0$. This proves our first conjecture, i.e. that the occurrence of negative values of λ is impossible for the wave equation, under normal physics. As we've pointed out before, this seemed physically obvious, since otherwise we would have had exponentially growing solutions for $T(t)$. What is reassuring, in terms of the validity of our mathematical model, is that the mathematical formulation, via Sturm-Liouville theory, reaches the same conclusion.

However, we're not done yet! There's still more information we can extract from this last equation. For suppose $\lambda = 0$. Then the left-hand side is zero. But every term on the right is non-negative! This is possible *if and only if* every one of those terms is zero, i.e. if

$$\int_0^L p[y']^2 dx = \int_0^L qy^2 dx = -p(x)y(x)y'(x)\Big|_0^L = 0 \quad.$$

But $p(x) > 0$. Thus

$$\int_0^L p[y']^2 dx = 0 \Rightarrow [y']^2 = 0 \quad,$$

or $y(x)$ must be *a constant* (which cannot be zero if $y(x)$ is a non-trivial solution). Note this immediately guarantees that the third term, $-pyy'|_0^L$, is identically zero as well . However, if a non-zero constant is a solution to the entire problem, it must satisfy the Sturm-Liouville boundary conditions. But substituting any constant into these conditions immediately implies $\alpha_1 = \alpha_2 = 0$. Therefore, the original boundary conditions must have been:

$$\begin{aligned} u_x(0,t) &= 0 \\ u_x(L,t) &= 0 \end{aligned} \quad.$$

Lastly, if $y(x)$ is a non-zero constant and λ is zero, then the basic Sturm-Liouville differential equation in (3.10.52) reduces to,

$$-q(x)y(x) = 0 \Rightarrow \quad q(x) = 0 \quad.$$

This result verifies our conjecture that the appearance of the zero eigenvalue is unlikely. It also allows us to decide, *by inspection*, whether $\lambda = 0$ is an eigenvalue in any Sturm-Liouville problem. Furthermore, this test for the zero eigenvalue can be interpreted either mathematically or physically. For example, for the uniform string, this test can either be stated as the mathematical condition

The zero eigenvalue will exist if and only if $q(x) \equiv 0$ and $\alpha_1 = \alpha_2 = 0$

or interpreted as the physical statement:

The homogeneous wave equation with homogeneous boundary conditions will have a non-trivial, rigid body motion solutions, if and only if both ends of the string are free, and there are no internal springs.

Again, we see that the mathematical restrictions make physical "sense." This may seem quite amazing - this degree to which the mathematics and physics reinforce each other. It should also be quite reassuring, in that we can feel confident that the mathematics and physics are not disjoint, but are really working hand in hand, and our models appear to have captured the essence of the physics involved.

Mathematically, we can sum up the results so far with

Theorem 2.1 The regular Sturm-Liouville problem:

$$
\begin{aligned}
[p(x)y'(x)]' + [\lambda w(x) - q(x)]y(x) &= 0 , \quad 0 < x < L \\
\alpha_1 y(0) - \beta_1 y'(0) &= 0 \\
\alpha_2 y(L) + \beta_2 y'(L) &= 0 ,
\end{aligned}
$$

where,

$p'(x)$, $w(x)$ and $q(x)$ are continuous for $0 < x < L$,

$p(x) > 0$ and $q(x) \geq 0$ for $0 \leq x \leq L$ and $\alpha_i \beta_i \geq 0$, $i = 1, 2$

$w(x) > 0$ for $0 < x < L$

has non-trivial solutions only if $\lambda \geq 0$. Furthermore, solution for $\lambda = 0$ occurs if and only if

(a) $q(x) \equiv 0$

(b) $y(x) = \text{const} \neq 0$ satisfies the boundary conditions.

This theorem is, of course, somewhat negative in that it tells us certain choices for λ which will **not** produce eigenvalues, but, unfortunately, not which ones **will**. In fact, it does not even assure us that there will be any eigenvalues. (Although we have always found them in our examples.) Or, if so, how many. (We always found an infinite number.) Again, this should not surprise us, for, as we have already discussed, the values of the actual eigenvalues and eigenfunctions depend totally on the specific problem one is solving, and must be determined by solving that problem in detail. There are no short cuts! However, Sturm-Liouville theory still does yield some insights into both the number of eigenvalues and the qualitative properties of the associated eigenfunctions.

The following theorem, whose proof is far beyond the scope of this course, is one general result that Sturm-Liouville does provide about the eigenvalues and eigenfunctions for the general case:

Theorem 2.2 Given a regular Sturm-Liouville problem, then there exist an infinite number of eigenvalues λ_n, and eigenfunctions $y_n(x)$ for the problem. The eigenvalues $\lambda_n \to +\infty$ as $n \to \infty$, and each $y_n(x)$ has one more zero in the interval $0 < x < L$ than its predecessor.

While not proven here, this theorem is easy to believe based on our examples. For our first problem (3.6.12) the eigenvalues were $\lambda_n = (n\pi/L)^2$, The corresponding eigenfunctions

(which we happened to call $X_n(x)$ there) were $\sin(n\pi x/L)$, and have $(n-1)$ zeros in the interval $0 < x < L$. Our second problem (3.7.31) had the same positive eigenvalues, but eigenfunctions equal to $\cos(n\pi x/L)$, which have n zeros in the interval $0 < x < L$. Theorem 2.2 guarantees that solutions to any Sturm-Liouville problem will act the same general way.

We next turn to our conjecture about orthogonality, where we observed from the behavior in our earlier examples that it appeared as if separation of variables were producing an orthogonal set of functions ("basis") "natural" to a given problem. We now show that, in fact, this is effectively the case.

Suppose λ_n and λ_m denote two *different* eigenvalues of

$$
\begin{aligned}
[p(x)y'(x)]' + [\lambda w(x) - q(x)]y(x) &= 0 , \quad 0 < x < L \\
\alpha_1 y(0) - \beta_1 y'(0) &= 0 \\
\alpha_2 y(L) + \beta_2 y'(L) &= 0 ,
\end{aligned}
$$

where,

$$p'(x), \ w(x) \text{ and } q(x) \text{ are continuous for } 0 < x < L,$$

$$p(x) > 0 \text{ and } w(x) > 0 \text{ for } 0 < x < L.$$

(Note we are not using *all* of the Sturm-Liouville restrictions.) By definition, associated with each eigenvalue will be some non-trivial solution, which we shall denote by $y_n(x)$ and $y_m(x)$, respectively. Thus, we have, since each is a solution

$$
\begin{aligned}
[p(x)y_n'(x)]' + [\lambda_n w(x) - q(x)]y_n(x) &= 0 \\
[p(x)y_m'(x)]' + [\lambda_m w(x) - q(x)]y_m(x) &= 0 ,
\end{aligned}
$$

and also

$$
\begin{aligned}
\alpha_1 y_n(0) - \beta_1 y_n'(0) &= 0 \\
\alpha_2 y_n(L) + \beta_2 y_n'(L) &= 0 , \\
\alpha_1 y_m(0) - \beta_1 y_m'(0) &= 0 \\
\alpha_2 y_m(L) + \beta_2 y_m'(L) &= 0 ,
\end{aligned}
\tag{3.10.55}
$$

But if we multiply the upper equation in (3.10.55) by $y_m(x)$ and the lower by $y_n(x)$, and then subtract, we have

$$[p(x)y_n'(x)]'y_m(x) - [p(x)y_m'(x)]'y_n(x) + (\lambda_n - \lambda_m)w(x)y_n(x)y_m(x) = 0 .$$

Integrating from 0 to L and dropping the explicit reference to dependence on x, yields

$$\int_0^L \{[py_n']'y_m - [py_m']'y_n\} \, dx + (\lambda_n - \lambda_m) \int_0^L w y_n y_m dx = 0 , \tag{3.10.56}$$

But note the following identity can be verified

$$
\begin{aligned}
\frac{d}{dx}[p(y_n'y_m - y_m'y_n)] &= [py_n']y_m' + [py_n']'y_m - [py_m']y_n' - [py_m']'y_n \\[2mm]
&= [py_n']'y_m - [py_m']'y_n ,
\end{aligned}
$$

125

since the first and third terms cancel. Thus the first integral in (3.10.56) has an exact antiderivative, and so,

$$\int_0^L \{[py_n']'y_m - [py_m']'y_n\}\, dx = [p(y_n'y_m - y_ny_m')]\Big|_0^L$$

$$= p(L)[y_n'(L)y_m(L) - y_n(L)y_m'(L)]- \tag{3.10.57}$$

$$p(0)[y_n'(0)y_m(0) - y_n'(0)y_m'(0)] \quad .$$

But, since both $y_n(x)$ and $y_m(x)$ are solutions to the boundary conditions, as well as to the differential equation, then we must also have

$$\alpha_2 y_n(L) + \beta_2 y_n'(L) = 0$$
$$\alpha_2 y_m(L) + \beta_2 y_m'(L) = 0 \quad .$$

Now wait! This is simply two *homogeneous* algebraic equations in two unknowns - α_2 and β_2. Furthermore, we have assumed that α_2 and β_2 are not *both* zero! (Otherwise, we wouldn't have a boundary condition in the original problem at $x = L$.) Therefore, by Cramer's rule, the determinant of the coefficients must be zero, i.e.

$$\begin{vmatrix} y_n(L) & y_n'(L) \\ y_m(L) & y_m'(L) \end{vmatrix} \equiv y_n(L)y_m'(L) - y_n'(L)y_m(L) = 0 \quad .$$

(In ordinary differential equation, this determinant is called the *Wronskian*.) But, except for a minus sign, this determinant is exactly the first term in brackets on the right hand side in (3.10.57). Therefore, since minus zero is still zero, we have

$$y_n'y_m - y_ny_m' = 0 \quad \text{at} \quad x = L.$$

Similarly, the boundary condition at $x = 0$ implies

$$y_n'y_m - y_ny_m' = 0 \quad \text{at} \quad x = 0.$$

Thus,

$$\int_0^L [y_m[p(x)y_n'(x)]' - y_n[p(x)y_m'(x)]']\, dx \equiv [p(y_n'y_m - y_ny_m')]\Big|_0^L = 0 \quad ,$$

i.e. the first integral in (3.10.56) must vanish. Hence

$$(\lambda_n - \lambda_m)\int_0^L wy_ny_m dx = 0.$$

But wait again! We said λ_n and λ_m were different. Thus

$$(\lambda_n - \lambda_m) \neq 0 \quad ,$$

and therefore

$$\int_0^L w(x)y_n(x)y_m(x)dx = 0 \quad . \tag{3.10.58}$$

126

So our two solutions, $y_n(x)$ and $y_m(x)$ are orthogonal! Or are they? What about the $w(x)$ term in this integral.

The relationship expressed in (3.10.58) is, in fact, a type of orthogonality. It's called orthogonality with respect to a *weighting function* - in this case $w(x)$. We will show shortly that this orthogonality provides all of the same benefits in terms of being able to compute coefficients that we observed in the simpler Fourier series orthogonality. For Sturm-Liouville problems in which $w(x) \neq 1$, the $w(x)$ term in (3.10.58) is in fact absolutely necessary, since for these problems the "other" integral,

$$\int_0^L y_n(x)y_m(x)dx \tag{3.10.59}$$

will not provide the correct relationships. (Although we shall not develop this aspect of the theory further at this point, we could show, at a quite higher mathematical level, that what we are essentially doing here is defining a "dot product" between functions by

$$f \cdot g = \int_0^L w(x)f(x)g(x)dx, \quad . \tag{3.10.60}$$

Here, mathematically, is where the restriction that $w(x) > 0$ in $0 < x < L$ becomes pivotal, for we know that for any "dot product", one vital property that must hold is

$$f \cdot f > 0 \quad \text{unless} \quad f = 0.$$

But if (3.10.60) is really, in some sense, analogous to a dot product, then, by definition,

$$f \cdot f = \int_0^L w(x)[f(x)]^2 dx \quad ,$$

and we see $w(x) > 0$ in $0 < x < L$ is necessary to guarantee that this integral is positive unless $f(x)$ is identically zero.)

We can summarize this result as

Theorem 2.3 If $y_n(x)$ and $y_m(x)$ are solutions of the regular Sturm-Liouville problem (3.10.52) corresponding to two different values of λ, then

$$\int_0^L w(x)y_n(x)y_m(x)dx = 0.$$

Note, interestingly enough, that for this result, we did not need to assume either $q(x) > 0$ or $\alpha_i \beta_i \geq 0, i = l, 2$. Thus, orthogonality holds even in the face of negative eigenvalues. Also, note, as we observed earlier, each of our specific examples of the one-dimensional wave equation dealt with the differential equation

$$X'' + \lambda X = 0 \quad ,$$

which, in terms of the Sturm-Liouville form, is equivalent to

$$y(x) = X(x) \quad , \quad p(x) = w(x) = 1 \quad , \text{and} \quad q(x) = 0 \quad .$$

This explains why the Fourier orthogonality integral always worked in our examples. According to our theorem, however, (3.10.59) apparently will not be the proper integral in cases where $w(x)$ is not a constant.

There is one final aspect of the Sturm-Liouville problem we wish to address here. This last item will not, by any means, exhaust the study of this quite fascinating equation. However, we shall by then have covered the high points, and those points of primary practical importance in solving partial differential equations by the separation of variables/eigenvalue-eigenfunction approach. This last item addresses the question of completeness. As we have seen in our example cases, we often wish to be able to write any reasonable function (e.g., the initial conditions) as an infinite series, using the eigenfunctions as a "basis." For example, this occurs because looking for the linearly independent solutions of

$$\rho(x)u_{tt} = \frac{\partial}{\partial x}\left[\tau(x)\frac{\partial u}{\partial x}\right] - q(x)u \quad , \quad 0 < x < L , \quad 0 < t$$

$$\begin{aligned} \alpha_1 u(0,t) - \beta_1 u_x(0,t) &= 0 \\ \alpha_2 u(L,t) + \beta_2 u_x(L,t) &= 0 \\ u(x,0) &= f(x) \\ u_t(x,0) &= g(x) \end{aligned} \tag{3.10.61}$$

by assuming

$$u(x,t) = X(x)T(t)$$

leads to, as we have seen, a regular Sturm-Liouville problem for $X(x)$, with orthogonal eigenfunctions

$$X_n(x) \quad , \quad n = 1, 2, \ldots \quad ,$$

$$\int_0^L w(x)X_n(x) \quad X_m(x)dx = 0 \quad , \quad m \neq n$$

while the $T(t)$ must satisfy the equation

$$T_n'' + \lambda_n T_n = 0 \quad .$$

Assuming only positive values of λ occur,

$$T_n(t) = A_n \cos\left(\sqrt{\lambda_n}t\right) + B_n \sin\left(\sqrt{\lambda_n}t\right) \quad , \tag{3.10.62}$$

and therefore the general solution to (3.10.61) can be formally written

$$u(x,t) = \sum_{n=1}^{\infty}\left\{A_n \cos\left(\sqrt{\lambda_n}t\right) + B_n \sin\left(\sqrt{\lambda_n}t\right)\right\} X_n(x) \quad . \tag{3.10.63}$$

Trying to fit the initial conditions in (3.10.61) with this solution leads to:

$$\begin{aligned} u(x,0) &= \sum_{n=1}^{\infty} A_n X_n(x) = f(x) \quad , \\ u_t(x,0) &= \sum_{n=1}^{\infty} \sqrt{\lambda_n}B_n X_n(x) = g(x) \quad . \end{aligned} \tag{3.10.64}$$

128

Then, since the $X_n(x)$ are orthogonal, we could, in the usual way, multiply both sides of the first equation in (3.10.64) by $w(x)X_m(x)$, where m denotes some fixed integer, and then integrate, to yield

$$\int_0^L w(x)f(x)X_m(x)dx = \sum_{n=1}^\infty A_n \underbrace{\int_0^L w(x)X_n(x)X_m(x)dx}_{= \ 0 \ , \ \ m\neq n},$$

and so,

$$\int_0^L w(x)f(x)X_m(x)dx = A_m \int_0^L w(x)\left[X_m(x)\right]^2 dx,$$

or

$$A_n = \frac{\int_0^L w(x)f(x)X_n(x)dx}{\int_0^L w(x)\left[X_n(x)\right]^2 dx}. \tag{3.10.65}$$

So we can formally compute the A_n! (A similar set of steps would produce formulas for the B_n.) But, mathematically, the question remains of whether we can ensure (3.10.64) is really a valid set of equalities. We have seen by example that simply using orthogonality to compute coefficients does not constitute a validation of the original equality. And as we have commented, just because we have an infinite number of functions may not guarantee that they are enough to represent an arbitrary $u(x,t)$. Also, as we have seen, either of the initial conditions in (3.10.61) may be discontinuous, at least in terms of their periodic extensions. So the right-hand sides in (3.10.64) may be discontinuous. But all the functions in the series on the left-hand side are continuous. So in what sense do these series converge to $f(x)$ and $g(x)$? Can we guarantee that we can always find a formal solution like (3.10.62) and that computing the coefficients using (3.10.65) will be valid?

The answer is yes, provided the $X_n(x)$ include *all* of the linearly independent eigenfunctions of the associated Sturm-Liouville problem which comes from separation of variables. We shall not prove this statement (such proof is far outside the scope of this course), but only present the appropriate theorem.

Theorem 2.4 If $f(x)$ is at least piecewise smooth on $0 < x < L$, and if $X_n(x)$, $n = 1, 2, \ldots$ are the full set of eigenfunctions to a regular Sturm-Liouville problem (3.10.52), then the series

$$f(x) = \sum_{n=1}^\infty A_n X_n(x) \ ,$$

where the A_n are determined by

$$A_n = \frac{\int_0^L w(x)f(x)X_n(x)dx}{\int_0^L w(x)\left[X_n(x)\right]^2 dx},$$

converges pointwise to $f(x)$ at any point where f is continuous, and to the mean (average) value where f is discontinuous.

(Since we do not expect to encounter functions at this level which are not at least piecewise smooth, we shall not discuss what occurs for functions that are less than piecewise smooth.)

This theorem concludes our discussion of the Sturm-Liouville problem. In closing, we would again emphasize that we have not exhausted the results known about this fascinating problem. However, we have covered the "high points", and those most relevant to the problems we shall cover.

PROBLEMS

1. For each of the following problems, determine if the given equation is in Sturm-Liouville form, and, if so, identify the values of the appropriate function $p(x)$, $q(x)$, $w(x)$, and the values of α_i and β_i:

    a. $[(x+1)y']' + \lambda(x+1)y - y = 0$

$$y(1) = 0$$
$$y(2) = 0$$

    b. $[(x^2 - 1)u']' + 3\lambda u = 0$

$$u(0) = 0$$
$$u(1/2) = 0$$

    c. $y'' + \lambda x y = 0$

$$y(0) = 0$$
$$y(3) + 2y'(3) = 0$$

    d. $y'' + xy' + \lambda y = 0$

$$y(0) = 0$$
$$y(1) = 0$$

2. Following similar steps to those used in class, show that the eigenfunctions of the singular Sturm-Liouville problem:

$$[p(x)y']' + \lambda w(x)y + q(x)y = 0 \quad , \quad a < x < b$$
$$y(a), y'(a) \text{ finite} \quad ,$$
$$\alpha_2 y(b) + \beta_2 y'(b) = 0$$

where $p'(x)$, $w(x)$, and $q(x)$ are continuous, and

$$p(x) > 0 \quad , \quad a < x \leq b \quad ,$$
$$p(a) = 0$$
$$w(x) > 0 \quad , \quad a < x < b$$
$$q(x) \geq 0 \quad , \quad a \leq x \leq b$$

corresponding to different eigenvalues are orthogonal with respect to the weighting function $w(x)$.

131

3.11 The Frequency Domain Interpretation of the Wave Equation

With the analysis of the Sturm-Liouville problem, we have in essence completely described the *mechanics* of solving the one-dimensional wave equation. By now, hopefully, you are beginning to feel at least somewhat comfortable in carrying out the steps necessary to generate the series solution to any one-dimensional, constant-coefficient wave equation. However, in mathematics, stopping with a grasp of merely the mechanics of the steps which generate solutions is rarely advisable. We don't mean that these mechanics are unimportant. Quite to the contrary! But a mastery of **only** the computations in a problem misses a fundamental purpose of mathematics - to provide *insights* beyond just the numbers that may emerge in various solutions. In this section and the following one, therefore, we seek to go beyond the mechanics of separation of variables in the one-dimensional wave equation, and bring to light some (more) fundamental insights about the nature of waves. We shall see that these insights are actually embodied in our one-dimensional string model, but have been obscured up until now by the rather imposing notation of the infinite series produced by separation of variables.

We start, as it were, back at the very beginning, with a special case of the first wave equation we solved - the string with fixed ends, an initial displacement, but no initial velocity. As we now know, the solution to this problem is

$$u(x,t) = \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi ct}{L}\right) \sin\left(\frac{n\pi x}{L}\right)$$

where,

$$A_n = \frac{2}{L} \int_0^L u(x,0) \sin\left(\frac{n\pi x}{L}\right) dx \ ,$$

and the eigenvalues and eigenfunctions are

$$\begin{aligned} \lambda_n &= \left(\frac{n\pi}{L}\right)^2 \\ X_n(x) &= \sin\left(\frac{n\pi x}{L}\right) \end{aligned} \quad , \quad n = 1, 2, 3, \dots \ ,$$

respectively. Observe however that a simple interchange of the order in which the functions within the terms within this series appear allows us to rewrite this solution as

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} \left[A_n \sin\left(\frac{n\pi x}{L}\right) \right] \cos\left(\frac{n\pi ct}{L}\right) \\ &= \left[A_1 \sin\left(\frac{\pi x}{L}\right) \right] \cos\left(\frac{\pi ct}{L}\right) + \left[A_2 \sin\left(\frac{2\pi x}{L}\right) \right] \cos\left(\frac{2\pi ct}{L}\right) \\ &\quad + \left[A_3 \sin\left(\frac{3\pi x}{L}\right) \right] \cos\left(\frac{3\pi ct}{L}\right) + \cdots \ , \\ &= A_1(x) \cos\left(\frac{\pi ct}{L}\right) + A_2(x) \cos\left(\frac{2\pi ct}{L}\right) + A_3(x) \cos\left(\frac{3\pi ct}{L}\right) + \cdots \ , \end{aligned} \qquad (3.11.66)$$

where

$$A_n(x) = \left[A_n \sin\left(\frac{n\pi x}{L}\right) \right] \ .$$

132

What, if any, new information or viewpoint does this alternative formulation provide?

Well, the rewritten form clearly still represents a sum (superposition) of different terms. However, now recall that for any ω, $\cos(\omega t)$ represents a *single temporal vibration* (i.e. a "pure" tone), such as one would hear from a tuning fork or electronic tuner. Therefore, we may view each separate term in this expression as representing a single individual tone, with frequency

$$f_n = \frac{c}{2\pi}\sqrt{\lambda_n} = \frac{nc}{2L} \quad,$$

but a *spatially-varying amplitude* as defined by $A_n(x)$. (Furthermore, since $\sqrt{\lambda_n} = n\sqrt{\lambda_1}$, these frequencies may be viewed as just multiples of the square root of the lowest eigenvalue. Moreover, the amplitude functions are really only multiples of the eigenfunctions.) One such vibration, corresponding to $n = 2$, is shown in Figure 33.
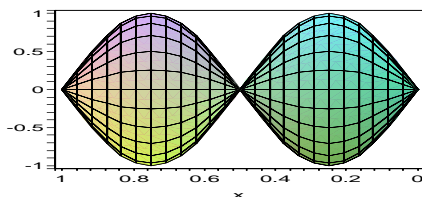
Figure 33: $A_2(x)\cos\left(\frac{2\pi ct}{L}\right)$

These individual frequencies that contribute to the solution (3.11.66),

$$f_1 = \frac{c}{2L} \,, \quad f_2 = \frac{2c}{2L} \,, \quad f_3 = \frac{3c}{2L} \,, \quad \cdots \quad,$$

must represent *natural frequencies* of the string, in that they will arise in the absence of any external stimulus other than that provided by the initial conditions. They are similar to the natural frequency ($\omega = \sqrt{k_s/m}$) which you should have already seen arise in the solution of the mechanical system

$$m\frac{d^2u}{dt^2} + k_s u = 0 \quad.$$

(Analogous natural frequencies are also found in electrical circuits.) There is one significant difference with the frequency structure in our model and that of the mechanical or electrical models commonly studied in ordinary differential equations. The mechanical system or the electrical circuit has only a *single* natural frequency. The vibrating string, by contrast, has an infinite number, but with a very definite structure. They're all multiples of f_1 - the so-called *fundamental* frequency, where

$$f_1 = \frac{c}{2L} = \frac{1}{2L}\sqrt{\frac{\tau}{\rho}} \quad,$$

133

where τ denotes the internal tension in the string and ρ its density per unit length. Moreover, this last equation clearly implies that there are exactly three ways to raise the fundamental frequency of a guitar string -

(1) Shorten it (i.e. decrease L), or

(2) Tighten it (i.e. increase τ), or

(3) Make it thinner (i.e. decrease ρ).

If you have never looked closely at a stringed instrument before, this would be a good time! Which strings give the bass notes? What happens to the sound when you press down on one of the strings somewhere along the neck? Or if you increase the tension by tightening one of the pegs? Don't these observations agree precisely with the above analysis?

The higher frequencies - f_2, f_3, ... - are called the *overtones* or the *harmonics*. Their presence is the principal contributor to the customary "rich" and full sound of a quality stringed musical instrument when compared to the sound of a tuning fork, for example. The degree to which any particular harmonic is present in any particular solution depends, of course, on the value of the A_i, which in turn depends directly on the initial conditions. This explains the slight variation in the perceived sound of a guitar when a string is plucked in the center as opposed to right next to the bridge.

Now, lets return to the amplitudes associated with each of these vibrations. The shapes defined by the $A_n(x)$, which are of course just scaled multiples of the eigenfunctions, are called the *modes* of vibration of the string. The first few are plotted below.
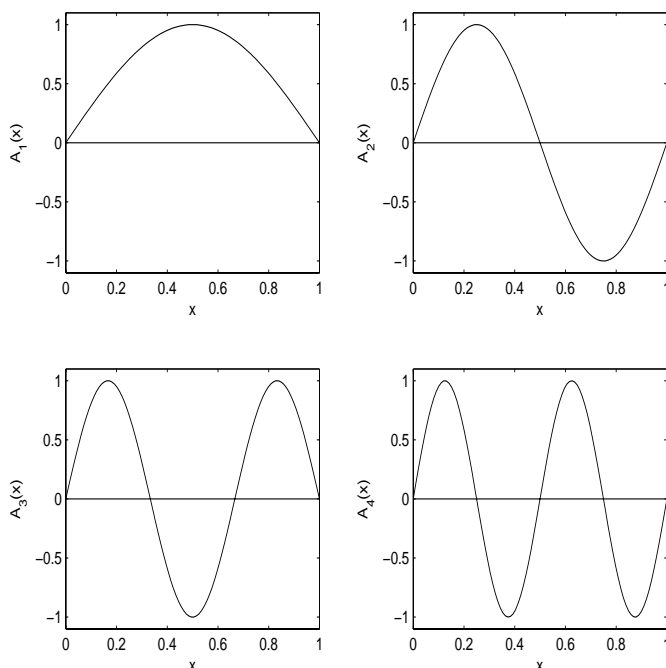

Figure 34: Various Modes of Vibration

Keep in mind that, in addition to having different shapes, these modes also vibrate at different rates - the higher the mode, the faster the rate of vibration. Physicists frequently refer to these vibrating modes as *standing waves*, since their basic shape never changes except for the continually changing scale introduced by the $\cos\left(\frac{n\pi ct}{L}\right)$ term. (The actual maximum displacements, as noted above are determined by the initial conditions through their effect on the values of the A_i.)

The concept of modes occurs not only in the theory of musical instruments, but also in communications. One of the simplest radio antennas is basically just a vertical wire. Although we shall not prove so here, the response of such an antenna to a radio signal satisfies the same wave equation as the vibrating string. Therefore, for the electrical analog of fixed ends, the antenna's natural frequency will be determined by its lowest eigenvalue, and the shorter the antenna, the higher the fundamental frequency. This fact helps to explain why AM broadcast stations require in general very long antennas, while most mobile car radios (cellular phones) use short ones - the car radios are operating at much higher frequencies. Furthermore, in a fixed end condition antenna, the fundamental mode will be $\sin(\pi x/L)$. This is exactly a half-cycle of the sine wave, and explains why such antennas are commonly called half-wave antennas by communications engineers. (There is also an antenna design that leads to mixed fixed-free boundary conditions. For hopefully obvious reasons, these antennas are referred to as quarter-wave antennas.)

We would close this part of the discussion with one last observation. We have just been discussing solution properties in terms of what happens at particular individual frequencies. In other words, in Fourier analysis terminology, we have been considering the solution in the *frequency domain*. Such an analysis is entirely appropriate and natural for musical instruments, since our primary "sensor" for evaluating those is the human ear, and, as we have alluded to before, the nature of the human ear is to try to decompose sounds into their component frequencies, and then match those against some "library" stored in the human brain. But we would also recall that the frequency domain is only one of the two domains in which Fourier analysis tells us we may view signals. The other, of course is the time domain. Therefore, we shall next turn our attention to what may be the time domain interpretations of solutions to the wave equation.

PROBLEM

1. Find the three lowest natural frequencies, and sketch the associated modes, for the equation:

$$u_{tt} = c^2 u_{xx}$$
$$u(0, t) = u_x(L, t) = 0$$

Plot, in the frequency domain, the natural frequencies of this "string."

3.12   The D'Alembert Solution of the Wave Equation

In the previous section, we developed a frequency domain analysis of the solution to the one-dimensional wave equation with fixed-end conditions, and observed that this solution could be interpreted as a superposition of standing waves on the string, each with its own unique shape and frequency. Furthermore, the frequencies of these waves were just some constant multiples of the square roots of the eigenvalues, and each shape was a multiple of the corresponding eigenfunction. This interpretation of the solution is valuable, since it can explain many of the simple observations that one can make about stringed musical instruments, not to mention about transmission lines and long-wire, high frequency radio antennas. In this section, we shall show that this *same* solution has a simultaneous, dual, time-domain interpretation - an interpretation that explains other observed wave phenomena that we would miss by considering the frequency domain alone.

Our analysis starts with the same solution as in the last section - the solution to the fixed end, zero initial velocity, initially displaced string -

$$u(x,t) = \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi ct}{L}\right) \sin\left(\frac{n\pi x}{L}\right) \tag{3.12.67}$$

where,

$$A_n = \frac{2}{L} \int_0^L u(x,0) \sin\left(\frac{n\pi x}{L}\right) dx \quad .$$

Using standard trigonometric identities, this reduces to

$$u(x,t) \;\; = \;\; \sum_{n=1}^{\infty} \frac{A_n}{2} \left[\sin\left(\frac{n\pi}{L}(x+ct)\right) + \sin\left(\frac{n\pi}{L}(x-ct)\right) \right]$$

$$= \frac{1}{2} \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{L}(x+ct)\right) + \frac{1}{2} \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{L}(x-ct)\right) \quad .$$

But now observe that the first of these two series is actually a function of only the single *combined* argument, $(x+ct)$, whereas the second is a function of the combined argument $(x-ct)$. Therefore, we can alternatively represent the solution (3.12.67) as

$$u(x,t) = F(x+ct) + F(x-ct) \quad , \tag{3.12.68}$$

where

$$F(x) = \frac{1}{2} \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right) \equiv \frac{1}{2} u(x,0) \quad .$$

This is truly intriguing! Our infinite series solution is in fact equal to the sum of *exactly two* functions, each of which appears to be just a half-size copy of the initial displacement. (Actually, this last statement is not quite accurate. To be precise, a close examination of the series shows that $F(x)$ is the odd periodic extension, of period $2L$, of a half size copy of the initial condition.) But will this always be the case? And exactly how do we interpret the $(x+ct)$ and $(x-ct)$ dependency?

It can be shown that (3.12.68) is actually a special case of a more general result. (This shouldn't necessarily be too surprising. After all, (3.12.68) only represents the solution in one special case - when the ends are fixed and the initial velocity zero.) This more general result applies (at least initially) to the wave equation in an *unbounded* (in other words, one without any boundaries) region and for *any* initial conditions, i.e. to the solution of

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} , \quad -\infty < x < \infty \\
u(x,0) &= f(x) \\
u_t(x,0) &= g(x) ,
\end{aligned}
\tag{3.12.69}
$$

and states that the solution to this problem can always be written in the form

$$
u(x,t) = F(x - ct) + G(x + ct) ,
\tag{3.12.70}
$$

for some suitable (and yet to be determined) functions $F(\ )$ and $G(\ )$. This form is commonly known as the *D'Alembert solution*[11]. That it satisfies the partial differential equation can be shown by straightforward substitution. The key step in showing this is the recognition that the function $F(x - ct)$ depends only on the *single* combined value $x - ct$. Thus, providing $F(\ )$ is suitably differentiable, the basic rules of partial differentiation yield

$$
\frac{\partial}{\partial x}[F(x - ct)] = \left\{ \frac{d}{d(x - ct)}[F(x - ct)] \right\} \left\{ \frac{\partial}{\partial x}[x - ct] \right\} = F'(x - ct)
$$

where $F'(\ )$ denotes the *ordinary* derivative function, in the sense that cosine is the derivative function for sine. Conversely, if we take a partial derivative with respect to t, we obtain

$$
\frac{\partial}{\partial t}[F(x - ct)] = -cF'(x - ct) .
$$

Second partial derivatives can be computed similarly. When we do so, we can easily show that

$$
\frac{\partial^2}{\partial t^2}[F(x - ct)] = (-c)^2 F''(x - ct) = c^2 \frac{\partial^2}{\partial x^2}[F(x - ct)] ,
$$

or, in other words, $F(x - ct)$ is always a solution of the partial differential equation in (3.12.69). A virtually identical set of calculations verifies that the same holds true for any other function of the single combined variable $(x + ct)$. But then, since the basic partial differential equation is linear and homogenous, the principal of superposition of homogeneous solutions implies that the combination (3.12.70) also solves the equation.

What is not quite so obvious is that the $F(\ )$ and $G(\ )$ in the D'Alembert solution can always be determined so that this form also satisfies the initial conditions of this problem, regardless of what $f(x)$ and $g(x)$ may be. (On the other hand, the initial conditions really represent only two equations to be solved. Therefore, since we have two unknown functions ($F(\ )$ and $G(\ )$) available to use, this shouldn't seem too surprising either.)

---

[11] Jean Le Rond d'Alembert,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/D'Alembert.html

We shall return momentarily to look at the relationship between specific initial conditions and the functions ($F(\ )$ and $G(\ )$) in the D'Alembert solution. But first we'll investigate an extremely interesting physical interpretation of this solution. Recall that, in general, the expression $F(x - a)$ simply represents the basic "shape" $F(x)$ shifted *to the right by the amount a.* Therefore, by plotting $F(x - ct)$ at several different values of t, we can easily see (Figure 35) that $F(x - ct)$ corresponds to the basic "shape" (wave) $F(x)$ *moving to the right with velocity c.*
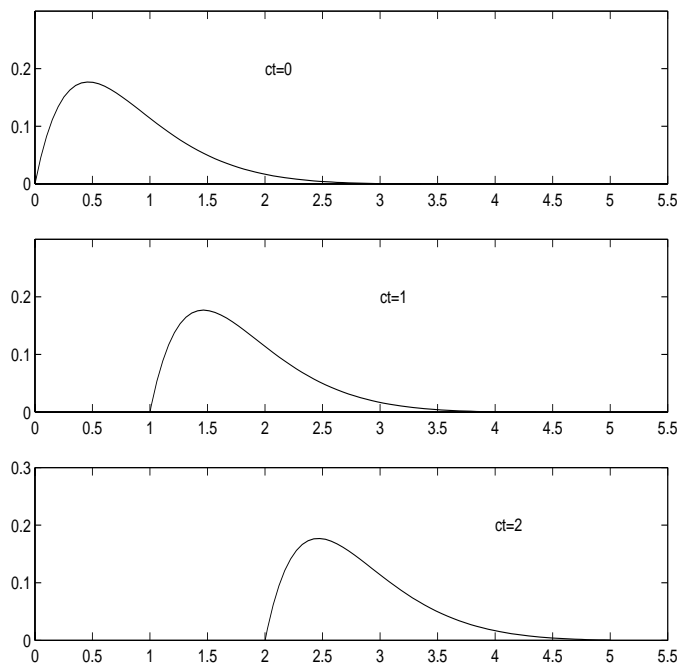
Figure 35: The Moving Function

Similarly, $G(x + ct)$ can be shown to imply a leftward moving wave. Therefore, looking back at the D'Alembert form (3.12.70) and incorporating these last observations, we see that the essence of, and the fundamental insight offered by the D'Alembert solution is that the solution to the one-dimensional wave equation can be visualized as the sum of two *traveling waves* - one ("$F(\ )$") moving to right with velocity c, and the other ("$G(\ )$") moving to the left with the same speed. (Furthermore, in the case of zero initial velocity, (3.12.68) seems to imply that $F(x) = G(x)$.) But think about these words carefully. What we have just described is how the solution evolves *with time*, almost in the manner of considering a sequence of successive snapshots! In other words, the D'Alembert solution is really a *time domain* interpretation of the solution of the wave equation. And, as with the earlier frequency domain interpretation, there are instances when this is clearly the proper interpretation to use. For example, while the frequency domain interpretation is the most useful for explaining the way a musical instrument "sounds," the time domain view best explains how the sound propagates from the instrument to the listener's ear.

We shall not actually prove that for a proper choice of $F(\ )$ and $G(\ )$ the D'Alembert solution (3.12.70) can satisfy arbitrary initial conditions in the general wave equation (3.12.69).

139

But we shall now show that it always reduces to our earlier solution (3.12.68) when the initial velocity is zero, i.e. when $g(x) = 0$.

As noted above, the general D'Alembert form always satisfies the partial differential equation in (3.12.69). Substituting that form into the initial position and velocity (with $g(x) = 0$) equations thus leads to

$$\begin{aligned} u(x,0) &= F(x) + G(x) = f(x) \\ u_t(x,0) &= -cF'(x) + cG'(x) = 0 \end{aligned} \quad . \tag{3.12.71}$$

But dividing the second equation by c, and then integrating, yields

$$-F(x) + G(x) = K \ , \tag{3.12.72}$$

where K is an arbitrary constant. But the first in (3.12.71) and (3.12.72) are just a system of two equations in two unknowns! Solving this system produces

$$\begin{aligned} F(x) &= \tfrac{1}{2}f(x) - K \\[2mm] G(x) &= \tfrac{1}{2}f(x) + K \ , \\[2mm] F(x - ct) &= \tfrac{1}{2}f(x - ct) - K \ , \quad \text{and} \\[2mm] G(x + ct) &= \tfrac{1}{2}f(x + ct) + K \quad . \end{aligned} \tag{3.12.73}$$

Finally, substituting these last two expressions for $F(x-ct)$ and $G(x+ct)$ in the D'Alembert solution (3.12.70) yields

$$\begin{aligned} u(x,t) &= F(x - ct) + G(x + ct) \\ &= \tfrac{1}{2}\left\{f(x - ct) + f(x + ct)\right\} \quad . \end{aligned} \tag{3.12.74}$$

Observe that the arbitrary constant K has canceled. So its value is totally irrelevant and we may therefore consider it to be zero, and ignore it in any further discussions of the D'Alembert solution. But with $K = 0$, then according to (3.12.73), $F(x) = G(x)$. Thus, at least in the case of the zero initial velocity wave equation in an unbounded region, our initial representation (3.12.68) and (3.12.74) are identical. (As alluded to above, more detailed calculations will show that the general D'Alembert solution (3.12.70) can satisfy (3.12.69) even when $g(x) \neq 0$, but in this case $G(\ )$ will no longer be identical to $F(\ )$.)

The graphical construction of a D'Alembert solution in the infinite domain for any particular zero initial velocity case consists of the fairly straightforward sequence of steps:

1. Sketch the initial displacement $f(x)$.

2. Sketch the function $F(x) = \tfrac{1}{2}f(x)$.

3. Determine the time t at which the solution $u(x,t)$ is to be determined.

4. Shift the curves for $F(x)$ right and left, respectively, by an amount ct.

5. Add the two shifted curves to give $u(x,t)$.

Figure 36 shows all of the steps in this process, applied at $t = 0.5$ and $t = 1.5$ to solution of the problem:

$$u_{tt} = c^2 u_{xx} , \quad -\infty < x < \infty$$
$$u(x,0) = \begin{cases} 1 - |x| & , -1 < x < 1 \\ 0 & , \text{ otherwise} \end{cases}$$
$$u_t(x,0) = 0 .$$
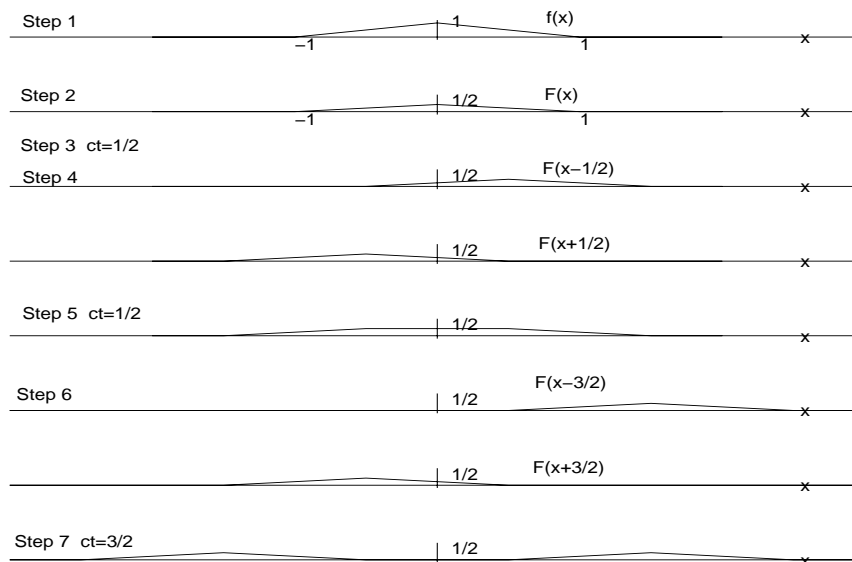
Figure 36: Constructing the D'Alembert Solution in the Unbounded Region

But wait! Our initial separation of variables solution was for a problem with boundary conditions! How do we reconcile that result with a D'Alembert solution in a region with no boundaries? The next section addresses this intriguing question.

3.13 The Effect of Boundary Conditions

As we noted during our derivation above, the general D'Alembert solution (3.12.70) does have one fundamental limitation - it describes only properties of the solution of the one dimensional wave-equation in an unbounded region. Yet, as we have commented earlier, and as our separation of variables solution reflected, all "real" strings are of finite length. So what happens to the D'Alembert solution when boundary conditions are introduced? The answer, as we shall now show, is that, with one new "wrinkle", the solution remains valid. A complete analysis, however, with general boundary conditions is beyond the scope of our study, so we shall instead look at special cases, starting with the Dirichlet problem we have

already solved by separation of variables

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \\
u(0,t) &= 0 \\
u(L,t) &= 0 \\
u(x,0) &= f(x) \\
u_t(x,0) &= 0
\end{aligned}
\tag{3.13.75}
$$

The D'Alembert solution form (3.12.70) clearly still satisfies this partial differential equation for **any** functions $F(\ )$ and $G(\ )$. However, it is not obvious under what conditions we can guarantee it will also satisfy the boundary conditions. Furthermore, since the initial conditions are now, strictly speaking, valid only for $0 < x < L$, we also need to question how to evaluate the terms in the D'Alembert solution when the values of $(x - ct)$ or $(x + ct)$ lie outside this range. Ignoring this last mathematical nicety *for the moment*, we plow on and see what happens. We first observe that since our initial velocity condition is zero, then, as we've discussed, the general D'Alembert solution (3.12.70) reduces in this case to (3.12.68), i.e.

$$
u(x,t) = F(x - ct) + F(x + ct) \ , \tag{3.13.76}
$$

with the initial displacement condition implying

$$
F(x) = \tfrac{1}{2} f(x) \ , \quad 0 < x < L \ .
$$

With the initial conditions thus accounted for, we now consider the effect of the left-hand boundary condition. Substituting $x = 0$ into (3.13.76) yields

$$
u(0,t) = F(-ct) + F(ct) = 0 \Rightarrow F(-ct) = -F(ct) \ , \quad t > 0 \ .
$$

But this equation, which must hold at all positive values of t, implies that $F(\ )$ is an odd function, i.e. that $F(x) = -F(-x)$. Similarly at the right-hand end, (3.13.76) reduces to

$$
u(L,t) = F(L - ct) + F(L + ct) = 0 \Rightarrow F(L - ct) = -F(L + ct) \ , \quad t > 0 \ .
$$

However, by the same logic, this is therefore equivalent to

$$
F(L + x) = -F(L - x) \ .
$$

There are now two alternative ways to interpret this. We can simply say that $F(\ )$ is *odd about x=L*. However, we can also use the fact that $F(\ )$ is odd (about zero) to give

$$
\begin{aligned}
F(L + x) &= -F(L - x) \\
&= F(x - L) \ ,
\end{aligned}
$$

and therefore (substituting $[L + x]$ for x on both sides)

$$
F(L + [L + x]) = F([L + x] - L)
$$

or

$$
F(x + 2L) = F(x) \ ,
$$

142

i.e. $F(\ )$ must be *periodic* of period $2L$. Combined with our earlier observation that $F(\ )$ was odd, this implies that we really need only the values of $F(x)$ for $0 < x < L$ in order to completely define $F(\ )$, and removes our earlier concerns about how to evaluate the solution when $x - ct$ or $x + ct$ was outside this interval. Furthermore, with this final set of restrictions, (3.13.76) satisfies every part of the original problem! In other words, it is the solution if we simply consider $F(x)$ to be the odd, periodic (of period $2L$) extension of a half-height copy of the original displacement. Equivalently, taking some liberty, if we now view $f(x)$ as the odd periodic extension of the original displacement, we see that we can represent the complete solution of the original boundary value problem as

$$u(x,t) = \tfrac{1}{2}\left\{ f(x - ct) + f(x + ct) \right\} \quad .$$

Thus, for the wave equation with fixed ends and no initial velocity, the general D'Alembert solution reduces to the sum of two waves, one moving to the right, the other to the left, both with velocity c, and each a perfect half-height copy of the odd periodic extension of the initial displacement. This, of course, is exactly the same interpretation that we discovered by applying trigonometric identities to the separation of variables solution. The important point here is that we have just rediscovered this fact, directly from the partial differential equation, without ever referring to, or even needing to know, the separation of variables solution.

Before considering an example of the D'Alembert solution for a string of fixed length, we would just note that in this last case, even though *mathematically* the D'Alembert solution exists for all x and t when $F(x)$ is viewed as the periodic extension of $\tfrac{1}{2}f(x)$, the only *physically real* solution exists for $0 < x < L$! (You might wish to imagine this as if there were opaque curtains covering the rest of the line, and that the parts of the D'Alembert solution only can be seen when one enters either "stage left" of "stage right.") Anything outside of this region is really a purely mathematical construct. This really should not concern us - provided we obtain the correct answer in the physical region $0 < x < L$.

The steps involved in the graphical construction of a D'Alembert solution for a region with boundaries are almost identical to those described above for the unbounded region, with the exception that the second step described earlier must be replaced by

2'. Sketch the function $F(x)$ as the extension, with appropriate symmetry and period, of $\tfrac{1}{2}f(x)$.

In Figure 37, we construct, according to the above discussion, the D'Alembert solution at $t = 1.5$ for the problem

$$
\begin{aligned}
u_{tt} &= u_{xx} \ , \quad 0 < x < 4 \ , \quad 0 < t \\
u(0,t) &= 0 \\
u(4,t) &= 0 \\
u(x,0) &= \begin{cases} x & , 0 < x < 1 \\ 2 - x & , 1 < x < 2 \\ 0 & , otherwise \end{cases} \\
u_t(x,0) &= 0
\end{aligned}
$$
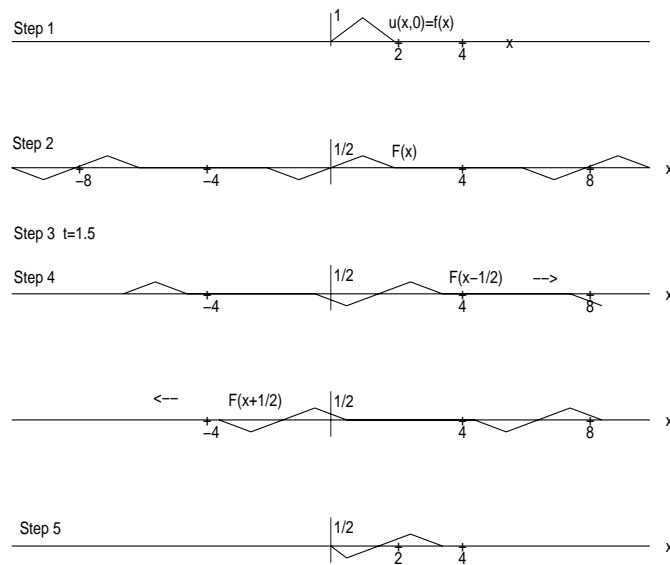
Figure 37: The D'Alembert Solution With Boundary Conditions

There is one last important insight that comes from this problem. Consider Figure 38, which plots the result of the D'Alembert solution to this problem at two different times. Look carefully what has happened! The initial shape, as we should have expected, started moving to the right with unit velocity (since $c^2 = 1$ for this problem). But then it ran into the right-hand boundary at $t = 3$! And reflected! (Mathematically of course, what happened is a portion of the leftward moving wave which originated outside the interval $0 < x < L$ entered the region. But remember the only physical reality is inside that interval. And physically what we would observe is in fact a reflection.) But not only has the shape reflected - but it's upside down. Looked at another way, either its phase has changed by 180°, or it reflected oddly. So here is another fundamental insight about waves - fixed boundaries produce odd reflections or phase reversals. And this is an insight which would never come from the separation of variables solution or any frequency domain analysis. It only comes from the time domain analysis implicit in the D'Alembert solution. (A similar analysis can show that free boundaries produce even reflections.)

144
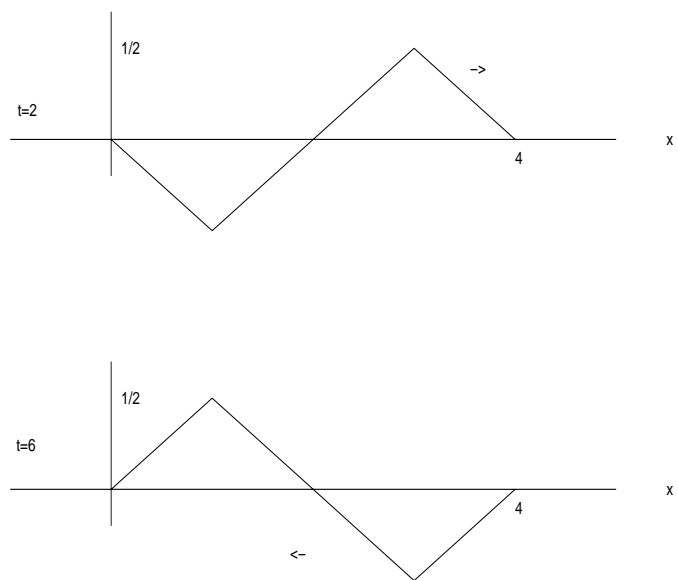
Figure 38: Boundary Reflections via The D'Alembert Solution

PROBLEM

1. Sketch the D'Alembert solutions at $t = 0, 1, 2.5$ and 4 to:

    a. $u_{tt} = u_{xx}$

            $u(0, t) = u(3, t) = 0$

$$u(x, 0) = \begin{cases} 2x & , & 0 < x < 1/2 \\ 1 & , & 1/2 < x < 3/2 \\ 4 - 2x & , & 3/2 < x < 2 \\ 0 & , & 2 < x < 3 \end{cases}$$

            $u_t(x, 0) = 0$

    b. $u_{tt} = u_{xx}$

            $u(0, t) = u_x(3, t) = 0$

$$u(x, 0) = \begin{cases} 2x & , & 0 < x < 1/2 \\ 1 & , & 1/2 < x < 3/2 \\ 4 - 2x & , & 3/2 < x < 2 \\ 0 & , & 2 < x < 3 \end{cases}$$

            $u_t(x, 0) = 0$

    c. $u_{tt} = 4u_{xx}$

            $u_x(0, t) = u(1, t) = 0$

$$u(x, 0) = \begin{cases} 1 & , & 0 < x < 1/2 \\ 2 - 2x & , & 1/2 < x < 1 \end{cases}$$

            $u_t(x, 0) = 0$

4 The Two-Dimensional Wave Equation

4.1 Introduction

As we remarked at the close of the previous chapter, we have effectively reached the point of diminishing returns with regard to our study of the one-dimensional wave equation. Outside of solving additional special cases, or perhaps considering non-uniform media, we have seen the major fundamental results - the separation of variables method, the eigenvalue/eigenfunction structure, Sturm-Liouville theory, and the complementary time and frequency domain interpretations. Therefore, the time has come to move on. The direction we will choose to proceed is to increase the number of spatial dimensions while still focusing primarily on partial differential equations whose solutions have wave interpretations. Hence, the most natural class of problems we should study next are those involving two spatial dimensions, say x and y. As we study these, one of our primary intents, as it was in the last chapter, will be to discern the basic principles or "rules" of this "game," as opposed to the specific, problem-dependent "players." Furthermore, we shall try not only to develop the mechanics of calculating the solutions to the problems, but also to look beyond these mechanics to the physical principles embodied in the solutions.

The two-dimensional analog of the vibrating string is the *membrane*, a thin, tightly stretched elastic sheet, such as would be found in a drum. As we did when we derived the partial differential equation for the vibrating string, we will assume the membrane is of uniform density, that a uniform internal tension produces motions which are effectively only in the vertical direction, that there are no external forces and that all displacements and angles are "small." Then the governing partial differential equation (which we shall not derive) for this vertical displacement (u) at any point becomes

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u \ \ ,$$

where ∇^2 is the operator commonly called the *Laplacian*. With the appearance of the Laplacian comes the first significant difference from the one-dimensional problem. For while there is only one representation, i.e. $\dfrac{\partial^2 u}{\partial x^2}$, for the second derivative in one dimension, the form of the Laplacian in two or more dimensions changes depending on the geometry (coordinate system) adopted. For example, in rectangular coordinates, the two-dimensional wave equation would become

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left\{ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right\} \ \ ,$$

while in cylindrical (polar) coordinates, it would be

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right\} \ \ .$$

We shall ignore for the moment the minor detail that cylindrical coordinates are actually more appropriate if we wish to model a drum, and consider the rectangular form of the two

147

dimensional wave equation. In terms of the subscript notation for partial derivatives, this becomes

$$u_{tt} = c^2 \left\{ u_{xx} + u_{yy} \right\} \quad .$$

By now, of course, we should recognize that this equation alone does not constitute a complete problem, i.e. it does not provide enough information to determine a unique solution. Specifically, the differential equations involve second derivatives each with respect to x, y and t. Therefore, we would expect to need two additional side conditions in each of these variables before we could reasonably expect unique solutions. Furthermore, based on our experience with the vibrating string, the most natural additional conditions to expect are two boundary conditions each in x and y, and two initial conditions in t. Moreover, while we expect that the exact boundary conditions will vary from problem to problem, the initial conditions should always involve precisely the displacement ($u(x, y, 0)$) and velocity ($u_t(x, y, 0)$).

4.2   The Rigid Edge Problem

We begin our study of the two-dimensional wave equation in rectangular coordinates, as we began our study of the vibrating string, with a special case, albeit the one which would seem physically to be the most likely. This case occurs when the edges of the membrane are rigidly attached to some immovable, rectangular frame of length L and width W, creating what is, in effect, a rectangular drum. (As we noted above, this is not quite as "normal" as a circular drum. However, as we shall see later, certain of its mathematical aspects are far simpler to treat than the circular case. Therefore, it is a much more appropriate pedagogical starting point.) The rigid frame of course implies zero displacement along the edges, i.e. homogeneous Dirichlet boundary conditions. With the addition of generic initial conditions, the complete statement of this problem then becomes

$$
\begin{aligned}
u_{tt} &= c^2 \left\{ u_{xx} + u_{yy} \right\} \\
u(0, y, t) &= u(L, y, t) = 0 \\
u(x, 0, t) &= u(x, W, t) = 0 \\
u(x, y, \ 0) &= f(x, y) \\
u_t(x, y, \ 0) &= g(x, y)
\end{aligned}
\tag{4.2.1}
$$

To solve this problem, we shall try, as we done all along, to build as much as possible on earlier insights. So the appropriate place to start is by reflecting on our experience with one-dimensional wave equation,

$$u_{tt} = c^2 u_{xx} \quad .$$

The key to that problem, obviously, was separation of variables, which was motivated by the observation that the partial differential equation possessed a natural separation into operations on t and operations on x and which then reduced the homogeneous problem to two ordinary differential equations. Some similar approach, then, would seem advisable here, and the first step would seem to be to identify the appropriate *homogeneous problem* for the two-dimensional case. A straightforward extrapolation from the comparison of (4.2.1) with

the one-dimensional problem indicates the most reasonable choice is

$$
\begin{aligned}
u_{tt} &= c^2 \left\{ u_{xx} + u_{yy} \right\} \\
u(0,y,t) &= u(L,y,t) = 0 \\
u(x,0,t) &= u(x,W,t) = 0
\end{aligned}
\tag{4.2.2}
$$

Clearly, as in the one-dimensional wave equation, this problem has at least the trivial solution, $u(x,y,t) \equiv 0$. Thus, the important question will again be what are its non-trivial solutions, how do we find them, and how many of them are linearly independent?

We begin to answer these questions with the observation that the partial differential equation in (4.2.2) again possesses a natural "splitting" - into operations on t, x, and y. This observation, coupled with our experience in the one-dimensional problem, suggests that the two-dimensional homogeneous solutions might also admit an analogous splitting. Thus we shall assume the homogeneous solutions have the form:

$$
u(x,y,t) = X(x)Y(y)T(t) \quad . \tag{4.2.3}
$$

As before, you should realize this is only an assumed solution form. There are clearly no *a priori* guarantees it will work. On the other hand, it certainly seems like a reasonable choice, and furthermore, at this point, it really seems like the "only game in town." Whether it will actually produce the necessary solutions depends on what happens when we substitute it into the problem and see what happens.

We shall start by substituting this product solution into the partial differential equation. When we do so, taking the appropriate derivatives, we obtain

$$
X(x)Y(y)T''(t) = c^2 \left\{ X''(x)Y(y)T(t) + X(x)Y''(y)T(t) \right\} \quad ,
$$

or, dividing both sides by the product $c^2 XYT$,

$$
\frac{T''(t)}{c^2 T(t)} = \frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} = -\lambda \quad . \tag{4.2.4}
$$

In this last equation, of course, we have simply invoked a similar argument to the one we used in the one-dimensional case. Specifically, the left-hand side here depends only on time and is therefore independent of x and y, while the right-hand side is independent of time, and yet both sides are equal. Therefore, they must both equal a constant. We chose to represent this constant by the symbol $-\lambda$, where we include the minus sign based on our experience with the one-dimensional problems and Sturm-Liouville theory. We thus arrive at the two separated problems

$$
T''(t) + \lambda c^2 T(t) = 0 , \qquad \text{and} \qquad \frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} = -\lambda
$$

However, unlike the one-dimensional case, we're not finished yet, since we still haven't uncoupled the behavior of $X(x)$ and $Y(y)$ in

$$
\frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} = -\lambda \quad ,
$$

and such an uncoupling would seem to be necessary if we are to continue with this approach. Achieving this uncoupling is not really difficult. We simply have to move either term in the last equation to the right-hand side of the equality. For example, moving the terms depending on $Y(y)$ to the right hand side yields

$$\frac{X''(x)}{X(x)} = -\frac{Y''(y)}{Y(y)} - \lambda = -\mu \quad .$$

Introducing the new separation constant μ involves again essentially only a repetition of the argument from the one-dimensional case, except that here the left-hand side is independent of y while the right-hand side is independent of x. (Since we had already used λ, we must represent this new separation constant by another symbol. The choice of the letter μ is totally arbitrary, however we choose the minus sign in order to, as much as possible, follow the one-dimensional case.) With this new constant, dropping the explicit dependencies, and a minimal amount of algebra, we then arrive at the fully separated ordinary differential equations

$$T'' + \lambda c^2 T = 0 \qquad X'' + \mu X = 0 \qquad Y'' + (\lambda - \mu)Y = 0 \quad ,$$

At this point, you should realize that we really can't say anything about the value(s) of λ or μ. (Furthermore, we should not even think of λ or μ as representing only a single value. Reasonably, we should expect there will be an infinite number of each of them.) That we can't yet specify either λ or μ at this point shouldn't be too surprising, since we haven't yet separated the boundary conditions. Therefore, if we really want our product solution (4.2.3) to satisfy the entire homogenous problem (4.2.2), we must substitute that form into the boundary conditions as well. When we do this at the left-hand boundary, we obtain

$$u(0, y, t) = X(0)Y(y)T(t) = 0 \quad .$$

This condition should again be familiar ground, since $X(0)$ is still the value of a function at one point, i.e. a number. Therefore, either $X(0) = 0$ or $Y(y)T(t) \equiv 0$. But, as in the one-dimensional case, the second alternative would yield only trivial solutions, and thus we must require that

$$X(0) = 0 \ .$$

Similarly, the other boundary conditions imply

$$
\begin{aligned}
u(L, y, t) &= X(L)Y(y)T(t) = 0 \Rightarrow X(L) = 0 \\
u(x, 0, t) &= X(x)Y(0)T(t) = 0 \Rightarrow Y(0) = 0 \\
u(x, W, t) &= X(x)Y(W)T(t) = 0 \Rightarrow Y(W) = 0 \ .
\end{aligned}
$$

Now, collecting all the information we have so far deduced about $X(x)$, $Y(y)$, and $T(t)$ as defined by (4.2.2) and (4.2.3), we have:

$$
\begin{array}{llll}
T'' + \lambda c^2 T = 0 & X'' + \mu X & = 0 & Y'' + (\lambda - \mu)Y & = 0 \\
& X(0) & = 0 & Y(0) & = 0 \qquad (4.2.5) \\
& X(L) & = 0 & Y(W) & = 0
\end{array}
$$

The equation here for $T(t)$ is solvable for any value of λ, as was the case in the one-dimensional problem, since there are no conditions on $T(t)$ other than the differential equation. By contrast, **both** $X(x)$ and $Y(y)$ now satisfy Sturm-Liouville problems! Fortunately, however, they are Sturm-Liouville problems that we have already solved before. (The second of these problems does have one small subtlety - it is the combined quantity $(\lambda - \mu)$ that actually plays the role of the eigenvalue in the equation for $Y(y)$. Thus one might easily argue that we should replace this combined quantity by another, single symbol. We shall not do that, however, since we prefer to keep the number of different symbols to a minimum.) Note that the problems for $X(x)$ and $Y(y)$ are essentially *independent* in that the eigenfunctions of each can be determined without any reference to the other. For example, the problem for $X(x)$ immediately leads to

$$\begin{aligned} \mu_n &= \left(\tfrac{n\pi}{L}\right)^2 , \\ X_n(x) &= \sin\left(\tfrac{n\pi x}{L}\right) , \end{aligned} \qquad n = 1, 2, 3, \ldots \quad . \tag{4.2.6}$$

(We again may omit any arbitrary constants, since each eigenvalue corresponds to only a single independent eigenfunction.)

The solution for $Y(y)$ is almost as direct, although properly representing it requires a bit more precise notation. The reason for this is that, as noted above, one may perfectly well take the view that it is the *combined* quantity $(\lambda - \mu)$ that plays the role of the eigenvalue in the second eigenvalue problem. Therefore, it might seem reasonable to write eigenvalues and eigenfunctions for $Y(y)$ as

$$\begin{aligned} (\lambda - \mu)_m &= \left(\tfrac{m\pi}{W}\right)^2 , \\ Y_m(y) &= \sin\left(\tfrac{m\pi y}{W}\right) , \end{aligned} \qquad m = 1, 2, 3, \ldots \quad . \tag{4.2.7}$$

(Note we must use a different index variable here than we used for $X_n(x)$ in order to emphasize the basic independence of the problems for $X(x)$ and $Y(y)$.) However, this representation does not really adequately convey one important point - because of (4.2.5), the μ in (4.2.7) may not be arbitrarily selected, but must *correspond to one of the eigenvalues of the problem for $X(x)$*. We could denote this by writing

$$(\lambda - \mu_n)_m = \left(\frac{m\pi}{W}\right)^2 , \qquad m, n = 1, 2, 3, \ldots \quad .$$

But this representation clearly implies that the values of λ then in fact depend on both m and n. Probably the "cleanest" way to denote this is

$$\lambda_{nm} = \left(\frac{n\pi}{L}\right)^2 + \left(\frac{m\pi}{W}\right)^2 \equiv \nu_{nm}^2 , \qquad m, n = 1, 2, 3, \ldots \quad , \tag{4.2.8}$$

where we introduce the symbol ν solely to avoid having to write square roots later on. This formula effectively validates our earlier decision to write the separation constant as $-\lambda$, since it is now clear from this representation and the values for m and n that λ can never be negative.

151

With the $X(x)$ and $Y(y)$ now determined, we must return and solve equation (4.2.5) for $T(t)$. Because the equation is second order, each value of λ_{nm} as given above will generate two linearly independent solutions, which we can represent as

$$T_{nm}(t) = A_{nm} \cos\left(\nu_{nm}ct\right) + B_{nm} \sin\left(\nu_{nm}ct\right) , \quad m, n = 1, 2, 3, \dots .$$

(For notational consistency, T, A and B should depend on the same subscripts as ν_{nm}.) Finally, combining the various pieces, we see that each value of λ_{nm} given by (4.2.8) produces two non-trivial, linearly independent solutions to (4.2.2) given by:

$$u_{nm}(x, y, t) = [A_{nm} \cos\left(\nu_{nm}ct\right) + B_{nm} \sin\left(\nu_{nm}ct\right)] \ \sin\left(\tfrac{n\pi x}{L}\right) \sin\left(\tfrac{m\pi y}{W}\right) ,$$
$$m, n = 1, 2, 3, \dots . \tag{4.2.9}$$

Actually, some pairs n, m may produce same eigenvalue but different eigenfunctions. Therefore, since we expect the general solution to be a linear combination of all the linearly independent homogenous solutions, the appropriate representation for the general solution would be

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [A_{nm} \cos\left(\nu_{nm}ct\right) + B_{nm} \sin\left(\nu_{nm}ct\right)] \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) . \tag{4.2.10}$$

Assuming that our logic up to this point is correct, then this series should also solve (4.2.1), once the constants A_{nm} and B_{nm} have been properly selected to satisfy the initial conditions. But can we, in fact, so select them? There is really only one way to find out, and that is let $t = 0$ in this expression, which is supposed to represent $u(x, y, t)$ at *all* times and substitute the resulting quantity into the initial conditions.

If we let $t = 0$ in (4.2.10), we have,

$$u(x, y, 0) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{nm} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) = f(x, y) . \tag{4.2.11}$$

Similarly, if we differentiate once with respect to t and then let $t = 0$,

$$u_t(x, y, 0) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \nu_{nm}cB_{nm} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) = g(x, y) . \tag{4.2.12}$$

But can we now find the A_{nm} and B_{nm} so that these equations are satisfied? Formally, the answer, is yes, provided $f(x, y)$ and $g(x, y)$ are at least piecewise smooth on the region $0 < x < L$, $0 < y < W$. In this case, we then can either appeal to Fourier series results or apply the orthogonality property directly. For example, we can write (4.2.11) as

$$u(x, y, 0) = \sum_{m=1}^{\infty} \underbrace{\left\{\sum_{n=1}^{\infty} A_{nm} \sin\left(\frac{n\pi x}{L}\right)\right\}}_{b_m(x)} \sin\left(\frac{m\pi y}{W}\right) = f(x, y) ,$$

which can be viewed, for any fixed value of x, as just a standard Fourier sine series in y, with coefficient $b_m(x)$. Thus, the usual formula gives

$$b_m(x) \equiv \sum_{n=1}^{\infty} A_{nm} \sin\left(\frac{n\pi x}{L}\right) = \frac{2}{W} \int_0^W f(x, y) \sin\left(\frac{m\pi y}{W}\right) dy \ .$$

But this is now just a Fourier sine series in x, and so a second application of the basic Fourier coefficient formula yields

$$
\begin{aligned}
A_{nm} \ &= \ \frac{2}{L} \int_0^L b_m(x) \sin\left(\frac{n\pi x}{L}\right) dx \\
&= \frac{2}{L} \int_0^L \left\{ \frac{2}{W} \int_0^W f(x, y) \sin\left(\frac{m\pi y}{W}\right) dy \right\} \sin\left(\frac{n\pi x}{L}\right) dx \qquad (4.2.13) \\
&= \frac{4}{LW} \int_0^L \int_0^W f(x, y) \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) dy dx \ .
\end{aligned}
$$

(You should convince yourself that the identical formula would result if you were to multiply both sides of (4.2.11) by

$$\sin\left(\frac{k\pi x}{L}\right) \sin\left(\frac{l\pi y}{W}\right) \ ,$$

integrate over the rectangle, separate the double integral under the summation into a product of iterated integrals, and apply orthogonality.)

A virtually identical set of steps leads to

$$B_{nm} = \frac{4}{\nu_{nm} c L W} \int_0^L \int_0^W g(x, y) \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) dy dx \ . \qquad (4.2.14)$$

Since it is difficult to conceive of a membrane in the "real world" whose initial displacements and velocities are not at least piecewise smooth, then the basic Fourier series results imply that all of these integrals and series exist. We have therefore solved the problem!

As an example of this, consider the problem

$$
\begin{aligned}
u_{tt} &= \ 4\left\{u_{xx} + u_{yy}\right\} \\
u(0, y, t) &= u(2, y, t) = 0 \\
u(x, 0, t) &= u(x, 1, t) = 0 \\
u(x, y, \ 0) &= 0 \\
u_t(x, y, \ 0) &= 1
\end{aligned}
$$

A review of the previous development shows this problem corresponds to the case where $c = 2$, $L = 2$, $W = 1$ and $f(x, y) \equiv 0$. Therefore, according to (4.2.11), (4.2.12) and (4.2.13), the solution should be

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} B_{nm} \sin(2\nu_{nm}t) \sin\left(\frac{n\pi x}{2}\right) \sin(m\pi y) \ ,$$

where

$$\nu_{nm} = \sqrt{\left(\frac{n\pi}{2}\right)^2 + (m\pi)^2} \quad ,$$

and

$$
\begin{aligned}
B_{nm} &= \frac{1}{\nu_{nm}} \int_0^2 \int_0^1 \sin\left(\frac{n\pi x}{2}\right) \sin\left(m\pi y\right) dy dx \\[2mm]
&= \frac{1}{\nu_{nm}} \int_0^2 \sin\left(\frac{n\pi x}{2}\right) dx \int_0^1 \sin\left(m\pi y\right) dy \\[2mm]
&= \frac{2}{mn\pi^2 \nu_{nm}} [1 - \cos(n\pi)][1 - \cos(m\pi)] \quad .
\end{aligned}
$$

The development of the series solutions for other appropriate combinations of boundary conditions closely parallels the development for the one-dimensional string, and is left for the exercises.

4.3 Frequency Domain Analysis

In the previous section, we developed the basic mechanics of constructing solutions to two-dimensional rectangular wave equation problems. However, as was also the case with the vibrating string, there is much more to the vibrating membrane problem than the mere mechanics (impressive as they may be) of writing doubly infinite series solutions. Equally important are the physical principles incorporated in those solutions. In this section and the following one we show that the solutions for a rectangular membrane can be analyzed in a manner similar to the solutions for a vibrating string - although there are some striking differences in the resulting properties.

The starting point for our analysis is the membrane with rigidly fixed edges and with zero initial velocity. (As was true for the vibrating string with fixed ends, this is simply the easiest case to treat algebraically. Other cases will not differ significantly in terms of their *qualitative* behavior.) According to the methods described above, the solution for this problem is

$$
\begin{aligned}
u(x,y,t) &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{nm} \cos\left(\nu_{nm} ct\right) \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) \\[2mm]
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \left[A_{nm} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) \right] \cos\left(\nu_{nm} ct\right) \\[2mm]
&= \left[A_{11} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{W}\right) \right] \cos\left(\nu_{11} ct\right) \\[2mm]
&\quad + \left[A_{12} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi y}{W}\right) \right] \cos\left(\nu_{12} ct\right) \\[2mm]
&\quad\quad + \left[A_{21} \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi y}{W}\right) \right] \cos\left(\nu_{21} ct\right) + \cdots \quad .
\end{aligned}
$$

(4.3.15)

(We have chosen to order the terms in this summation so that those corresponding to the same value of $m + n$ are written together. This is a fairly common convention, and generally results in terms of approximately the same order being grouped together.)

154

As was the case with the vibrating string, the above solution is clearly a sum of different shapes (modes),

$$A_{nm}(x,y) = A_{nm} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{W}\right) \ , \quad m, n = 1, 2, 3, \ldots \quad ,$$

each vibrating separately at its own natural frequency given by

$$f_{nm} = \frac{\nu_{nm} c}{2\pi} \ , \quad m, n = 1, 2, 3, \ldots \quad .$$

However, unlike the modes of the vibrating string, these modes depend on **two** independent variables, x and y, i.e. these shapes represent *surfaces*, not curves. Hence, both plotting and visualizing them is a bit more difficult than was the case with the string. Two alternative methods of displaying these modes are commonly used. One is to effectively project them onto a plane, as is done in any painting. The other is to draw their contour lines, as is done with maps. Figures (39) and (40) use each of the techniques, respectively, to display the first few modes for a rectangle with fixed edges.
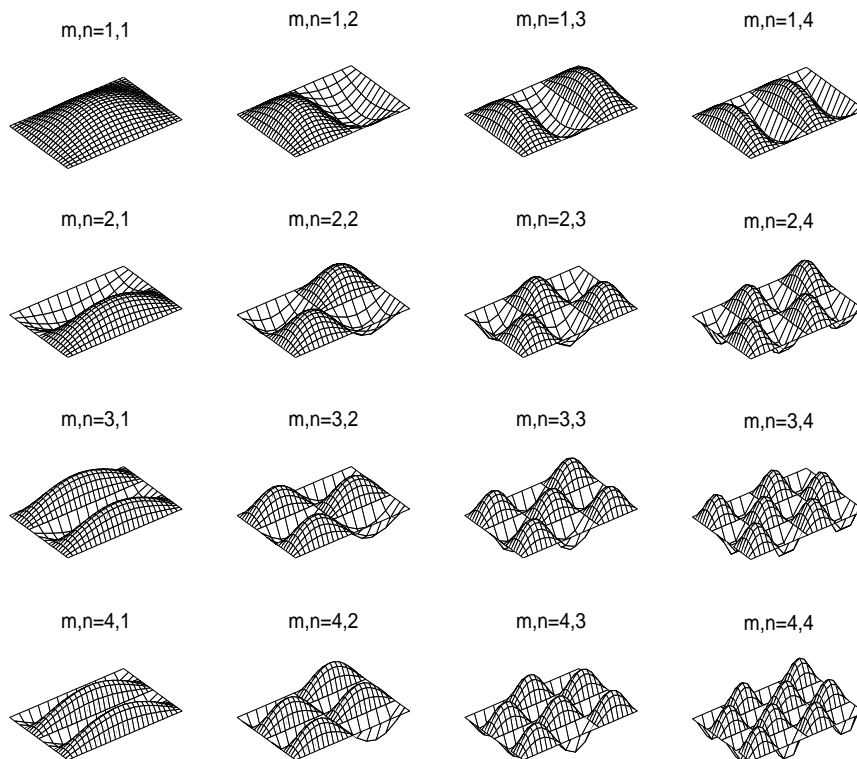

Figure 39: Modes of A Vibrating Rectangle

The rectangular membrane and the vibrating string, however, differ in more ways that
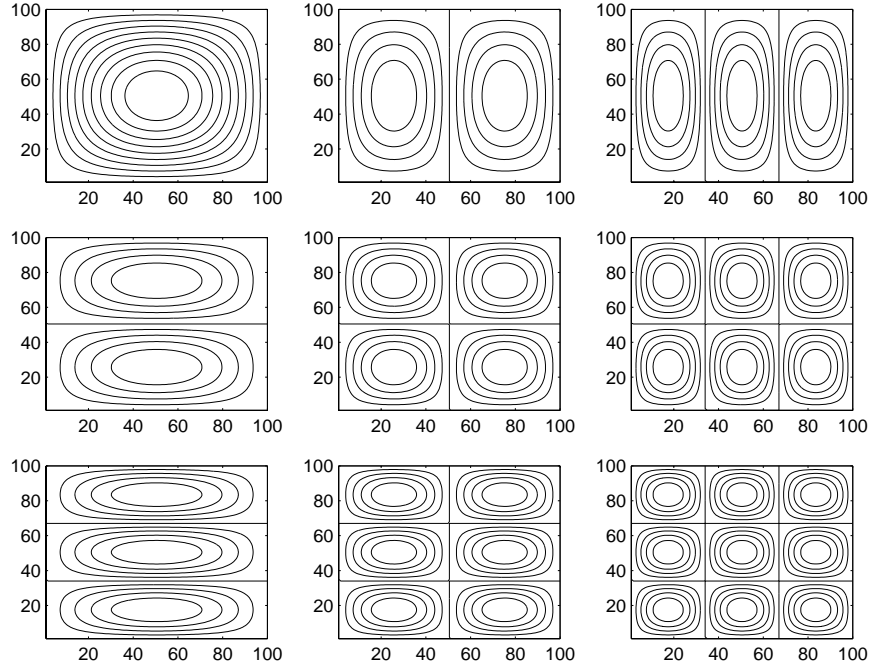
Figure 40: Contour Lines for Modes of a Vibrating Rectangle

just the nature of their modes. The structure of the natural frequencies of the membrane,

$$ f_{nm} = \frac{\nu_{nm}c}{2\pi} \quad = \frac{c}{2\pi}\sqrt{\left(\frac{n\pi}{L}\right)^2 + \left(\frac{m\pi}{W}\right)^2} $$

$$ = \frac{c}{2}\sqrt{\left(\frac{n}{L}\right)^2 + \left(\frac{m}{W}\right)^2} \;, \quad m, n = 1, 2, 3, \ldots \quad , $$

also differs significantly from the structure of those for the string,

$$ \tilde{f}_n = \frac{nc}{2L} \;, \quad n = 1, 2, 3, \ldots \quad , $$

for two principal reasons. The first is obviously that the frequencies for the membrane depend on two subscripts, not one. Thus listing them requires a table rather than simply listing them in order. This leads to a situation which is best illustrated by the following table, which shows the first few frequencies for a rectangle of length three and width two (with $c = 2\pi$):

$n\backslash m$	1	2	3	4
1	1.888	3.312	4.827	6.370
2	2.618	3.776	5.157	6.623
3	3.512	4.443	5.564	7.025
4	4.474	5.236	6.305	7.551
5	5.467	6.106	7.044	8.179
6	6.477	7.025	7.854	8.886

Carefully study this table for a moment. Note that if we wished to list these frequencies in strictly increasing order (certainly not an unreasonable idea), we should have to jump around in the table a significant amount, since

$$f_{11} < f_{21} < f_{12} < f_{31} < f_{22} < f_{32} < f_{41} < f_{13} < f_{23} < f_{42} < \ldots \quad .$$

Moreover, also unlike the frequency structure of the vibrating string, the same frequency in the vibrating membrane may correspond to two or more distinct eigenfunctions. For example, in the table we just created, note that

$$f_{34} = f_{62} \qquad !$$

A further significant difference between the membrane and the string with respect to the structure of their natural frequencies arises because of the fact that the square root of a sum is not the sum of the square roots. Therefore, as a simple calculation with the above table will show, the natural frequencies of the membrane (f_{nm}) are not constant multiples of the fundamental frequency (f_{11}). Therefore the spectrum of a rectangular drum will lose the even, "picket fence" appearance that characterized both the string and the ordinary Fourier series. Figure 41 displays this for the above table of frequencies. (Note that in this figure we have plotted all of the amplitudes as identical, a situation which would almost never occur in practice. Don't dwell on that - the purpose of this figure is solely to emphasize the uneven spacing of the frequencies in the rectangular membrane.) This picture clearly indicates that, while the issue of which sounds better - a drum or a violin - would have to be left to the individual listener, there is no question that they should not sound at all alike.

Figure 41: The Spectrum of the Rectangular Drum
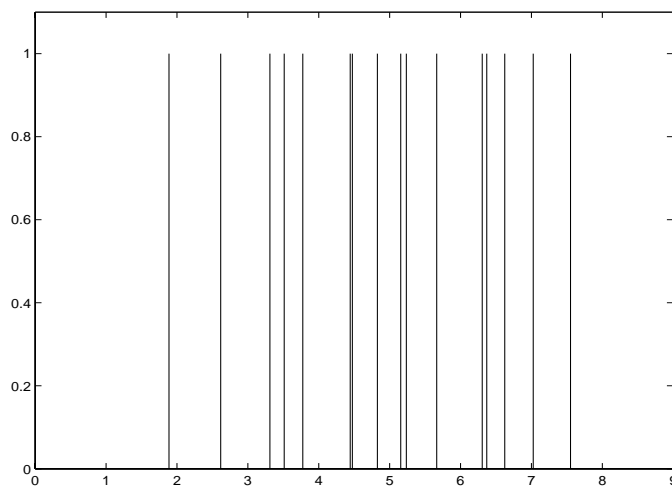
We close this section by simply noting that what we have demonstrated here is that the same kind of frequency domain arguments that apply to "normal" Fourier series can also be applied to two (and by implication three) dimensional wave equation solutions, at least in rectangular geometries, and that this analysis leads to additional insights into the workings of the "real world."

4.4   Time Domain Analysis

In the last section we analyzed the solution to the two-dimensional wave equation in terms of its component frequencies. But recall that we previously saw that the solution for the vibrating string also could be interpreted in terms of traveling waves via the D'Alembert solution. We now show, although we shall skip most of the algebra, that traveling wave solutions also exist for the two-dimensional wave equation. Specifically, suppose that k_1 and k_2 are any constants such that

$$k_1^2 + k_2^2 = 1 \ .$$

Then, if $F(\ )$ is any suitably differentiable function, the standard rules for partial derivatives show that

$$\frac{\partial^2}{\partial t^2} F(k_1 x + k_2 y - ct) \ = c^2 F''(k_1 x + k_2 y - ct)$$

$$\frac{\partial^2}{\partial x^2} F(k_1 x + k_2 y - ct) \ = k_1^2 F''(k_1 x + k_2 y - ct)$$

$$\frac{\partial^2}{\partial y^2} F(k_1 x + k_2 y - ct) \ = k_2^2 F''(k_1 x + k_2 y - ct) \ ,$$

and therefore

$$u(x, y, t) = F(k_1 x + k_2 y - ct)$$

is a solution of

$$u_{tt} = c^2 \left\{ u_{xx} + u_{yy} \right\} \ .$$

But how should we interpret this solution. Well, recall that we chose $u(x, y, t)$ to denote the vertical displacement of a two-dimensional surface. Therefore, $u(x, y, t) = F(k_1 x + k_2 y - ct)$ must represent a vertical displacement which is of constant height along any curve on which $k_1 x + k_2 y - ct$ is constant. But, of course, at any fixed time t, the equation

$$k_1 x + k_2 y - ct = constant$$

simply defines a straight line, normal to the vector

$$\vec{k} = k_1 \vec{i} + k_2 \vec{j} \ ,$$

i.e., at any fixed time the displacement described by $F(k_1 x + k_2 y - ct)$ would look like a set of parallel ridges.

Equally interesting is the interpretation of how this picture of the displacement will change with time. It is fairly easily shown that if $k_1 x + k_2 y - ct$ is equal to any particular value at some time t, then at time $t + 1$, $k_1 x + k_2 y - ct$ is equal to the *same* value on a line parallel to the first, and located exactly c units away in the direction of $\vec{k}$. Therefore, whatever feature was originally located on the first line has in fact propagated at velocity c in the direction of $\vec{k}$. (See Figure 42.) Thus whatever features may happen to be described by $F(k_1 x + k_2 y - ct)$ move in parallel lines, like ranks of marchers in a parade. For obvious
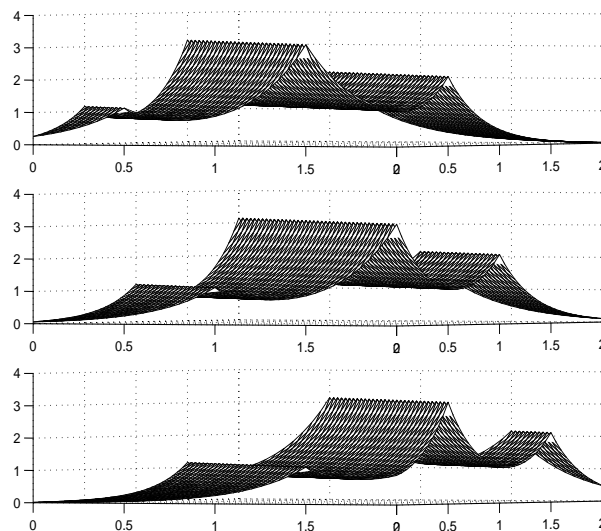
Figure 42: A Traveling Plane Wave

reasons, such motion is commonly called a *plane wave*. That such solutions should exist for the two-dimensional wave equation should not actually be that surprising. It's really just the analog of the D'Alembert solution for the one-dimensional string - with one fundamental difference. On the one-dimensional string there are exactly two traveling waves - one moving to the left, and the other to the right. By contrast, the two-dimensional wave equation can have an *infinite number* of simultaneous independent plane wave solutions, since there are an infinite number of different possible choices for $\vec{\mathbf{k}}$. Figure 43 displays a solution consisting of the superposition of two wave traveling in different, but not opposite directions. One of the waves consists of two crests, and is moving generally from left to right. The second consists of a single crest, and is moving more or less toward the reader.

We shall not pursue this line of investigation any further. The actual mechanics of decomposing an arbitrary two-dimensional initial displacement into a sum of plane waves is far more involved than the one-dimensional D'Alembert decomposition into two half-height copies of the initial displacement, and beyond the scope of these notes. We are simply content to have observed that the rectangular membrane enjoys the same dual frequency domain, time-domain nature that we discovered for the vibrating string. Furthermore, as with the string, which nature dominates in a particular problem depends on both the problem and the phenomenon being considered. The frequency domain view answers why drums and violins do not sound alike. The time domain view explains how the sound from either instrument propagates across a room to a listener's ear.

4.5 The Wave Equation in Circular Regions

We turn now to the physically more realistic, and mathematically more challenging problem of describing the vibrations of a circular membrane (i.e. a standard drum). The natural
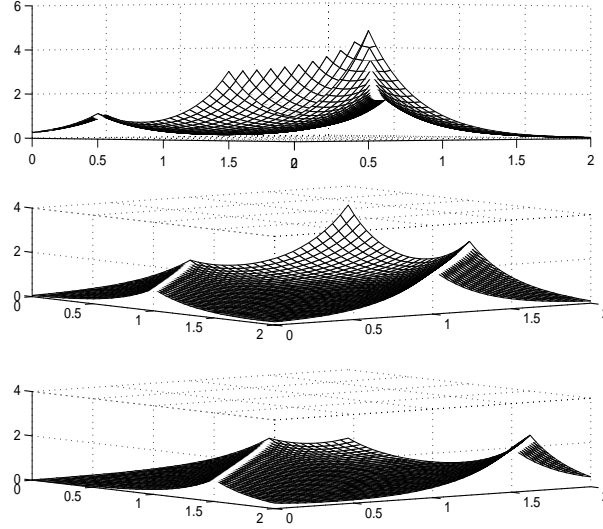
159

Figure 43: Two Plane Waves Traveling in the Directions $\hat{\mathbf{k}}$ and $\tilde{\mathbf{k}}$

coordinate system, of course, in which to formulate this problem is polar (or cylindrical) coordinates. Furthermore, as we have already noted, in this coordinate system the two-dimensional wave equation becomes

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right\} \quad . \tag{4.5.16}$$

Qualitatively, this problem retains some similarities to the rectangular membrane, most notably in that it is still second order with respect to t, r and θ. However, it has one new, striking dissimilarity. The original two-dimensional wave equation, written in terms of the Laplacian, represented a uniform membrane and had constant coefficients. In (4.5.16), however, as a direct consequence of our expressing the Laplacian in cylindrical coordinates, we have a *variable coefficients* partial differential equation. Although we shall not prove so, the appearance of such variable coefficients can be shown to be not an isolated occurrence, but an expected outcome of expressing the Laplacian in other than rectangular coordinates.

The appearance of these variable coefficients is not particularly auspicious! After all, our experience with ordinary differential equations should have shown that moving from constant coefficient problems to variable coefficient ones caused a significant increase in computational difficulty, since the solutions were no longer simple exponentials, but usually some kind of general infinite series. (This difficulty, however, was still far less than that encountered with nonlinear differential equations.) But, first things first! Before we can even talk about solving **this** problem, we must again address the fact that any differential equation, in and of itself, represents an incomplete problem. Moreover, based on now familiar arguments, it seems that to convert (4.5.16) into a complete problem will require defining two boundary conditions in r, two boundary conditions in θ, and two initial conditions in t. Specifying the latter (i.e. appropriate initial conditions for this problem) is not a challenge. As before,

160

they will simply be the initial position and velocity, i.e.

$$\begin{aligned} u(r, \theta, 0) &= f(r, \theta) \\ u_t(r, \theta, 0) &= g(r, \theta) \end{aligned}$$

However, deducing the appropriate boundary conditions for this problem is not quite so straightforward. After all, the circular drum has *only one physical boundary*, the outer edge $(r = L)$! For any normal drum, of course, this outer edge is firmly clamped to the frame of the drum, which clearly implies that the boundary condition

$$u(L, \theta, t) = 0$$

is appropriate. But where is the other boundary in r? And where do the two boundary conditions in θ come from?

In order to understand how these "missing" boundary conditions are developed, we need to first appreciate one fundamental and basic fact. This fact is that while this problem has only one physical boundary, there are really three other *mathematical boundaries* - places which, while they do not coincide with a physical ending of the medium, nevertheless cannot be crossed, any more than we can cross the boundary at $r = L$ without leaving the drum. (We shall see precisely where these boundaries are shortly.) Physically, these other boundaries are artifices of the particular coordinate system involved. Nevertheless, the differential equation sees them as completely real. Furthermore, as we shall now see, these mathematical boundaries in fact introduce precisely the conditions necessary to ensure that (4.5.16) have a unique solution.

The first such mathematical boundary occurs at $r = 0$, for, by convention the radius cannot take on negative values in cylindrical coordinates. Any situation that might mathematically, lead to a negative radius is taken care of by keeping the radius positive and changing the angle (θ) by 180°. (In nautical terms, for example, one would never hear of another ship being sighted at "bearing twenty, range minus ten miles.") But since $r = 0$ is a boundary, what conditions are appropriate there? The answer to this question is fairly easy to deduce, at least for the circular drum. The fundamental requirement, here and in all similar situations, is that *mathematical boundaries may not introduce physically unrealistic solutions* into the problem. But how does this produce a boundary condition at $r = 0$?

Well, as we noted earlier, the partial differential equation (4.5.16) becomes singular at the origin. Moreover, as is well-known, at least in the study of ordinary differential equations, singular points in the differential equation may cause corresponding singularities in the solutions themselves. However a singular solution would imply an infinitely large vertical displacement at the center of the drum, which would clearly be physically unrealistic. Equally unacceptable would be a solution whose gradient became singular at the center. Therefore, we must now pose a boundary condition which ensures that such solutions will not occur mathematically in the solution to (4.5.16). One such condition is

$$u(0, \theta, t), \quad \frac{\partial u}{\partial r}(0, \theta, t) \quad \text{are} \quad \text{finite} \quad .$$

(This condition actually involves a bit of "overkill." Simply requiring that $u(0, \theta, t)$ be finite will turn out to have been sufficient to produce the correct solutions in all cases we shall

161

consider.) Thus we now have our second boundary condition in r, and so may turn to developing the appropriate boundary conditions in angle.

The mathematical boundaries in θ also arise because of convention related to the coordinate system. Specifically, in polar coordinates, the values of θ are restricted to lie in a 360° range, e.g. from $-\pi$ to π. Any mathematical calculation which would produce an angle outside of this range is taken care of by increasing or decreasing the argument by a sufficient multiple of 2π to bring it back into the acceptable range. (For example, $\theta = 760^o$ is reset to 40°.) This restriction in effect introduces boundaries at $\theta = -\pi$ and $\theta = \pi$. (Actually, we could equally well place these boundaries at 0 and 2π. This would make no difference in the condition we shall develop.) But, mathematically, how do we ensure the solutions to (4.5.16) recognize this? One way is to impose a periodicity restriction

$$u(r, \theta, t) = u(r, \theta + 2\pi, t) \qquad . \qquad (4.5.17)$$

(This condition simply says that because the points (r, θ) and $(r, \theta + 2\pi)$ are physically identical, then the solutions there must also be identical.) But wait! Didn't we say we needed **two** boundary conditions in θ ?

Well, we did. And this periodicity condition actually does lead to two conditions in θ, although this last fact is probably not obvious! The reason is that we can evaluate (4.5.17) at $\theta = -\pi$, yielding

$$u(r, \pi, t) = u(r, -\pi, t) \quad ,$$

and then take a partial derivative of the same condition with respect to θ, and evaluate that derivative also at $\theta = -\pi$ to yield the second condition:

$$u_\theta(r, \pi, t) = u_\theta(r, -\pi, t) \qquad .$$

Thus we in fact have two boundary conditions in θ, although we shall choose to still use the periodicity condition since it is slightly simpler to write and to understand.

Since we now appear to have developed the requisite number of boundary conditions, we should therefore have enough information to ensure the existence of a unique solution to (4.5.16). So we simply combine all of these various pieces into the following complete problem for the circular drum

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right\}$$

$$u(0, \theta, t), \ \frac{\partial u}{\partial r}(0, \theta, t) \ \text{ are } \ \text{finite}$$

$$\qquad (4.5.18)$$

$$u(L, \theta, t) = 0$$
$$u(r, \theta, t) = u(r, \theta + 2\pi, t)$$
$$u(r, \theta, 0) = f(r, \theta)$$
$$u_t(r, \theta, 0) = g(r, \theta)$$

As it turns out, this is still a fairly difficult conceptual and computational problem. Therefore, in the next section, we shall restrict ourselves to investigating the solution of a slightly simpler special case of it.

162

4.6 Symmetric Vibrations of the Circular Drum

In these notes, we shall solve in detail only one special case of the wave equation for circular regions. This case nevertheless exhibits the most important aspects of this problem - especially those aspects which are new and different compared to properties of solutions in rectangular regions. The special case we shall study is that of symmetric vibrations, i.e. those which are independent of angle. (Such vibrations should arise, due to symmetry, if, for example, a drum were struck precisely in its center.) The boundary value problem which describes such vibrations may be found simply by dropping all of the angle-related dependence in (4.5.18) and setting all derivatives with respect to angle equal to zero. This yields

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right]$$

$$u(0,t), \ \frac{\partial u}{\partial r}(0,t) \ \text{ are } \ \text{finite} \tag{4.6.19}$$
$$u(L,t) = 0$$
$$u(r,0) = f(r)$$
$$u_t(r,0) = g(r)$$

Like the simple vibrating string, this problem involves precisely two independent variables, r and t. Furthermore, like the string, the partial differential equation here possess a natural splitting - only time operations on the left, and only spatial operations on the right. Therefore, using the by-now familiar logic of separation of variables, we shall assume that the homogenous solutions, i.e. the solutions to

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right]$$

$$u(0,t), \ \frac{\partial u}{\partial r}(0,t) \ \text{ are } \ \text{finite} \tag{4.6.20}$$

$$u(L,t) = 0$$

have the form

$$u(r,t) = R(r)T(t) \qquad .$$

Substituting this product into the partial differential equation and then dividing by the product $c^2 R(r)T(t)$ yields

$$\frac{T''(t)}{c^2 T(t)} = \frac{(rR'(r))'}{rR(r)} = -\lambda \qquad .$$

Here, as before, we introduce the separation constant (λ) because of the standard argument - the left hand side is independent of r and the right hand side is independent of t. (The minus sign is included to conform to our earlier experience.) This last equation can then be written as the two equivalent equations

$$T'' + \lambda c^2 T = 0 \qquad \text{and} \qquad (rR')' + \lambda rR = 0 \quad ,$$

where we have again chosen to suppress the explicit dependence on the variables.

At this point, both of these differential equations are, in theory, solvable for any value of λ. But, of course, as we have also seen before, that is because we have not yet applied our separation assumption to the full homogeneous problem, i.e. we must still separate the boundary conditions. In terms of the separated form, the condition at the origin becomes

$$
\begin{aligned}
u(0,t) &= R(0)T(t) \text{ is finite, and} \\
\frac{\partial u}{\partial r}(0,t) &= R'(0)T(t) \text{ is also finite,}
\end{aligned}
$$

which clearly implies

$$R(0), R'(0) \text{ are finite,}$$

or stated more compactly

$$R(0), R'(0) \text{ finite.}$$

The boundary condition at the outer rim of the drum is identical to that in Dirichlet conditions on the string, and separates similarly

$$u(L,t) = R(L)T(t) = 0 \Rightarrow R(L) = 0 \ .$$

Thus, the complete homogenous problem for the separation form solutions becomes

$$
\begin{array}{ll}
T'' + \lambda c^2 T = 0 & (rR')' + \lambda r R = 0 \\
& R(0), R'(0) \text{ finite} \\
& R(L) = 0
\end{array}
\qquad (4.6.21)
$$

By now, the next step should be fairly obvious. We must determine the eigenvalues and eigenfunctions of

$$
\begin{array}{l}
(rR')' + \lambda r R = 0 \\
R(0), R'(0) \text{ finite} \\
R(L) = 0
\end{array}
\qquad (4.6.22)
$$

Before we look at the details of solving this problem, there are a few general observations we should make. First of all, in one important aspect this problem is radically different from all of our earlier cases - the ordinary differential equation is **not** constant coefficient. (As we noted earlier, this is a direct consequence of the representation of the Laplacian in nonrectangular coordinates.) Equally important, however is the fact that, even with this added twist, the problem is still Sturm-Liouville, although of a slightly different type than we have encountered before. (The correspondence to the standard Sturm-Liouville form

$$[p(x)y'(x)]' + [\lambda w(x) - q(x)]y(x) = 0 \ , \quad 0 < x < L$$

is $x = r$, $y(r) = R(r)$, $p(r) = r$, $w(r) = r$, and $q(r) \equiv 0$.) It is, however, not a regular problem, in the sense we defined in the previous chapter, because it violates the condition that $p(x)$ should be positive on the entire interval, including the end points, but is an example of what is called a *singular* Sturm-Liouville problem. Nevertheless, its solutions can be shown to possess essentially all of the same properties as we found for solutions of the

164

regular problem. Thus, for example, there will be no negative eigenvalues. Furthermore, zero will not be an eigenvalue, because a nonzero constant cannot satisfy the boundary condition at $r = L$. Moreover, there will be an infinite number of positive eigenvalues, $\lambda_n = \xi_n^2 \to \infty$ as $n \to \infty$. Lastly, each eigenfunction will have one more zero in the interval $0 < r < L$ than did its predecessor. And all of these conclusions can be made even before we know the form of the general solution to the variable coefficient ordinary differential equation involved!

We are however, still left with the problem of calculating the specific eigenvalues and eigenfunctions of (4.6.22) when $\lambda > 0$. Conceptually, this could be reasonably difficult given the variable coefficient nature of the differential equation. According to the theory of ordinary differential equations, we would have to assume a general power series solution, i.e.

$$R(r) = \sum_{n=0}^{\infty} a_n r^{n+p} \quad ,$$

and then substitute this form into the differential equation to determine the value(s) of p and the recurrence relation for the coefficients. Fortunately, we will not have to carry out all of these steps here, because the ordinary differential equation which appears in this problem has already been extensively studied. This equation, which we shall rewrite in the equivalent form

$$r^2 R'' + r R' + \lambda r^2 R = 0 \qquad (4.6.23)$$

is called Bessel's[12] equation of order zero. (Appendix A.) The general solution to this problem involves two linearly independent functions (as does the general solution to any second order linear ordinary differential equation). These solutions are normally represented by the symbols $J_0(\xi r)$ and $Y_0(\xi r)$, and called the ordinary Bessel functions of the first and second kinds, of order zero, respectively. Both of them can be calculated using infinite series which converge for all values of r, and most standard computer systems include programs which will compute their values. Their graphs are shown at Figure 44. For our purposes, we shall treat these functions as somewhat like "black boxes," in that every time we encounter (4.6.23), we shall simply proceed to write its general solution as a linear combination of $J_0(\xi r)$ and $Y_0(\xi r)$ with the full assurance that these functions could be calculated whenever required, even though we may not at this point fully appreciate the exact mechanisms by which those calculations would be made. (This approach may seem a bit troubling, and perhaps should be, at least initially. However, you might reflect that, to a fair degree, we really do little more than this when we write the solution to

$$y'' + \xi^2 y = 0 \qquad \text{as} \qquad y(x) = a_1 \cos(\xi x) + a_2 \sin(\xi x) \quad ,$$

for, after all, what do we really immediately know about the sine and cosine except

- What their graphs look like,

- That their values can be obtained either by using a calculator (or computer), or from tables, and

---

[12]Friedrich Wilhelm Bessel,
see http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Bessel.html

- That they satisfy certain identities?

For example, do you really know how that chip inside your pocket calculator arrives at a value when you hit the **SIN** key? You probably don't, and it's really not all that important. What is crucial is that those values are accurate and you can use them in calculations! We shall adopt a similar view toward Bessel functions.)
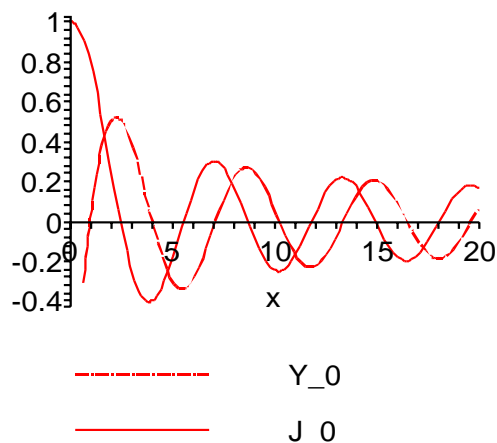

Figure 44: The Ordinary Bessel Functions $J_0(r)$ and $Y_0(r)$

Since, as noted above, we can now represent the general solution to the differential equation in (4.6.22) (with $\lambda = \xi^2$) as

$$R(r) = a_1 J_0(\xi r) + a_2 Y_0(\xi r) \qquad . \tag{4.6.24}$$

We must now determine the exact eigenvalues by applying the boundary conditions to this solution. Referring to Figure 44, we see that $Y_0(\xi r)$ is singular at the origin, therefore it cannot satisfy the boundary condition at $r = 0$, and thus it cannot appear in the solution i.e.

$$R(0), R'(0) \text{ finite} \Rightarrow a_2 = 0 \quad .$$

(Note we could **not** make a similar conclusion in any case which did not include the origin as part of the physical region, e.g. if we were to build a "donut-shaped" drum. But in such a region, there would be a second real physical boundary (at the inner radius), and hence we would not have needed the finiteness condition anyway.)

Finally, the remaining boundary condition in (4.6.22) yields

$$R(L) = 0 \Rightarrow a_1 J_0(\xi L) = 0 \Rightarrow J_0(\xi L) = 0 \quad . \tag{4.6.25}$$

Sturm-Liouville theory and the general shape of the graph of $J_0(r)$ in Figure 44 assure us that there will be an infinite number of different values of ξ which satisfy this condition. In fact, every axis crossing of $J_0(r)$ in that figure will generate another solution of (4.6.25), in the same way that every axis crossing of the curve of $\sin(x)$ generated another integer multiple

166

of π, and another solution of $\sin(\xi L) = 0$ in the vibrating string problem. Really the only difference between solving (4.6.25) and determining the eigenvalues of the one dimensional string problem is that, unfortunately, there is no convenient general formula (such as $n\pi$) that describes the roots of $J_0(r) = 0$, and therefore no convenient formula for representing the solutions of $J_0(\xi L) = 0$. They can only be listed, e.g.

$$\xi_1 L = 2.4048 \;\; , \quad \xi_2 L = 5.5201 \;\; , \quad \xi_3 L = 8.6537 \;\; , \quad \cdots$$

or

$$\xi_1 = \frac{2.4048}{L} \;\; , \quad \xi_2 = \frac{5.5201}{L} \;\; , \quad \xi_3 = \frac{8.6537}{L} \;\; , \quad \cdots \tag{4.6.26}$$

Again, we choose to view this as more of a notational inconvenience than a computational drawback, and therefore conclude that we have found the eigenfunctions for the problem

$$R_n(r) = J_0(\xi_n r) \;\; , \;\; n = 1, 2, \ldots \tag{4.6.27}$$

with the ξ_n as defined above. (Furthermore, and we shall need to use this shortly, Sturm-Liouville theory also assures that these eigenfunctions will be orthogonal with respect to the weighting function $w(r) = r$, i.e. that

$$\int_0^L r J_0(\xi_n r) J_0(\xi_m r) dr = 0 \;\; , \;\; m \neq n \;\; . \tag{4.6.28}$$

Lastly, these eigenfunctions are guaranteed to be complete in the sense that we can use them to expand any piecewise smooth function defined on $0 < r < L$ in a Fourier-like series.)

Of course, with the eigenfunctions in hand, we have still not completely identified the linearly independent solutions to the homogeneous partial differential equation. We need also to solve for the $T(t)$. But this isn't really very difficult, since we know, according to our above development and (4.6.21) that the $T(t)$ must satisfy the equation

$$T_n'' + \xi_n^2 c^2 T_n = 0 \;\; . \tag{4.6.29}$$

Therefore, as before, there will be two linearly independent time solutions, which we represent as

$$T_n(t) = A_n \cos\left(\xi_n ct\right) + B_n \sin\left(\xi_n ct\right) \;\; . \tag{4.6.30}$$

Thus, in the usual way, the general solution to (4.6.20) can be formally written

$$u(r, t) = \sum_{n=1}^{\infty} \left\{ A_n \cos\left(\xi_n ct\right) + B_n \sin\left(\xi_n ct\right) \right\} J_0(\xi_n r) \;\; . \tag{4.6.31}$$

Of course, by calling this the general solution we mean it satisfies not only the homogeneous partial differential equation and the boundary conditions, but also should, for the correct choice of A_n and B_n, be able to satisfy any appropriate initial conditions.

Trying to fit the initial conditions in the original boundary value problem (4.5.18) with this solution requires substituting $t = 0$ into (4.6.31) and its partial derivative with respect

to t, and leads to:

$$u(r,0) \;=\; \sum_{n=1}^{\infty} A_n J_0(\xi_n r) = f(r) \quad,$$

(4.6.32)

$$u_t(r,0) \;=\; \sum_{n=1}^{\infty} \xi_n c B_n J_0(\xi_n r) = g(r) \quad.$$

These two equations really represent the same general problem. The only difference is that the effective coefficient in the second initial condition is not A_n, but $\xi_n c B_n$. Thus we shall focus on determining the coefficients in the first initial condition, with the assurance that whatever approach will find the coefficients in that problem could be immediately applied to find the coefficients in the initial velocity condition.

Of course, were it not for Sturm-Liouville theory, finding a formula for the coefficients in (4.6.32) could be a formidable task. But because this theory assures us that since the $J_0(\xi_n r)$ are orthogonal, we can apply the general Sturm-Liouville coefficient formula (given by (3.10.57) in Chapter 3) directly. Equivalently, we could in the usual way multiply both sides of (4.6.32) by $r J_0(\xi_m r)$, where m denotes some fixed integer, and then integrate yielding

$$\int_0^L r f(r) J_0(\xi_m r) dr = \sum_{n=1}^{\infty} A_n \underbrace{\int_0^L r J_0(\xi_n r) J_0(\xi_m r) dr}_{=\ 0\ ,\quad m\neq n},$$

and then apply the orthogonality property. Both approaches will lead to the same formula

$$A_n = \frac{\displaystyle\int_0^L r f(r) J_0(\xi_n r) dr}{\displaystyle\int_0^L r \left[J_0(\xi_n r)\right]^2 dr} \quad.$$

(4.6.33)

So we can formally compute the A_n! (A similar set of steps can be shown to produce the following formula for the B_n

$$B_n = \frac{\displaystyle\int_0^L r g(r) J_0(\xi_n r) dr}{\displaystyle \xi_n c \int_0^L r \left[J_0(\xi_n r)\right]^2 dr} \qquad .)$$

(4.6.34)

You should appreciate (and almost certainly do) that these integrals will generally not be elementary, i.e. the antiderivatives will **not** be found in most standard calculus texts. In most cases, the antiderivatives may not even be known, and therefore the values of the various coefficients would have to be done *numerically*, e.g. by Simpson's rule. Nevertheless, mathematically, the results of Sturm-Liouville theory ensure us that (4.6.32) is really a valid set of equalities, and that, using the coefficients computed by (4.6.33)-(4.6.34) will satisfy the initial condition, provided the $J_0(\xi_n r)$ include *all* of the linearly independent eigenfunctions of the associated Sturm-Liouville problem which comes from separation of variables. This effectively completes the solution of (4.5.18).

We close this section with a brief example illustrating the above method

$$\frac{\partial^2 u}{\partial t^2} = \frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial u}{\partial r}\right] \quad , 0 < r < 4 , \quad 0 < t$$

$$u(0,t), \ \frac{\partial u}{\partial r}(0,t) \text{ finite}$$

$$u(4,t) = 0$$

$$u(r,0) = \left(1 - \frac{r^2}{16}\right)$$

$$u_t(r,0) = 0$$

This represents a circular drum, of radius four, with a tension and density such that $c^2 = 1$. There is an initial displacement, but no initial velocity. According to the above development, the eigenfunctions are

$$R_n(r) = J_0(\xi_n r) , \ n = 1, 2, \ldots \quad ,$$

where the eigenvalues are determined by

$$J_0(4\xi_n) = 0 , \ n = 1, 2, \ldots \quad ,$$

and thus have the values

$$\xi_1 = 0.6012 \ , \quad \xi_2 = 1.3800 \ , \quad \xi_3 = 2.1634 \ ,$$

$$\xi_4 = 2.9479 \ , \quad \xi_5 = 3.7327 \ , \quad \xi_6 = 4.5178 \ , \quad \cdots \quad .$$

The general solution is then (since $c^2 = 1$)

$$u(r,t) = \sum_{n=1}^{\infty} \{A_n \cos(\xi_n t) + B_n \sin(\xi_n t)\} J_0(\xi_n r) \quad ,$$

with the ξ_n as given. The initial conditions reduce to

$$u(r,0) = \sum_{n=1}^{\infty} A_n J_0(\xi_n r) = 1 - \frac{r^2}{16} \quad ,$$

$$u_t(r,0) = \sum_{n=1}^{\infty} \xi_n B_n J_0(\xi_n r) = 0 \Rightarrow B_n = 0 \quad ,$$

or,

$$u(r,t) = \sum_{n=1}^{\infty} A_n \cos(\xi_n t) J_0(\xi_n r) \quad ,$$

where, according to (4.6.33)

$$A_n = \frac{\displaystyle\int_0^4 r(1 - r^2/16) J_0(\xi_n r) dr}{\displaystyle\int_0^4 r\left[J_0(\xi_n r)\right]^2 dr} \quad .$$

Numerical integration could show

$$A_1 = \frac{\displaystyle\int_0^4 r(1 - r^2/16)J_0(.6012r)dr}{\displaystyle\int_0^4 r\left[J_0(.6012r)\right]^2 dr} = 1.1080$$

$$A_2 = \frac{\displaystyle\int_0^4 r(1 - r^2/16)J_0(1.3800r)dr}{\displaystyle\int_0^4 r\left[J_0(1.3800r)\right]^2 dr} = -0.1398$$

$$A_3 = \frac{\displaystyle\int_0^4 r(1 - r^2/16)J_0(2.1634r)dr}{\displaystyle\int_0^4 r\left[J_0(2.1634r)\right]^2 dr} = 0.0455$$

$$A_4 = \frac{\displaystyle\int_0^4 r(1 - r^2/16)J_0(2.9479r)dr}{\displaystyle\int_0^4 r\left[J_0(2.9479r)\right]^2 dr} = -0.0210$$

$$A_5 = \frac{\displaystyle\int_0^4 r(1 - r^2/16)J_0(3.7327r)dr}{\displaystyle\int_0^4 r\left[J_0(3.7327r)\right]^2 dr} = 0.0116$$

$$A_6 = \frac{\displaystyle\int_0^4 r(1 - r^2/16)J_0(4.5178r)dr}{\displaystyle\int_0^4 r\left[J_0(4.5178r)\right]^2 dr} = -0.0072 \ , \quad \dots \ .$$

4.7 Frequncy Domain Analysis of the Circular Drum

In the previous section, we showed that the vibrations of a circular membrane, fixed at the outer edge and with a non-zero initial displacement but no initial velocity, are given by

$$u(r,t) = \sum_{n=1}^{\infty} A_n \cos\left(\xi_n ct\right) J_0(\xi_n r) \quad ,$$

where the ξ_n are defined by

$$J_0(\xi_n L) = 0 \ , \ n = 1, 2, \dots \ .$$

As was the case with both the vibrating string and the rectangular membrane, we may rewrite this solution as

$$u(r,t) = \left[A_1 J_0(\xi_1 r)\right]\cos\left(\xi_1 ct\right) + \left[A_2 J_0(\xi_2 r)\right]\cos\left(\xi_2 ct\right) + \left[A_3 J_0(\xi_3 r)\right]\cos\left(\xi_3 ct\right) + \cdots \quad ,$$

170

i.e. as a sum of standing vibrations, each vibrating with a different frequency

$$f_n = \frac{\xi_n c}{2\pi} \quad , \quad n = 1, 2, \dots \quad .$$

The various modes of this vibration are, as before, simply the eigenfunctions of the problem

$$J_0(\xi_n r) \quad , \quad n = 1, 2, \dots \quad .$$

The first four of these are plotted in Figure 45. (This figure actually plots a full cross section of the membrane, where the negative points, in polar coordinates, really mean a change in angle of 180°.) Their shape is similar to that of the modes in the fixed end vibrating string, although not all of the "peaks" in any one mode have the same height.
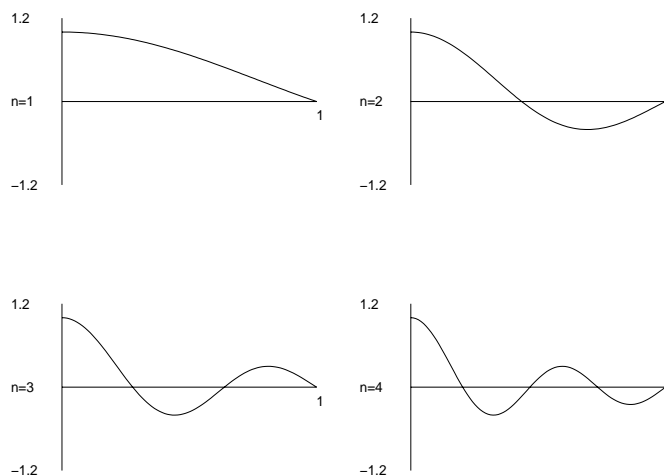

Figure 45: Modes of the Circular Membrane

The natural frequencies of the circular membrane are determined, except for a scale factor of $c/2\pi$, by the zeros of the Bessel function. But as we have already discussed at some length, these zeros are not uniformly spaced along the horizontal axis. Therefore, the spectrum of the circular membrane, like that of the rectangular membrane, will lack the regular picket fence structure of the spectrum of the vibrating string. However, the degree to which the circular membrane's frequencies are not evenly spaced is not as pronounced as is the case with the rectangular membrane. This is illustrated by Figure 46, which plots the frequencies of the circular membrane. One must look fairly closely at this figure to see the uneven spacing. Nevertheless, they are not evenly spaced and therefore, while which would sound "better" is clearly a matter of personal taste, the unmistakable conclusion is that a circular drum will not sound like a violin!

4.8   Time Domain Analysis of the Circular Membrane

A complete analysis of the structure of traveling waves in cylindrical coordinates is beyond the scope of our discussion here. Nevertheless, we claim that such an interpretation is possible, and that circular traveling waves in fact exist. For the proof of this, one need only drop a small stone into a small, flat pond of water and observe the resulting motion.
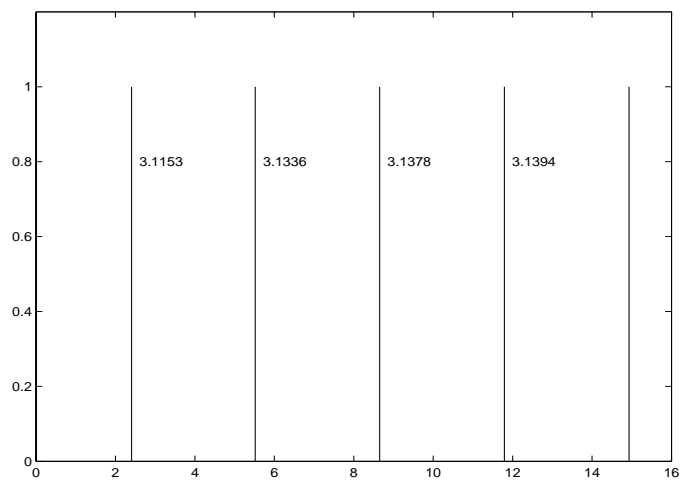
Figure 46: Spectrum of the Circular Membrane

PROBLEMS

1. It can be shown that the small free vertical vibrations of a uniform beam (e.g. a bridge girder) are governed by the fourth order partial differential equation:

$$\frac{\partial^2 u}{\partial t^2} + c^2 \frac{\partial^4 u}{\partial x^4} = 0$$

where c^2 is a constant involving the elasticity, moment of inertia, density and cross sectional area of the beam. If the beam if freely supported at both ends, e.g. sitting on a piling, the boundary conditions for this problem become:
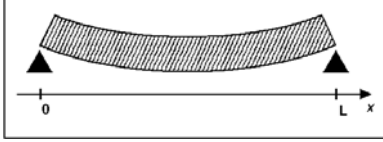
$$u(0, t) = u(L, t) = 0$$
$$u_{xx}(0, t) = u_{xx}(L, t) = 0$$

Show that separation of variables "works" in this problem, and, in case the beam is initially at rest, i.e.

$$u_t(x, 0) = 0$$

produces a general solution of the form:

$$\sum_{n=1}^{\infty} A_n \cos\left(\frac{n^2 \pi^2 ct}{L^2}\right) \sin\left(\frac{n\pi x}{L}\right)$$

2. Solve the two-dimensional rectangular wave equation:

$$u_{tt} = u_{xx} + u_{yy}$$
$$u(0, y, t) = u(1, y, t) = 0$$
$$u(x, 0, t) = u(x, 1, t) = 0$$
$$u(x, y, 0) = .01xy(1 - x)(1 - y)$$
$$u_t(x, y, 0) = 0$$

3. Solve the two-dimensional rectangular wave equation:

$$u_{tt} = 16 \left(u_{xx} + u_{yy}\right)$$
$$u(0, y, t) = u(3, y, t) = 0$$
$$u(x, 0, t) = u(x, 2, t) = 0$$
$$u(x, y, 0) = y(2 - y) \sin\left(\frac{2\pi x}{3}\right)$$
$$u_t(x, y, 0) = 0$$

173

4. Find the eigenvalues and the form of the eigenfunctions for:

$$u_{tt} = 9 \left(u_{xx} + u_{yy} \right)$$
$$u(0, y, t) = u(4, y, t) = 0$$
$$u_y(x, 0, t) = u(x, 1, t) = 0$$

Calculate the actual values of the four lowest natural frequencies.

5. One of the "quirks" of the two-dimensional wave equation in rectangular coordinates is that, unlike the one-dimensional problem, two different values of n and m may yield the same natural frequency, and therefore this single natural frequency may have two (or more) independent modes ("shapes") associated with it. For example, if $L = 2$ and $W = 1$, the eigenvalues and eigenfunctions are,

$$\lambda_{nm} = \left[\left(\tfrac{n}{2} \right)^2 + m^2 \right] \pi^2$$

and

$$u_{nm} = \sin(m\pi y) \sin \left(\tfrac{n\pi x}{2} \right) \quad .$$

Show that the following eigenvalues are in fact equal:

$$\lambda_{41} = \lambda_{22} \ ; \quad \lambda_{61} = \lambda_{23} \ ; \quad \lambda_{62} = \lambda_{43} \ ; \quad \lambda_{72} = \lambda_{14}$$

6. Show that in the square membrane, certain natural frequencies may have four independent modes ("shapes") associated with them.

PROBLEMS

1. Show that separation of variables $(u(r, \theta, t) = R(r)\Theta(\theta)T(t))$, applied to the wave equation in a circular region of radius A,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right\}$$

$$u(A, \theta, t) = 0$$

$$u(0, \theta, t) , \quad \frac{\partial u}{\partial r}(0, \theta, t) \quad \text{finite}$$

$$u(r, \theta, t) = u(r, \theta + 2\pi, t)$$

leads to

$$T'' + \lambda c^2 T = 0 \qquad\qquad r\left[rR'\right]' + (\lambda r^2 - \mu)R = 0 \qquad\qquad \Theta'' + \mu\Theta = 0$$
$$R(0) , \quad R'(0) \quad \text{finite} \qquad\qquad \Theta(\theta) = \Theta(\theta + 2\pi)$$
$$R(A) = 0$$

2. Explain the mathematical and physical significance of the condition

$$u(r, \theta, t) = u(r, \theta + 2\pi, t).$$

3. Find the three lowest natural frequencies of

$$u_{tt} = \frac{6}{r} \frac{\partial}{\partial r} \left[r \frac{\partial u}{\partial r} \right]$$

$$u(4, t) = 0$$
$$u(0, t) , \quad \frac{\partial u}{\partial r}(0, t) \quad \text{finite}$$

$$u(r, 0) = f(r)$$
$$u_t(r, 0) = 0$$

You need the Appendix to solve the following two problems.

4. Using the recurrence formulas, and a table of values for $J_0(x)$ and $J_1(x)$, find
   a. $J_1'(x)$ in terms of $J_0(x)$ and $J_1(x)$
   b. $J_2(2.0)$
   c. $J_3'(1.0)$

5. Write, in terms of $J_n(x)$ and $Y_n(x)$, the general solution to

a. $x^2 y'' + xy' + 4x^2 y = 0$

b. $x^2 y'' + xy' + (9x^2 - 4)y = 0$

c. $4x^2 y'' + 4xy' + (x^2 - 1)y = 0$

6. Solve the following problems. (*Non-zero* coefficients may be left in terms of definite integrals of known functions.)

a.

$$u_{tt} = \frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial u}{\partial r}\right]$$

$$u(2, t) = 0$$
$$u(0, t), \quad \frac{\partial u}{\partial r}(0, t) \quad \text{finite}$$

$$u(r, 0) = \sin(\pi r)$$
$$u_t(r, 0) = 0$$

b.

$$u_{tt} = \frac{4}{r}\frac{\partial}{\partial r}\left[r\frac{\partial u}{\partial r}\right]$$

$$u(1, t) = 0$$
$$u(0, t), \quad \frac{\partial u}{\partial r}(0, t) \quad \text{finite}$$

$$u(r, 0) = 1 - r^2$$
$$u_t(r, 0) = 0$$

c.

$$u_{tt} = c^2\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial u}{\partial r}\right]$$

$$u(2, t) = 0$$
$$u(0, t), \quad \frac{\partial u}{\partial r}(0, t) \quad \text{finite}$$

$$u(r, 0) = 0$$
$$u_t(r, 0) = 1$$

7. Solve the following problem. (*Non-zero* coefficients may be left in terms of definite integrals of known functions.) Physically interpret the boundary conditions, and relate this

to the properties of the solution:

$$u_{tt} = c^2 \frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial u}{\partial r}\right]$$

$$u_r(L,t) = 0$$
$$u(0,t) \, , \quad \frac{\partial u}{\partial r}(0,t) \quad \text{finite}$$
$$u(r,0) = f(r)$$
$$u_t(r,0) = 0$$

5 Introduction to the Fourier Transform

5.1 Periodic and Aperiodic Functions

Thus far in our study we have focused on developing an understanding of the properties and uses of the Fourier series (and its immediate relatives). Almost from the beginning we have seen the power of Fourier series as a tool for both decomposing and constructing general (usually complicated) *periodic* functions in terms of "pure" sines and cosines, and we have seen how these series provide significant insights into the physical world. The physical world, however, is full of interesting *aperiodic* functions as well as periodic ones. In fact, a cynic might well argue "**Only** the aperiodic functions are real. A truly periodic function must continue, unabated, for **all** time, which clearly requires **infinite energy**, and therefore the existence of such functions is impossible." (For example, as any astronomer will attest, even the Sun itself is actually "running down" - losing energy every day!) The pragmatist, of course, would likely counter "So what! Mathematical physics is based on modeling approximations, neglecting small terms, etc. As long as we can *model* an aperiodic function sufficiently closely with a periodic one (or vice versa), what's the harm? Moreover, if it takes a million years to tell the difference, who's going to notice!"

The degree to which the Fourier series results we've studied so far do agree with observed physical behavior argues strongly that we have not in fact made any unwarranted approximations. Therefore, since aperiodic functions are a very real fact of life, in the coming sections we shall study how they and their properties relate, in the Fourier sense, to those of the periodic functions.

We start by considering how we might represent an arbitrary, aperiodic function. We shall denote this function as $h(t)$. (We choose not to call it $f(t)$ for two reasons. First, we wish to reinforce the notion that not all functions use the letter f. Secondly, we are later on going to want to use the letter f to stand for a different physical quantity.) But since $h(t)$ is not periodic, we cannot represent it in terms of a Fourier series - at least not immediately. So how do we start?

We shall take an approach which generally produces valid results in applied mathematics - combine a bit of common sense with an appeal to physical "reasonableness." Specifically, we introduce a "new" function, denoted $\tilde{h}(t)$, with the properties that

(1) $\tilde{h}(t)$ is periodic, of period $2L$,

(2) L is "large," and

(3) $\tilde{h}(t)$ and $h(t)$ are identical for $-L < t < L$.

Our idea, of course, is that if L is sufficiently large, say a million years, then for all *practical* purposes $\tilde{h}(t)$ and $h(t)$ will be indistinguishable. Moreover, reason says that the larger L becomes, the closer their mathematical representations should become. This suggests that, mathematically, we consider the limit of $\tilde{h}(t)$ as $L \to \infty$. (After all, this is a mathematics text!) In this limit, $\tilde{h}(t)$ and $h(t)$ should become physically identical, and therefore, by all

reason, the limit of any mathematical representation we've developed for $\tilde{h}(t)$ should also represent $h(t)$. But, of course, if we are to do all of this, we must first have a representation for $\tilde{h}(t)$! (You should appreciate that, in practical problems, we would also first need to have a strong understanding of the time scales of the dominant physical processes, since terms such as "large" or "small" really have meaning only in a *relative* sense. For example, a hour would clearly be a very large time for the analysis of a single radar pulse, while a million years might actually be too short to examine an astronomical phenomenon!)

5.2   Representation of Aperiodic Functions

Deriving a representation for $\tilde{h}(t)$ is not difficult. After all, we've assumed that it is periodic! We can therefore simply express it as Fourier series, and this is exactly what we shall do, albeit with one slight "twist" - we'll use the *complex* form of the series. (The reason for this is primarily convention. Almost all other texts develop a complex-valued representation for aperiodic functions. You should rest assured that a completely identical, if somewhat more cumbersome representation could be derived from the real form of the Fourier series.) The complex form of the Fourier series, as we have already shown, is:

$$\tilde{h}(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\pi t/L} \tag{5.2.1}$$

where (in deference to the convention of most signal processing texts) we now use j to stand for $\sqrt{-1}$, and the coefficients,

$$c_n = \frac{1}{2L} \int_{-L}^{L} \tilde{h}(t) e^{-jn\pi t/L} dt = \frac{1}{2L} \int_{-L}^{L} h(t) e^{-jn\pi t/L} dt \tag{5.2.2}$$

represent the (complex) amplitudes of the various component frequencies which make up $\tilde{h}(t)$. (Note the use of our assumption that $h(t)$ and $\tilde{h}(t)$ are identical for $-L < x < L$.) If we now use (5.2.2) to substitute for c_n in (5.2.1) (naturally changing to some dummy variable of integration other than t), we have:

$$\tilde{h}(t) = \sum_{n=-\infty}^{\infty} \frac{1}{2L} \left\{ \int_{-L}^{L} h(u) e^{-jn\pi u/L} du \right\} e^{jn\pi t/L} \tag{5.2.3}$$

We now modify this representation slightly by introducing the change of variables

$$\Delta \omega = \frac{\pi}{L} \qquad \text{and} \qquad \omega_n = \frac{n\pi}{L} \qquad . \tag{5.2.4}$$

(This change really involves nothing particularly new - ω_n is simply the nth radian frequency and $\Delta \omega$ the difference between two such adjacent frequencies in the representation for $\tilde{h}(t)$.) With these new variables, the series representation now becomes

$$\tilde{h}(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left\{ \int_{-L}^{L} h(u) e^{-j\omega_n u} du \right\} e^{j\omega_n t} \Delta \omega \tag{5.2.5}$$

179

But wait! This series is now of the form,

$$\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(\omega_n) \Delta\omega \qquad (5.2.6)$$

where

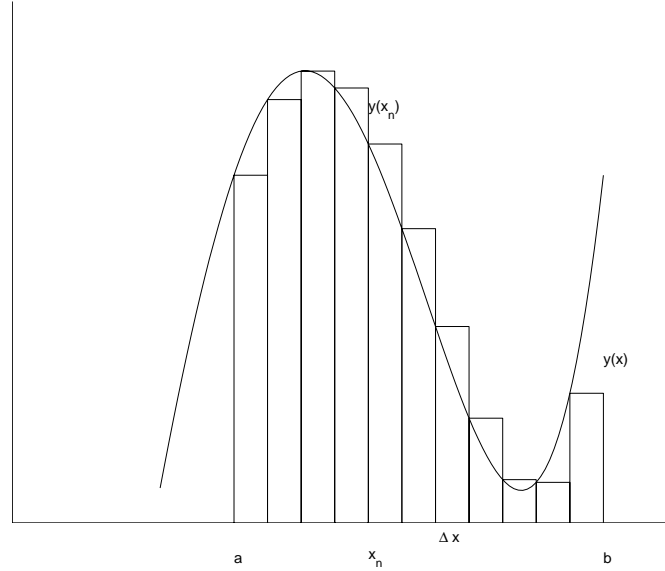$$F(\omega_n) = \left\{ \int_{-L}^{L} h(u) e^{-j\omega_n u} du \right\} e^{j\omega_n t}$$

Figure 47: Approximation of the Definite Integral

and the form of (5.2.6) should look at least a bit familiar. For it's very close (although as we shall see not quite identical) to the Riemann[13] sum used in virtually every calculus text as part of the introduction of the definite integral. In the Riemann sum, an interval of interest (say $a \le x \le b$) is first divided up into N equal segments of length Δx ($= (b-a)/N$). Then the area under some arbitrary continuous function $y(x)$ is approximated by the area under a set of rectangles, each of which has a base length of Δx, and a height equal to the height of the curve at some value of x (denoted x_n) inside the nth rectangle (Figure 47). The Riemann sum, which is just the total area under all of these rectangles, then clearly approximates the area under the curve. Moreover, as Δx becomes progressively smaller, this sum should become closer and closer to the actual area. Thus, since by definition the area under the curve is the definite integral, we then have that in Figure 47

$$\lim_{\Delta x \to 0} \sum_{n=1}^{N} y(x_n) \Delta x = \int_{a}^{b} y(x) dx \qquad (5.2.7)$$

---

[13]Georg Friedrich Bernhard Riemann,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Riemann.html

180

We now choose to view this last formula as simply a formal operational procedure - i.e. to arrive at the definite integral, start with the appropriate Riemann sum, then, in the limit, replace the summation by an integral sign with the proper limits, drop the subscript on (x_n), and finally replace Δx with dx. (Of course, significant and nontrivial mathematical results precede this formal procedure. One of the hardest of these is proving the limit even exists! And while we shall not concern ourselves with repeating the proofs of these results here, we must still recognize their crucial importance to the formal procedure we deduce from (5.2.7).)

Returning now to our attempt to derive a representation for $h(t)$, recall that we expect the representations for $\tilde{h}(t)$ and $h(t)$ to become identical in the limit as $L \to \infty$. However, according to (5.2.4), we see that

$$L \to \infty \qquad \text{is equivalent to} \qquad \Delta\omega \to 0 \quad ,$$

and so it seems perfectly reasonable to write

$$h(t) = \lim_{L\to\infty} \tilde{h}(t) \;\; = \lim_{L\to\infty} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left\{ \int_{-L}^{L} h(u)e^{-j\omega_n u}du \right\} e^{j\omega_n t} \Delta\omega$$

$$= \frac{1}{2\pi} \lim_{\Delta\omega\to 0} \sum_{n=-\infty}^{\infty} F(\omega_n)\Delta\omega$$

(5.2.8)

But, since the last term in (5.2.8) now looks like the left-hand side of (5.2.7), we should be able to apply the formal procedure we've just described to compute these limits! (Actually, this argument glosses over a great deal of "thin" theory. The interval of integration associated with (5.2.8) is $-\infty < \omega < \infty$, whereas (5.2.7) involves an interval of only finite length. Thus our integral will be improper, and so (5.2.8) represents a very non-trivial extension of the Riemann sum in somewhat the same way that the infinite series itself is a non-trivial extension of the usual sum. As with other similar concerns in this text, we shall not dwell on this one. We prefer, *for the moment*, to assume this extension will work, and see whether the results seem either useful or interesting, or both. If no useful results emerge from this approach, then whether they are valid or not becomes in essence moot. If, however, our formal procedure does produce apparently valuable results, we should then rigorously verify the validity of the various steps we're taking. Such verification, however, is far beyond the level of this text.)

If we do assume that (5.2.8) can be treated as the limit of a "normal" Riemann sum (5.2.7), then all that remains is the mechanics. As noted in our earlier comments, the first step of these mechanics is to replace the summation sign in (5.2.8) with an integral. Furthermore, since the values of ω_n in our summation ranges over all values from $-\infty$ to ∞, these values should become the limits on this integral. The second formal step in (5.2.7) is to drop the subscripts in what will become the integrand. In our present case, however, the integrand-to-be is

$$F(\omega_n) = \left\{ \int_{-L}^{L} h(u)e^{-j\omega_n u}du \right\} e^{j\omega_n t}$$

and therefore, since we are really taking the limit as L tends to ∞, we should not only drop the subscript (on ω_n), but also replace the upper and lower limits on this inner integral by ∞

181

and $-\infty$, respectively. (Applying this argument to the inner integral here again involves a step beyond the replacements in (5.2.7), and represents another part of our derivation that properly requires later theoretical analysis.) The final step of our formal procedure, which we must apply to (5.2.8), is to replace $\Delta\omega$ by $d\omega$. When we have completed all these steps as described, (5.2.8) becomes

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} h(u)e^{-j\omega u} du \right\} e^{j\omega t} d\omega \tag{5.2.9}$$

With (5.2.9) we have reached our goal of finding a representation for a general aperiodic function. But we're not quite done! The reason we say this is not that (5.2.9) is not a completely valid expression. It is! There are, however, other *equivalent* forms that are more commonly seen. One of these alternative forms, defined by the change of variables

$$\omega = 2\pi f \ , \tag{5.2.10}$$

is

$$h(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} h(u)e^{-2\pi j f u} du \right\} e^{2\pi j f t} df \tag{5.2.11}$$

This is the form we shall study for the rest of this text.

5.3   The Fourier Transform and Inverse Transform

We closed the previous section when we had derived (5.2.11) as the Fourier representation for a general aperiodic function, $h(t)$. We now consider some of the implications of this identity, and especially the similarities and differences between it and the Fourier series representation for a periodic function. Therefore, before continuing here, you may first want to review the previous section, concentrating especially on how each of the particular parts in (5.2.11) arose from the Fourier series for $\tilde{h}(t)$ in the course of our derivation.

One immediate and striking difference between the representation of an aperiodic function as given by (5.2.11) and the Fourier series representation of a periodic function, such as the complex form given by (5.2.1) is simply that representing an aperiodic function apparently requires an integral instead of a series. This is consistent with our observation, made during the derivation in the previous section, that the difference between neighboring radian frequencies in the series representation for $\tilde{h}(t)$ was π/L. But the limiting process of letting L go to infinity forces this difference to approach zero, in effect squeezing the neighboring frequencies of $\tilde{h}(t)$ closer and closer together until finally they become in essence a continuous "smear" extending across the entire frequency spectrum. Sums, of course, even infinite series, are inadequate to combine values of something with respect to a continuously varying argument - this is precisely why the definite integral appears in basic calculus courses.

In addition, we have very deliberately placed brackets inside the integral (5.2.11). These brackets effectively separate that integral into two parts, and allow us, if we choose, to easily replace this single identity with the **two** separate formulae

$$H(f) \ = \int_{-\infty}^{\infty} h(t)e^{-2\pi jft} \ dt$$

$$(5.3.12)$$

and

$$h(t) \ = \int_{-\infty}^{\infty} H(f)e^{2\pi jft} \ df$$

(Note that this separated form also removes any need to use the dummy variable of integration u in the first integral. Furthermore, since t is a dummy variable in that integral, the result must be, as we have denoted, a function of only the variable f.) Choosing to express the identity (5.2.11) in two parts should not seem either unusual or surprising, since this is virtually the same thing we routinely do with Fourier series. For example, don't we normally express the complex Fourier series with (5.2.1) and (5.2.2), rather than the combined form (5.2.3)?

The first formula (5.3.12) in this two-part representation for aperiodic functions is commonly called the **Fourier transform** (or Fourier integral transform) of $h(t)$. Moreover, as a review of the derivation of the previous section will show, the integral in this formula is essentially the direct descendent, via the limiting process as $L \rightarrow \infty$, of the Fourier series coefficient (5.2.2). Therefore we shall interpret it similarly - i.e. as describing, in some sense, the amplitudes of the continuum of component frequencies that make up the aperiodic function $h(t)$, i.e. $H(f)$ represents the (complex) amplitude of the "pure tone" of frequency f found in $h(t)$. Clearly then, in terms of our terminology thus far, $H(f)$ represents the frequency domain analysis of the time domain signal $h(t)$.

The second formula (5.3.12) above is commonly referred to as the **inverse Fourier transform** of $H(f)$. Since we are interpreting the Fourier transform as the analog in the case of aperiodic functions to the Fourier series coefficients for periodic functions, we then naturally view the inverse transform as describing how to reconstruct or assemble the aperiodic time domain signal, $h(t)$, from the (frequency domain) knowledge of the amplitudes of its component frequencies. In other words, the second of (5.3.12) performs the identical function for aperiodic functions that the Fourier series (5.2.1) itself does for periodic ones.

Calling the first of (5.3.12) the Fourier transform may bring to mind another transform you have undoubtedly seen - the *Laplace*[14] transform

$$\mathcal{L}[h(t)] = \int_0^{\infty} h(t)e^{-st}dt \ .$$

There are several similarities between the Fourier and Laplace transforms, some of which we shall investigate later in more detail. For the moment, we would simply note that the Fourier and Laplace transforms both involve an integral depending on another variable - s for the Laplace transform and f for the Fourier. (We must also note one significant difference between these two transforms - because Euler's identity implies that the complex exponential has both real and imaginary parts, **the Fourier transform will generally have a non-zero imaginary part** even when the time-domain function ($h(t)$) is real, while the Laplace transform of a real-valued function is real as long as s is.)

---

[14]Pierre-Simon Laplace,
see http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Laplace.html

The Fourier and inverse Fourier transforms are sufficiently important in applications that, like the Laplace transform, they are generally given their own symbol, usually a script capital letter. We shall follow this convention, and write

$$\mathcal{F}[h(t)] = H(f) = \int_{-\infty}^{\infty} h(t)e^{-2\pi jft} \, dt \tag{5.3.13}$$

and

$$\mathcal{F}^{-1}[H(f)] = h(t) = \int_{-\infty}^{\infty} H(f)e^{2\pi jft} \, df \tag{5.3.14}$$

You should recognize that any specific aperiodic function $h(t)$ is uniquely identified by its Fourier transform, just as any periodic function is uniquely identified by Fourier series coefficients. Therefore, for any particular function, (5.3.13) and (5.3.14) should (in fact must) be viewed as an inseparable *pair*, knowledge of either member of which serves to uniquely determine the other. (The same statement is true for Laplace transforms, and is precisely why tables of that transform are so useful in finding inverse transforms!) In fact, this pairing metaphor is so central to the concept of Fourier transforms that many texts use a special notation, such as

$$h(t) \leftrightarrow H(f)$$

to symbolize it.

The Fourier transform is, however, unlike the Laplace transform in one important respect. This difference arises because, due to their limits of integration, both the Fourier and Laplace transforms are improper integrals. Therefore, according to the standard arguments from the calculus, whether the transforms of particular functions actually even exist depends on the asymptotic behavior of the integrands at these limits. However, the $e^{-2\pi jft}$ term inside the Fourier transform, does **not** decay to zero as $t \to \pm\infty$, but continually oscillates. Therefore, the Fourier transform integral as we have defined it will exist only for time-domain functions $h(t)$ that themselves decay to zero as $t \to \pm\infty$ strongly enough to force convergence of the integral. (In contrast, the rapid decay at infinity of e^{-st} guarantees convergence of the Laplace transform for all but the most rapidly growing functions!) Proving the exact conditions under which a function will have a Fourier transform is beyond the scope of this text. Therefore we will simply state two common tests, either of which being satisfied are sufficient to ensure that the Fourier transform exists, at least in the usual calculus sense. (The reason for this last qualification is that, later on, we shall extend the notion of the Fourier transform to "functions" that are, in the view of the usual calculus, not well-behaved.) These tests are contained in the following

Theorem 5.1: The Fourier integral transform of the function $h(t)$,

$$H(f) = \int_{-\infty}^{\infty} h(t)e^{-2\pi jft} \, dt$$

will exist if either

$$\int_{-\infty}^{\infty} |h(t)| \, dt \; < \; \infty \qquad \text{or} \qquad \int_{-\infty}^{\infty} |h(t)|^2 \, dt \; < \; \infty$$

Functions which satisfy the second of these two tests are commonly referred to as *square integrable*. In light of our earlier discussions for Fourier series, we shall interpret square integrable functions are being those having only finite energy over all time. This theorem further implies that functions with infinite energy, e.g. truly periodic ones, will probably not have Fourier transforms, at least, again, in the sense of the usual calculus.

We close this section with one last insight about the Fourier transform. Observe that our fundamental identity (5.2.11) involves a double integral. Therefore, we may, at least formally, interchange the order of integration there and rewrite the identity as

$$h(t) = \int_{-\infty}^{\infty} h(u) \left\{ \int_{-\infty}^{\infty} e^{-2\pi j f u} e^{2\pi j f t} \, df \right\} du.$$

Something about this formula should look familiar! Specifically it seems to say that the value of a general aperiodic function at any specific point may be obtained as the integral of the product of (all values of) that function with the "function"

$$\int_{-\infty}^{\infty} e^{-2\pi j f u} e^{2\pi j f t} \, df \ = \ \int_{-\infty}^{\infty} e^{2\pi j f(t-u)} \, df \ .$$

(Note we call this last integral a function simply because it contains variables t and u which are not the variable of integration.) But only one "function," which you should have already encountered in conjuction with the Laplace transform, has the property of being so able to "punch out" a single value of the remaining parts of an integrand. This function is, of course, the *delta* function or *unit impulse*, normally symbolized $\delta(t - u)$. Thus we formally appear to have

$$\int_{-\infty}^{\infty} e^{2\pi j f(t-u)} \, df = \delta(t - u) \ . \tag{5.3.15}$$

This interpretation raises other questions - for example, does (5.3.15) embody a generalization of the concept of orthogonality to aperiodic functions? Such a question is intriguing. But before considering it, or any other questions, we wish to develop some feel for the mechanics of the Fourier transform by computing the transforms of some specific functions.

5.4 Examples of Fourier Transforms and Their Graphical Representation

In the previous section, we discussed the formulas for the Fourier transform. We shall now apply these to compute the transform of a specific function,

$$h(t) = \left\{ \begin{array}{ccc} 1 & , & |t| < \frac{1}{2} \\ 0 & , & \text{otherwise} \end{array} \right.$$

This function is commonly called a square pulse (Figure 48). (Note further it is not a square wave, which is periodic, but only one single, isolated pulse!) Since such a single pulse clearly contains only finite energy, we are assured by our earlier result that its Fourier transform exists. By definition (5.3.13), this transform is
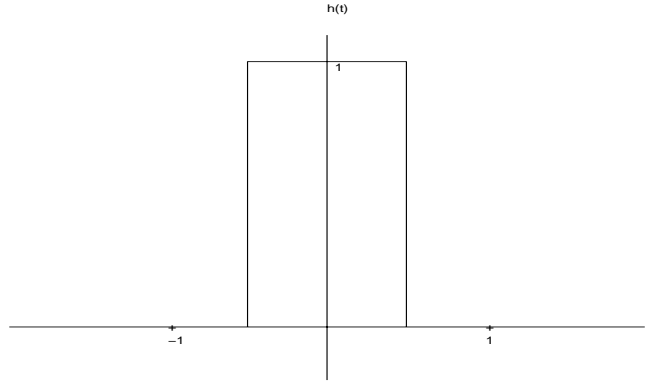
Figure 48: The Square Pulse

$$H(f) \ = \int_{-\infty}^{\infty} h(t)e^{-2\pi jft} \ dt = \int_{-1/2}^{1/2} 1 \cdot e^{-2\pi jft} \ dt$$

$$= -\left.\frac{e^{-2\pi jft}}{2\pi jf}\right|_{-1/2}^{1/2} = \frac{e^{j\pi f} - e^{-j\pi f}}{2\pi jf}$$

$$= \frac{\sin(\pi f)}{\pi f}$$

(where we simplified the result by using Euler's identity to replace the two comples exponentials). The graph of this $H(f)$ is shown in Figure 49. This curve represents one of the classic Fourier transforms, and many texts introduce the special function name of **sinc** just to describe it. More specifically,

$$\text{sinc}(x) \equiv \frac{\sin(x)}{x}$$

and therefore in this example $H(f) = \text{sinc}(\pi f)$. However, the name we choose to call this function is really secondary. What is important, and what we should focus on at this moment are its properties when we think of it as the frequency domain description of the square pulse, i.e. according to our earlier discussion, when viewed as displaying the amplitudes of the various frequencies necessary to (re)construct $h(t)$. One property is clearly that $H(f)$ seems to exist for *all* values of f. Therefore, we conclude that constructing the square pulse in the time domain requires energy at all frequencies (except $\pm 1, \pm 2, \ldots$). Secondly, as f approaches $\pm\infty$, this transform decays like $1/f$, or, in terms of our order notation from Chapter 1,

$$H(f) = \mathbf{O}\left(\frac{1}{f}\right)$$

On reflection, this asymptotic behavior should not be surprising, since $h(t)$ here is only piecewise continuous, and we know that amplitudes in the Fourier series for piecewise continuous periodic functions only decay like $1/n$, so this behavior is similar. Lastly, the shape of this graph seems to indicate that the large amplitudes, and hence the bulk of the energy in this

signal, occur in the lower frequencies, an insight which again parallels our experience with most Fourier series.
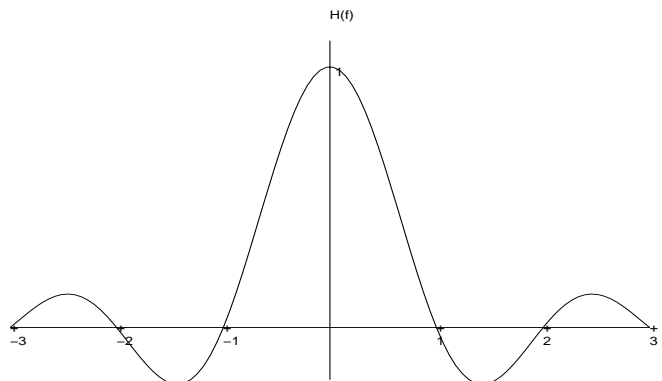

Figure 49: The Fourier Transform of the Square Pulse

You should realize that graphing $H(f)$ for the square pulse was quite straightforward since the sinc function has only real values. However, based on our earlier discussion about the general properties of the Fourier Transform, you should also appreciate that this last example was probably somewhat atypical. Specifically, due to the presence of the complex exponential in the Fourier transform integrand, we have already concluded that most functions will not have a purely real Fourier transform. So we shall now consider, again by example, how to graphically best convey the frequency domain information when the Fourier transform is complex-valued.

We choose the time-domain function

$$h(t) = \begin{cases} e^{-t} & , \quad 0 < t < \infty \\ 0 & , \quad otherwise \end{cases} \tag{5.4.16}$$

(Figure 50). According to our definition (5.3.13), its Fourier Transform is

$$H(f) \;\; = \int_{-\infty}^{\infty} h(t) e^{-2\pi j f t} dt$$

$$= \int_{0}^{\infty} e^{-t} e^{-2\pi j f t} dt = \left. -\frac{e^{-(1+2\pi j f)t}}{1 + 2\pi j f} \right|_{0}^{\infty}$$

$$= \frac{1}{1 + 2\pi j f}$$

$$= \frac{1 - 2\pi j f}{1 + (2\pi f)^2} = \frac{1}{1 + (2\pi f)^2} - j\frac{2\pi f}{1 + (2\pi f)^2}$$

As our discussion has led us to expect, and unlike the square pulse example, this transform is a true complex-valued function, with nonzero real and imaginary parts given respectively
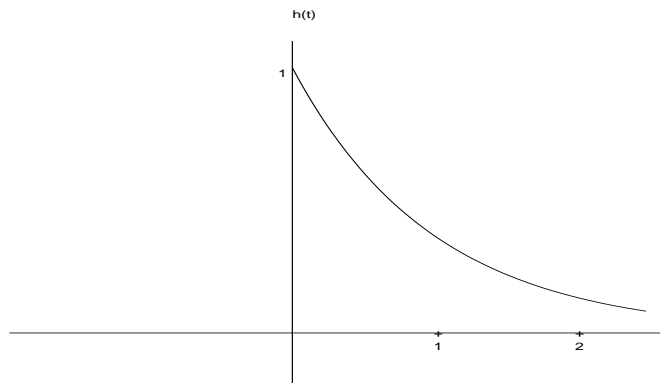
187

Figure 50: The Function $h(t)$ given by 5.4.16

by

$$\Re[H(f)] = \frac{1}{1 + (2\pi f)^2} \qquad \text{and} \qquad \Im[H(f)] = -\frac{2\pi f}{1 + (2\pi f)^2} \quad .$$

Therefore, graphing this transform provides a bit more of a challenge than did graphing the sinc function. Various authors use different approaches here. Some try a three-dimensional perspective, with f along one axis, the real part of $H(f)$ along a second, and the imaginary part along the third. A graph of this type for the transform we have here is shown in Figure 51-a. In general, however, such three-dimensional plots are not widely used, largely because of the difficulties many people have visualizing them. Instead, most authors choose simply to display two different two-dimensional graphs. Even here, however, there is not complete uniformity. Certain texts display $\Re[H(f)]$ versus f on one plot, and $\Im[H(f)]$ versus f on the other, as shown in Figure 51-b. However, most authors do not favor this approach either, but rather plot the *magnitude* of $H(f)$

$$|H(f)| = \sqrt{\Re[H(f)]^2 + \Im[H(f)]^2} \equiv \sqrt{H(f)H^*(f)}$$

versus f in one plot and the *argument* of $H(f)$

$$\Theta\left(H(f)\right) = \arctan\left[\frac{\Im[H(f)]}{\Re[H(f)]}\right] \quad ,$$

versus f in the other. These plots are commonly referred to as the *Amplitude Spectrum* and the *Phase Spectrum*, respectively. In our view, this approach, which is very similar to what we did with Fourier series, produces the most physically illuminating description, since it fairly directly addresses the first question that usually arises during the analysis of any signal - how is the energy in that signal apportioned between the different component frequencies? (In saying this, we are actually using the fact that there is a direct parallel in Fourier transforms to Parseval's theorem in basic Fourier series. Parseval's theorem, as discussed in Chapter 2, states that the total energy in a periodic signal is proportional to the sum of the squares of the amplitudes at each frequency.) In this example, the magnitude

188

and argument of $H(f)$ are, respectively

$$|H(f)| = \sqrt{H(f)H^*(f)} = \frac{1}{\sqrt{1 + (2\pi f)^2}}$$

and

$$\Theta\left(H(f)\right) = \arctan\left[\frac{-2\pi f}{1}\right] = -\arctan\left[2\pi f\right] \quad .$$

The graphs of these two quantities are plotted in Figure 51-c. Again, as expected, we see that the larger amplitudes occur at the lower frequencies. Furthermore, as in the square pulse, $H(f) = \mathbf{O}(1/f)$, a result which again seems reasonable in light of our Fourier series experience, since $h(t)$ is discontinuous at $t = 0$.
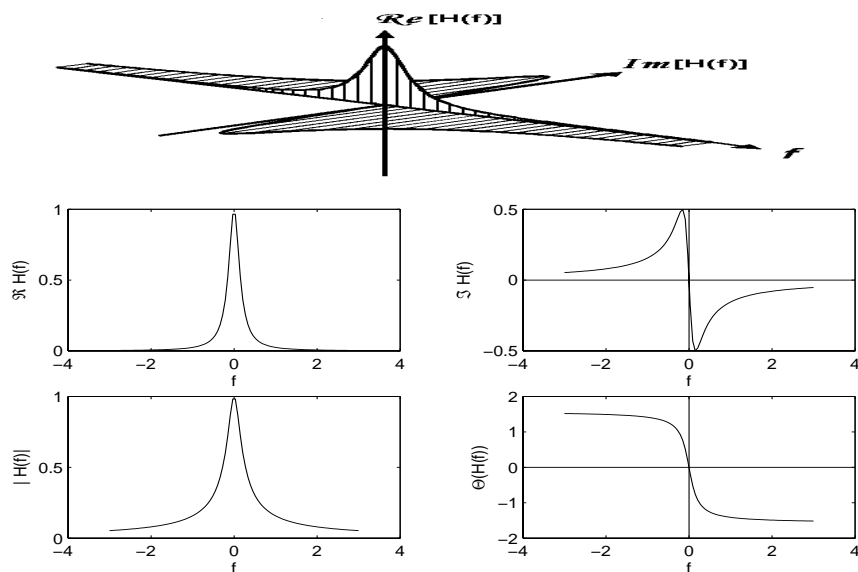

Figure 51: Alternative Graphical Descriptions of the Fourier Transform

These two examples have introduced some of the basic considerations involved in interpreting Fourier transforms. We shall continue to use these ideas as we now turn to study the mechanics of computing these transforms in some more detail.

5.5   Special Computational Cases of the Fourier Transform

Computing Fourier transforms involves, as did computing Fourier series, evaluating integrals - an often tedious and time-consuming process. Tables help, but only if the function you are looking for is in the tables. In some cases, including developing the tables in the first place, one has no choice but to compute the transform by direct integration. Fortunately

however, with Fourier transforms, as with the Fourier series, there are certain special cases when we can exploit particular properties of $h(t)$ in order to simplify, or even totally avoid, having to fully compute of these integrals. In this section we consider the most important of these cases - when the time domain function is either even or odd, or when it is zero for $t < 0$.

In the case of an even or odd time-domain function, simplification is possible because of Euler's identity, and the fact that the integral of a sum is the sum of the integrals. These allow us, respectively, to replace the complex exponential in (5.3.13) with trigonometric terms, and then simplify the result as

$$H(f) = \int_{-\infty}^{\infty} h(t) \cos(2\pi ft)dt - j \int_{-\infty}^{\infty} h(t) \sin(2\pi ft)dt \quad .$$

(Note that this formula implies, among other things, that for purely real-valued time domain functions $(h(t))$

$$\Re[H(f)] = \int_{-\infty}^{\infty} h(t) \cos(2\pi ft)dt \qquad \text{and} \qquad \Im[H(f)] = -\int_{-\infty}^{\infty} h(t) \sin(2\pi ft)dt \quad .$$

These formulas, unfortunately, are not correct if $h(t)$ is complex.) But, since $\cos(2\pi ft)$ is even and $\sin(2\pi ft)$ is odd, then when $h(t)$ is itself real and either even or odd, we can simplify these integrals, just as we did for the a_n and b_n in the Fourier series. Specifically, by essentially repeating the derivation in Chapter 2, we can show here that

If h(t) is real and even, then

$$H(f) = 2 \int_{0}^{\infty} h(t) \cos(2\pi ft)dt \qquad (5.5.17)$$

and furthermore $H(f)$ is real and even, and

If h(t) is real and odd, then

$$H(f) = -2j \int_{0}^{\infty} h(t) \sin(2\pi ft)dt \qquad (5.5.18)$$

and moreover $H(f)$ is purely imaginary, with an odd imaginary part.

Thus, for example, since the function

$$h(t) = e^{-|t|}$$

is clearly even (Figure 52), then according to (5.5.17) we could equally well compute its Fourier transform as

$$H(f) = 2 \int_{0}^{\infty} e^{-t} \cos(2\pi ft)dt = \frac{2}{1 + (2\pi f)^2}$$

(Note however, that in contrast with our two earlier examples, $h(t)$ here is continuous, but $h'(t)$ is not, and therefore, apparently as a consequence of this, $H(f) = \mathbf{O}(1/f^2)$.)
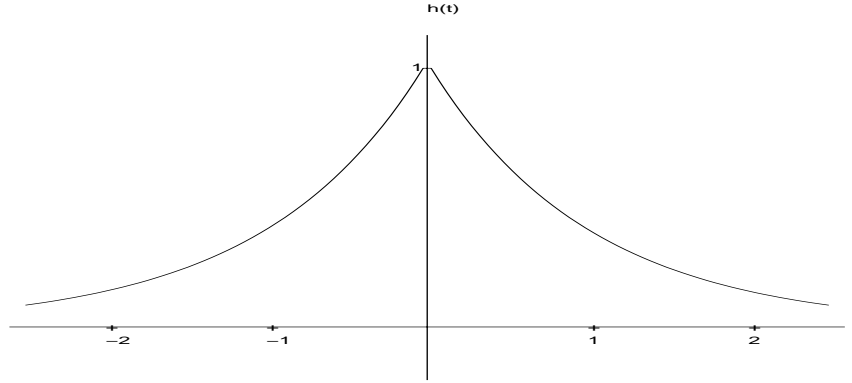
190

Figure 52: The Function $e^{-|t|}$

We would add that these results about even and odd functions can actually also be used to generate what are frequently called, respectively, Fourier cosine and sine transforms. (These transforms are analogous to Fourier cosine and sine series for periodic functions.) We shall not, however, discuss the properties of the sine and cosine transforms any further, but be content with the knowledge that, in appropriate cases, our results may produce faster, simpler calculation of the Fourier transform. These results also imply that the fact that the sinc function was real and even was a predictable consequence of its being the transform of a real and even time domain function - the square pulse.)

A second simplification of the Fourier transform occurs when $h(t)$ is a so-called **causal** function, that is when $h(t) = 0$ for $t < 0$. (The terminology causal arises because most physical applications assume that whatever process or system is being studied "starts" at $t = 0$ (with perhaps some initial conditions given). Therefore, any such system which would exhibit a nonzero response or behavior at negative values of t would, in essence, be anticipating stimuli which do not occur until $t = 0$, i.e. acting without a valid cause.) But now observe that for such a causal function, the Fourier transform (5.3.13) reduces to

$$H(f) = \int_0^\infty h(t)e^{-2\pi jft}\ dt\ \ .$$

If we now compare this formula to the Laplace transform of the *same* $h(t)$,

$$\mathcal{L}[h(t)] = \int_0^\infty h(t)e^{-st}dt\ \ ,$$

we see that the two are identical, except that the Fourier transform replaces the Laplace transform variable (s) by the quantity $2\pi jf$, i.e. for causal functions

$$H(f) = \mathcal{L}[h(t)]\Big|_{s=2\pi jf}\ \ . \tag{5.5.19}$$

Thus, if $h(t)$ were a causal function whose Laplace transform were known, e.g. from a table of transforms, we could write its Fourier transform by simply replacing the s in the Laplace table with $2\pi jf$. This approach would have worked in one of our earlier examples, since,

according to standard LaPlace tables

$$\mathcal{L}[e^{-t}] = \frac{1}{s+1}$$

and therefore the Fourier transform of the function shown in Figure 50 can be found by simply letting $s = 2\pi jf$, i.e.

$$H(f) = \frac{1}{2\pi jf + 1} \equiv \frac{1}{1 + 2\pi jf} \quad .$$

(Actually, one needs to be quite careful in applying this result! Specifically, it may produce erroneous results unless the function $h(t)$ is not only causal, but also has a Fourier transform in the "usual" sense, i.e when $h(t)$ satisfies the conditions of Theorem 5.1, e.g. by having finite energy. Attempting to apply this method to functions which do not satisfy one of these conditions may produce incorrect Fourier transforms. One such causal function with infinite energy, for which the above procedure does not compute the correct transform, is

$$h(t) = \left\{ \begin{array}{lll} 1 & , & 0 \leq t \\ 0 & , & otherwise \end{array} \right.$$

We shall show how to correctly compute the transform of this function directly, by a different approach, in a later section.)

There are more "tricks" which exploit other special properties of $h(t)$ in order to simplify computing Fourier transforms, but the above are the principal ones. As with the special formulas for calculating the Fourier series coefficients of even and odd functions, however, the results we've presented in this section do not really provide any new information which could not be obtained by computing the integrals from the standard Fourier transform formula. They only generate what are usually computationally more efficient ways to determine that information in certain special cases.

5.6   Relations Between the Transform and Inverse Transform

Thus far we have defined the Fourier and inverse Fourier integral transforms and computed the transforms of some example functions. By now, you hopefully also appreciate that computing transforms directly from the definition, i.e. by actually doing the integration in (5.3.13), is generally **not** the preferred approach! (After all, who computes derivatives directly from their limit definition?) In practice, most Fourier transforms are computed just as are most Laplace transforms - by using tables, formulas and rules that exploit, as much as possible, already computed transforms. In this section, we shall determine certain relationships between the transform and its inverse that can be used to simplify such Fourier transform computations.

These relationships depend in part on adopting a view which we have perhaps already alluded to, although not strongly. The central idea of this view is that transforms are, by their very nature, *functions themselves*. The difference is that for a transform, the inputs and outputs are not numbers, but other functions. Thus, in light of our discussion in the
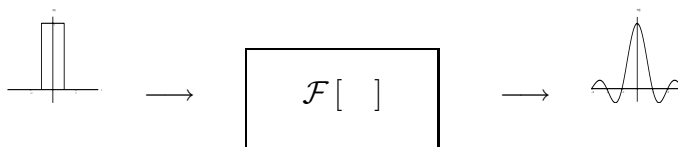
first chapter, we could also view the Fourier transform as a "black box," which inputs time-domain functions and outputs frequency domain ones, i.e.

$$h(t) \quad \longrightarrow \quad \boxed{\mathcal{F}[\ ]} \quad \longrightarrow \quad H(f)$$

But functions can be equally well specified either analytically (i.e. by a formula) or graphically! Therefore, we may equally well think of the Fourier transform as a black box whose inputs and outputs are graphs. For example, we could view the relationship between the square pulse and its Fourier transform as either

$$\left\{ \begin{array}{ll} 1 & ,\ |t| < 1/2 \\ 0 & ,\ \text{otherwise} \end{array} \right. \quad \longrightarrow \quad \boxed{\mathcal{F}[\ ]} \quad \longrightarrow \quad \frac{\sin(\pi f)}{\pi f}$$

or

 $\longrightarrow$ $\boxed{\mathcal{F}[\ ]}$ $\longrightarrow$

Stated another way, the transform pair is just as completely defined by displaying the respective graphs of the time (input) and frequency (output) domain functions as by giving their formulas. (Of course, based on our earlier discussion, we might prefer to use multiple graphs if either the input or output (or both) were complex.)

With this view in mind, we shall now perform a little formal algebra. According to the transform formulas, any time-domain signal $h(t)$ can be recovered from its frequency-domain transform $H(f)$ according to the formula

$$h(t) = \int_{-\infty}^{\infty} H(f)e^{2\pi jft}df$$

But if, for the moment, we treat the variables t and f as simply symbols, without physical interpretation, we can algebraically manipulate them. Specifically, if we replace t by $(-f)$, and f by t, on both sides, we obtain the identity

$$h(-f) = \int_{-\infty}^{\infty} H(t)e^{-2\pi jft}dt \tag{5.6.20}$$

But look carefully at the right-hand side of this identity and compare it to (5.3.13)! This integral is really nothing more than, by definition, the Fourier transform of the time-domain function whose graph is $H(t)$, i.e. the function whose graph has the same shape as a function of t that $H(f)$ has as a function of f. (For example, in terms of the time-domain square pulse and its transform, we would have

$$H(t) = \frac{\sin(\pi t)}{\pi t} \quad .)$$

Interpreting the left-hand side of (5.6.20) is only slightly more difficult, and this difficulty is due solely to the presence of the minus sign. A little elementary curve sketching should convince you that, for any f, the value on the curve $h(-f)$ is the same as the value on the curve for $h(t)$ at $t = -f$. In other words, the curve for $h(-f)$ as a function of f is the *mirror image* (with respect to the vertical axis) of the curve for $h(t)$ - i.e. we simply take the curve for $h(t)$, swap the positive and negative ends of the horizontal axis, and relabel the horizontal axis as f (Figure 53). This rather intriguing result can be summarized with

Figure 53: The Relationship of $h(t)$ and $h(-f)$

the formula

$$\mathcal{F}[H(t)] = h(-f) \tag{5.6.21}$$

or, in terms of our black boxes,

$$H(t) \quad \longrightarrow \quad \boxed{\mathcal{F}[\ \ ]} \quad \longrightarrow \quad h(-f)$$

This result further implies, among other things, that every Fourier transform actually generates *two* transform pairs, or two entries in any table of transforms. For example, the square pulse generates both

$$\begin{cases} 1 & , \ |t| < 1/2 \\ 0 & , \ otherwise \end{cases} \qquad \longleftrightarrow \qquad \frac{\sin(\pi f)}{\pi f}$$
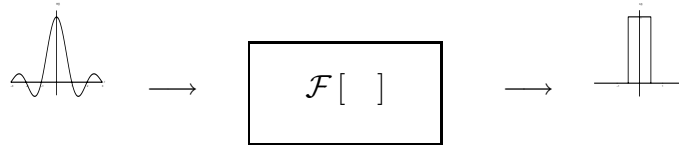
and

$$\frac{\sin(\pi t)}{\pi t} \qquad \longleftrightarrow \qquad \begin{cases} 1 & , \ |f| < 1/2 \\ 0 & , \ \text{otherwise} \end{cases}$$

(Of course, since $h(t)$ here is an even function, swapping the positive and negative axes to produce $h(-f)$ produces no perceptible effect.) In terms of our black box model, we could equivalently portray the result of (5.6.21) for these specific functions as

and

194

(We would add that this result has no counterpart in Laplace transform theory.)

There is one further interesting result which may be arrived at by manipulating the Fourier inverse formula

$$h(t) = \mathcal{F}^{-1}\left[H(f)\right] = \int_{-\infty}^{\infty} H(f)e^{2\pi jft}df$$

If we again algebraically interchange the variables, this time replacing t by f, and f by $(-t)$, we find

$$h(f) = \int_{\infty}^{-\infty} H(-t)e^{-2\pi jft}(-dt) \quad ,$$

or, equivalently,

$$h(f) = \int_{-\infty}^{\infty} H(-t)e^{-2\pi jft}dt \quad . \tag{5.6.22}$$

If we now compare this integral with the definition of the Fourier transform (5.3.13), we see that we have here nothing more than the calculation of the Fourier transform of $H(-t)$, i.e. the function whose graph (as a function of t) is the same as the mirror image of the graph of $H(f)$ as a function of f. While at first this may seem to be not all that surprising, it nevertheless has a fundamental practical impact. This impact is very simply that, because of (5.6.22), one needs only to design a piece of computer hardware and software that can calculate the Fourier transform. The identical device can then be used to compute the inverse transform, provided only that one must enter the data in reverse order. The potential economic savings of such an observation are immense.

5.7 General Properties of the Fourier Transform - Linearity, Shifting and Scaling

As we've already alluded to, one seldom computes transforms directly from the definition, i.e. by actually integrating (5.3.13). Generally, the preferred approach is to use tables of already computed transforms (some of which may have been computed using rules such as those discussed in the previous section). However, creating and using tables involves certain tradeoffs - the more complete the table, the longer and more costly it becomes, and the harder it often becomes to find any particular formula. Therefore, any rules or other properties which can cut down on the length of tables, without compromising their effectiveness are potentially valuable. In this section, we shall develop some common rules that can be used to simplify the computation of Fourier transforms, given information about the transforms of related functions.

The first of these rules is a direct and straightforward consequence of the facts that the Fourier transform is defined in terms of an integral and that the integral as an operation has certain properties. Specifically, given two (different) time domain functions, say $h(t)$

and $g(t)$, and constants a and b, then, according to the definition of the transform (5.3.13) and the fact that the integral of a sum is the sum of the integrals

$$\mathcal{F}\left[ah(t) + bg(t)\right] \quad = \int_{-\infty}^{\infty}[ah(t) + bg(t)]e^{-2\pi jft}dt$$

$$= a\int_{-\infty}^{\infty} h(t)e^{-2\pi jft}dt + b\int_{-\infty}^{\infty} g(t)e^{-2\pi jft}dt$$

$$= a\mathcal{F}\left[h(t)\right] + b\mathcal{F}\left[g(t)\right]$$

or, equivalently

$$\mathcal{F}\left[ah(t) + bg(t)\right] = aH(f) + bG(f) \tag{5.7.23}$$

Stated slightly differently, the Fourier transform is a *linear* operation (as most of the common operations of the calculus, such as the derivative and the integral).

The second rule we shall consider involves functions of the form $h(at)$, where $h(t)$ is a function whose transform, denoted $H(f)$, we already know. (In the following development, we shall assume that $a > 0$. We invite the reader to ponder what will change when $a < 0$.) However, in order to be able to better interpret our result, we shall first review, in graphical terms, the meaning of the expression $h(at)$. Specifically, in terms of our input/output model of functions, $h(at)$ is the function whose output value, given an input of say t_0, is the value of the original function $h(t)$, but for an input of (at_0), not t_0. Thus, for an input value of 1, the function $h(at)$ produces the output value of $h(a)$. This interpretation is depicted in Figure 54, for both the cases $a < 1$ and $a > 1$. (Of course, if $a = 1$, the graphs of $h(t)$ and $h(at)$ are identical.) Note that if $a < 1$, the graph of $h(at)$ is a spread out version of the graph of $h(t)$, while if $a > 1$, it is compressed, i.e. replacing t by at to create $h(at)$ effectively *rescales* the horizontal (t) axis. If we think of t as a physical time variable, then we can interpret $a < 1$ as effectively "slowing down" time, while $a > 1$ "speeds it up." Viewed another way, an event that occurs at $t = 1$ in the function $h(t)$ occurs before $t = 1$ in the function $h(at)$ if $a > 1$, and after $t = 1$ if $a < 1$. A realistic physical model for this situation is a variable speed tape recorder, where $a > 1$ corresponds to playing the tape faster than the speed it was recorded at, and $a < 1$ to playing it slower.
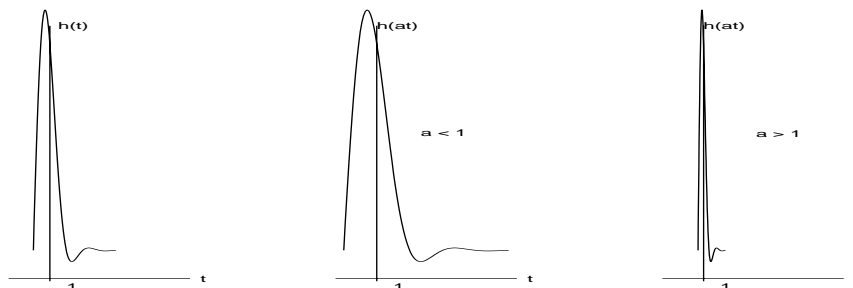

Figure 54: The Relationship of $h(t)$ and $h(at)$

With this model in mind, we now consider the algebra of computing the Fourier transform

of $h(at)$. We start at the beginning, with the definition

$$\mathcal{F}[h(at)] = \int_{-\infty}^{\infty} h(at)e^{-2\pi jft}dt$$

If we now change variables inside the integral, replacing t by (t/a), we have

$$\mathcal{F}[h(at)] = \int_{-\infty}^{\infty} h(t)e^{-2\pi jf(t/a)}d(t/a)$$

$$= \frac{1}{a}\int_{-\infty}^{\infty} h(t)e^{-2\pi j(f/a)t}dt$$

(Note that this procedure normally also changes the limits on the integral, but because they are infinite in this case and a is positive, that change will be unnoticeable here.) But a careful look at the last integral here shows that it is, by definition (5.3.13), nothing more than the Fourier transform of the original, unscaled function $h(t)$, except that this integral is evaluated at f/a, not f. Equivalently, we can now state

$$\mathcal{F}[h(at)] = \frac{1}{a}H\left(\frac{f}{a}\right) \tag{5.7.24}$$

Now note the effect of the right hand side of this equality. It is nothing more than a scaled version of $H(f)$, but with two important distinctions relative to our interpretation of $h(at)$ as a scaled version of $h(t)$. First of all, the vertical scale of $H(f/a)$ is different than that of $H(f)$, by the factor $1/a$. Secondly, the horizontal axis is also scaled by the reciprocal of a. In other words, if $a > 1$, then in the frequency domain the maximum amplitudes are reduced, and the amplitude spectrum is spread out. Conversely, if $a < 1$, the maximum amplitudes are both increased and pushed toward the lower frequencies. This behavior is shown in Figure 55. This figure clearly supports common sense, which tells us that when $a > 1$, the effect of time scaling on the Fourier transform is to push the energy of $h(t)$ up into the higher frequencies. In other words, if we try to play our tape recorder in "fast forward," we will hear the "Mickey Mouse" sound. The opposite, of course, happens when $a < 1$. (Don't forget however that (5.7.24) was derived assuming that a is positive! You should consider how this result would change if a were negative instead.)

Example: We already have shown that for $h(t) = e^{-|t|}$

$$H(f) = \mathcal{F}\left[e^{-|t|}\right] = \frac{2}{1 + (2\pi f)^2}$$

Therefore, according to (5.7.24)

$$\mathcal{F}\left[e^{-2|t|}\right] = \tfrac{1}{2}H\left(\tfrac{f}{2}\right) = \frac{1}{2}\left\{\frac{2}{1 + (2\pi(f/2))^2}\right\}$$

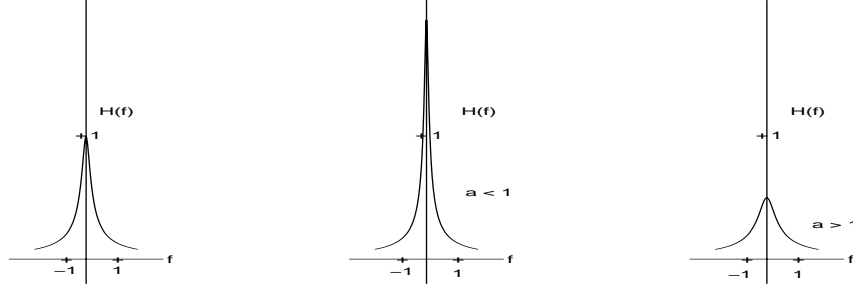$$= \frac{1}{1 + (\pi f)^2}$$

Figure 55: The Relationship of $H(f)$ and $\dfrac{1}{a}H\left(\dfrac{f}{a}\right)$

A third rule for simplifying the computation of Fourier transforms involves functions of the form $h(t - b)$, which are nothing more than identical copies of the function $h(t)$, but shifted to the right (assuming b is positive) by the amount b (Figure 56). Physically, such shifts generate exactly the same signal, only later in time. To determine the effect of this time shift on the Fourier transform, we start again with the definition (5.3.13), but then change the variables by replacing t by $(t + b)$ inside the integral. This leads to

$$\mathcal{F}\left[h(t - b)\right] = \int_{-\infty}^{\infty} h(t - b)e^{-2\pi jft}dt$$

$$= \int_{-\infty}^{\infty} h((t + b) - b)e^{-2\pi jf(t+b)}d(t + b)$$

$$= e^{-2\pi jfb}\int_{-\infty}^{\infty} h(t)e^{-2\pi jft}dt \quad .$$

But, of course, by definition the last integral in the above formula is nothing more than the transform of the unshifted function. Therefore we can summarize this rule as

$$\mathcal{F}\left[h(t - b)\right] = e^{-2\pi jfb}H(f) \tag{5.7.25}$$

As was the case with our rule for the transform of a scaled function, this rule also has a clear interpretation in terms of the amplitude and phase spectra in the transform (frequency) domain. Specifically

$$\left|e^{-2\pi jfb}H(f)\right| = \left|e^{-2\pi jfb}\right||H(f)| = |H(f)| \quad ,$$

since $\left|e^{-2\pi jfb}\right| = 1$, and

$$\Theta\left(e^{-2\pi jfb}H(f)\right) = \Theta\left(e^{-2\pi jfb}\right) + \Theta(H(f)) = -2\pi fb + \Theta(H(f)) \quad .$$

In other words, a pure time shift introduces absolutely no change in the amplitude spectrum, but only affects the phase. Since a time shift does not really alter the signal (for example, your tape will sound the same tomorrow as it does today, assuming you play it at the same speed), this result seems perfectly reasonable.

Figure 56: The Relationship of $h(t)$ and $h(t - b)$

Example: Consider the function

$$g(t) = \left\{ \begin{array}{ll} e^{1-t} & , \quad 1 < t \\ 0 & , \quad \text{otherwise} \end{array} \right.$$

A quick sketch should convince you that this function is nothing more than the function of Figure 50, shifted to the right by one unit, i.e. $g(t) = h(t - 1)$. Therefore, taking $b = 1$,

$$\mathcal{F}[g(t)] = e^{-2\pi j f(1)} H(f) = e^{-2\pi j f} \frac{1}{1 + 2\pi j f}$$

You may also observe that we can also combine (5.7.24) and (5.7.25) into a single rule, as follows

$$\begin{aligned} \mathcal{F}[h(at - b)] \; &= \mathcal{F}[h(a(t - b/a))] = e^{-2\pi j f(b/a)} \mathcal{F}[h(at)] \\[2mm] &= \frac{e^{-2\pi j f(b/a)}}{a} H\left(\frac{f}{a}\right) \end{aligned} \tag{5.7.26}$$

The rules and properties presented thus far are by no means exhaustive. However they are sufficient to compute the transforms of many interesting and commonly occurring functions. Equally importantly, they may be used to begin to generate *tables* of Fourier transforms, in much the same way that tables of Laplace transforms are produced (and then later to extend such tables). Transform tables can significantly streamline and speed up transform computations, especially when used in conjunction with linearity, shifting, etc.

5.8 The Fourier Transform of Derivatives and Integrals

The previous section introduced rules which we could use to simplify the computation of the Fourier transform of shifted or rescaled versions of a given time-domain function. In this section we shall develop some further rules that apply to a somewhat different situation - when the function we wish to transform is the derivative or integral of a function whose transform we know.

You should recall that a very specific relationship exists between the Laplace transforms of a function and its derivative, specifically

$$\mathcal{L}[h'(t)] = s\mathcal{L}[h(t)] - h(0) \quad .$$

199

(Furthermore, this relationship allows us to convert linear, constant coefficient ordinary differential equations into algebraic ones which include all the effects of the initial conditions. Unfortunately, as you should also recall, that the resulting algebraic equation is in terms of the **transform** variable, not the original time-domain one. Therefore, we must still invert the solutions to this algebraic equation in order to recover the solution to the differential equation.) Given our earlier comments on the similarities between the Fourier and Laplace transforms, we should not be surprised then if a similar, but perhaps not identical relation occurs with the Fourier transform.

In the following discussion, as before, we shall denote the particular time-domain function whose transform we know by $h(t)$, and its transform by $H(f)$. In addition, we assume that the Fourier transform of $h'(t)$ also exists and furthermore that $h(t) \to 0$ as $t \to \pm\infty$. (This last restriction is actually less of an assumption that an almost virtually guaranteed consequence if $h(t)$ is any function which satisfies Theorem 5.1.) With these assumptions, we start with the definition of the transform and integrate by parts one time, to yield

$$\mathcal{F}\left[h'(t)\right] = \int_{-\infty}^{\infty} h'(t)e^{-2\pi jft}dt$$

$$= h(t)e^{-2\pi jft}\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} h(t)\left[-(2\pi jf)e^{-2\pi jft}\right]dt$$

$$= h(t)e^{-2\pi jft}\Big|_{-\infty}^{\infty} + 2\pi jf \int_{-\infty}^{\infty} h(t)e^{-2\pi jft}dt$$

But since, by assumption, $h(t)$ vanishes at $\pm\infty$, the first term in the last equation drops out. Moreover, the integral in the second term is, by definition, nothing more than the Fourier transform of the original, undifferentiated time-domain function. Therefore, we have shown

$$\mathcal{F}\left[h'(t)\right] = 2\pi jf H(f) \tag{5.8.27}$$

Note that, hopefully not unexpectedly, this is very similar to the result for the Laplace transform of the derivative. The major difference, outside of the replacement of the Laplace variable (s) by the Fourier term $(2\pi jf)$, is that no initial value data appears in the Fourier transform of the derivative. On reflection however, this lack of initial data in the Fourier transform should not seem all that surprising. After all, the Laplace transform only integrates values for $t > 0$. Therefore, in line with the basic theory of ordinary differential equations, some initial data would seem to be mandatory when transforming the derivative. On the contrary, the Fourier transform should not need initial data, since its interval of integration assumes the values of the function are known "for all time."

Later we shall see how (5.8.27) can be used in the solution of differential equations. For the moment however, we consider it as just another rule that may be invoked, if required, to simplify the computations. Thus consider the following

Example: While we could compute the Fourier transform of the function

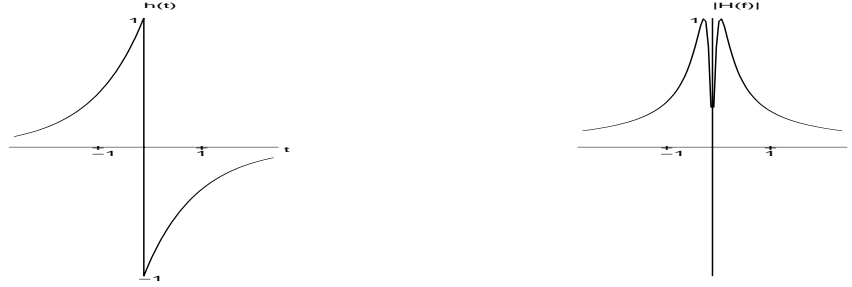$$g(t) = \begin{cases} -e^{-t} & , & 0 \le t < \infty \\ e^t & , & -\infty < t < 0 \end{cases}$$

200

Figure 57: A Fourier Transform Computed Using the Derivative Rule

(Figure 57) directly (and simplify that computation by using the fact that $g(t)$ is an odd function), note that

$$g(t) = \frac{d}{dt}\left[e^{-|t|}\right] \quad .$$

Since we have already computed the Fourier transform of $e^{-|t|}$, we can thus use that earlier computation, plus this last rule to state

$$\mathcal{F}\left[g(t)\right] = \mathcal{F}\left[\frac{d}{dt}\left[e^{-|t|}\right]\right] \quad = 2\pi j f \mathcal{F}\left[e^{-|t|}\right]$$

$$= 2\pi j f \left\{\frac{2}{1+(2\pi f)^2}\right\} = \frac{4\pi j f}{1+(2\pi f)^2}$$

(Again, as we already observed, the time-domain function $g(t)$ here was odd. Therefore, as we should have expected, based on our earlier discussion, this transform is purely imaginary and the imaginary part itself is an odd function. Furthermore, the transform is $\mathbf{O}(1/f)$ as f→ ±∞, apparently as a consequence of the jump in $g(t)$ at $t = 0$.)

Finally, you should note that we can apply (5.8.27) repeatedly, as we commonly do with the corresponding rule for Laplace transforms, and show that

$$\mathcal{F}\left[\frac{d^n h}{dt^n}\right] = (2\pi j f)^n H(f) \tag{5.8.28}$$

Derivatives also figure prominently in the next Fourier transform rule we consider, but from a different perspective. Specifically, the transform $H(f)$ is itself a function of f. Therefore, it seems reasonable to see if there is any valuable information in the derivative of $H(f)$ with respect to f. To determine this, we again start with the definition of the transform,

then apply the usual Liebnitz rule for differentiating under an integral sign

$$\frac{d}{df}[H(f)] \equiv \frac{d}{df}\left[\int_{-\infty}^{\infty} h(t)e^{-2\pi jft}dt\right]$$

$$= \int_{-\infty}^{\infty}\frac{\partial}{\partial f}\left[h(t)e^{-2\pi jft}\right]dt$$

$$= \int_{-\infty}^{\infty}h(t)\left[-2\pi jte^{-2\pi jft}\right]dt$$

$$= -2\pi j\int_{-\infty}^{\infty}[th(t)]\,e^{-2\pi jft}dt$$

But this last integral is simply by definition the Fourier transform of the quantity $[th(t)]$. Therefore, after dividing both sides by the $(-2\pi j)$ factor, we have

$$\mathcal{F}[th(t)] = -\frac{1}{2\pi j}\frac{d}{df}[H(f)] \tag{5.8.29}$$

or, when applied repeatedly

$$\mathcal{F}[t^n h(t)] = \frac{(-1)^n}{(2\pi j)^n}\frac{d^n}{df^n}[H(f)] \tag{5.8.30}$$

(where, of course, for this to be valid, $t^n h(t)$ must also satisfy the conditions of Theorem 5.1). The use of this rule is demonstrated by the following

Example: Consider

$$g(t) = te^{-|t|}, \quad -\infty < t < \infty$$

Now, according to our last result, since we already have computed the Fourier transform of $e^{-|t|}$,

$$G(f) = \mathcal{F}\left[te^{-|t|}\right] = -\frac{1}{2\pi j}\frac{d}{df}\left[\mathcal{F}\left[e^{-|t|}\right]\right]$$

$$= -\frac{1}{2\pi j}\frac{d}{df}\left[\frac{2}{1+(2\pi f)^2}\right] = \frac{-8\pi jf}{(1+(2\pi f)^2)^2}$$

(Again note how the oddness of $g(t)$ evidences itself in $G(f)$ being purely imaginary with an odd imaginary part, and the fact that $H(f) = \mathbf{O}(1/f^3)$ relates to the fact that $h(t)$ and $h'(t)$ here are both continuous.)

The last rule we shall consider in this section is basically the reverse of the derivative rule. Specifically, from the fundamental theorem of the integral calculus,

$$h(t) = \frac{d}{dt}\left[\int_{-\infty}^{t}h(s)ds\right]$$

and therefore, by (5.8.27)

$$H(f) = \mathcal{F}[h(t)] = 2\pi jf\,\mathcal{F}\left[\int_{-\infty}^{t}h(s)ds\right]$$

202

or equivalently

$$\mathcal{F}\left[\int_{-\infty}^{t} h(s)ds\right] = \frac{H(f)}{2\pi j f} \quad, \tag{5.8.31}$$

provided that $\int_{-\infty}^{\infty} h(t)dt = 0$ (This last restriction is necessary in order to satisfy our earlier requirement that any function whose derivative we are transforming must vanish at $\pm\infty$.)

This section completes the first part of our introduction to definition, interpretation and computation of the Fourier transform. We could perhaps continue and derive a few other rules, but the ones we have covered so far represent the principal ones needed to compute the transform of "normal" functions, i.e. those that satisfy one of the two tests of Theorem 5.1. Unfortunately, many real-world applications involve signals that are most easily modeled by functions that do not satisfy this theorem. (The periodic functions are one obvious example.) Therefore, rather than continue to try to find more rules for normal functions, we shall turn instead in the next section to extending the definition and computation of the Fourier transform to a broader class of physically interesting functions.

5.9   The Fourier Transform of the Impulse Function and Its Implications

We have already briefly mentioned the impulse or delta function, which we assume you have already been introduced to in Laplace transforms. This function (which is really not a function in the sense of the usual calculus) is normally symbolized by $\delta(t)$ and represents an "instantaneous" force, such as occurs in the inelastic collision between billiard balls. (Of course, no truly instantaneous forces exist, just as there are no truly periodic ones. What there are, however, are forces which act over such short periods of time, relative to the time scales of the rest of a particular physical process, that they appear to be effectively instantaneous *in that process.*) However, instantaneous forces create a minor mathematical dilemma. The reason is that a force, to be instantaneous, must be zero at all times other than the instant of application. But, mathematically, a force also produces no noticeable effect unless it accomplishes physical work, and work is related to the *integral* of the magnitude of the force over its duration of application. Therefore, since in the usual calculus sense the area under a point of finite height is zero, a force of *finite* amplitude and instantaneous duration can do no physical work. Hence, to have any physical effect, an instantaneous force must be modeled as having an infinite amplitude at the instant of application. Such functions clearly cannot be treated by the normal calculus.

Physically, a model instantaneous force is completely determined once we know *when it occurs* and *the value of its integral.* (This integral, which we will colloquially refer to as the area under the instantaneous function, of course effectively determines the amount of work the force can perform.) Thus, the simplest instantaneous force to consider would seem to one acting at $t = 0$ and with unit area - this, of course, becomes what we call $\delta(t)$, or the unit impulse. Most texts usually generate $\delta(t)$ as the limit of a set of even rectangular pulses of unit area and continually smaller duration (and hence proportionally growing amplitudes). As their limit, $\delta(t)$ will thus have infinite amplitude at $t = 0$ and be zero elsewhere. Graphically, it is usually portrayed as a bold-face vertical arrow, or "spike" of unit height located

at $t = 0$ (Figure 58). In line with the above discussion, its basic properties are usually stated as

$$\delta(t) = 0 \quad, \quad t \neq 0$$

and
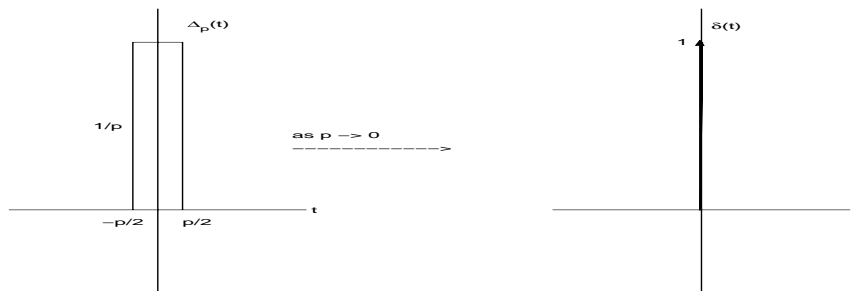
$$\int_{-\infty}^{\infty} \delta(t - t_0) \, dt = 1$$

Figure 58: The Graphical Interpretations of $\delta(t)$

Of course, not all instantaneous forces need occur at $t = 0$. Therefore, we must also be able to represent an instantaneous force of unit area and located at some $t = t_0 \neq 0$. Notationally, we do this in the "natural" way, i.e. with the symbol $\delta(t - t_0)$. In terms of this notation, we can then show that the delta function also satisfies a third critical equality, commonly called the *sifting property*. The sifting property states that, for any *continuous* function $h(t)$,

$$\int_{-\infty}^{\infty} h(t)\delta(t - t_0) \, dt = h(t_0) \quad.$$

(Note that while we mentally view $\delta(t - t_0)$ as an infinitely strong pulse of unit area located at $t = t_0$, we shall graphically portray it as simply a shifted copy of $\delta(t)$, i.e. as a bold spike of unit height at $t = t_0$.) Furthermore, since the sifting property implies that

$$\int_{-\infty}^{\infty} c\delta(t - t_0) \, dt = c \quad,$$

where c is any constant, we shall therefore graph an instantaneous force with area (work) c which occurs at $t = t_0$ as a spike of height c located at t_0. Moreover, since the area of the unit impulse does not change, we can easily adopt the view that its location coincides with the point where the overall argument (e.g. $t - t_0$) is zero, and state the following equalities

$$\begin{aligned} \delta(t_0 - t) &= \delta(t - t_0) \\ \delta(-t) &= \delta(t) \end{aligned}$$

(Alternatively, according to this last equation, we may say that Figure 58 implies that the delta function is even.) Finally, since Figure 58 seems clearly to allow us to consider the impulse as a real and non-negative valued function, we shall also accept

$$|\delta(t)| = \delta(t)$$

204

Before continuing, we must reemphasize that, because of its infinite height, the delta function is not a function in the sense of the ordinary calculus. Moreover, not all of its properties are intuitively clear. (For example, as we will show later

$$\delta(2t) = \frac{1}{2}\delta(t) \quad ,$$

a result which seems, at face value, to be inconsistent with the graphical interpretation of the time scaling of "normal" functions as portrayed in Figure 54.) Therefore, any mathematics involving the delta function must really be justified by arguments beyond those of the ordinary calculus, e.g. the sifting property cannot really be justified in terms of normal integrals. Such justifications do exist, and form the core of a rather elegant section of mathematics called distribution or generalized function theory. This theory is, however, well beyond our level. Therefore much of the following development will be formal rather than rigorous. Fortunately, we can formally derive all of the results using little more than the sifting property and the concept of the Fourier transform as a pair.

We start by considering the Fourier transform of the shifted delta function. Assuming that this transform actually exists, we can then compute it formally from the definition of the transform and the sifting property as

$$\mathcal{F}\left[\delta(t - t_0)\right] \equiv \int_{-\infty}^{\infty} \delta(t - t_0)e^{-2\pi jft}dt = e^{-2\pi jft_0}$$

Furthermore, if we then evaluate this transform at $t_0 = 0$, we have the second transform pair

$$\mathcal{F}\left[\delta(t)\right] = 1$$

(Figure 59). But according to the symmetry property of the transform, i.e. the property that $\mathcal{F}\left[H(t)\right] = h(-f)$, we immediately have still another transform pair

Figure 59: The Transform Pair for $\mathcal{F}\left[\delta(t)\right] = 1$

$$\mathcal{F}\left[1\right] = \delta(-f) = \delta(f)$$

(where we now apply our earlier view of the delta function as even.) This pair immediately leads to a fairly powerful insight on the nature of the transform. A physical model for the constant function in the time domain is a direct current (DC) battery that never runs down.

(Such a battery is, of course, unrealizable since it must store infinite energy.) But recall the Fourier transform supposedly identifies the amplitudes of the component frequencies of a time-domain signal! Therefore, it seems quite reasonable that the transform of the constant function should be a delta function located at the origin in the frequency domain. Its "real" height is infinite, because, of course, the battery apparently has infinite energy, while its area (as denoted by the height of the spike) corresponds exactly to the amplitude of the time domain signal. The location of the spike also seems appropriate, since by definition direct current does not oscillate.

The symmetry property of the transform and the fact that the unit impulse is uniquely characterized by the location of the zero of its argument also allow us to compute another new pair

$$\mathcal{F}\left[e^{-2\pi j f_0 t}\right] = \delta(-f - f_0) = \delta(f + f_0) \quad .$$

This particular pair, however, is not very illuminating in and of itself, since complex-valued time-domain signals are somewhat unrealistic. However, we can use it to compute several more relevant transforms. The first of these uses this transform, along with the identity,

$$\cos(t) = \frac{e^{jt} + e^{-jt}}{2}$$

and the linearity of the transform, as follows:

$$
\begin{aligned}
\mathcal{F}\left[\cos(2\pi f_0 t)\right] &= \mathcal{F}\left[\frac{1}{2}e^{2\pi j f_0 t} + \frac{1}{2}e^{-2\pi j f_0 t}\right] \\
&= \frac{1}{2}\mathcal{F}\left[e^{2\pi j f_0 t}\right] + \frac{1}{2}\mathcal{F}\left[e^{-2\pi j f_0 t}\right] \\
&= \frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)
\end{aligned}
$$

This result yields still another fairly valuable insight into the general nature of the transform. The cosine represents a single pure frequency (f_0) in the time domain, which, because it is periodic, also contains an infinite amount of energy. Its transform consists of exactly two frequency domain impulses, located at $(\pm f_0)$, i.e. at precisely the frequency (and its negative) of the time domain oscillation. These impulses apparently have a similar interpretation to the impulse at the origin for the direct current battery, i.e. they connote infinite energy in the time domain exists at the real frequency f_0. Furthermore, each of these impulses has an area exactly half the amplitude of the real time-domain periodic function. This again is behavior we have previously encountered, for the complex Fourier series also divides the amplitudes of a real signal evenly between the positive and negative frequencies in frequency domain.

Using essentially the same steps as we used above, we can also show that

$$\mathcal{F}\left[\sin(2\pi f_0 t)\right] = \frac{j}{2}\delta(f + f_0) - \frac{j}{2}\delta(f - f_0) ,$$
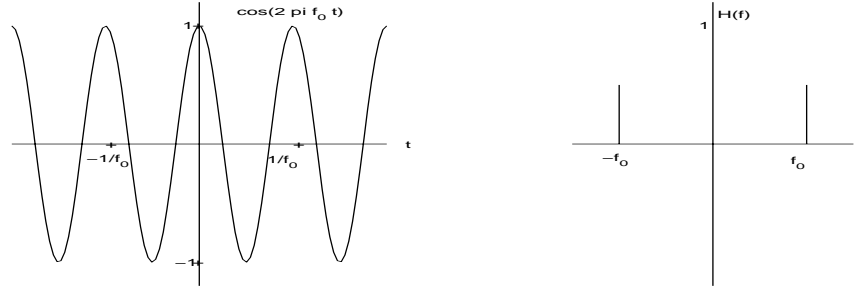
Figure 60: The Transform Pair for $\mathcal{F}\left[\cos(2\pi f_0 t)\right]$
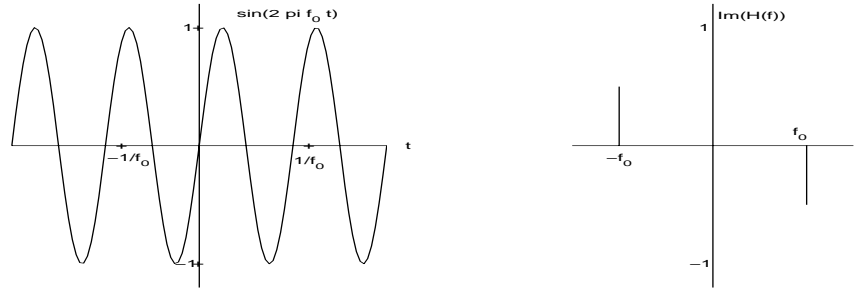

Figure 61: The Transform Pair for $\mathcal{F}\left[\sin(2\pi f_0 t)\right]$

and, except for the fact that this transform is purely imaginary (apparently a direct consequence of the oddness of the sine function), make similar conclusions that were made about the transform of the cosine (Figure 60).

The transform of the complex exponential provides at least one additional interesting insight. We have already seen that a general periodic function ($h(t)$) in the time domain can always be represented in the complex Fourier series form

$$h(t) = \sum_{n=-\infty}^{\infty} c_n e^{2jn\pi f_0 t} \quad ,$$

where, in terms of our earlier notation $f_0 = 1/(2L)$. Therefore, provided we assume that the linearity of the Fourier transform applies to the sum of convergent infinite series as well as to the sum of just two terms (a **very non-trivial** assumption), we have, formally,

$$\mathcal{F}\left[h(t)\right] = \sum_{n=-\infty}^{\infty} c_n \delta(f - n f_0) \quad .$$

In other words, the transform of a general periodic function is an infinite "train" of impulses, located at all multiples of the fundamental frequency (plus of course the constant (DC) term). Each impulse has an area equal to (i.e. is drawn graphically as a spike with an amplitude of) the value of the corresponding complex Fourier series coefficient. Thus, except for the inclusion of the "arrowheads" to signify impulses, the graph of $|H(f)|$ for this transform would look identical to the amplitude spectrum for the original Fourier series, as developed in Chapter 2.
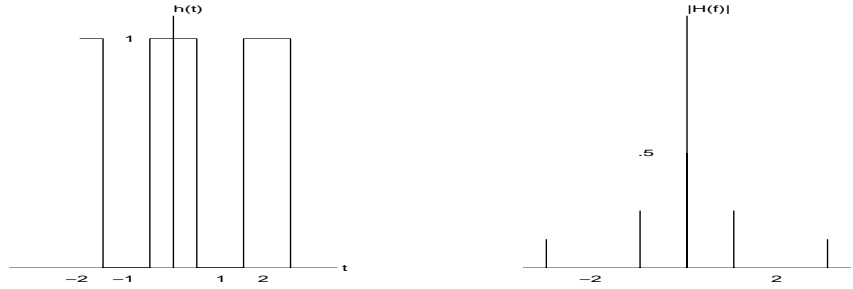
Figure 62: The Transform Pair for a Periodic Function

We close this section with one last insight about impulses. Somewhat in contrast to the ones thus far, this one highlights one important difference between "normal" functions and impulses - the effect of time scaling. Consider $\delta(2t)$. According to our earlier discussions, this is still clearly an impulse and is located at $t = 0$. But is it still a **unit** (i.e. unit area) impulse? Interesting enough, if we assume that impulses obey the basic "rules" derived thus far for Fourier transforms, it is not! Specifically, according to (5.7.24), with $a = 2$,

$$\mathcal{F}\left[\delta(2t)\right] = \frac{1}{2}\mathcal{F}\left[\delta(t)\right]\Big|_{\frac{f}{2}} = \frac{1}{2}\left[\ 1\ \right]\Big|_{\frac{f}{2}} = \frac{1}{2} = \frac{1}{2}\mathcal{F}\left[\delta(t)\right] = \mathcal{F}\left[\frac{1}{2}\delta(t)\right]$$

Thus, we must conclude that exactly one of the following is true

- $\delta(2t) = \frac{1}{2}\delta(t)$,

- Time-domain impulses and their transforms cannot be considered as inseparable pairs, or

- Impulses do not satisfy all the rules for Fourier transforms.

Since the last two choices are very unattractive, we must basically accept the first one, i.e. define

$$\delta(2t) = \frac{1}{2}\delta(t)$$

(Actually, this is not really a very painful choice to have to make. Furthermore, this is the only interpretation that is consistent with the "natural" view that $\delta(2t)$ should be the limit, as $p \rightarrow 0$, of the square pulses given by $\Delta_p(2t)$, where $\Delta_p(t)$ is defined in Figure 58.)

This completes our introduction to the relationships between impulses and Fourier transforms. Most of the insights developed here should be physically quite satisfying. Moreover, because some of these results seem to imply that one can meaningfully talk about the transforms of at least some time-domain functions which represent an infinite amount of energy, we shall next try to extend the notion of the transform to other, non-periodic, functions with infinite energy.

5.10 Further Extensions of the Fourier Transform

In previous sections we have developed the Fourier transform and its properties for both "normal" functions (i.e. those that satisfy one of the two tests of Theorem 5.1), and impulses and periodic functions. The fact that we can extend the transform to the periodic functions strongly implies that, if we are willing to accept impulses or other singularities in the transform domain, we may extend the transform even further to include other infinite energy functions. This is in fact the case, although any rigorous showing of this would require concepts for distribution theory that are well beyond the level of this text. Therefore, as with other similar properties, we shall proceed in a rather formal manner. Our basic approach however will be fairly intuitive in that we shall start by trying to approximate such functions with other functions that are "close" in some sense, but also have Fourier transforms in the normal sense.

Throughout the following discussion, we will assume that two restrictions hold true

1. $h(t)$ is bounded, i.e. $|h(t)| < M$ for some constant M and all values of t, and

2. $h(t)$ contains no periodic components

Under these restrictions, we now proceed to consider the transforms of two special cases of functions which do not satisfy Theorem 5.1. The first of these is

Case I: $$\lim_{L \to \infty} \left[\frac{1}{2L} \int_{-L}^{L} h(t)dt \right] = 0$$

In terms of electric circuits, Case I functions can be interpreted physically as having no long-term average DC level. We shall define the Fourier transform of such functions to be

$$H(f) = \lim_{\alpha \to 0^+} \left[\int_{-\infty}^{\infty} \left\{ e^{-\alpha|t|} h(t) \right\} e^{-2\pi jft} dt \right] \tag{5.10.32}$$

provided such a limit exists. (If no such limit were to exist, we would assume that the function did not have a transform.) Observe carefully what we have in effect done here. Physically, we claim it seems to be the most reasonable thing to do. First of all, since we have assumed that the function $h(t)$ is bounded, then multiplying it by the decaying exponential $e^{-\alpha|t|}$ produces an integrand which clearly satisfies the conditions of Theorem 5.1, for all $\alpha > 0$. Thus the transform of $h(t)e^{-\alpha|t|}$ should be a perfectly "well-behaved" function for all $\alpha > 0$, and therefore we should be able to formally take the limit as $\alpha \to 0^+$ fairly easily. However, for "small" values of α (i.e. for very weak decay), $h(t)$ and $h(t)e^{-\alpha|t|}$ should be reasonably close to each other, at least for a fairly "long" period of time. Furthermore, the smaller α becomes, the closer $h(t)$ and $h(t)e^{-\alpha|t|}$ should become. Therefore, it seems very plausible that, as α decreases their transforms should also become closer, and, in the limit as $\alpha \to 0^+$, become identical. (We would again note however, that in dealing with improper integrals, which include the Fourier transform, as in dealing with infinite series, plausible results are not always correct - witness Zeno's paradox. Therefore we strongly advise the reader who

may need to extend these results beyond the examples presented here to first consult a text which contains a more complete discussion of generalized functions and their transforms.)

Computing the transform of a Case I type function is illustrated by the following

Example:

$$\text{sgn}(t) = \begin{cases} 1 & , & 0 < t < \infty \\ 0 & , & t = 0 \\ -1 & , & -\infty < t < 0 \end{cases}$$

(Figure 63). This function, which really represents nothing more than the algebraic sign

Figure 63: The Transform Pair for the Function $\text{sgn}(t)$

of t, clearly does not satisfy the conditions of Theorem 5.1, but does satisfy the conditions of Case I. Therefore, using the fact that $\text{sgn}(t)$ is obviously odd, we may compute its transform according to the above procedure as follows

$$\mathcal{F}\left[\text{sgn}(t)\right] = \lim_{\alpha \to 0^+} \mathcal{F}\left[e^{-\alpha|t|}\text{sgn}(t)\right]$$

$$= \lim_{\alpha \to 0^+} \left[\int_{-\infty}^{\infty} \left\{e^{-\alpha|t|}\text{sgn}(t)\right\} e^{-2\pi jft} dt\right]$$

$$= -2j \lim_{\alpha \to 0^+} \left[\int_{0}^{\infty} e^{-\alpha t} \sin(2\pi ft) dt\right]$$

$$= -2j \lim_{\alpha \to 0^+} \left[\frac{2\pi f}{\alpha^2 + (2\pi f)^2}\right] = -\frac{j}{\pi f}$$

Note that, hopefully not unexpectedly, this transform is singular at the origin.

Functions which do not satisfy the restriction of Case I, but do satisfy the other restrictions of this section will be called Case II type functions. Specifically, Case II functions occur when

Case II: $\qquad \lim_{L \to \infty}\left[\frac{1}{2L}\int_{-L}^{L} h(t)dt\right] = A \neq 0$ (In line with our earlier discussion,

we shall then interpret Case II functions as those which, in the electrical sense, have a long-term average DC level. Note that, based on our earlier discussion of the impulses, we now clearly expect the transforms of such functions, if they can be computed, to have an

210

impulse at the origin in the frequency domain.) We treat Case II functions by a standard mathematical ploy - reducing them to a problem we've already solved. Specifically, if $h(t)$ is any Case II function, then we can write

$$h(t) = A + g(t)$$

(Graphically, $g(t)$ is simply $h(t)$ shifted down by the amount A, i.e. it's $h(t)$ but with the DC level removed.) But as we have already discussed, the Fourier transform is a linear operation and therefore we have

$$\mathcal{F}[h(t)] = \mathcal{F}[A] + \mathcal{F}[g(t)] = A\delta(f) + \mathcal{F}[g(t)]$$

However, $g(t)$ as constructed is clearly a Case I type function, since all of the long-term DC level in $h(t)$ is contained in A. Therefore, if we don't already know the transform of $g(t)$, we can compute it by the procedure we just finished describing for Case I functions. Thus the transform of $h(t)$ is completely determined.

Example: Consider the function commonly called the unit step (Figure 64).

$$u(t) = \left\{ \begin{array}{ccc} 1 & , & 0 \le t < \infty \\ 0 & , & -\infty < t < 0 \end{array} \right.$$

It's quite easy to show that this function has a long-term average DC level of $1/2$, and
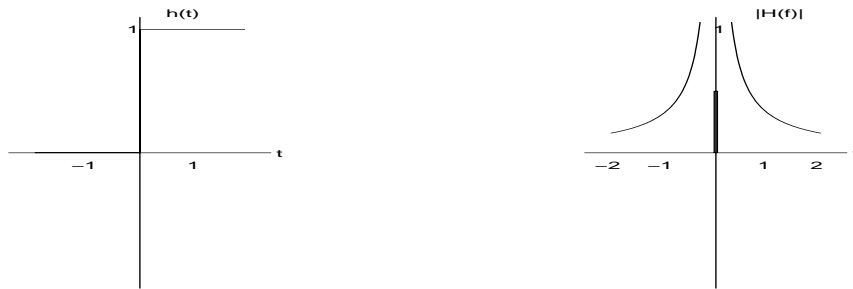

Figure 64: The Transform Pair for the Unit Step Function

therefore, based on the immediately previous discussion, we shall represent it as

$$u(t) = \frac{1}{2} + \frac{1}{2}\,\text{sgn}(t)$$

But now, also according to that discussion, and the transform example computed just before this one, we have

$$\begin{aligned} \mathcal{F}[u(t)] &= \mathcal{F}\left[\tfrac{1}{2} + \tfrac{1}{2}\,\text{sgn}(t)\right] = \tfrac{1}{2}\mathcal{F}[1] + \tfrac{1}{2}\mathcal{F}[\text{sgn}(t)] \\ &= \frac{1}{2}\delta(f) - \frac{j}{2\pi f} \end{aligned}$$

(Note the presence of the expected impulse at the origin with a height equal to the DC level in the time domain. This pattern by now should be expected. Note also that even though

the time domain function in this example is causal and Laplace transformable, we would obtain an erroneous result, missing the impulse at the origin, if we simply replaced the s in its Laplace transform by $2\pi jf$. This is why we cautioned you earlier about converting the Laplace transforms of functions that did not satisfy the conditions of Theorem 5.1.)

This last example completes our discussion of the transform of functions with infinite energy. While there are undoubtedly more rules and properties associated with computing the transform which we could study, we hope you appreciate that our interest in this transform is not primarily as a formula to be evaluated. On the contrary, our interest in this transform, as in any transform, is primarily in its utility as a **tool** which we can use to solve physical problems. Therefore, in the next chapter, we turn to investigating some of the various problems to which the Fourier transform may be applied, and consider some of the insights into physical processes that it and its dual time-domain/frequency-domain view of the world provide.

PROBLEMS

1. Compute, from the definition, and using the properties of even and odd functions where appropriate, the Fourier transform of each of the following functions. In each case, plot $h(t)$ and the Amplitude spectrum and phase angle graphs.

a. $h(t) = e^{-\alpha|t|}$, $-\infty < t < \infty$, $\alpha > 0$.

      (Plot for $\alpha = 1$ and $\alpha = .05$)

b. $h(t) = \begin{cases} 1 & , & 0 \le t \le 1 \\ 0 & , & \text{otherwise} \end{cases}$

c. $h(t) = \begin{cases} te^{-t} & , & 0 < t < \infty \\ 0 & , & \text{otherwise} \end{cases}$

d. $h(t) = \begin{cases} (1 - t^2) & , & -1 < t < 1 \\ 0 & , & \text{otherwise} \end{cases}$

e. $h(t) = \begin{cases} (1 - t)^2 & , & -1 < t < 1 \\ 0 & , & \text{otherwise} \end{cases}$

f. $h(t) = Ae^{-\alpha|t|}\cos(2\pi t)$, $-\infty < t < \infty$, $\alpha > 0$.

      (Plot for $\alpha = 1$ and $\alpha = .05$)

g. $h(t) = \begin{cases} (1 + t) & , & -1 < t < 0 \\ 1 & , & 0 \le t \le 1 \\ (2 - t) & , & 1 < t < 2 \\ 0 & , & \text{otherwise} \end{cases}$

h. $h(t) = Ate^{-\alpha|t|}$, $-\infty < t < \infty$, $\alpha > 0$

i. $h(t) = \begin{cases} t & , & -1 < t < 1 \\ 0 & , & \text{otherwise} \end{cases}$

2. Find, directly from the definition, the inverse of the following Fourier transforms, and plot $h(t)$ and the amplitude and phase graphs:

a. $H(f) = \begin{cases} (1 - f^2)^2 & , & -1 < f < 1 \\ 0 & , & \text{otherwise} \end{cases}$

b. $H(f) = |f|e^{-2|f|}$, $-\infty < f < \infty$.

213

PROBLEMS

1. Compute the Fourier transform of each of the following functions, using tables, shifting and scaling, etc. when appropriate. In each case, plot $h(t)$ and the Amplitude spectrum and phase angle graphs.

a. $h(t) = \begin{cases} 2 & , \quad -1 < t < 5 \\ 0 & , \quad \text{otherwise} \end{cases}$

b. $h(t) = \begin{cases} t & , \quad 0 \le t \le 2 \\ 4 - t & , \quad 2 < t \le 4 \\ 0 & , \quad \text{otherwise} \end{cases}$

c. $h(t) = \sin\left(\dfrac{t}{3}\right)$

d. $h(t) = \begin{cases} 2 & , \quad 3 < t < \infty \\ 0 & , \quad \text{otherwise} \end{cases}$

e. $h(t) = \dfrac{1}{4 + t^2}$

f. $h(t) = \dfrac{\sin^2(3t)}{6t^2}$

g. $h(t) = \begin{cases} e^{-t} & , \quad 0 < t < 2 \\ 0 & , \quad \text{otherwise} \end{cases}$

2. Find, using tables, shifting and/or scaling, etc., the inverse of each of the following Fourier transforms, and plot $h(t)$ and the amplitude and phase graphs:

a. $H(f) = \dfrac{1}{1 + f^2}$, $\quad -\infty < f < \infty$

b. $H(f) = e^{-3j\pi f} e^{-2|f|}$, $\quad -\infty < f < \infty$

c. $H(f) = \begin{cases} 2 & , \quad -3 < f < 3 \\ 0 & , \quad \text{otherwise} \end{cases}$

6   Applications of the Fourier Transform

6.1   Introduction

The Fourier transform has become widely used as a tool in mathematics, physics and engineering - especially in the solution of differential equations and in the design of signal processing systems. We have already encountered some of the properties that have made it so popular - its linearity, the manner in which it relates the transform of a derivative to the transform of the original function, and its ability to describe the behavior of systems on a frequency-by-frequency basis.

In this chapter we shall briefly examine some of the more commonly found applications of the Fourier transform. Our intent is neither to be all-inclusive nor to conduct in-depth engineering analyses. On the contrary, we simply want to show, in the context of some of its more classic applications, how the Fourier transform is used and how it helps to provide unique and powerful insights into the behavior of physical systems.

However, before we can do this we must develop one more property of the transform. This property involves the operation called convolution, and may be the transform's single most important one in relation to the applications we will consider. In one sense, the behavior of the Fourier transform relative to convolution may be viewed as simply one more rule, like those of the previous chapter. However, convolution is so central to applications that we chose not to present it earlier, but instead waited to consider it in this chapter. (There is actually also a second operation, called correlation, which is perhaps almost as important in other applications as is the convolution. We shall briefly note its properties in relation to Fourier transforms at the end of this chapter.)

6.2   Convolution and Fourier Transforms

We define the **convolution** of two functions, say $h(t)$ and $g(t)$ as

$$g(t) * h(t) = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)d\tau \quad . \tag{6.2.1}$$

The order in which we write $g(t)$ and $h(t)$ here is actually irrelevant, since it is fairly easily shown that

$$\int_{-\infty}^{\infty} g(t-\tau)h(\tau)d\tau = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)d\tau$$

and therefore

$$g(t) * h(t) \equiv h(t) * g(t) \quad ,$$

i.e. convolution is a commutative operation. It is, however, crucial to recognize that while the variable τ in (6.2.1) is a dummy variable, the variable t is not! Therefore the convolution of two functions is itself a function, i.e. $h(t) * g(t)$ is a *graph*, not a number!

You should have already encountered the term convolution in your study of Laplace transforms. If so, parts of the above definition should seem familiar. But (6.2.1) is **not** identical to the definition of convolution for Laplace transform functions, since the lower

limit of the Laplace convolution is different than the lower limit in (6.2.1). However, we can quite easily show that when both functions are causal (i.e. when both $h(t) = 0$ and $g(t) = 0$ for $t < 0$), then the Fourier and the Laplace forms of convolution become identical. Therefore, since the Laplace transform treats the functions it deals as effectively having no values when $t < 0$, we may then view Laplace convolution as simply a special case of the Fourier form, not a different definition.

Computing the convolution of two functions from the definition as given by (6.2.1) is reasonably straightforward, provided one first carefully considers all the aspects of that integral. Specifically, since it is expressed as an integral, the convolution involves the area under some curve. However, as noted before, t is a free variable in that integration and therefore in most cases both the particular curve being integrated **and** the limits of integration will vary with t. Moreover, since the integral is with respect to τ, not t, one must understand how the graphs of $g(\tau)$ and $h(t - \tau)$ should be interpreted as functions of τ.

The graphical interpretation of $g(\tau)$ is simple - its the same graph as $g(t)$ except that the horizontal axis is relabeled as τ. However, because of the minus sign, $h(t - \tau)$, viewed as a function of τ is **not** simply a shifted copy of $h(t)$. Rather, $h(t - \tau)$ represents $h(t)$ first shifted so that the original origin becomes located at $\tau = t$, but then with the shape of the original curve **reversed** (Figure 65). In other words, $h(t - \tau)$ will look as if time were being "run backwards." The convolution of $g(t)$ and $h(t)$ then becomes the area under the product of this backwards curve and the curve for $g(\tau)$.
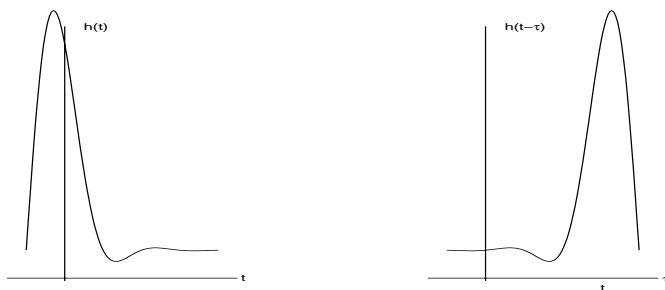

Figure 65: The Relation of $h(t)$ as a Function of t and $h(t - \tau)$ as a Function of τ

As we have already discussed, the convolution of two functions is itself a function. In terms of our just-completed graphical description, this occurs because the amount by which the curve for $h(t)$ is shifted before it is reversed depends on the variable t, and so the particular pictures used to describe the convolution change with t. Efficient computation of convolutions therefore requires having a reasonably structured procedure for generating the proper pictures. We would suggest the following sequence of steps

1. Draw $g(\tau)$ and $h(\tau)$ as a function of τ.

2. Reverse the graph for $h(\tau)$ to give $h(-\tau)$. Shift this curve some arbitrary amount and label the point on the reversed curve which corresponds to the original origin as t. (This is now the graph of $h(t - \tau)$.)

216

3. Lay the graph for $h(t - \tau)$ on top of that for $g(\tau)$, with the point labeled with t located near $\tau = -\infty$ on the graph for $g(\tau)$.

4. Slowly slide the graph for $h(t - \tau)$ to the right. At any point, the area under the *product* of the two graphs is the convolution for that value of t.

To demonstrate this procedure, consider the following

Example:
$$g(t) = \left\{ \begin{array}{cc} 1 & , \quad 0 < t < 1 \\ 0 & , \quad \text{otherwise} \end{array} \right.$$

$$h(t) = \left\{ \begin{array}{cc} e^{-t} & , \quad 0 \leq t < \infty \\ 0 & , \quad -\infty < t < 0 \end{array} \right.$$

(Figure 66) Notice that after we apply the second step, we can think of the reversed exponential as having a "leading edge" at $\tau = t$. Therefore, initially (when $t < 0$) it has no nonzero values in common with the square pulse, and so their product (and the integrand for the convolution) would be zero (Case I). Eventually, however, as the leading edge of the reversed exponential moves further to the right, it begins to intersect the pulse, causing a nonzero integrand between the origin and the point $\tau = t$ (Case II). Finally, the leading edge of the reversed exponential will move to the right of the leading edge of the pulse and there will be a nonzero integrand only between the origin and $\tau = 1$ (Case III). Computationally, we can express these three cases as

Case I : $\quad -\infty < t \leq 0$,
$$g(t) * h(t) = 0$$
Case II : $\quad 0 < t < 1$,
$$g(t) * h(t) = \int_0^t 1 \cdot e^{-(t-\tau)} \, d\tau = e^{-(t-\tau)} \Big|_0^t = 1 - e^{-t}$$

Case III : $\quad 1 \leq t < \infty$,
$$g(t) * h(t) = \int_0^1 1 \cdot e^{-(t-\tau)} \, d\tau = e^{-(t-\tau)} \Big|_0^1 = e^{-(t-1)} - e^{-t}$$

Lastly, we can combine these three cases into a single graph (Figure 67) which represents the convolution of these two functions.

Now, with hopefully a firm(er) understanding of the mechanics of computing convolutions, we consider how the Fourier transform of a convolution relates to the Fourier transforms of the two individual functions which produced that convolution. We start by assuming that $g(t)$ and $h(t)$ both have transforms, which we denote respectively as $G(f)$ and $H(f)$. Then, as we have in most previous discussions, we write the formal definition for the transform of their convolution, i.e.

$$\mathcal{F}\left[g(t) * h(t)\right] = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\tau)h(t - \tau)d\tau \right\} e^{-2\pi jft} dt \ .$$
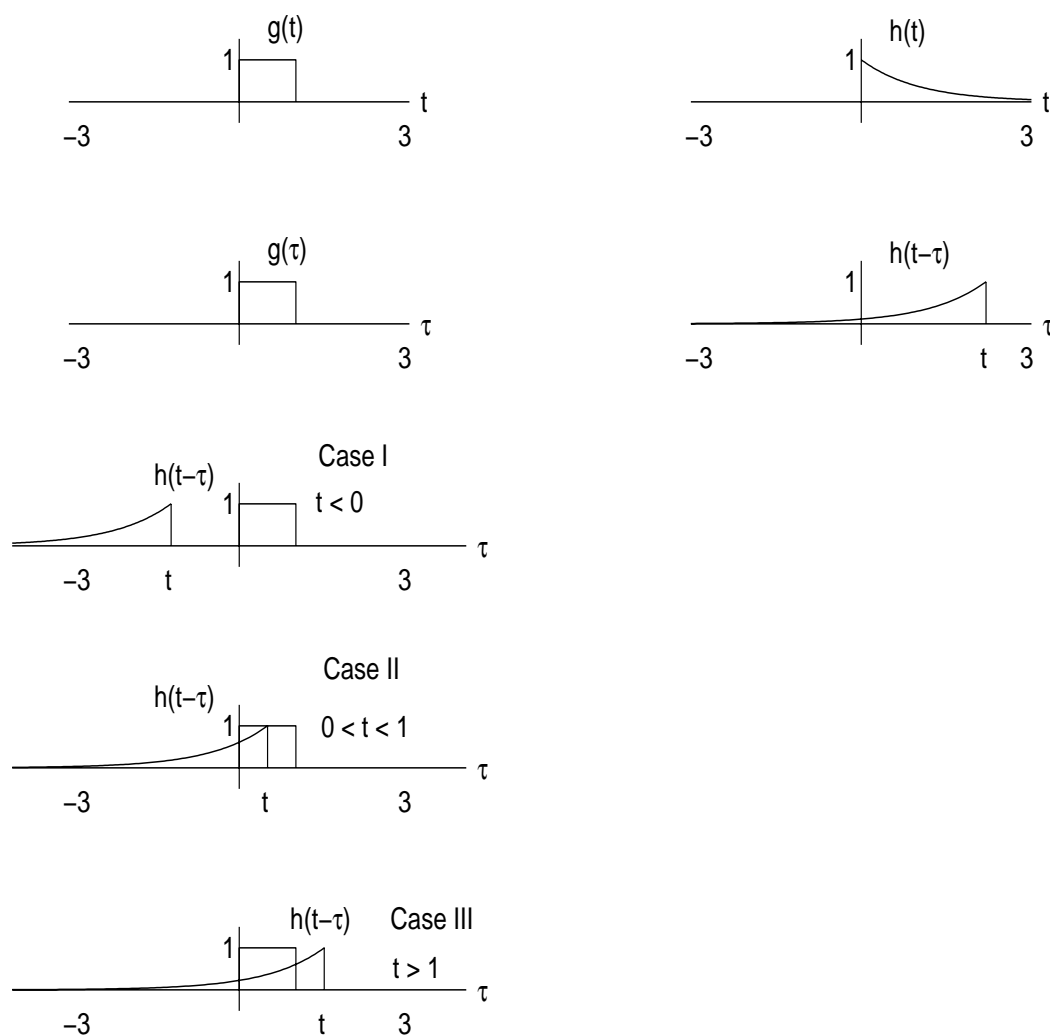
Figure 66: The Graphical Description of a Convolution

However, since this is now a double integral, we may at least formally interchange the order of integration and write

$$\mathcal{F}\left[g(t) * h(t)\right] = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\tau)h(t-\tau)e^{-2\pi jft}dt \right\} d\tau$$

(Note, as in most of our other discussions of the Fourier transform, we call this interchange formal because we really have not rigorously justified extending the basic calculus result to improper integrals such as these.) But, with the integral now written this way, we may move $g(\tau)$ outside of the new inner integral, since it does not depend on the variable of integration (t) there. This yields

$$\mathcal{F}\left[g(t) * h(t)\right] = \int_{-\infty}^{\infty} g(\tau) \left\{ \int_{-\infty}^{\infty} h(t-\tau)e^{-2\pi jft}dt \right\} d\tau$$
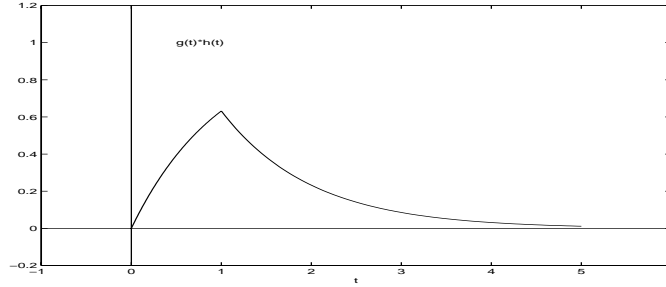
218

Figure 67: The Graph of a $g(t) * h(t)$ from the Example

Now, by definition, the remaining inner integral is just the Fourier transform with respect to t of $h(t - \tau)$, i.e.

$$\mathcal{F}\left[g(t) * h(t)\right] = \int_{-\infty}^{\infty} g(\tau)\mathcal{F}\left[h(t - \tau)\right] d\tau \ .$$

Therefore, by the shifting property of the transform we can replace $\mathcal{F}\left[h(t - \tau)\right]$ by

$$\mathcal{F}\left[g(t) * h(t)\right] = \int_{-\infty}^{\infty} g(\tau)e^{-2\pi j f \tau}H(f) \, d\tau \ .$$

However, in this form, $H(f)$ does not depend on τ and therefore we may take it completely outside the integral, i.e.

$$\mathcal{F}\left[g(t) * h(t)\right] = H(f)\int_{-\infty}^{\infty} g(\tau)e^{-2\pi j f \tau} d\tau \ .$$

But when we do this, the remaining integral is, by definition, nothing more than just the transform of $g(t)$, and so our transform simply reduces to

$$\mathcal{F}\left[g(t) * h(t)\right] = H(f)G(f)$$

To summarize, we have just shown that

$$\mathcal{F}\left[g(t) * h(t)\right] = G(f)H(f) \ , \qquad (6.2.2)$$

or, in words, *a convolution in the time domain corresponds to multiplication of the respective transforms in the frequency domain.* This result is commonly referred to as the Convolution Theorem. (As you may recall, basically the same property holds for the Laplace transform of the Laplace convolution, except, of course, we do not interpret the Laplace transform domain in terms of frequencies. In view of our earlier comments about the similarities between the Laplace and Fourier transforms, and of the identity of the Laplace and Fourier convolutions for causal functions, this should not be too surprising.)

To demonstrate the convolution theorem, consider the following

Example:

$$h(t) = \left\{ \begin{array}{ccc} 1 & , & -\dfrac{1}{2} < t < \dfrac{1}{2} \\ 0 & , & \text{otherwise} \end{array} \right.$$

219

We can show fairly straightforwardly (using Figure 68) that

$$h(t) * h(t) = \begin{cases} (1+t) & , & -1 < t < 0 \\ (1-t) & , & 0 < t < 1 \\ 0 & , & \text{otherwise} \end{cases}$$
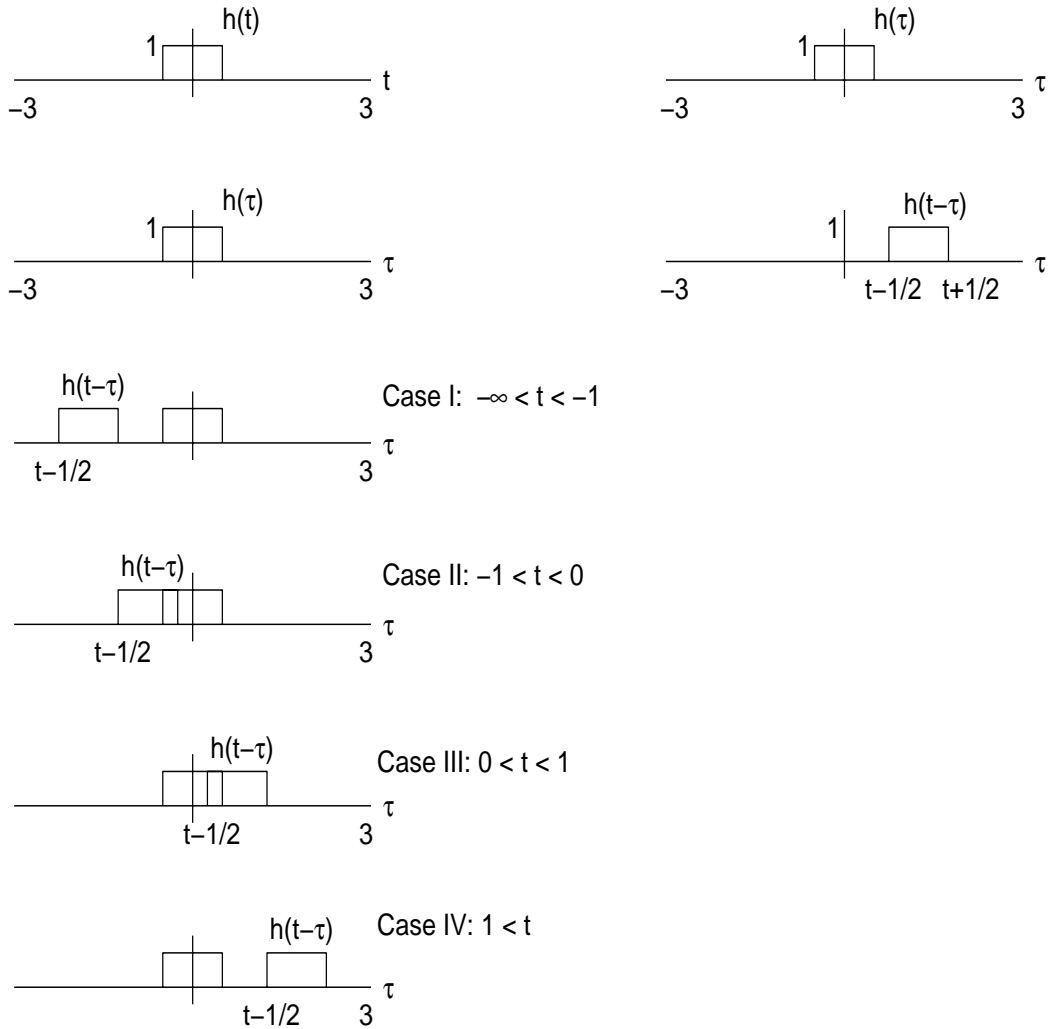

Figure 68: The Graphical Description of a Second Convolution

Since the resulting convolution is an even function, we can then compute its Fourier transform

directly as follows

$$\mathcal{F}\left[h(t) * h(t)\right] \ = 2 \int_0^1 (1-t)\cos(2\pi f t)dt$$

$$= \left[\frac{(1-t)}{2\pi f}\sin(2\pi f t) - \frac{1}{(2\pi f)^2}\cos(2\pi f t)\right]\Bigg|_0^1$$

$$= \frac{1-\cos(2\pi f)}{2(\pi f)^2} = \frac{\sin^2(\pi f)}{(\pi f)^2}$$

$$= \frac{\sin(\pi f)}{(\pi f)}\cdot\frac{\sin(\pi f)}{(\pi f)}$$

This result, of course, is exactly what the convolution theorem predicts. (Note our use of the trigonometric identity

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

Therefore, on the one hand, we might say the convolution theorem, like any other transform rule, does not really provide any information that we could not obtain by direct computation. On the other hand, if all we had needed was $\mathcal{F}\left[h(t) * h(t)\right]$, using the convolution theorem would have saved us significant computational effort. But, as we shall see in subsequent sections, the importance of the convolution theorem is not solely as a computational tool. The convolution theorem also provides fundamental insights into the behavior of physical systems, insights that would be extremely difficult to find without it.

We would close this section with one further thought. One of the powerful aspects of Fourier transform analysis is that, because of the similarity of the transform and its inverse, most time domain properties produce related frequency domain properties. This is also the case with the convolution theorem, for we can fairly straightforwardly show that the Fourier transform of the product of two functions produces a convolution of their transforms in the frequency domain, i.e.

$$\mathcal{F}\left[g(t)\cdot h(t)\right] = G(f) * H(f) \equiv \int_{-\infty}^{\infty} G(\phi)H(f-\phi)\ d\phi \quad . \qquad (6.2.3)$$

We shall also see applications of this result in later sections.

PROBLEMS

1. Compute, using the definition, the convolution ($h(t) * g(t)$) in the following cases. Then, in each case, compute the Fourier transform of the convolution and verify the result agrees with the convolution theorem:

a. $h(t) = \begin{cases} 2 & , \quad 0 < t < 2 \\ 0 & , \quad \text{otherwise} \end{cases}$

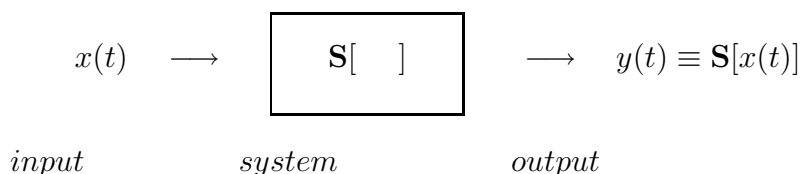$g(t) = \begin{cases} e^{-t} & , \quad 0 < t \\ 0 & , \quad \text{otherwise} \end{cases}$

b. $h(t) = g(t) = \begin{cases} 2 & , \quad -2 < t < 2 \\ 0 & , \quad \text{otherwise} \end{cases}$

c. $h(t) = e^{-|t|}$, $\quad -\infty < t < \infty$

$g(t) = \cos(2\pi t)$, $\quad -\infty < t < \infty$

6.3   Linear, Shift-Invariant Systems

One of the primary uses of the Fourier transform is in the analysis of what several authors (e.g. Gaskill) refer to as linear, shift-invariant systems. (So-called *filters*, which arise in signal processing, are one especially important instance of such systems.) Linear, shift-invariant systems are, as we shall see, fundamental to much of engineering design. However, before proceeding any further, we first need to specify exactly what we mean by the term linear, shift-invariant system.

When we speak of a system, we mean any mathematical or physical process that can be essentially described by an input/output relationship. (Many control engineering texts use the terms "plant" and system interchangeably.) In other words, a system is basically any entity that can be described by the "black box" model

$$x(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad y(t) \equiv \mathbf{S}[x(t)]$$

$$\textit{input} \qquad\qquad \textit{system} \qquad\qquad \textit{output}$$

Systems such as this arise any time we subject a physical object to some kind of force or control mechanism, and are often described by differential equations where $y(t)$ represents the solution, and $x(t)$ the applied force. Examples of systems defined by ordinary differential equations are
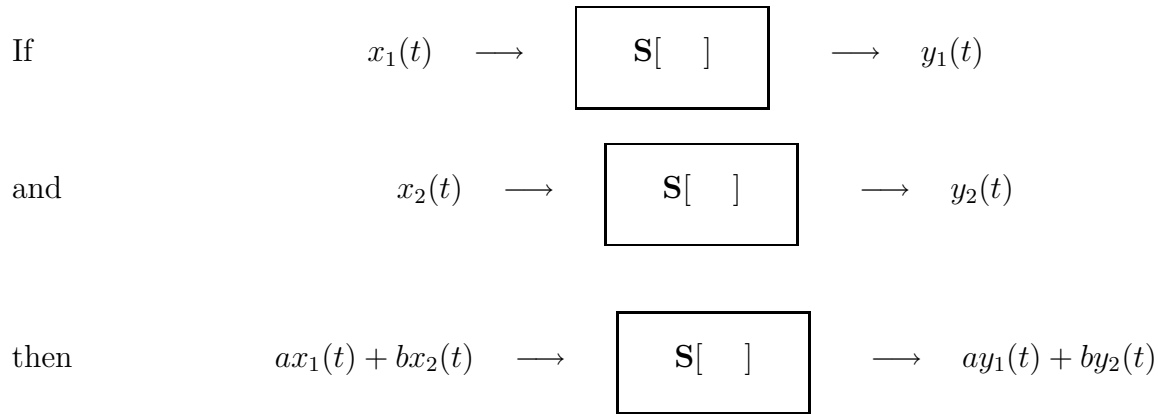
$$y''(t) + 3y'(t) + 2y(t) = x(t)$$
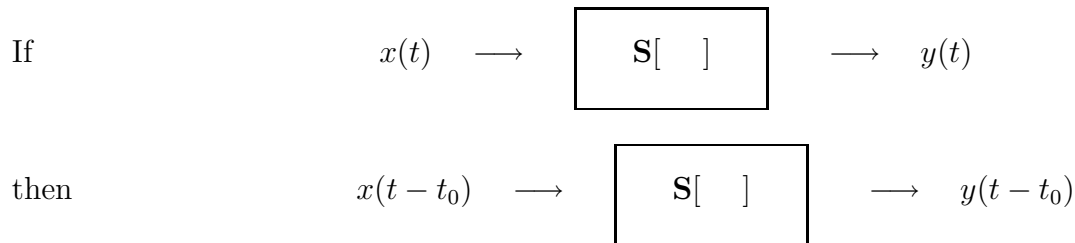
and

$$[y'(t)]^2 + y(t) = x^2(t) \quad .$$

(Note that the above diagram indicates that the input and output for a system both occur in the **same** domain (in this case the "time" domain.) This contrasts with the Fourier transform box we used earlier, where the inputs and outputs were in different (time and frequency, respectively) domains. Furthermore, the inputs and outputs ($x(t)$ and $y(t)$, respectively) to a system need not be limited to only scalars, but may be vectors as well.)

Systems may be either linear or nonlinear. A linear system, is one which obeys the principal of superposition, i.e. one in which the response (output) of that system to a sum of inputs is the sum of the responses (outputs) to the corresponding individual inputs. In terms of our black box model, a linear system must behave as follows

223

If $\qquad x_1(t) \longrightarrow \boxed{\mathbf{S}[\ \ ]} \longrightarrow y_1(t)$

and $\qquad x_2(t) \longrightarrow \boxed{\mathbf{S}[\ \ ]} \longrightarrow y_2(t)$

then $\qquad ax_1(t) + bx_2(t) \longrightarrow \boxed{\mathbf{S}[\ \ ]} \longrightarrow ay_1(t) + by_2(t)$

where $x_1(t)$ and $x_2(t)$ represent arbitrary inputs and a and b are arbitrary scalars.

Lastly, a system is shift-invariant if delaying the input by an arbitrary amount produces the identical response as to the undelayed input, except that the response is delayed by the same amount as was the input, or, in terms of our black box model:

If $\qquad x(t) \longrightarrow \boxed{\mathbf{S}[\ \ ]} \longrightarrow y(t)$

then $\qquad x(t - t_0) \longrightarrow \boxed{\mathbf{S}[\ \ ]} \longrightarrow y(t - t_0)$

Shift-invariant behavior is so common that we perhaps take it for granted. But, anything else would make chaos of almost all engineering design! For example, how could we possibly build automobiles if turning the steering wheel clockwise produced right turns on Mondays, left turns on Tuesdays, no effect on Wednesdays, etc.?

With regards to the preceding discussion, you should realize that as far as systems are concerned, linearity and shift invariance are independent considerations. For example, a system may be linear, but not shift-invariant (e.g. the system represented by

$$y'' + ty = x(t)\ \ )$$

or neither linear nor shift invariant (e.g. the system represented by

$$y'' + te^y = x(t)\ \ .)$$

In light of these comments, you should not be surprised to find that when linear, shift-invariant systems are expressed in terms of differential equations, the equations will be linear and constant coefficient.

Most engineering design focuses, at least initially, on linear, shift-invariant systems. Whether this is because most real-world systems can at least be approximated by them,

or whether it is just because these are the only ones for which explicit solutions are generally possible is arguable, but the result is the same. Linear, shift-invariant systems are simply fundamental to real applications.

For the purposes of our study, the mathematical importance of linear, shift-invariant systems is that we can show, at least formally, that a system is linear and shift-invariant if and only if its output is a convolution of the input with some other, system specific function, i.e., in terms of our black box model, a linear, shift-invariant system must obey the relation

$$x(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad y(t) \equiv \mathbf{S}[x(t)] = \int_{-\infty}^{\infty} x(u)h(t-u)du$$

where $h(t)$ is some function that depends solely on the particular system being considered, and not on $x(t)$.

Formally showing that systems whose outputs are convolutions must be linear and shift-independent is extremely straightforward. Linearity follows immediately from basic calculus principles, since the integral of a sum is the sum of the integrals. Showing that the convolution is also shift-invariant requires only a simple change of variables in the integral, i.e.

$$y(t - t_0) \quad = \int_{-\infty}^{\infty} x(u)h((t-t_0) - u)du$$

$$= \int_{-\infty}^{\infty} x(u - t_0)h(t-u)du \equiv \mathbf{S}[x(t-t_0)]$$

where we convert between the first and second integrals by replacing u in the first integral by $(u - t_0)$ everywhere.

Proving that a linear, shift-invariant systems must satisfy a convolution relationship is a bit more involved. We start by considering some arbitrary interval, say from T_0 to T_1. We next divide this interval up into subintervals of equal length, denoting the endpoints of these subintervals as $u_0, u_1, u_2, \ldots, u_N$, where $u_0 = T_0$ and $u_N = T_1$. We also let $\Delta u = u_{i+1} - u_i$. Then, because of its linearity, given *any* functions $x(t)$ and $w(t)$, the system must obey

$$\sum_{i=1}^{N} x(u_i)w(t-u_i)\Delta u \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \sum_{i=1}^{N} x(u_i)\mathbf{S}[w(t-u_i)]\Delta u$$

where $\mathbf{S}[w(t)]$ denotes the response (output) of the system to the input $w(t)$. (The output must have this form because the u_i are all constant values. Therefore the $x(u_i)$ are just numbers, i.e. coefficients, and so the input is just a linear combination of the $w(t - u_i)$. Hence, since the system is linear, the output must be just the sum of the corresponding individual responses.) However, both this input and output are in precisely the form of a Riemann sum approximation to a definite integral. Therefore, applying the same arguments as we used in deriving the original Fourier transform identity in the previous chapter, we claim that as $\Delta u \to 0$, the system input/output relationship must become

$$\int_{T_0}^{T_1} x(u)w(t-u)du \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \int_{T_0}^{T_1} x(u)\mathbf{S}[w(t-u)]du$$

But now, in the special case where $T_0 = -\infty$ and $T_1 = \infty$ (a case which again really requires some very delicate consideration of limiting processes), we can formally write that

$$\int_{-\infty}^{\infty} x(u)w(t-u)du \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \int_{-\infty}^{\infty} x(u)\mathbf{S}[w(t-u)]du$$

You should appreciate that we have derived this last relationship using no information except that this system was linear, and assuming nothing about $x(t)$ and $w(t)$. We shall now proceed, continuing to assume that $x(t)$ is arbitrary, but limiting our attention to a very special $w(t)$ - an impulse - i.e. $w(t) = \delta(t)$. In this case, the input/output black box relationship immediately above can be written

$$\int_{-\infty}^{\infty} x(u)\delta(t-u)du \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \int_{-\infty}^{\infty} x(u)\mathbf{S}[\delta(t-u)]du$$

or, using the sifting property of the delta function,

$$x(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \int_{-\infty}^{\infty} x(u)\mathbf{S}[\delta(t-u)]du$$

But we have also assumed that the system is shift-invariant. Therefore, if we denote the response of the system to the basic impulse $\delta(t)$ as $h(t)$, i.e. if

$$h(t) = \mathbf{S}[\delta(t)]$$

or equivalently, if

$$\delta(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad h(t)$$

then, because of shift invariance, $\mathbf{S}[\delta(t-u)] = h(t-u)$. Therefore, in summary, we have shown that, for an arbitrary linear, shift invariant system,

$$x(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \int_{-\infty}^{\infty} x(u)h(t-u)du$$

Since $\delta(t)$ is an impulse, then $h(t)$ as defined above is commonly referred to as the *impulse response* of the system. Therefore we may restate the conclusion of this discussion as being that *the output of any linear, shift-invariant system is the convolution of the input with the system impulse response.*

This last result expresses the fundamental behavior of linear, shift-invariant systems *in the time domain.* But, as we have seen many times before, one of the central ideas of Fourier

analysis is that phenomena exist simultaneously in **both** the time and frequency domain. Therefore, for completeness, we should also interpret the input/output relationship of linear, shift-invariant systems in the frequency domain. But finding this interpretation is simple - the convolution theorem (6.2.2) tells us that, in the transform domain,

$$X(f) \quad \longrightarrow \quad \boxed{\ \mathbf{S}[\quad] \ } \quad \longrightarrow \quad X(f)H(f)$$

In this representation, of course, $H(f)$ simply denotes the Fourier transform of the system's impulse response, i.e.

$$H(f) = \int_{-\infty}^{\infty} h(t)e^{-2\pi j f t} dt \ \ .$$

This transform is commonly also called the *transfer function* or *frequency response* of the system.

The fact that a linear, shift-invariant system's output is a product in the frequency domain turns out to be of immense value! Analyzing the responses of such systems to general inputs in the time domain would require computing convolutions and, as you hopefully appreciate from the examples we have already presented, both the mechanics of performing such computations and the subsequent visualization of the result can be fairly involved. By contrast, as a consequence of the convolution theorem, analyzing a system's general behavior in the frequency domain simply requires multiplying two curves together. This is usually very straightforward and the result is generally quite easy to picture. For example, consider Figure 69. This figure clearly implies that the output of the system under consideration will be a signal with all of the "low" frequencies from the original input unaffected, and all of the "high" frequencies lost. (Whether such an output is "good" or "bad" is a design issue, and not the point of the discussion here. The important point for us is that the frequency domain/transfer function formulation of the system response lets us very quickly capture the essence of how this particular system affects input signals.)

The use of the frequency domain/transfer function formulation can also greatly simplify the analysis and construction of complicated systems, by allowing a modular, building block approach. Specifically, suppose we have a system consisting of two serial subsystems, denoted $\mathbf{S}_1$ and $\mathbf{S}_2$. Each of these subsystems will have its own impulse response and transfer function, which we denote in the natural way, e.g. $H_2(f)$ would be the transfer function for subsystem $\mathbf{S}_2$. In the frequency domain, the input/output diagram for this system becomes

$$X(f) \quad \longrightarrow \quad \boxed{\ \mathbf{S}_1[\quad] \ } \quad \longrightarrow \quad H_1(f)X(f) \quad \longrightarrow$$

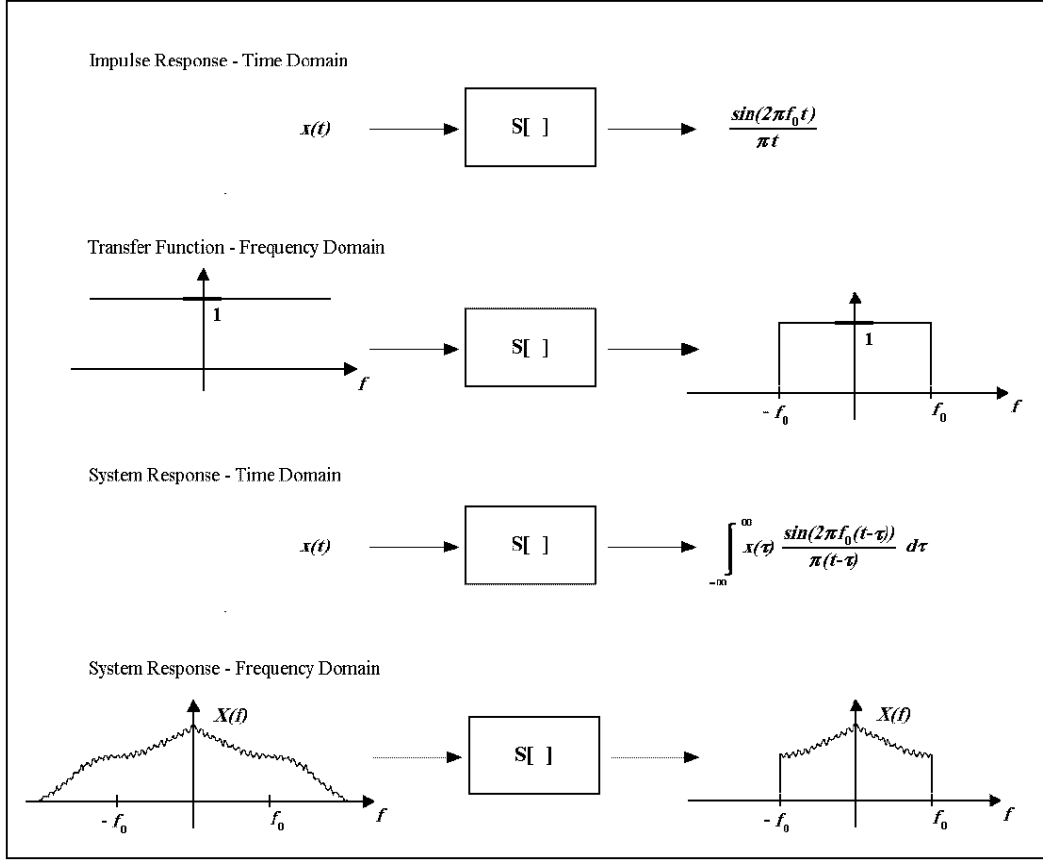$$\longrightarrow \quad \boxed{\ \mathbf{S}_2[\quad] \ } \quad \longrightarrow \quad H_2(f)H_1(f)X(f)$$

Figure 69: The Graphical Description of a System Output in the Transform Domain

But this diagram clearly implies that the resultant output would be the same as from a single system whose transfer function were $H_1(f)H_2(f)$. Even more importantly, it implies that we can replace any single, complicated system with a system made up of a number of simpler components, provided the product of the transfer functions of the components of the simpler systems is identical to the transfer function of the more complicated system. This conclusion has far-reaching implications in terms of both design complexity and cost!

As the preceding discussion indicates, the impulse response and transfer function of a system are pivotal to describing the behavior of that system and for designing systems to behave in certain desired ways. (Of course, since the impulse response and transfer function are a Fourier transform pair, knowledge of either uniquely determines the other and uniquely and completely defines the system itself.) At this point though, the procedures for finding either one are not clear. Therefore, we shall consider that question next.

6.4 Determining a System's Impulse Response and Transfer Function

Given the fundamental roles of the impulse response and transfer function in defining the behavior of any given system, determining them becomes a crucial part of system design and

228

analysis. (Actually, as noted above, determining either one is, in concept, sufficient.) We therefore now consider the various ways in which either may be found, with special emphasis on those techniques that apply not only to model (mathematical) systems, but also in the "real" world.

Theoretically, finding the impulse response for a system should be quite straightforward. After all, the impulse response is, by definition, just the system behavior resulting from an impulse force at $t = 0$. So why not just "hit" the system with an impulse and see what happens? Mathematically, this is quite easily done - at least for systems which are modeled by constant coefficient ordinary differential equations. All we need to do is just replace the forcing function in the system's differential equation by $\delta(t)$, zero out any initial conditions and then solve, e.g. by Laplace transforms. For example, consider the RC circuit shown in Figure 70. At any time, the charge on the capacitor ($Q(t)$) and the applied external voltage ($E(t)$) are related by
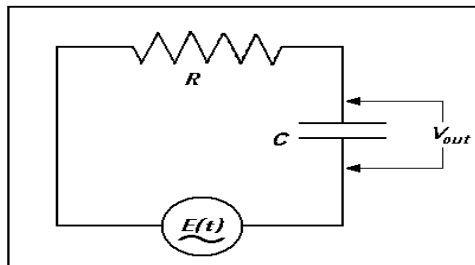
Figure 70: A Sample RC Circuit

$$R \frac{dQ}{dt} + \frac{1}{C} Q = E(t)$$

Therefore, if we take the "input" as the applied external voltage and the "output" as the voltage measured across the capacitor, we can represent this system by the black box

$$E(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad V_{\text{out}}(t) = \frac{1}{C} Q(t)$$

But the charge on the capacitor produced by a unit impulse voltage at $t = 0$ is simply the solution to

$$R \frac{dQ}{dt} + \frac{1}{C} Q = \delta(t)$$
$$Q(0) = 0$$

Now, as noted earlier, constant coefficients and impulsive forcing terms generally dictate solving by Laplace transforms. In this instance, taking the transform of both sides leads to

$$\left(Rs + \frac{1}{C}\right) \mathcal{L}\left[Q(t)\right] = 1$$

or (using tables as appropriate)

$$\mathcal{L}\left[Q(t)\right] = \frac{1}{R}\left[\frac{1}{s + \frac{1}{RC}}\right] \Rightarrow Q(t) = \begin{cases} \dfrac{1}{R}\, e^{-t/RC} & , \quad t \geq 0, \\[2ex] 0 & , \quad \text{otherwise} \end{cases}$$

Converting the charge to the voltage on the capacitor shows that, by definition, the impulse response of this circuit is

$$h(t) = \frac{1}{C}\, Q(t) = \begin{cases} \dfrac{1}{RC}\, e^{-t/RC} & , \quad t \geq 0, \\[2ex] 0 & , \quad \text{otherwise} \end{cases} \tag{6.4.4}$$

With the impulse response now available, we could now find this system's transfer function by Fourier transforming $h(t)$. But note that this $h(t)$ is clearly causal and has only finite energy. (You should view this as a fairly natural consequence of the fact that this was a "real" physical system.) Therefore, because we already have the Laplace transform of $h(t)$, we can find its Fourier transform by one of the rules of the previous chapter, i.e. by simply replacing the variable s in the Laplace transform by the Fourier term $(2\pi jf)$,

$$H(f) = \frac{1}{C}\, Q(f) = \frac{1}{C} \cdot \frac{1}{R} \cdot \frac{1}{2\pi jf + \frac{1}{RC}} = \frac{1}{1 + 2\pi jf RC} \tag{6.4.5}$$
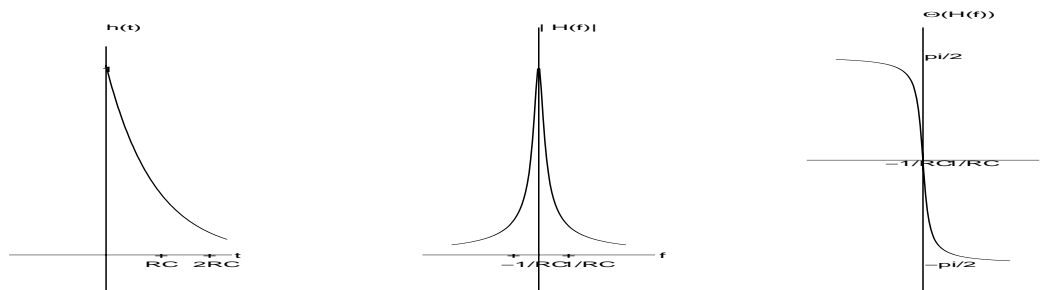
Figure 71: An Example Impulse Response and Transfer Function

(Both the impulse response and transfer function for this system are plotted in Figure 71.)

Actually, since we know how to relate the Fourier transform of a derivative to the Fourier transform of the original function, we could equally well have attacked this problem in the reverse order. That is, we could have first determined the transfer function by Fourier transforming the differential equation. This would yield

$$\left(2\pi jf R + \frac{1}{C}\right) Q(f) = 1 \Rightarrow Q(f) = \frac{1}{2\pi jf R + \frac{1}{C}}$$

Because of the linearity of the transform, we could then divide by C to obtain the transform of the output (capacitor) voltage. This would, of course, yield the same $H(f)$ as we found

above. Finally, inverse Fourier transforming this transfer function would produce the impulse response, $h(t)$.

Mathematically, this approach - subjecting the system to an impulse - is direct, simple, and clearly applicable to any system described by a constant coefficient ordinary differential equation. In general, the worst that we expect to encounter would be "nasty" algebra as part of inverting the Laplace transform, or computing Fourier transform or inverse transform integrals. However, subjecting a *real* system to an impulse may not be such a good idea! Real impulses are not that easy to generate - they can't be truly instantaneous, but still must be of extremely short duration in order to adequately approximate a true impulse. Furthermore, because of their short duration, real impulses must also have very high amplitudes. Unfortunately, high amplitudes, even when of short duration, can easily damage a real system - especially if sensitive electronics are involved. Therefore, some less stressful testing method than impulse loading seems preferable for real systems. But what method should we then use?

One of the easiest forcing functions to generate, and one of the more harmless, is a simple sinusoid, e.g. AC current. (Of course, sinusoids are periodic, and there are no truly periodic functions. However, common sense argues that, provided we subject a system to a force which is periodic over a sufficient length of time for all natural transients to die out, the (steady-state) response of the system should be essentially the same as the response to a truly periodic force.) But does the response of a system to a sinusoidal input shed any light on the response of that same system to an impulse?

To find out, we select as our forcing function the complex exponential $e^{2\pi j f_0 t}$, where f_0 is some fixed frequency. In terms of our black box model and the properties of the impulse response, this produces the following input/output situation

$$e^{2\pi j f_0 t} \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad \int_{-\infty}^{\infty} h(t-\tau)e^{2\pi j f_0 \tau}\ d\tau$$

But now, if we make the change of variables $u = t - \tau$ in the output integral (remembering that τ, not t is the variable of integration, we have

$$\int_{-\infty}^{\infty} h(t-\tau)e^{2\pi j f_0 \tau}\ d\tau\ =\ \int_{\infty}^{-\infty} h(u)e^{2\pi j f_0 (t-u)}\ (-du)$$

$$=\ e^{2\pi j f_0 t}\int_{-\infty}^{\infty} h(u)e^{-2\pi j f_0 u}\ du$$

But this last integral is nothing more than, by definition, the Fourier transform of $h(t)$ evaluated at the frequency f_0! Thus, in terms of our black box model

$$e^{2\pi j f_0 t} \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad e^{2\pi j f_0 t}H(f_0)$$

i.e. the output is simply the input sinusoid, multiplied by the value of transfer function at that frequency. In other words, we can directly *measure* a system's transfer function, just by

measuring the system's response to a periodic (and almost certainly nondamaging) forcing function! (More precisely, $|H(f_0)|$ is the ratio of the amplitude of the output to the amplitude of the input, and $\Theta\left(H(f_0)\right)$ is the difference in phase between the output and input. Of course, strictly speaking, such a measurement only tells us the transfer function at the single forcing frequency f_0. But we simply need to repeat the test, over any appropriate range of frequencies, in order to find the transfer function over that range. This testing method helps give the transfer function its other commonly-used name - the frequency response.) Moreover, once we have obtained the transfer function by this method, we can then find the impulse response mathematically by just taking an inverse transform, thus completely avoiding having to load the system with an impulse.

For systems described by constant coefficient ordinary differential equations, we can also implement this last approach mathematically. We simply replace the forcing function in the differential equation with a complex exponential, then solve by undetermined coefficients for the steady-state solution. For example, for the circuit shown in Figure 70, with an input signal of $E(t) = e^{2\pi j f_0 t}$, the differential equation becomes

$$R\,\frac{dQ}{dt} + \frac{1}{C}\,Q = e^{2\pi j f_0 t}$$

By undetermined coefficients, the steady-state particular solution to this problem is

$$Q(t) = \frac{1}{2\pi j f_0 R + \frac{1}{C}}\,e^{2\pi j f_0 t} = \frac{C}{2\pi j f_0 RC + 1}\,e^{2\pi j f_0 t}$$

and therefore the response across the capacitor is

$$V_{\text{out}}(t) = \frac{1}{C}\,Q(t) = \frac{1}{2\pi j f_0 RC + 1}\,e^{2\pi j f_0 t}$$

Thus,

$$H(f_0) = \frac{V_{\text{out}}(t)}{e^{2\pi j f_0 t}} = \frac{1}{2\pi j f_0 RC + 1} \Rightarrow H(f) = \frac{1}{1 + 2\pi j f RC}$$

This of course, is exactly the same transfer function as we found earlier, and we could now again determine the impulse response of this system (if we did not already know it) simply by computing the inverse Fourier transform of this $H(f)$.

There is one small practical problem with the above development. Complex-valued functions are really only mathematical artifices, and not producible in the "real" world. So what do we use as the actual input signal?

The answer is quite simple - we use $\cos(2\pi f_0 t)$! As you should be able to verify from

basic trigonometric identities,

$$h(t) * \cos(2\pi f_0 t) = \int_{-\infty}^{\infty} h(\tau) \cos(2\pi f_0(t-\tau)) \, d\tau$$

$$= \cos(2\pi f_0 t) \int_{-\infty}^{\infty} h(\tau) \cos(2\pi f_0 \tau) \, d\tau$$

$$+ \sin(2\pi f_0 t) \int_{-\infty}^{\infty} h(\tau) \sin(2\pi f_0 \tau) \, d\tau$$

$$= \cos(2\pi f_0 t) \, \Re\mathbf{e} \, [H(f_0)] - \sin(2\pi f_0 t) \, \Im\mathbf{m} \, [H(f_0)]$$

Thus the real and imaginary parts of $H(f_0)$, respectively, can be found directly, with a real input signal, just by measuring the amplitude of the output at two times - one in phase with the input and the other 90° out of phase with the input.

We would close this section with one last point. The results presented here provide the basic methods for computing the impulse response and transfer function of a given system. We would reemphasize that finding these is not just a mathematical exercise. The transfer function is fundamental to analyzing the behavior of systems, since our earlier analysis of linear, shift-invariant systems showed that the convolution theorem expresses the output of such systems as a product of the transform of the input with the transfer function. For example, consider the transfer function we have calculated in this section (Figure 71). Fairly clearly

$$\begin{aligned} |H(f)| &\approx 1 \quad , \; f << 1/RC \\ |H(f)| &\to 0 \quad , \; f \to \infty \end{aligned}$$

and so this system will pass "low" frequency components in the input relatively unaffected, but severely attenuate "high" frequencies. (Furthermore, because $\Theta\left(H(f)\right)/f$ is not constant, different frequency components in the input signal will be delayed by different amounts and become "out of step.") In fact, if

$$X(f) = \mathbf{O}\left(\frac{1}{f}\right) \qquad \text{then} \qquad X(f)H(f) = \mathbf{O}\left(\frac{1}{f^2}\right)$$

In other words, this system will apparently sufficiently attenuate the high frequencies which are necessary to produce discontinuities that, given a discontinuous input signal, the output will be continuous, with only a discontinuous derivative. Therefore, we expect this system will generally distort input signals. We may confirm this analysis by considering a sample input function, e.g. the one shown, together with the corresponding system output (Figure 72). (This output was computed directly as the convolution of the input with (6.4.4). The computation is not presented here because, except for scaling terms, it is identical to the first example in this chapter.) As expected, the output, while clearly related to the input, is also not a "clean" copy, but a distorted one. (The similarities between this analysis and the one we conducted for the LRC circuit with periodic input in Chapter 2 are **not** coincidental!)
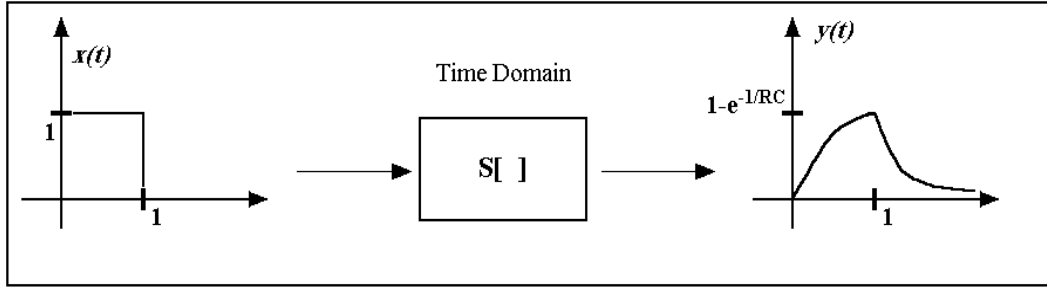
Figure 72: An Example Input and Output for an *RC* Circuit

6.5   Applications of Convolution - Signal Processing and Filters

As indicated at the close of the last section, the importance of the Fourier transform, of convolution, of the transfer function, etc., is in their value as tools for analyzing physical systems. One of the most important application areas where they are used is in so-called signal processing. In signal processing, we start by assuming that some information-carrying signal is initially transmitted into a physical medium. (For example, either an antenna may transmit a radio signal into the atmosphere or a transducer may transmit a sonar signal into water.) The medium is commonly referred to as the *channel*. After passing through the medium, the signal is received at some location. (This location is usually, but not always, somewhere other than the transmitting location.) However, because of physical effects encountered while passing through the channel, or perhaps because of the presence of more than one signal in the channel at the same time, the received signal is generally some distorted or otherwise affected version of the original transmitted signal (in somewhat the same way the output at the capacitor in our last example was a distorted version of the input pulse). Schematically, we may represent this process as

$$x(t) \quad \longrightarrow \quad \boxed{\mathcal{C} \ [\quad]} \quad \longrightarrow \quad y(t)$$

Transmitted Signal   *Channel*   *Received Signal*
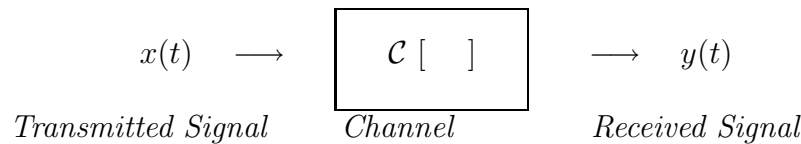
Signal processing then tries to design some circuit, etc., which, if we pass the received signal ($y(t)$) through it, will "undo" the effects introduced by the channel, resulting in an output that is reasonably close to the original transmitted signal, i.e.

$$y(t) \quad \longrightarrow \quad \boxed{\mathbf{S}[\quad]} \quad \longrightarrow \quad p(t) \ \ (\approx x(t))$$

Received Signal   *Processing*   *Output*
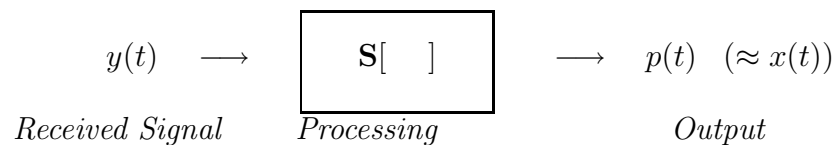
(Of course, ideally we would like to have $p(t)$ equal $x(t)$.)

   The simplest model of such a situation occurs when the channel is linear and shift-invariant. Then the channel has its own transfer function, which we denote $H_c(f)$, and therefore the transform of the received signal would be

$$Y(f) = H_c(f)X(f)$$

But theoretically, if this is true, all we need to do is design the processing system so that its transfer function is the reciprocal of that of the channel, i.e.

$$H_s(f) = \frac{1}{H_c(f)} \Rightarrow P(f) = H_s(f)Y(f) = \frac{1}{H_c(f)} \cdot H_c(f)X(f) \equiv X(f)$$

Unfortunately, for several reasons, only some of which we will touch on here, we usually can't reach this ideal in real problems. For one reason, in many cases we can't even completely specify all the effects of the channel - which makes it very difficult to remove them! However, we may nevertheless often reverse the primary corrupting effects of the channel, and recover an effectively fully usable version of the transmitted signal.

Perhaps the most basic type of signal processing involves *filters*. As their name implies, filters let some "things" through and reject others. The simplest filters are designed to let certain *frequencies* through and block (or attenuate) the remaining ones. Such filters fall into three general categories, which are almost self-defining

- Low-pass Filters - which allow low frequencies through but block high ones;

- High-pass Filters - which allow high frequencies through and block low ones; and

- Band-pass Filters - which block both very low and very high frequencies, but pass those in some intermediate range
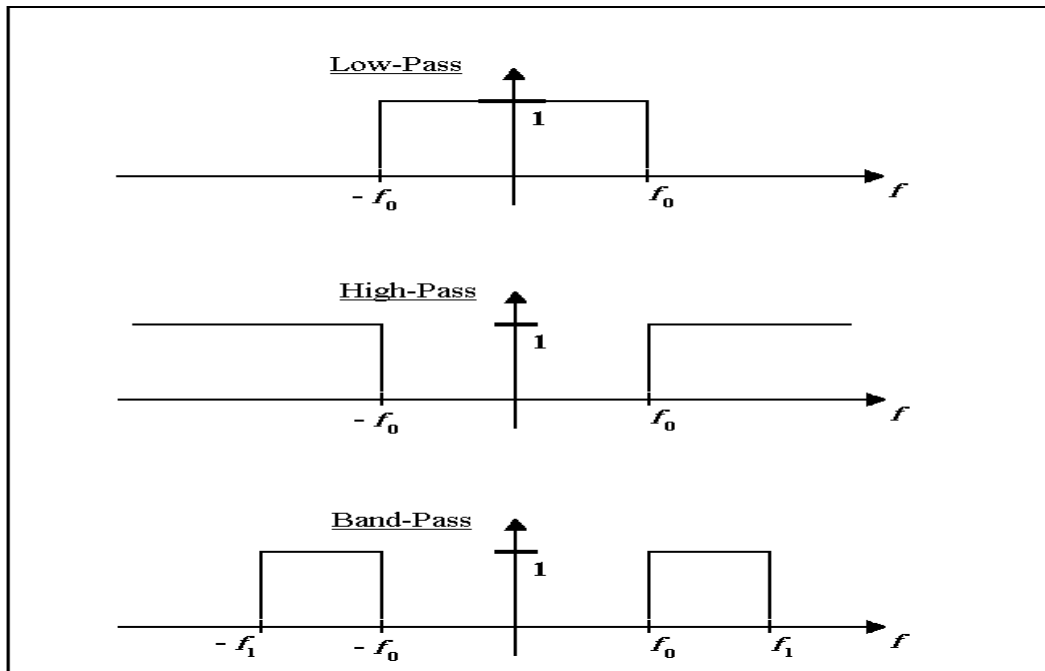
Figure 73: Transfer Functions for Ideal Filters

Figure 73 displays the transfer functions for the ideal cases of each of these. We call these ideal cases because, none of them are in fact physically achievable. The reason they are not is essentially the same in each case - they represent systems that are **not causal**. For example, the transfer function for the ideal low-pass filter is

$$H_{LP}(f) = \left\{ \begin{array}{ll} 1 & , \quad |f| < f_0 \\ 0 & , \quad \text{otherwise} \end{array} \right.$$

Unfortunately, this means that the impulse response for this filter is

$$h_{LP}(t) = \frac{\sin(2\pi f_0 t)}{\pi t}$$

and, as already indicated, this is not a causal system. On the other hand, causal systems with reasonably the same performance are realizable. For example, the RC system described in Figure 70 is easily to build, quite inexpensive, and, according to Figure 71, reasonably able to pass low frequencies but block high ones, provided we select R and C appropriately. High-pass and band-pass filters can be similarly realized by simply combining appropriately-valued resistors, capacitors and inductors, and measuring the output voltage on the appropriate device. The (magnitudes of the) transfer functions for all of these circuits are shown in Figure 74. Similar filters are common in virtually all home electronics. For example, a band-pass filter, consisting of an LRC circuit with a *variable* capacitor, is found in the tuning control of virtually every radio. Varying the capacitor adjusts the frequency at center of the pass-band and effectively selects, from all the frequencies received, signals with only that frequency.
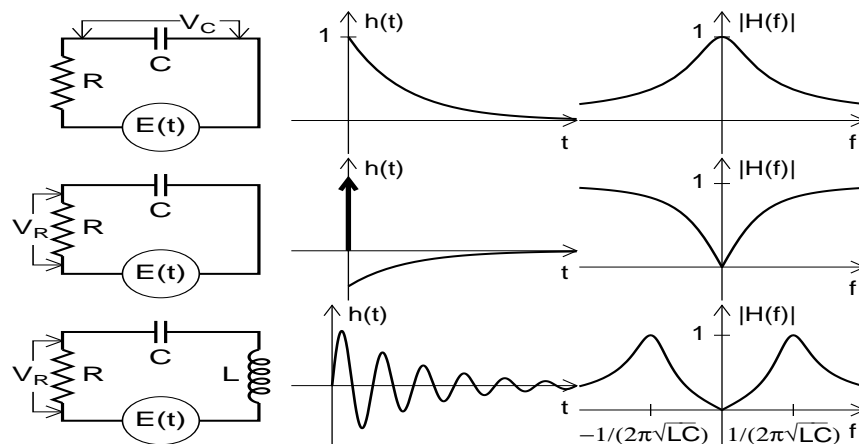
Figure 74: Real Filters With Their Impulse Responses and Transfer Functions. Top show RC filter (lo-pass),middle is RC filter (high-pass), and bottom is LRC filter (band-pass)

At this juncture, we shall close our discussion of filters and, to a reasonable degree, of linear, shift-invariant systems as well. We hope we have at least reasonably well set out their importance in engineering systems design, and the pivotal role the transfer function plays in their use. We are not however done with the Fourier transform yet, but will examine instead some of its other applications.

6.6 Applications of Convolution - Amplitude Modulation and Frequency Division Multiplexing

As we have previously alluded to, one attractive feature of the Fourier transform is the manner in which the near-symmetry between the transform and inverse transform operations implies that almost every time-domain property has a near mirror-image in the frequency domain. One instance of this is our earlier-mentioned converse to the convolution theorem (3), which we repeat here

$$\mathcal{F}[g(t) \cdot h(t)] = G(f) * H(f)$$

We now investigate one of the most common applications of this result.

Commercial radio broadcasts originally utilized what we call amplitude modulation. (The first commercial radio stations were also limited to using only a fairly narrow band of broadcast frequencies in the low MHz range. Because of the modulation used by virtually all of these early stations, this band of frequencies is generally still referred to as the AM band.) In amplitude modulation, the actual information signal, e.g. a talk show or music, constitutes what is generally referred to as the *baseband* signal, which we denote here as $m(t)$. We commonly refer to the range of frequencies which comprise the baseband, i.e. which are present in its Fourier transform ($M(f)$), as its *bandwidth*, which we denote by B. (In AM radio, the baseband bandwidth is generally limited to a relatively narrow range of low (less than 5 KHz) frequencies.) Amplitude modulation consists of electronically multiplying this baseband signal by a high frequency sinusoid - the so called *carrier* signal, which we shall represent as $\cos(2\pi f_0 t)$ - to produce the broadcast signal, which we denote $s(t)$ (Figure 75). (The carrier frequency (f_0) is typically in the MHz range.) In the time domain then, the resulting broadcast signal may be viewed as a very rapid oscillation within the envelope defined by the slowly varying baseband signal. The term amplitude modulation aptly describes this process because, as the figure shows, the amplitude of the carrier is modified (modulated) according to the amplitude of the baseband signal.
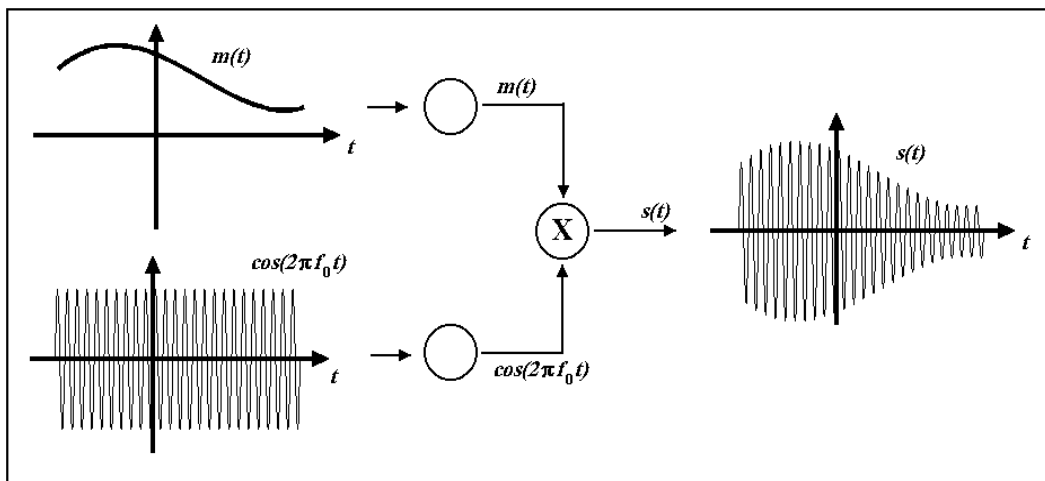

Figure 75: Amplitude Modulation - The Time Domain View

The frequency domain interpretation of amplitude modulation turns out to be quite intriguing. Specifically, according to (6.2.3), in the transform domain

$$S(f) \;=\; \mathcal{F}\left[s(t)\right] = \mathcal{F}\left[m(t) \cdot \cos(2\pi f_0 t)\right] = M(f) * \mathcal{F}\left[\cos(2\pi f_0 t)\right]$$

$$= M(f) * \left[\frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)\right]$$

$$= \int_{-\infty}^{\infty} M(f - \phi)\left[\frac{1}{2}\delta(\phi - f_0) + \frac{1}{2}\delta(\phi + f_0)\right] d\phi$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} M(f - \phi)\delta(\phi - f_0) \, d\phi + \frac{1}{2}\int_{-\infty}^{\infty} M(f - \phi)\delta(\phi + f_0) \, d\phi$$

But the sifting property of the delta function simplifies the two last integrals here, and shows that the transform of the broadcast signal is simply

$$S(f) = \frac{1}{2}M(f - f_0) + \frac{1}{2}M(f + f_0) \tag{6.6.6}$$
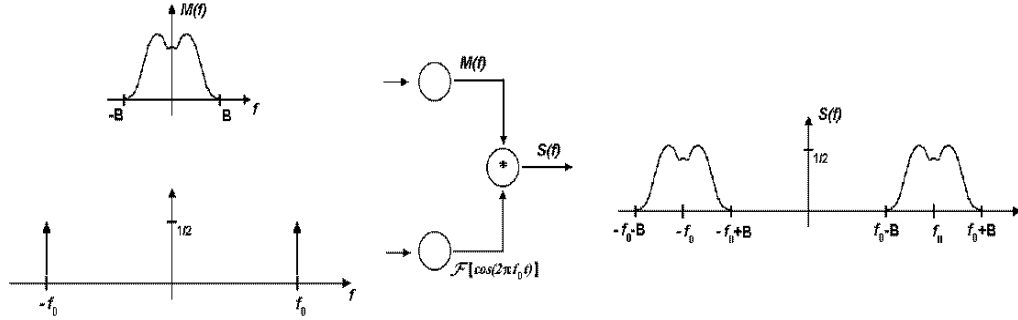
Figure 76: Amplitude Modulation - The Frequency Domain View

The graphical interpretation of this result in the frequency domain is both straightforward and elegant. Specifically, (6.6.6) say that $S(f)$ consists of precisely two half-size copies of the transform of the original baseband signal. One copy is centered about f_0, the other about $-f_0$ (Figure 76). (For graphical convenience, we have chosen to represent $M(f)$ in this figure as a real, even function. While the following discussion uses that fact, our conclusions remain valid, with at most minor modifications, in the case of a complex $M(f)$.)

Figure 76 also helps to explain another commonly occurring variant of amplitude modulation. In this figure, because of the even symmetry of $M(f)$, the part of the spectrum of $S(f)$ between $(f_0 - B)$ and f_0 is clearly totally redundant of that part between f_0 and $(f_0 + B)$. (These two parts are referred to, respectively, as the lower and upper sidebands of the transmitted signal.) But if the lower sideband contains no information not also available in the upper sideband, why bother to transmit both? The answer is we don't need to. We could

simply band-pass filter $S(f)$ before transmitting it, as in Figure 77, where the filter selects only the upper sideband. (Note that in Figure 77 we plot only the positive frequencies, since, due to symmetry, the spectrum for negative frequencies would be a mirror image.) This method of transmission is usually referred to as *single sideband* modulation. (In this terminology then, transmitting the full signal, as in Figure 76, may properly be referred to as double sideband modulation, although it is just as often simply called amplitude modulation.) Single sideband modulation is attractive since it can transmit the same information as a double sideband signal, but in only half the spectrum. As a practical matter, however, the additional circuitry costs associated with single sideband modulation frequently make it an uneconomical alternative to double sideband modulation, even though it more efficiently utilizes the available frequencies. Amplitude modulation has one additional impor-
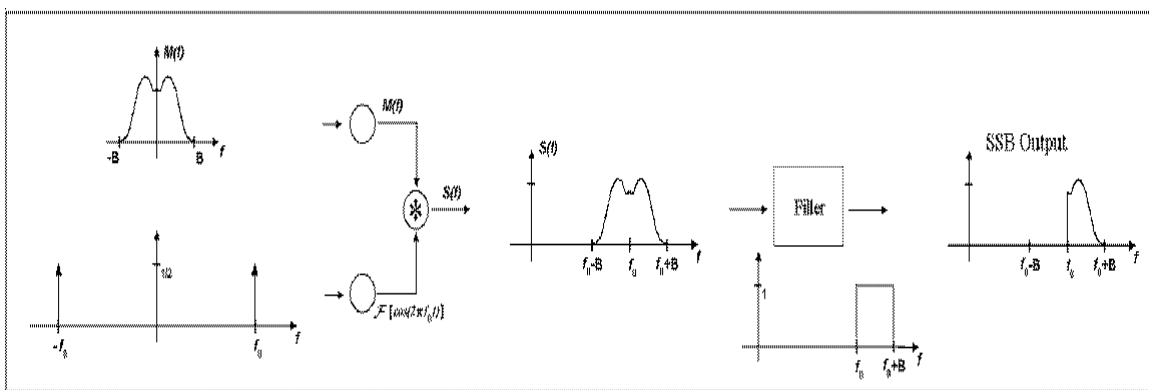
Figure 77: Single Sideband Modulation - The Frequency Domain View

tant feature. Specifically, suppose we have two *different* information-carrying signals, both of bandwidth B, and denoted respectively as $m_0(t)$ and $m_1(t)$, which we wish to transmit *simultaneously.* (We shall further assume that we wish to transmit both from the same transmitter, although the identical argument would apply were we to use different transmitters.) We now amplitude modulate both signals, but at *different* carrier frequencies (denoted f_0 and f_1 respectively), to produce the modulated output signals $s_0(t)$ and $s_1(t)$. We then add the modulated signals to form the single output signal $s(t)$. If we assume that f_0 and f_1 are chosen so that $(f_0 + B)$ is less than $f_1 - B$, then, based on our above discussion, the time and frequency domain representations of this process must then be as shown in Figure 78. But look carefully at this figure! The two output signals, $S_0(f)$ and $S_1(f)$ occupy totally *distinct, non-overlapping* regions of the frequency spectrum. Therefore, if we wish to recover either one alone, we need only band-pass filter the total received signal to eliminate the unwanted part (Figure 79). This method of transmission, often referred to as *frequency division multiplexing*, is the "secret" behind commercial broadcasting. After all, there is only one atmosphere, and all commercial stations must share it. But when they all use different carrier frequencies, their signals can propagate together, and yet a listener can select the station they wish to listen to by simply using the tuner on their radio receiver. For that tuner is nothing more, as we have noted earlier, than a band pass filter with a movable band!

We shall now leave this extremely brief introduction to amplitude modulation, although

not because the topic is uninteresting, and certainly not because we have fully covered all of its interesting aspects. In this section, and the earlier one on filters, we have in fact barely scratched the surface of what is the central focus of signal processing - the transmission of information and its subsequent recovery from the received signal. We are stopping at this point because this is simply not a text on signal processing, and we only intended to give a brief appreciation for its flavor and for the central role the Fourier transform plays in it. This we hope we have done.
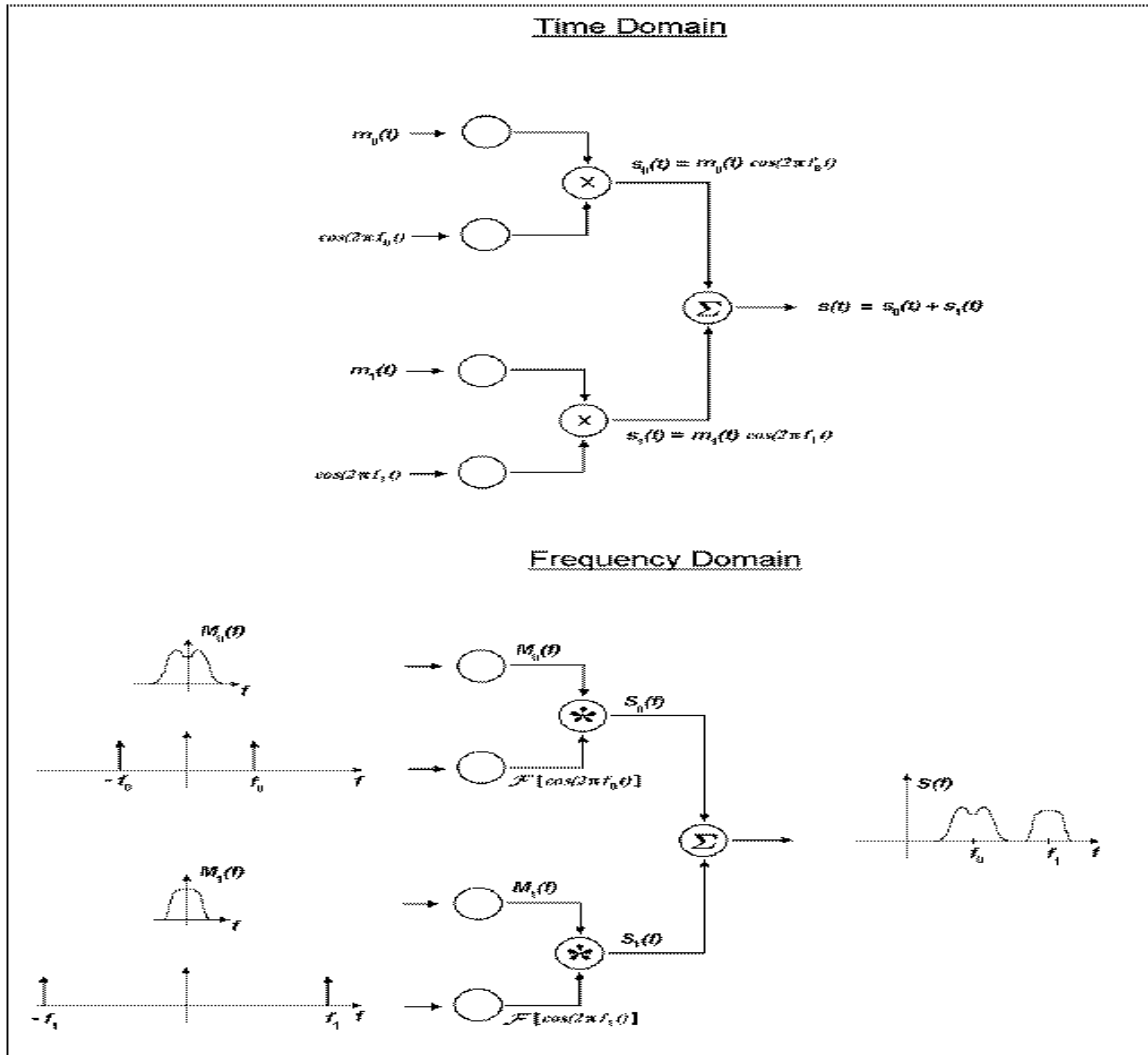
Figure 78: Frequency Division Multiplexing - The Time and Frequency Domain Views
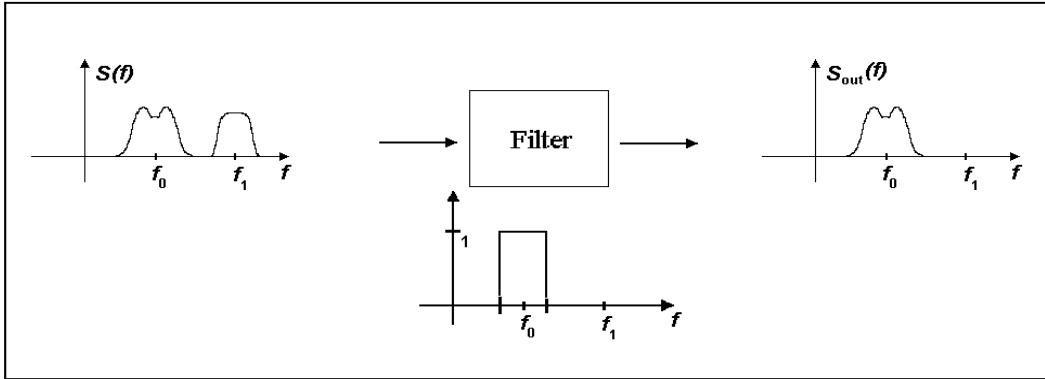
Figure 79: Recovering a Frequency Division Multiplexed Signal

6.7   The D'Alembert Solution Revisited

We now turn to the application of the Fourier transform in a totally different context than signal processing - the study of wave propagation. (Actually, we can also apply several of the ideas which we shall discuss here to image processing, which can be loosely thought of as the signal processing of photographic or other images in order to improve their quality.) The biggest single difference between this application and our earlier discussion on signal processing is that we are now going to be computing transforms with respect to a different variable - a spatial one (x) rather than time (t).

Mathematically, changing the symbol of the variable we consider makes effectively no difference. For example, if we have a function, denoted $f(x)$, we could equally well define a spatial Fourier transform by rewriting the fundamental Fourier transform relationships ((13) and (14) in Chapter 5) as

$$\tilde{F}(k) = \tilde{\mathcal{F}}\left[f(x)\right] = \int_{-\infty}^{\infty} f(x)e^{-2\pi jkx}dx \tag{6.7.7}$$

$$f(x) = \tilde{\mathcal{F}}^{-1}\left[\tilde{F}(k)\right] = \int_{-\infty}^{\infty} \tilde{F}(k)e^{2\pi jkx}dk \tag{6.7.8}$$

Formally, all this requires is using the symbol k instead of f, and x instead of t. (We also choose to write the overhead tilde here just to continually remind ourselves that this is a spatial, rather than a temporal Fourier transform. Other than serving as this reminder, the tilde is extraneous.) Moreover, with the proper change of symbols, all of the rules of Fourier transforms are still valid, e.g.

$$\tilde{\mathcal{F}}\left[\frac{df}{dx}\right] = (2\pi jk)\tilde{F}(k) \quad ,$$

etc. Physically, of course, using the spatial Fourier transform will require that several interpretations change. For example, the discussion will now often be in terms of wavelengths, rather than frequencies.

One additional important factor comes into play when the spatial Fourier transform is applied to the study of wave phenomena. This addition is necessary because, as we saw in the earlier chapters of this text, most wave phenomena involve functions of two or more independent variables, and are described by partial differential equations. Therefore we do have to consider, briefly, how to interpret the transform of a function of two variables with respect to *only one* of them, i.e. how to interpret

$$\tilde{\mathcal{F}}\left[u(x,t)\right]$$

This is actually not difficult, provided we carefully keep track of the variables involved, and remember that the spatial Fourier transform is with respect to x, not t. Specifically, by definition, we may then write

$$\tilde{\mathcal{F}}\left[u(x,t)\right] \equiv \tilde{U}_k(t) = \int_{-\infty}^{\infty} u(x,t)e^{-2\pi jkx}\ dx$$

Note that we have also slightly modified our notation here by writing the transform variable k as a subscript. We do this to emphasize that, in problems of this type, k is generally treated as a parameter rather than as an independent variable. (In other words, transforming the problem corresponds in effect to looking at how the solution evolves with respect to time on a frequency-by-frequency basis.) The result of the spatial transform of partial derivatives can be determined in the same manner we investigated the time transform of ordinary time derivatives, i.e. by starting with the definition and integrating by parts. The only difference is that we now must carefully distinguish between the transforms of partials with respect to the variable we are transforming on and those of partials of other variables, i.e. the transform of a partial with respect to x (the transform variable) behaves like

$$\tilde{\mathcal{F}}\left[\frac{\partial u}{\partial x}(x,t)\right] = \int_{-\infty}^{\infty} \frac{\partial u}{\partial x}(x,t)e^{-2\pi jkx}\ dx$$

$$= u(x,t)e^{-2\pi jkx}\Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} u(x,t)\frac{\partial}{\partial x}\left[e^{-2\pi jkx}\right]\ dx$$

$$= -\int_{-\infty}^{\infty} u(x,t)(-2\pi jk)\left[e^{-2\pi jkx}\right]\ dx \qquad (6.7.9)$$

$$= (2\pi jk)\int_{-\infty}^{\infty} u(x,t)e^{-2\pi jkx}\ dx$$

$$\equiv (2\pi jk)\tilde{U}_k(t)$$

but the transform of a partial with respect to t (which is not the transform variable) becomes

$$\tilde{\mathcal{F}}\left[\frac{\partial u}{\partial t}(x,t)\right] = \int_{-\infty}^{\infty} \frac{\partial u}{\partial t}(x,t)e^{-2\pi jkx}\ dx$$

$$= \frac{d}{dt}\left[\int_{-\infty}^{\infty} u(x,t)e^{-2\pi jkx}\ dx\right] \qquad (6.7.10)$$

$$\equiv \frac{d}{dt}\tilde{U}_k(t) = \frac{d\tilde{U}_k}{dt}$$

Of course, as with the Fourier time transform, these results can be extended to higher derivatives, e.g.

$$\tilde{\mathcal{F}}\left[\frac{\partial^2 u}{\partial x^2}(x,t)\right] = (2\pi jk)^2\tilde{U}_k(t) \qquad (6.7.11)$$

and

$$\tilde{\mathcal{F}}\left[\frac{\partial^2 u}{\partial t^2}(x,t)\right] = \frac{d^2\tilde{U}_k}{dt^2} \qquad (6.7.12)$$

With these preliminaries out of the way, we now consider the one-dimensional wave equation in a region without boundaries, with an initial displacement but no initial velocity,

i.e. the problem

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad , \quad -\infty < x < \infty \ , \ 0 < t$$

$$u(x, 0) = f(x)$$
$$\frac{\partial u}{\partial t}(x, 0) = 0$$

Of course, we have already seen that this problem can be solved by D'Alembert's principle. However, it is nevertheless instructive to also apply the spatial Fourier transform here. (Note that in terms of the time variable, this problem has the character of an initial value problem. Therefore, taking a Fourier transform with respect to time would not be sensible anyway, since initial conditions in the transform variable have no real meaning in the context of Fourier transforms. The only transform with respect to time that we could apply here would be a Laplace transform - but that's a different issue!) If we simply spatially Fourier transform both sides of the partial differential equation

$$\tilde{\mathcal{F}}\left[\frac{\partial^2 u}{\partial t^2}(x, t)\right] = \tilde{\mathcal{F}}\left[c^2 \frac{\partial^2 u}{\partial x^2}(x, t)\right] \tag{6.7.13}$$

and use (6.7.11) and (6.7.12) , we have

$$\frac{d^2 \tilde{U}_k}{dt^2} = -(2\pi ck)^2 \tilde{U}_k \ , \tag{6.7.14}$$

which is just an ordinary differential equation for $\tilde{U}_k(t)$. Since it is second-order however, we still need two initial conditions (in the *transform domain*) in order to find a unique solution. But where do these come from? The answer should be obvious. We had two initial conditions in the original (spatial) domain! So why not transform them? If we do, we obtain

$$\tilde{\mathcal{F}}[u(x, 0)] \equiv \tilde{U}_k(0) = \int_{-\infty}^{\infty} f(x)e^{-2\pi jkx} \ dx \equiv \tilde{F}_k$$

and

$$\tilde{\mathcal{F}}\left[\frac{\partial u}{\partial t}(x, 0)\right] \equiv \frac{d\tilde{U}_k}{dt}(0) = 0 \ ,$$

respectively. Combining all of these results in the single problem statement

$$\frac{d^2 \tilde{U}_k}{dt^2} = -(2\pi ck)^2 \tilde{U}_k$$

$$\tilde{U}_k(0) = \tilde{F}_k$$
$$\frac{d\tilde{U}_k}{dt}(0) = 0$$

The solution to this is easily seen to be

$$\tilde{U}_k(t) = \tilde{F}_k \cos(2\pi ckt)$$

But, of course, this is the solution in the transform domain. To find the solution in the original physical (spatial) domain, we must invert

$$u(x,t) \equiv \tilde{\mathcal{F}}^{-1}\left[\tilde{U}_k(t)\right] = \int_{-\infty}^{\infty} \tilde{U}_k(t) e^{2\pi j k x} \; dk$$

$$= \int_{-\infty}^{\infty} \tilde{F}_k \cos(2\pi c k t) e^{2\pi j k x} dk$$

But if we now apply Euler's formula to the cosine term, we have

$$u(x,t) = \int_{-\infty}^{\infty} \tilde{F}_k \left[\frac{1}{2}\left(e^{2\pi j c k t} + e^{-2\pi j k c t}\right)\right] e^{2\pi j k x} \; dk$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi j c k t} e^{2\pi j k x} \; dk + \frac{1}{2}\int_{-\infty}^{\infty} \tilde{F}_k e^{-2\pi j k c t} e^{2\pi j k x} \; dk \qquad (6.7.15)$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi j k (x+ct)} \; dk + \frac{1}{2}\int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi j k (x-ct)} \; dk$$

However, according to our formulation for the inverse transform (6.7.8),

$$f(x) = \int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi j k x} \; dk$$

Comparing this last equation to the two integrals in the previous one, we see that all three are identical, except that the two integrals comprising $u(x,t)$ are evaluated not at x, but at $(x+ct)$ and $(x-ct)$ respectively, i.e.

$$u(x,t) = \frac{1}{2}f(x+ct) + \frac{1}{2}f(x-ct)$$

This, of course, is just D'Alembert's principle. What should be interesting about the above development is the manner in which the spatial Fourier transform permits us to *derive* the principle and so better understand how it arises. Knowing how a formula arises is generally far preferable to simply stating the formula - as if it arose by magic. But, as we shall see in the next section, the utility of the spatial Fourier transform is not just limited to this derivation.

6.8   Dispersive Waves

In addition to providing a derivation of D'Alembert's principle, the spatial Fourier transform also allows us to study other wave propagation problems to which D'Alembert's principle simply does not even apply. For example, consider the problem

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

$$u(x,0) = f(x)$$
$$\frac{\partial u}{\partial t}(x,0) = 0$$

$$(6.8.16)$$

We shall not dwell on the physics that introduce the additional $(-u)$ term here. However there are real physical models, such as transmission lines, for which this equation is a model. Our intent is solely to demonstrate the information the spatial Fourier transform can provide about the solutions to this problem, and how these solutions may differ from those of the "pure" one-dimensional wave equation.

It is fairly easily shown that neither of the functions $F(x - ct)$ nor $G(x + ct)$ satisfy the partial differential equation in (6.8.16). Therefore, for this problem, a D'Alembert-type solution is not feasible. However, because this partial differential equation is linear and constant coefficient, transform methods still are. Therefore we shall simply proceed to directly transform the partial differential equations and boundary conditions. If we do this in a similar manner to our derivation of (6.7.13)-(6.7.13), and the associated initial conditions, we find

$$\frac{d^2 \tilde{U}_k}{dt^2} = -\left[1 + (2\pi ck)^2\right] \tilde{U}_k$$

$$\tilde{U}_k(0) = \tilde{F}_k$$
$$\frac{d\tilde{U}_k}{dt}(0) = 0$$

The solution to this may be written in the following form

$$\tilde{U}_k(t) = \tilde{F}_k \cos(2\pi ck\nu_k t) \quad , \tag{6.8.17}$$

where we define

$$\nu_k = \sqrt{1 + \frac{1}{(2\pi ck)^2}} \tag{6.8.18}$$

(The reason introducing ν_k is most likely not transparent at this point. However, as we shall see, it simplifies the following analysis somewhat.) Again, as in the previous problem, we recognize that (6.8.17) represents the solution in the transform domain, and so must be inverted. With the exception of the addition of the ν_k term, this process is identical to the previous inversion, and can be shown to lead to

$$u(x, t) = \frac{1}{2} \int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi jk(x + \nu_k ct)} \, dk + \frac{1}{2} \int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi jk(x - \nu_k ct)} \, dk$$

Here however, we cannot simplify the result as we did in the D'Alembert problem. The difficulty is the ν_k term, which makes neither integral a function of a simple expression like $(x - ct)$. (Of course, since k is the variable of integration, we cannot take ν_k outside the integral!) However, we can still obtain some insights. Specifically, consider the second integral

$$\frac{1}{2} \int_{-\infty}^{\infty} \tilde{F}_k e^{2\pi jk(x - \nu_k ct)} \, dk$$

Since this integral arose from an inverse transform, we must view it as the analog of the Fourier series, i.e. as the attempt to reconstruct the physical function by combining its

components over all (spatial) frequencies (or equivalently over all wavelengths). Any one of these particular components is described by

$$\frac{1}{2}\tilde{F}_k e^{2\pi jk(x-\nu_k ct)}$$

for a fixed value of k. This expression clearly represents a wave, of a single wavelength, and amplitude $\frac{1}{2}\tilde{F}_k$, moving to right, **but** with velocity $c\nu_k$, **not** c. (A similar analysis and conclusion can be made on the other integral, except of course the wave there moves to the left.) But ν_k depends on k. In other words, the frequency components that made up the initial displacement ($f(x)$) propagate at *different speeds*. Therefore, as time evolves, they will become progressively more "out of step," or disperse. (For this reason, this problem is commonly called the dispersive wave equation.) Thus, while we apparently still have propagation in this case, as time passes the dispersion will cause an increasing distortion of the shape of the propagating disturbance relative to the original displacement. By contrast, in the D'Alembert inverse (6.7.15), all components move at velocity c, and therefore remain "in step," causing the original displacement not to lose shape as it propagates. Figure 80 displays the graphs of the solution to a dispersive wave equation, computed by numerically evaluating the integrals involved. The important point of this figure is that the computed integrals substantiate our intuitive expectation, an expectation which came directly from Fourier transform analysis.
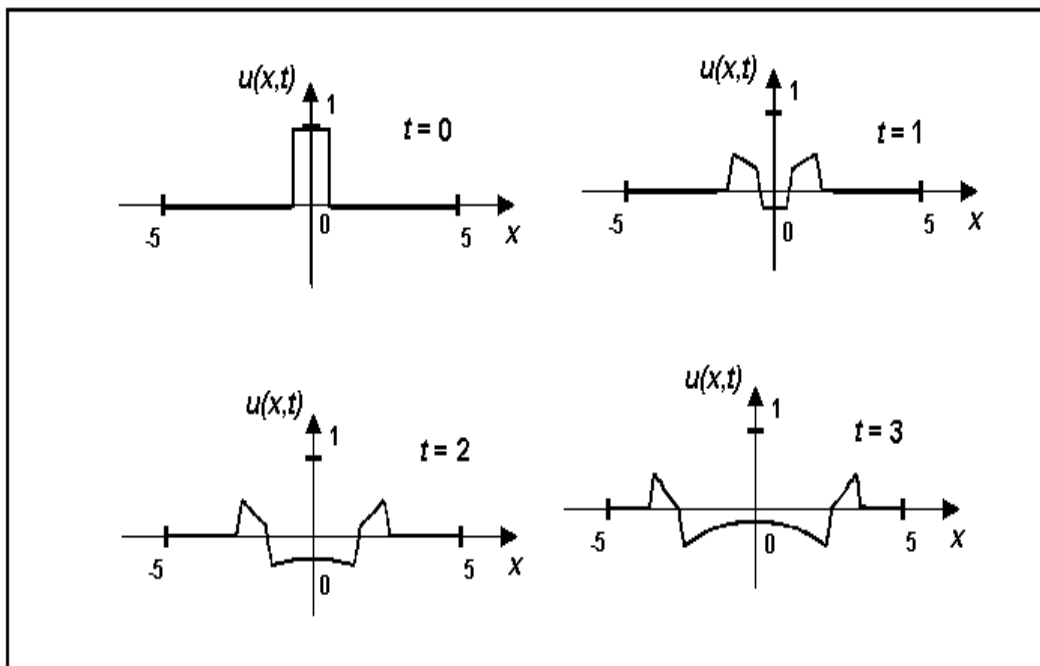
Figure 80: Solutions of a Dispersive Wave Equation at Different Times

With this section, we close our brief introduction to the spatial Fourier transform. As with our discussion of signal processing, this discussion was not intended to be either in-depth or all-inclusive. On the contrary, our only intention was to provide a brief flavor of

247

another instance where Fourier methods provide powerful insights into the workings of the physical world. As with our earlier introduction to signal processing, we hope these short sections have accomplished this for the spatial Fourier transform.

6.9   Correlation

There is one further time-domain operation that is commonly associated with the Fourier transform. This is the so-called correlation, which is defined as

$$x(t) \star y(t) \equiv \int_{-\infty}^{\infty} x(\tau)y(t+\tau) \ d\tau \tag{6.9.19}$$

In form, correlation looks quite close to convolution, except correlation uses the term $(t+\tau)$ in the integrand, vice $(t-\tau)$ in convolution. On the other hand, the uses and interpretations of correlation and convolution are widely different. Correlation is most often used to measure the similarity between signals, especially when one has been contaminated by noise. Because of the forms of convolution and correlation look quite close, the Fourier transform of a correlation produces a very similar, though not identical, result to the convolution theorem. For the sake of completeness, we shall present that result here. We start, as usual, with the formal definition

$$\mathcal{F}\left[x(t) \star y(t)\right] \equiv \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x(\tau)y(t+\tau) \ d\tau \right\} e^{-2\pi jft} \ dt$$

As with the convolution theorem, because this is a two-dimensional integral, we may at least formally interchange the order of integration, yielding

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x(\tau)y(t+\tau)e^{-2\pi jft} \ dt \right\} d\tau$$

But then, because the inner integral does not depend on τ, we may move $x(\tau)$ outside of the inner integral, yielding

$$= \int_{-\infty}^{\infty} x(\tau) \left\{ \int_{-\infty}^{\infty} y(t+\tau)e^{-2\pi jft} \ dt \right\} d\tau$$

and since the inner integral now corresponds to the Fourier transform of $y(t)$, shifted by $(-\tau)$, we may use the shift rule to write

$$= \int_{-\infty}^{\infty} x(\tau)Y(f)e^{-2\pi jf(-\tau)} \ d\tau$$

But since $Y(f)$ does not depend on the variable of integration, we may now take it completely outside the integral, and write

$$= Y(f) \int_{-\infty}^{\infty} x(\tau)e^{-2\pi j(-f)\tau} \ d\tau$$

However, by definition, the remaining integral is simply the Fourier transform of $x(t)$, but evaluated at $(-f)$, i.e.

$$\mathcal{F}\left[x(t) \star y(t)\right] = Y(f)X(-f) \qquad (6.9.20)$$

(It's fairly important to note here that, unlike the convolution, correlation is generally **not** commutative, i.e. in general

$$x(t) \star y(t) \neq y(t) \star x(t)$$

This is reflected in (6.9.20) because, in general

$$Y(f)X(-f) \neq X(f)Y(-f)$$

Dealing with transforms of correlations is thus quite a bit more delicate than dealing with the transforms of convolutions. Regrettably, we shall not pursue this any further at this point.)

Beyond presenting this result (6.9.20), we shall not delve further into the uses of correlation. Nevertheless, it does represent one more operation which, in the proper context, may be better understood in the frequency domain, via the Fourier transform, than in the time domain.

6.10   Summary

In this chapter, we have briefly covered some of the classic applications of the Fourier Transform - signal processing, modulation, and wave propagation. As we noted several times, our coverage of each was very abbreviated, and only designed to show how the transform produces insights that might be very difficult to arrive at by any other means. The key point here is the ability with the transform, as with the Fourier series, to provide two views of a physical phenomenon - one in the time domain, the other in the frequency domain. Depending on which aspect of a particular phenomenon is of primary interest will usually determine which domain is the most appropriate to study the problem in.

PROBLEMS

1. Consider the linear, shift-invariant system represented by the following circuit
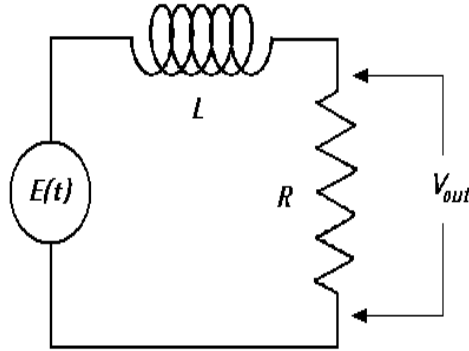
$$L\frac{dI}{dt} + RI = E(t)$$
$$V_{out} = RI$$

a. Directly determine, e.g. using the Laplace transform, the impulse response of this system. Sketch this response.

b. (1) Find the Transfer function of this system by computing the Fourier transform of the impulse response determined in part a. above.

b. (2) Show that the alternative method of finding the Transfer function, i.e. as the response of a system to the forcing function $e^{2\pi j f_0 t}$, produces the same result as in part (1),

c. Sketch the amplitude and phase spectra of the Transfer function computed in part b.

2. Repeat problem 1 for the same circuit, except with the output taken as the voltage across the inductor, i.e.
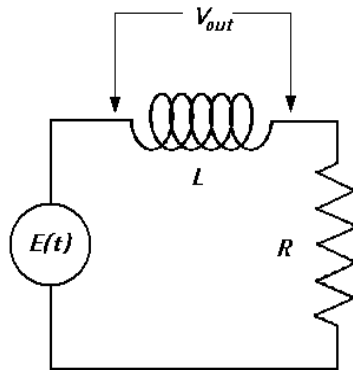
$$L\frac{dI}{dt} + RI = E(t)$$
$$V_{out} = L\frac{dI}{dt}$$

7 Appendix A - Bessel's Equation

7.1 Bessel's Equation

By definition, *Bessel's Equation of Order n*, commonly just called Bessel's Equation, is the ordinary differential equation

$$x^2 y'' + xy' + (x^2 - n^2)y = 0 \quad . \tag{7.1.1}$$

Full details of the solution of this equation are found in most standard ordinary differential texts, such as Boyce and DiPrima. We shall not repeat those in detail here, but only summarize the highlights. We start by observing that a cursory visual inspection of Bessel's equation reveals two primary attributes:

1. It has variable coefficients, and

2. It has a singular point at $x = 0$.

In standard ordinary differential terminology, the origin is a regular singular point for Bessel's equation. Because of this singularity, the standard solution technique (the so-called method of Frobenius[15]) is to assume a solution in the form of a general power series, i.e.

$$\sum_{k=0}^{\infty} a_k x^{k+p} \quad , \tag{7.1.2}$$

where a_0 is assumed to be non-zero, and the remainder of the a_n, as well as the value of p, are to be determined. Notice that with the exponent as written, the first term of this series is $a_0 x^p$. Therefore there is no real loss of generality in the assumption $a_0 \neq 0$, since p is not determined beforehand. This power series is, of course, just an assumed solution form. Whether or not any solutions in this form actually exist, and, if so, exactly how many, depends on what develops when one substitutes the series into the differential equation. This last comment is quite relevant for this equation, since, as you may recall, one "quirk" of regular singular point problems is that while one linearly independent solution will always exist in form (7.1.2), a second one need not. This is in contrast the behavior of solutions in the neighborhood of ordinary points, where two linearly independent solutions can always be found in terms of Taylor series.

As noted above, we shall not present the full solution of Bessel's equation here, but refer the interested reader to any standard ordinary differential equations text for full details. We simply assert that the conclusion after substituting (7.1.2) into the differential equation is that **only one** linearly independent power series solution exists, and that this solution can be expressed as

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{x}{2}\right)^{2k+n} \quad .$$

---

[15]Ferdinand Georg Frobenius,
see: http://turnbull.mcs.st-and.ac.uk/history/Mathematicians/Frobenius.html

The function represented by this series is conventionally referred to as the *Ordinary Bessel Function of the First Kind of Order n*, or more commonly as simply the Bessel Function of Order n. It is virtually universally represented by the symbol $J_n(x)$, i.e.

$$J_n(x) \equiv \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{x}{2}\right)^{2k+n} \quad . \tag{7.1.3}$$

Of course, like any second order linear equation, Bessel's equation requires two linearly independent solutions in order to construct the general solution. However, according the discussion just completed, any solution which is independent of $J_n(x)$ cannot be written as a "pure" power series, and therefore must be determined by some procedure other than the method of Froebenius. Fortunately, there are several such techniques available for determining a second, linearly independent solution, to an ordinary differential equation, once a first one is known. One general method is *reduction of order*. A second option, more tailored to this specific type of problem, involves careful differentiation of the series solution with respect to the parameter p. Either of these methods could be applied here to generate a solution which is linearly independent of $J_n(x)$.

As before, we omit any details of this actual construction of this second linearly independent solution to Bessel's equation. (Again, these can be found in standard texts. Before continuing, however, one slight word of caution, is in order. You should recall that there is no such thing as a *unique* set of linearly independent solutions to any second-order differential equations. For example, we could represent the solution to

$$y'' - y = 0$$

equally well as either

$$y(x) = c_1 e^x + c_2 e^{-x}$$

or

$$y(x) = d_1 e^x + d_2 \cosh(x) \quad .$$

Therefore, occasionally, other texts or papers may use different formulas or symbols than the ones we shall present. However our choice and notation are extremely common.)

Probably the most frequently used choice for the second solution to Bessel's equation is the function almost universally denoted by $Y_n(x)$, which has the rather imposing definition

$$Y_n(x) \;\; = \; -\frac{1}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!} \left(\frac{x}{2}\right)^{2k-n} + \frac{2}{\pi} \ln\left(\frac{x}{2}\right) J_n(x)$$

$$\tag{7.1.4}$$

$$-\frac{1}{\pi} \sum_{k=0}^{\infty} \{\psi(k+1) + \psi(n+k+1)\} \frac{(-1)^n}{k!(n+k)!} \left(\frac{x}{2}\right)^{2k+n}$$

where

$$\psi(1) \;\; = \; -\gamma$$

$$\psi(k) \;\; = \; -\gamma + \sum_{i=1}^{k-1} \frac{1}{i}$$

252

and γ is a constant, commonly called *Euler's*[16] *constant.* Like π, the value of Euler's constant cannot be determined exactly, but is approximately $\gamma = 0.57721566\ldots$. With the caveat then that our notation is almost, but not quite, universal, we finally arrive at the conclusion that the general solution to (7.1.1) may be represented

$$y(x) = c_1 J_n(x) + c_2 Y_n(x) \quad .$$

7.2   Properties of Bessel Functions

Bessel functions occur in a wide variety of applications. Because of their frequent appearance, they have been studied in great detail for well over a hundred years. These studies have uncovered many quite interesting properties and relationships, some of the more important of which we shall now briefly consider. (More comprehensive listings of these may be found in standard reference texts such as Abramowitz and Stegun, or Spanier and Oldham.)

At least initially, both the qualitative and quantitative properties of Bessel functions are obscured by the infinite series involved in their definitions ((7.1.3) and (7.1.4)). However, in one sense, these series are not as "bad" as they might be. Both series can be shown to converge at all values of x (by the ratio test), and both are *alternating term* series, i.e. they change sign every term. Furthermore, for "large" k, the denominators grow approximately as $(k!)^2$. Therefore, for most "reasonable" values of x, these series should converge fairly rapidly, and, because the terms alternate, the errors in approximating these series by partial sums are reasonably easy to monitor and control. Hence, especially given the power of modern computers, computing and graphing both $J_n(x)$ and $Y_n(x)$ then becomes fairly straightforward. Figures 81, display the graphs of $J_n(x)$ (top) and $Y_n(x)$ (bottom) for $n = 0, 1, 2, 3$.

Several qualitative properties of Bessel functions may be directly inferred from these graphs. Both types of Bessel functions oscillate and appear to resemble decaying trigonometric functions for large values of x. Each has an infinite number of axis crossings (roots or zeros), a property they share with the sine and cosine. (This property will be important for our purposes since the eigenvalues of any particular problem are always related to such axis crossings.) These graphs, however, also highlight one major qualitative distinction in the behavior of the $J_n(x)$ as contrasted to the $Y_n(x)$. This is that all of the $Y_n(x)$ appear to be singular at the origin, while $J_0(x)$ approaches unity at the origin and all of the other $J_n(x)$ appear to go to zero there. (These last two statements can be immediately verified from the series representation (7.1.3).)

Another intriguing and important property of Bessel functions can also be derived directly from the series representations, although the algebra involved is non-trivial. This property is the so-called *recurrence formulas*, which establish certain relationships between Bessel

---

[16]Leonhard Euler,
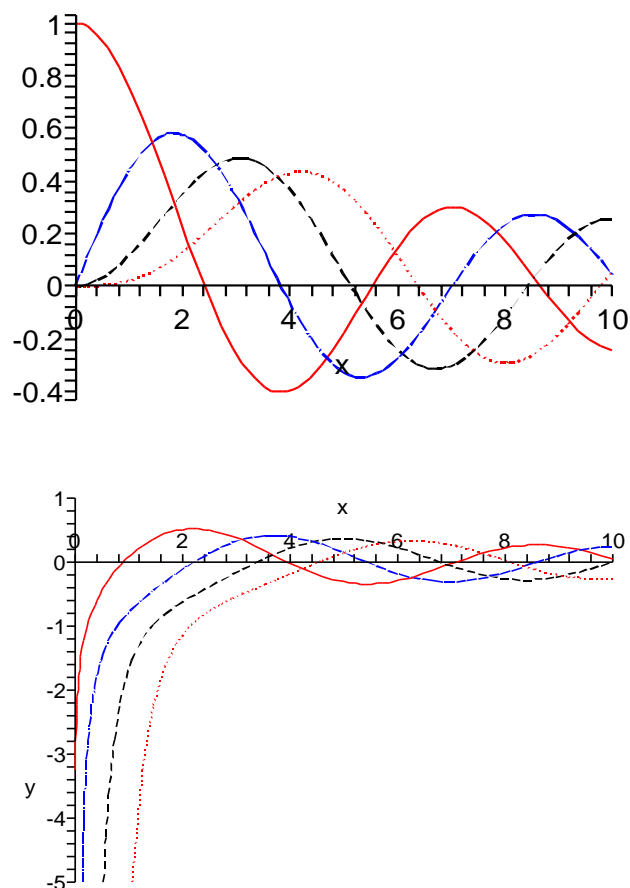see: http://www-gap.dcs.st-and.ac.uk/ history/Mathematicians/Euler.html

Figure 81: The Bessel Functions J_n (a) and Y_n (b).

functions of different orders and their derivatives. The most important of these are

$$nJ_n(x) + xJ'_n(x) = xJ_{n-1}(x) \qquad\qquad nY_n(x) + xY'_n(x) = xY_{n-1}(x)$$

and

$$nJ_n(x) - xJ'_n(x) = xJ_{n+1}(x) \qquad\qquad nY_n(x) - xY'_n(x) = xY_{n+1}(x) \qquad .$$

These equations can further be combined to yield others, perhaps the most useful of which comes from simply adding and then regrouping terms to yield

$$J_{n+1}(x) = \frac{2n}{x}J_n(x) - J_{n-1}(x) \qquad \text{and} \qquad Y_{n+1}(x) = \frac{2n}{x}Y_n(x) - Y_{n-1}(x) \qquad .$$

The value of these last two formulas is that their repeated application leads to the conclusion that determining the value of $J_n(x)$ (or $Y_n(x)$) at a particular value of x, for **any** n, requires only a table of values (or a computer program) that accurately gives $J_0(x)$ and $J_1(x)$ (or $Y_0(x)$ and $Y_1(x)$) at that same point. For example, evaluating the last recurrence relation involving

254

the ordinary Bessel functions for both $n = 2$ and $n = 1$ yields

$$J_3(x) = \frac{4}{x} J_2(x) - J_1(x)$$

$$J_2(x) = \frac{2}{x} J_1(x) - J_0(x)$$

,

which can be simplified (by eliminating $J_2(x)$) to

$$J_3(x) = \left(\frac{8}{x^2} - 1\right) J_1(x) - \frac{4}{x} J_0(x) \quad .$$

Unfortunately, from the point of view of realistic computations, there is at least one "minor" point we have glossed over. As we noted above, the series involved in the Bessel functions converge at all values of x. However, the presence of the x^{2k+n} term in these series does cause them to converge more slowly as the value of x increases. As a practical matter then, these series are useful primarily when x is small. Therefore, mathematicians have devoted significant efforts to develop simple expressions, commonly called *asymptotic formulas*, that give approximate values for the various Bessel functions - values that become more and more accurate the larger x is. Perhaps the most widely used of these approximations are

$$J_n(x) \approx \sqrt{\frac{2}{\pi x}} \, \cos\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right) \quad , \qquad x \to \infty \qquad (7.2.5)$$

and

$$Y_n(x) \approx \sqrt{\frac{2}{\pi x}} \, \sin\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right) \quad , \qquad x \to \infty \qquad . \qquad (7.2.6)$$

Note that both these two asymptotics confirm our earlier observation that the Bessel functions behave somewhat like decaying sinusoids. (They are however, and this is important, not equal to decaying sinusoids, since (7.2.5) and (7.2.6) are only approximate formulas.)

k	$J_0(x)$	$J_1(x)$	$J_2(x)$	$J_3(x)$
1	2.4048	3.8317	5.1356	6.3802
2	5.5201	7.0156	8.4172	9.7610
3	8.6537	10.1735	11.6198	13.0152
4	11.7915	13.3237	14.7960	16.2235
5	14.9309	16.4706	17.9598	19.4094

Table 1: Roots of the Ordinary Bessel Functions of the First Kind - J_n

Before closing this section, there is one final important point we must make about the qualitative behavior of Bessel functions. As our earlier discussion and the asymptotic formulas (7.2.5)-(7.2.6) have indicated, the Bessel functions, like the sine and cosine, have an

k	$Y_0(x)$	$Y_1(x)$	$Y_2(x)$	$Y_3(x)$
1	0.8936	2.1971	3.3842	4.5270
2	3.9577	5.4297	6.7938	8.0976
3	7.0861	8.5960	10.0235	11.3965
4	10.2223	11.7492	13.2100	14.6231
5	13.3611	14.8974	16.3790	17.8185

Table 2: Roots of the Ordinary Bessel Functions of the Second Kind - Y_n

infinite number of distinct zeros. Nevertheless, there is a fundamental difference in the structure of these zeros as opposed to the structure the zeros of trigonometric functions. This difference is that the zeros of the Bessel functions are **not** evenly spaced along the axis, and there is no convenient formula, such as integer multiples of π, that exactly describes them. The only way to express them is with a list or table. For example, Table 1 lists the first few zeros of the first three $J_n(x)$ functions, while Table 7.2 lists the first few zeros of the first three $Y_n(x)$ functions. (In most practical applications, the zeros of the $Y_n(x)$ are generally not used as much as the zeros of the $J_n(x)$). However, the situation is not quite as bleak as the above discussion might at first seem. This is because the asymptotic formulas imply that for large x the kth zero (large k) of $J_n(x)$ will approach $\left(\frac{n}{2} + k - \frac{1}{4}\right)\pi$, while the kth zero of $Y_n(x)$ will approach $\left(\frac{n}{2} + k - \frac{3}{4}\right)\pi$. Therefore, once one decides with how much accuracy they need to know the roots of a Bessel function, they can simply use tables until the asymptotic value agrees with the table value to the desired accuracy.)

7.3 Variants of Bessel's Equation

Bessel functions appear frequently in applications, and especially in the solution of partial differential equations in cylindrical coordinates. In addition (and sometimes during the solution of these same partial differential equations), several variants of Bessel's equation very frequently occur. Perhaps the most common of these variants is

$$x^2 y'' + x y' + (\xi^2 x^2 - n^2)y = 0 \quad . \tag{7.3.7}$$

In this equation, the change of variables $s = \xi x$, including the application of the chain rule, e.g. $\frac{dy}{dx} = \frac{dy}{ds}\frac{ds}{dx} = \xi \frac{dy}{ds}$, reduces the differential equation to

$$s^2 \frac{d^2 y}{ds^2} + s \frac{dy}{ds} + (s^2 - n^2)y = 0 \quad ,$$

which is clearly just Bessel's equation of order n, but with s as the independent variable instead of x. Therefore, the general solution to (7.3.7) can be represented

$$y(x) = c_1 J_n(\xi x) + c_2 Y_n(\xi x) \quad .$$

256

(It may help to view this last result as analogous to the fact that the general solution to

$$y'' + y = 0 \qquad \text{is} \qquad y(x) = c_1 \sin(x) + c_2 \cos(x) \ ,$$

while the general solution to

$$y'' + \xi^2 y = 0 \qquad \text{is} \qquad y(x) = c_1 \sin(\xi x) + c_2 \cos(\xi x) \ .)$$

A more impressive variant of Bessel's equation is the general form

$$x^2 y'' + \left[a + 2bx^p \right] xy' + \left[c + dx^{2q} + b(a + p - 1)x^p + b^2 x^{2p} \right] y = 0$$

where a, b, c, d, p and q are constants, which can be reduced to Bessel's equation by a fairly tedious change of variables. We shall omit the details here, and simply present the resulting general solution

$$y(x) = x^\alpha e^{-\beta x^p} \left[c_1 J_\nu \left(\lambda x^q \right) + c_2 Y_\nu \left(\lambda x^q \right) \right] \ ,$$

where

$$\alpha = \frac{1 - a}{2} \ , \qquad \beta = \frac{b}{p}, \qquad \lambda = \frac{\sqrt{d}}{q} \ , \qquad \text{and} \quad \nu = \frac{\sqrt{(1 - a)^2 - 4c}}{2q} \ .$$

One last fairly common variant of Bessel's equation that we shall mention is the so-called *modified Bessel's equation*

$$x^2 y'' + xy' - (x^2 + n^2)y = 0 \ . \tag{7.3.8}$$

This equation can actually be viewed as nothing more than the special case (7.3.7) which occurs when $\xi = j = \sqrt{-1}$. Its solution could therefore be written

$$y(x) = c_1 J_n(jx) + c_2 Y_n(jx) \ .$$

This however, would be as awkward as writing the solution to

$$y'' - y = 0 \qquad \text{as} \qquad y(x) = c_1 \sin(jx) + c_2 \cos(jx)$$

instead of

$$y(x) = d_1 e^x + d_2 e^{-x} \ .$$

Therefore, the conventional way to represent the solution to (7.3.8) is by introducing two additional Bessel functions, denoted by $I_n(x)$ and $K_n(x)$, and commonly referred to as the modified Bessel functions of the first and second kinds, respectively. The general solution to (7.3.8) is then represented

$$y(x) = c_1 I_n(x) + c_2 K_n(x) \ .$$

(Similarly, the solution to

$$x^2 y'' + xy' - (\xi^2 x^2 + n^2)y = 0$$

is then represented

$$y(x) = c_1 I_n(\xi x) + c_2 K_n(\xi x) \ .)$$

We shall not encounter the modified Bessel functions in this text, and simply refer the interested reader to Abramowitz and Stegun, or Spanier and Oldham for a complete discussion.

PROBLEMS

Use Bessel Functions (the big ugly equation) to find the *general solution* to each ODE below.

1.
$$x^2 y'' + 3\, xy' + \left(-3 + 4\, x^4\right) y = 0$$

2.
$$x^2 y'' + \left(x + 2\, x^2\right) y' + \left(-4 + 9\, x^{-2} + x + x^2\right) y = 0$$

3.
$$x^2 y'' - 5\, xy' + \left(9 + 4\, x^2\right) y = 0$$

4.
$$x^2 y'' + \left(x - 2\, x^3\right) y' + \left(-1/4 + x^{-2} - 2\, x^2 + x^4\right) y = 0$$

Answers

1.
$$y(x) = A\, J_1(x^2)/x + B\, Y_1(x^2)/x$$

2.
$$y(x) = e^{-x}\, \left(A\, J_2(3/x) + B\, Y_2(3/x)\right)$$

3.
$$y(x) = x^3\, \left(A\, J_0(2x) + B\, Y_0(2x)\right)$$

4.
$$y(x) = e^{x^2/2}\, \left(A\, J_{1/2}(1/x) + B\, Y_{1/2}(1/x)\right)$$

References

Abramowitz, M., and Stegun, I., *Handbook of Mathematical Functions*, Dover Pub. Co. New York, 1965.

Boyce, W. E. and DiPrima, R. C., *Elementary Differential Equations and Boundary Value Problems*, Fifth Edition, John Wiley & Sons, New York, 1992.

Spanier, J. and Oldham, K. B., *An Atlas of Functions*, Hemisphere Pub. Co., New York, 1987.

Index