Linear Methods of Applied Mathematics

Evans M. Harrell II and James V. Herod

(Some remarks for the instructor).

version of 12 August 1997.

Click here to switch to the HTML version of this chapter.

Some of the calculations of this chapter are available in a Maple worksheet or in a Mathematica notebook.

(Samples of the Maple worksheet and the Mathematica notebook are available here in a format which does not require either Maple or Mathematica.)

I. Linearity.

This course is concerned with some of the most important methods of applied mathematics, without which many technological developments of the nineteenth and twentieth centuries would have been impossible. When you have finished this course you will be able to solve most of the partial differential equations and integral equations that you are likely to encounter in engineering, and, just as importantly, you will understand what to expect from the solutions, which might emerge in a mysterious fashion from a computer at your workplace. You will also learn techniques that are useful in other contexts - Fourier series, special functions, and iteration.

The leitmotifs are linearity and orthogonality. By this stage in your education, you are familiar with vectors. You have added them by the head-to-tail construction, or by adding their components. You also know how to take their scalar (dot) product, producing a number equal to the product of the two lengths and the cosine of the angle between them:

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\,|\mathbf{w}| \cos(\theta)$$

One of the most powerful and astonishing ideas in modern mathematics is that in many regards functions can be treated just like vectors. There is a very close analogy between

sets of functions and spaces of vectors, which will guide us to striking solutions for several of the most important differential equations you will use in science and engineering, especially the wave equation, the heat equation, and the potential equation.

Mathematicians have a way of making analogies without the mushiness that so often afflicts analogies in so many other contexts. It is called *abstraction.* While abstraction may sometimes seem like theory divorced from reality, when used properly it is a wonderful, and very practical, tool for solving problems.

This is why we will begin by making vectors abstract. Let us recall some facts about everyday vectors. If we have a bag full of vectors, we can scale them and add them together. We call the result a *linear combination*:

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4$$

We'll normally use Greek letters for scalars (= ordinary real or complex numbers). It doesn't matter how many are in the combination, but unless we explicitly state otherwise, we will assume that it is only a finite sum, a *finite linear combination*. Of course, we can make linear combinations of functions, too:

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \ldots + \alpha_n f_n(x)$$

and the result is another function. In this way, the set of all functions is a vector space.

Definition I.1. More formally, a *vector space over the complex numbers* is a set of entities, abstractly called vectors, for which

1. Any finite linear combination of vectors is a member of the same set
2. The usual commutative rule holds for addition: $v + w = w + v$,
3. Just for consistency, the usual commutative, associative, and distributive laws hold for vector addition and scalar multiplication. In other words

$\alpha v + \alpha w = \alpha (v+w)$,
$\alpha v + \beta v = (\alpha+\beta)v$,
$(\alpha \beta) v = \alpha(\beta v)$.
Also, $1 v = v$.

In practice the rules in 3 are obvious and not very interesting. From these rules, you can show some other properties, such as:

There is a special element, which will be called the zero vector, equal to the scalar 0 times any vector whatsoever. It has the property that for any vector v, 0 + v = v.

For any vector v, there is a negative vector, called -v, such that v + (-v) = 0 (the zero vector).

Certain pretty reasonable conventions will be made, such as writing v - w instead of v + (-w). Great - all the usual stuff works, right? Well, not quite. We *don't* assume abstractly that we can multiply vectors by one another. Things like the dot product and the cross product work for some vector spaces but not for others. The other deep issue lurking in the shadows is infinite linear combinations, i.e., infinite series. Fourier liked them, and we'll see lots of them later. But if the vectors are functions, perhaps the infinite series converges for some x and not for others. For instance,

$$\sum_{n=1}^{\infty} \sin(n\,x) = 0$$

when x is a multiple of π, but it is certainly an improper sum when $x = \pi/2$, and it is not immediately clear what happens for other values of x. What does the series mean then? (We shall address this issue in Chapter III).

Examples I.2.

1. The usual two-dimensional or three-dimensional vectors. It makes little difference whether they are thought of as column vectors such as

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

or as row vectors (2,3). The set of all such vectors will be denoted C^2 or C^3 (assuming complex entries are allowed - otherwise R^2 or respectively R^3). Follow these links to review the usual vector operations in Mathematica or Maple.

2. The set of complex numbers. Here there is no difference between a vector and a scalar, and you can check all the properties pretty easily. We call this a one-dimensional vector space.

3. The set C^n of n numbers in a list. These are manipulated just like 2- or 3- vectors, except that the number of components is some other fixed number, n. For instance, with C^4, we might have elements such as (1,2,3,4) and (1,0,-1,-2), which can be added and multiplied as follows:

$$(1,2,3,4) + (1,0,-1,-2) = (2, 2, 2, 2)$$
$$(1,2,3,4) \cdot (1,0,-1,-2) = -10,$$

etc.

4. The set of continuous functions of a variable x, $0 < x < 1$. The rather stupid function $f_0(x) = 0$ for all x plays the rôle of the zero element.

5. A smaller vector space of functions. Instead of simply listing n numbers, let us multiply them to define another vector space of functions. For example, for some fixed n, consider

$$a_1 \sin(x) + a_2 \sin(2x) + \ldots + a_n \sin(nx),$$

where the numbers a_k can take on any value. Notice that this vector space is a part of the one of example 4. In other words, it is a *subspace* .

6. The set of, say, 2×3 matrices. Addition means addition component-wise:

$$\begin{bmatrix} 1 & 0 & i \\ -1 & \pi & 2+i \end{bmatrix} + \begin{bmatrix} 2 & -i & 0 \\ -1 & -\pi & 2-i \end{bmatrix} = \begin{bmatrix} 3 & -i & i \\ -2 & 0 & 4 \end{bmatrix}$$

and scalar multiplication affects all components:

$$5\begin{bmatrix} 1 & 0 & i \\ -1 & \pi & 2+i \end{bmatrix} = \begin{bmatrix} 5 & 0 & 5i \\ -5 & 5\pi & 10+5i \end{bmatrix}$$

7. The set of 3-component vectors (x,y,z) such that $x - 2y + z = 0$. This is a plane through the origin.

Definitions I.3. A set of vectors $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ is *linearly independent* if it is impossible to write any of the vectors in the set as a linear combination of the rest. *Linearly dependent* is the opposite notion. The *dimension* of a vector space V is the largest number n of linearly independent vectors $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ which are contained in V. The *span* of a set $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ is the vector space V obtained by considering all possible linear combinations

from the set. We also say that $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ *spans* V. A set $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ is a *basis* for a finite-dimensional vector space V if $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ is linearly independent and spans V.

Notice that the only way that two vectors can be linearly dependent is for them to be proportional (parallel) or for one of them to be $\mathbf{0}$. If the set has the same number of vectors as each vector has components, which frequently is the case, then there is a calculation to test for linear dependence. Array the vectors in a square matrix and calculate its determinant. If the determinant is 0, they are dependent, and otherwise they are independent. For example, consider the vectors {(1,2,3), (4,5,6),(7,8,9)}, which are not obviously linearly dependent. A calculation shows that

$$\det\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 0.$$

Indeed, we can solve for one of these vectors as a linear combination of the others:

$$(7,8,9) = (-1)\ (1,2,3) + 2\ (4,5,6)$$

Many of our vector spaces will be infinite-dimensional. For example, {sin(n x)} for n = 1, 2, ..., is an infinite, linearly independent set of continuous functions - there is no way to write sin(400 x), for example, as a linear combination of sine functions with lower frequencies, so you can see that each time we introduce a sine function of a higher frequency into the list, it is independent of the sines that were already there. An infinite linearly independent set is a basis for V is every element of V is a limit of finite linear combinations of the set - but we shall have to say more later about such limits. A vector space has lots of different bases, but all bases contain the same number of items.

For practical purposes, the dimension of a set is the number of degrees of freedom, i.e., the number of parameters it takes to describe the set. For example, C^n has dimension n, and the set of 2×3 matrices has six elements, so its dimension is 6.

Model problem I.4. Show that the plane x + y + z = 0 is a two-dimensional vector space.

The verification of the vector space properties will be left to the reader. Here is how to show that the dimension is 2:

Solution. The solution can't be 3, since there are vectors in R^3 which are not in the plane, such as (1,1,1). On the other hand, here are two independent vectors in the plane: (1,-1,0) and (1,1,-2).

- -

Further observations. The general vector in the plane can be written with 2 parameters multiplying these two vectors, and it is possible to directly find the formula expressing the general vector this way. First write the general vector in the plane as (x,y,-x-y) (by substituting for z - notice **2** parameters for **2** dimensions). It is a straightforward exercise in linear algebra - with or without software - to solve for α and β such that

$$(x,y,-x-y) = \alpha\ (1,-1,0) + \beta\ (1,1,-2).$$

We can solve this equation with the choices $\alpha = (x-y)/2$ and $\beta = (x+y)/2$.

Hence the two vectors we found in the solution are a basis for the plane. Any two independent vectors in the plane form a basis.

Definition I.5. A *linear transformation* is a function on vectors, where it doesn't matter whether linear combinations are made before or after the transformation. In other words,

$$F(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4)$$
$$= \alpha_1 F(v_1) + \alpha_2 F(v_2) + \alpha_3 F(v_3) + \alpha_4 F(v_4)$$

Linear transformations are also called *linear operators*, or just operators for short. You know plenty of examples:

Examples I.6.

1. Matrices. If M is a matrix and v, w etc. are column vectors, then

$$M(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4)$$
$$= \alpha_1 M v_1 + \alpha_2 M v_2 + \alpha_3 M v_3 + \alpha_4 M v_4$$

Think of rotation and reflection matrices here. If you put a bunch of vectors head-to-tail and rotate the assemblage, or look at it in a mirror, you get the same effect as if you first rotate or reflect the vectors, and then put them head-to-tail. It may be less geometrically obvious when the matrix distorts vectors in a trickier way, or there are more than three dimensions, but it is still true. In this example, algebraic intuition might be more convincing than geometric intuition

2. Derivatives and integrals. As we know,

$$\frac{d}{dx}(\alpha_1 f_1(x) + \alpha_2 f_2(x)) = \alpha_1 f_1'(x) + \alpha_2 f_2'(x),$$

$$\int_0^x (\alpha f(y) + \beta g(y))\, dy = \left(\alpha \int_0^x f(y)\, dy + \beta \int_0^x g(y)\, dy\right)$$

3. This may seem silly at first, but the *identity operator* Id, which just leaves functions alone is a linear transformation:

$$\mathrm{Id}(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 \mathrm{Id}\, f_1 + \alpha_2 \mathrm{Id}\, f_2,$$

since both sides are just round-about ways of writing $\alpha_1 f_1 + \alpha_2 f_2$. The identity operator is a useful bit of notation for much the same reason as the identity matrix,

$$\mathrm{Id} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

the effect of which on any vector is to leave it unchanged.

A linear transformation is a function defined on vectors, and the output is always a vector, but *the output need not be the same kind of vector as the input.* You should be familiar with this from matrices, since a 2×3 matrix acts on 3-vectors and produces 2-vectors. Similarly, the operator D acts on a vector space of functions assumed to be differentiable and produces functions which are not necessarily differentiable.

Example I.7. More exotic possibilities are possible, such as the operator which acts on 2×2 matrices by the rule:

$$F\left(\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}\right) = m_{11} \sin(x) - 3\, m_{22} \sin(2x).$$

The whole theory we are going to develop begins with the analogy between Example I.3.1 and Example I.3.2. We can think of the *linear operator* of differentiation by x as a kind of abstract matrix, denoted D. If we also think of the functions f and g as entities unto

themselves, without focusing on the variable x, then the expression in the first part of Example I.3.2 looks very much like Example I.3.1:

$$D(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 Df_1 + \alpha_2 Df_2.$$

The custom with linear transformations, as with matrices, is to do without the parentheses when the input variable is clear. It is tempting to manipulate linear operators in many additional ways as if they were matrices, for instance to multiply them together. This often works. For instance $D^2 = D\,D$ can be thought of as the second derivative operator, and expressions such as $D^2 + D + 2\,Id$ make sense. In passing, notice that if S is the integration operator of the second part of Example 2, then $D\,S\,f = f$, so D is an inverse to S in the sense that $D\,S = Id$. (But is $S\,D = Id$?)

Linear ODE's. In your course on ordinary differential equations, you studied linear differential equations, a good example of which would be

$$D^2y + Dy + 2y = 0$$

More specifically, this is an example of a "linear, homogeneous differential equation of second order, with constant coefficients". We can picture this equation as one for the *null space* of a *differential operator* A:

$$A\,y := (D^2 + D + 2\,Id)\,y = 0. \qquad (1.1)$$

By definition, the null space N(A) is the set of all vectors solving this equation; some texts refer to it as the *kernel* of A. There is no difference. You may remember that the null space of a matrix is always a linear subspace, and the same is true for the null space of any linear operator. E. g., the plane of Example 1.2.6 is the null space of the matrix

$$M = \begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix}. \qquad (1.2)$$

What does this abstract statement about subspaces mean in a concrete way for the problem of solving a linear homogeneous problem? It is the famous *superposition principle*: If we

can find two (or more) solutions of a homogeneous equation, then any linear combination of them is also a solution.

For matrices the general solution of the homogeneous equation is the set of linear combinations of a finite number of particular solutions. This will also be true for linear ordinary differential equations (the situation is a little more complicated for linear partial differential equations). Indeed, the number of independent functions in the null space of an ordinary differential operator is equal to its order (the highest power of D). This is true even when the coefficients are allowed to depend on x, so long as there are no singularities (such as values of x where the coefficients become 0 or ∞).

Let us illustrate the situation with the two operators mentioned above.

Model Problem I.8. Find the null space of M as given in (1.2), which is the same as finding the general solution of M v = 0.

In a linear algebra class you probably learned several techniques for finding the solution vectors v, and the way you describe the null space may depend on the technique chosen. We may prefer to use software, since Maple and Mathematica have commands to find the null space of a matrix.

Solution.

The nullspace can be found with software as a formula relating x, y, and z:

$$x = 2\,y - z$$

Equivalently, this is the plane satisfying $x - 2y + z = 0$. The plane passes through the origin, so it must be true that any linear combination of vectors in this plane again lies in the plane. (Why is it important that the plane passes through the origin?) A second way to describe the plane is to find a basis, i.e., two independent vectors such as

$v_1 = (1,1,1)$ and $v_2 = (1,0,-1)$

which are both in the plane and which form a basis for it.

Model Problem I.9. Find the general solution of

$$A\,u(x) = 0. \qquad (1.3)$$

Solution. The method here is to guess a solution of the form e^{mx}, and substitute to find what m must be. Usually, we find two possible values of m, and in a class on ordinary differential equations we learned that the general solution can be obtained from any two linearly independent solutions (for second order ODE's). If u(x) is of the form e^{mx}, then $A\,u = (m^2 + m + 2)\,u$, so if u is not the zero function, we must have

$$m = -\frac{1}{2} \pm \frac{\sqrt{-7}}{2} = -\frac{1}{2} \pm \frac{\sqrt{7}\,i}{2}.$$

Thus, two linearly independent solutions to (1.3) are

$$u_+(x) = \exp\left(\left(-\frac{1}{2} + \frac{\sqrt{7}\,i}{2}\right)x\right) \text{ and } u_-(x) = \exp\left(\left(-\frac{1}{2} - \frac{\sqrt{7}\,i}{2}\right)x\right).$$

(See the Maple worksheet or the Mathematica notebook.)

As these are complex valued - remembering Euler's formula that

$$\exp(i\gamma) = \cos(\gamma) + i\sin(\gamma)$$

- they are not necessarily in the most convenient form. However, they can be recombined easily into a pair of independent real functions,

$$u_1(x) = e^{-x/2}\cos\left(\frac{\sqrt{7}\,x}{2}\right) \text{ and } u_2(x) = e^{-x/2}\sin\left(\frac{\sqrt{7}\,x}{2}\right).$$

The general solution is the span of these two functions, i.e.,

$$\alpha_1 u_1(x) + \alpha_2 u_2(x).$$

Finally, let us recall the solution of nonhomogeneous linear problems, and observe how similarly it looks for matrix equations and differential equations.

Model Problem I.10. Find the general solution of

$$Mv = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Solution. The first step is to find a *particular solution* .A bit of trial and error leads to, perhaps,

$$v_p = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

There are many other, equally good choices for v_p. The general solution is the set of vectors of the form

$$v = v_p + \sum_j \alpha_j v_j$$

so we see that the homogeneous solution is just added onto v_p. The matrix M annihilates the terms $\alpha_j v_j$, so

$$Mv = Mv_p = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

The explicit answer is

$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Compare with the following:

Model Problem I.11. Find the general solution of

$$A\,u = \exp(2x).$$

Solution. Again, the first step is to find a particular solution. Trial and error with the guess $u_p(x) = C \exp(2x)$ leads to

$$A\,u_p = (4 + 2 + 2)\, C \exp(2x),$$

showing us that we must choose C = 1/8, making $u_p = \frac{e^{2x}}{8}$.

The general solution is

$$u(x) = \frac{e^{2x}}{8} + \alpha_1 u_1(x) + \alpha_2 u_2(x).$$

where $u_{1,2}$ are as given above. (Mathematical software can of course solve Model Problem I.11 directly; see the Mathematica notebook or the Maple worksheet for this chapter.)

Exercises.

Exercise I.1. For fluency with abstract vectors, use Definition I.1 to derive the rules that followed the definition.

Exercise I.2. Verify that the spaces in Examples I.2 are all vector spaces.

Exercise I.3. Find which of the following are vector spaces and which are not. In all cases, define addition and scalar multiplication in the standard way.

a) The plane of vectors (x,y,z) such that x - 2y + z = 1.

b) The set consisting of the zero function and all functions for $0 < x < 1$ which are not continuous at 1/2.

c) The set of 2 by 2 matrices with determinant 0.

d) The set of 2 by 2 matrices with trace 0. (Recall that the trace of a matrix is the sum of all of its diagonal elements from upper left to lower right.)

e) The set of 3 by 3 matrices with determinant 0.

f) The set of all polynomials in the variable x.

Exercise I.4. Show that a set $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ is linearly independent if and only if the statement that $\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n = 0$, implies that all the α's are zero.

Exercise I.5. With the operators D and S as defined in the text, acting on the vector space consisting of all polynomials, is S D = Id? If not, give a simple description of the operator SD.

Exercise I.6. (For review of ordinary differential equations.) Find the general solutions of the following ordinary differential equations.

a) $u'(t) = - 4 u(t)$

b) $u''(x) + (1/x) u'(x) - 9 u(x)/x^2 = 0$

c) $u'''(x) = u(x)$

Exercise I.7. Find the null space of the matrix

$$M = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

and find the general solution of

$$Mv = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Exercise I.8. Find the null space of the differential operator

$$A = D^2 - 2 D + Id$$

and find the general solution of

$$A u(x) = \cos(x).$$

Link to HTML version of Chapter I
Chapter II
Table of Contents
Evans Harrell's home page
Jim Herod's home page

Linear Methods of Applied Mathematics

Evans M. Harrell II and James V. Herod*

Some remarks for the instructor.

version of 28 September 1997

Click here to switch to the HTML version of this chapter.

Some of the calculuationsof this chapter are available in a Maple worksheet or in a Mathematica notebook..

II. The geometry of functions.†

The other familiar vector operation we shall use, besides sums and scalar multiples, is the dot product, which we abstractly call an *inner product.* (We won't be concerned with analogues of the cross product, although you may see these in other courses.) You probably learned about the dot product by beginning with trigonometry - the dot product of two vectors is given as the product of their lengths:

$$\mathbf{v}\,\mathbf{w} = |\mathbf{v}|\,|\mathbf{w}|\cos(\mathbf{v},\mathbf{w}) \qquad (2.1)$$

Later you learned that this could be conveniently calculated from the coordinates of the two vectors, by multiplying given components together and summing: If the components of the vector v are $v_1, \ldots, v_n$ and those of w are $w_1, \ldots, w_n$, then

$$v\,w = \sum_k v_k \overline{w}_k. \qquad (2.2)$$

From the abstract point of view it is best not to begin with angle, since we don't have a good intuitive picture of the angle between two functions or matrices. Instead, we will make sense of the "angle" between functions from the connection between these formulae.

In the abstract setting, we shall denote the inner product <v,w> rather than v w. Given an inner product, we may speak about the length, defined by $\|v\| = \sqrt{\langle v,v\rangle}$. We shall usually call this the *norm* rather than the length.

An abstract inner product will share most of the properties you learned about for the ordinary dot product:

1. The inner product is a mapping taking pairs of vectors and producing scalars
2. *Linearity* :

$$\langle \alpha_1 w_1 + \alpha_2 w_2, v\rangle = \alpha_1 \langle w_1, v\rangle + \alpha_2 \langle w_2, v\rangle.$$

Note: some texts define an inner product to be linear on the right side rather than the left side. This makes no practical difference, but if there are complex quantities, you should be careful to be consistent with whichever convention you follow.

3. *Symmetry* :

$\langle v,w\rangle = \overline{\langle w,v\rangle}$. (The bar denotes complex conjugate, in case these are complex numbers.)

4. *Positivity* :

$\langle v,v\rangle \geq 0$, and $\langle v,v\rangle = 0$ only if v is the zero vector.

5. The *Cauchy, Schwarz*, or *Buniakovskii inequality* (depending on nationality):

$$|\langle v,w\rangle| \leq \|v\|\|w\|. \qquad (2.3)$$

6. The *triangle inequality* :

$$\|v + w\| \leq \|v\| + \|w\|. \qquad (2.4)$$

We need Property 5 if an inner product is to correlate with Eq. (2.1), and we need Property 6 if the length $\|v\| = \sqrt{\langle v,v\rangle}$ is to make sense geometrically. We will shortly define an inner product for functions. The reason we can speak of the geometry of functions is that properties 5 and 6 follow automatically from Properties 1-4. Since this is not at all obvious, I shall prove it below.

Definition II.1. Given an abstract vector space, an *inner product* is any mapping satisfying properties 1-4 above. A vector space with an inner product is called an *inner product space*.

Examples II.2.

1. The usual dot product, as defined in (2.2).

2. A modified dot product on 2-vectors. A modification of the usual dot product can also produce an inner product defined by

$$\langle v,w\rangle_A := \sum_{j,k=1}^{2} v_j A_{jk} \overline{w}_k$$

where

$$A_{jk} := \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Actually, we could consider the vector space of n-vectors and let A be any positive n×n matrix. (By definition, a *positive matrix* is one for which the inner product so defined satisfies Property 4 of the inner product.)

3. The standard inner product for functions. Consider the set of functions which are continuous on an interval $a \le x \le b$. Then the *standard inner product* on them is the integral: $\langle f,g\rangle := \int_a^b f(x)\,\overline{g(x)}\,dx$. Another name for this is the L^2 inner product. We can use it for functions with discontinuities and even singularities, so long as the singularities are not too strong.

4. Other inner products for functions. We can generalize Example 3 in various ways. The first is to insert a positive weight function w(x):

$$\langle f,g\rangle := \int_a^b f(x)\,\overline{g(x)}\,w(x)\,dx.$$

Another generalization is to make the functions and integrals multidimensional, running over some region Ω:

$$\langle f,g\rangle := \int_\Omega f(x)\,\overline{g(x)}\;d^n x.$$

5. An inner product for matrices considered as vectors. Let the vector space V be the set of m×n matrices. Define the *adjoint* M* of a matrix M with entries m_{jk} to be the n×m matrix the entries of which are $\overline{m_{kj}}$. Define the trace of a matrix by $\mathrm{Tr}(M) = \sum_{k=1}^{n} m_{kk}$ (sum the diagonal elements). Then $\langle M,N\rangle := \mathrm{Tr}(MN^*)$ is an inner product.

Or - here is a really great one - we could have weight functions and lots of dimensions!

Theorem II.3. If V is an inner product space, then the CSB inequality (Property 5) and the triangle inequality (Property 6) hold.

proof (don't worry -it won't hurt you!): Because of the positivity property 4, the square length of any vector is ≥ 0, in particular for any linear combination $\alpha v+\beta w$,

$$\begin{aligned} 0 \le \langle \alpha v+\beta w, \alpha v+\beta w\rangle &= |\alpha|^2\|v\|^2 + |\beta|^2\|w\|^2 + \alpha\bar{\beta}\langle v,w\rangle + \alpha\beta\langle w,v\rangle \\ &= |\alpha|^2\|v\|^2 + |\beta|^2\|w\|^2 + 2\,\mathrm{Re}\left(\alpha\bar{\beta}\langle v,w\rangle\right) \end{aligned}$$

The trick now is to choose the scalars just right. Some clever person - perhaps Hermann Amandus Schwarz - figured out to choose $\alpha = \|w\|^2$ and $\beta = -\langle v,w\rangle$. The inequality then boils down to

$$0 \le \|w\|^4\|v\|^2 - \|w\|^2|\langle v,w\rangle|^2.$$

If we collect terms and divide through by $\|w\|^2$, we get the CSB inequality. (If $\|w\|=0$, the CSB inequality is automatic.)

For the triangle inequality, we just calculate:

$$\begin{aligned}\|v+w\|^2 &= \langle v+w,v+w\rangle = \langle v,v\rangle + \langle w,w\rangle + 2\,\mathrm{Re}\,(\langle v,w\rangle)\\ &\le \|v\|^2 + \|w\|^2 + 2\,\|v\|\,\|w\|\\ &\le \left(\|v\| + \|w\|\right)^2\end{aligned}$$

QED

Having the CSB inequality in hand, we may now define a strange but useful idea - the angle between two functions. Consider, for example, the interval $0\le x\le L$ and the functions $f(x) = 1$ and $g(x) = x$. With the standard inner product, we first calculate their norms:

$$\|1\| := \sqrt{\int_0^L 1^2\,dx} = \sqrt{L}, \text{ and } \|x\| := \sqrt{\int_0^L x^2\,dx} = \sqrt{\frac{L^3}{3}}.$$

Since their inner product is

$$\langle 1,x\rangle = \int_0^L x\,dx = \frac{L^2}{2},$$

the cosine of the angle between the two functions must be

$$\cos(\theta) = \frac{\frac{L^2}{2}}{\sqrt{L}\sqrt{\frac{L^3}{3}}} = \frac{\sqrt{3}}{2}.$$

Thus the "angle" between the functions 1 and x is $\pi/6$ radians. The most useful angle to deal with, however, is a right angle:

Definition II.4. Two functions f and g are said to be *orthogonal* if $\langle f,g\rangle = 0$. A set of functions $\{f_j\}$ is *orthogonal* if $\langle f_j,f_k\rangle = 0$ whenever $j \ne k$. The set is said to be *orthonormal* if it is orthogonal and $\|f_j\| = 1$ for all j.

With the Kronecker delta symbol, $\delta_{jk} = 0$ when $j\ne k$, and $\delta_{kk} = 1$, orthonormality can be expressed as $\langle f_j,f_k\rangle = \delta_{jk}$.

Examples II.5.

1. Integral tables, mathematical software, integration by parts (twice), substitution with the cosine angle-sum rule, and rewriting trigonometric functions as complex exponentials can all be used to evaluate integrals such as

$$\int_0^L \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right) dx.$$

Any or all of these methods will lead to the same conclusion, *viz.*:

$$\int_0^L \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right) dx = \frac{L}{2}\delta_{mn}. \qquad (2.5)$$

The set of functions $\left\{\sin\left(\frac{m\pi x}{L}\right)\right\}_{m=1}^{\infty}$ is orthogonal on the interval [0,L], and to turn it into an orthonormal set, we normalize the functions by multiplying by the appropriate constant:

$$\left\{\sqrt{\frac{2}{L}}\sin\left(\frac{m\pi x}{L}\right)\right\}_{m=1}^{\infty} \qquad (2.6)$$

2. Similarly,

$$\left\{\sqrt{\frac{2}{L}}\cos\left(\frac{m\pi x}{L}\right)\right\}_{m=1}^{\infty} \qquad (2.7)$$

is orthonormal on the interval [0,L], and we can even include another function, the constant:

$$\left\{\sqrt{\frac{1}{L}}\right\} \quad \left\{\sqrt{\frac{2}{L}}\cos\left(\frac{m\pi x}{L}\right)\right\}_{m=1}^{\infty}$$

3. We can mix the previous two sets to have both sines and cosines as long as we leave out all of the odd coefficients:

$$\left\{\sqrt{\frac{1}{L}}\right\} \quad \left\{\sqrt{\frac{2}{L}}\cos\left(\frac{2m\pi x}{L}\right)\right\}_{m=1}^{\infty} \quad \left\{\sqrt{\frac{2}{L}}\sin\left(\frac{2n\pi x}{L}\right)\right\}_{n=1}^{\infty} \qquad (2.8)$$

is also an orthonormal set. This one is the basis of the usual Fourier series, and is perhaps the most important of all our orthonormal sets. By the way, we do not claim that the functions in (2.6), (2.7), and (2.8) are orthogonal to functions in the other sets, but only separately among themselves. For instance, $\sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right)$ is not orthogonal to $\sqrt{\frac{1}{L}}$ onthe interval [0,L].

4. Recall that by Euler's formula, $\exp(i\alpha) := e^{i\alpha} := \cos(\alpha) + i\sin(\alpha)$. The complex trigonometric functions $\exp\left(\frac{2\pi inx}{L}\right)$ have many useful properties, including

$$\left|\exp\left(\frac{2\pi inx}{L}\right)\right| = 1 \text{ for all } x;$$

and

$$\left\{\frac{1}{\sqrt{L}}\exp\left(\frac{2\pi inx}{L}\right)\right\}_{n>-\infty}^{\infty} \tag{2.9}$$

is an orthonormal set on [0,L].

For later purposes, we observe here that the sets of functions (2.5)-(2.9) are each orthogonal on *any* interval [a,b] with L = b-a.

Before finishing this section we need two more notions about vectors and functions, thought of as abstract vectors. The first is distance. With the standard inner product, we would like to define the distance between two functions f and g as

$$\|f-g\| = \sqrt{\int_a^b |f(x)-g(x)|^2 dx}\ .$$

This turns out to be a familiar quantity in data analysis, called the *root-mean-square, or r.m.s., deviation*. It is a convenient way of specifying how large the error is when the true function f is replaced by a mathematical approximation or experimentally measured function g. It is always positive unless f = g *almost everywhere*.

Definition II.6. *Almost everywhere* is a technical phrase meaning that f and g differ for sufficiently few values of x that all integrals involving f have the same values as those involving g. For most practical purposes we may regard f and g as the same functions, and write:

$$f = g \text{ a.e.}$$

A typical example is where a function is discontinuous at a point and some arbitrary decision is made about the value of the function at the discontinuity. Different choices are still equal a.e. Much more exotic possibilities than this typical example can occur, but will not arise in this course.

The second notion we generalize from ordinary vectors is that of projection. Suppose that we have a position vector in three dimensions such as v = (3, 4, 5).

Model Problem II.7. We wish to find the vector in the plane spanned by (1,1,1) and (1,-1,0), which is the closest to v = (3, 4, 5) .

Solution. We solve this problem with a projection. The projection of a vector v onto the vector v_1 is given by the formula:

$$P_{v_1} v = \frac{(v \ v_1)\, v_1}{\|v_1\|^2} \tag{2.10}$$

Notice that this points in the direction of v_1 but has a length equal to $\|v\| \cos(\theta)$, where θ is the angle between v and v_1. The length of v_1 has nothing to do with the result of this projection - if we were being very careful, we would say that we were projecting v onto the *direction determined by* v_1, or onto the *line through* v_1. For similar reasons we notice that the vector v_1 could be normalized to have length 1, so that the denominator can be ignored - it is 1. In our example,

$$P_{\begin{bmatrix}1\\1\\1\end{bmatrix}} \begin{bmatrix}3\\4\\5\end{bmatrix} = \frac{(12)\begin{bmatrix}1\\1\\1\end{bmatrix}}{3} = \begin{bmatrix}4\\4\\4\end{bmatrix}.$$

(Here we write the vectors as column vectors because the projection operator is equivalent to a 3×3 matrix multiplying them.)

- -

If the basis for the plane consists of orthogonal vectors v_1 and v_2, as in our example, then the projection into the plane is just the sum of the projections onto the two vectors:

$$P_{\{v_1, v_2\}} v = \sum_{n=1}^{2} \frac{(v \ v_n)\, v_n}{\|v_n\|^2}. \tag{2.11}$$

In our example,

$$P_{\{v_1, v_2\}} v = \begin{bmatrix}4\\4\\4\end{bmatrix} + \begin{bmatrix}-1/2\\1/2\\0\end{bmatrix} = \begin{bmatrix}7/2\\9/2\\4\end{bmatrix}.$$

Calculations like these are easily automated with Mathematica or Maple (see the notebook or worksheet which accompany this text).

Model Problem II.8. We wish to find a) the vector in the plane spanned by (1,1,1) and (1,2,3), which is the closest to v = (3, 4, 5) and b) the vector in the same plane closest to (1,-1,0).

Solution. This is similar to the previous problem, except that the vectors defining the plane are not orthogonal. We need to replace them with a different pair of vectors, which are linear combinations of the first, but which are orthogonal. (We'll do this later systematically, with the *Gram-Schmidt* method.) The formula (2.11) is definitely wrong if the vectors v_n are not orthogonal. After finding the new pair of vectors, however, the solution will be as before - just sum the projections onto the orthogonal basis vectors.

There is more than one good choice for the pair of orthogonal basis vectors. If we decide that the pair of orthogonal vectors will include v_1 = (111), then we can look for a second vector of the form v_2 = (1,2,3) - α (1,1,1) which will be orthogonal to v_1, but will still lie in the plane. For orthogonality, we need to solve the vector equation

$$\begin{bmatrix}1\\1\\1\end{bmatrix}\left(\begin{bmatrix}1\\2\\3\end{bmatrix}-\alpha\begin{bmatrix}1\\1\\1\end{bmatrix}\right)=0$$

for α. We get α = 2, so a suitable second vector which is orthogonal to (1,1,1) is (-1,0,1).

The projection of (3,4,5) into the plane is the sum of its projections onto these vectors, i.e.,

$$\frac{3+4+5}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}+\frac{-3+0+5}{2}\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}3\\4\\5\end{bmatrix}.$$

Perhaps it looks strange to see that the projection of the vector (3,4,5) is itself, but the interpretation is simply that the vector was already in the plane before it was projected.

In general a vector will be moved (and shortened) when it is projected into a plane, and we can see this when we project (1,-1,0):

$$\frac{1-1+0}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}+\frac{-1+0+0}{2}\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}1/2\\0\\-1/2\end{bmatrix}.$$

Now that we have a vector in the plane, if we project again, it won't move. Algebraically, projections satisfy the equation

$$P^2 = P\ .$$

We shall next make these same calculations in function space to find the best mean-square fit of a function f(x) by a nicer expression, such as the functions in (2.8) or polynomials. (This can be automated with Mathematica or Maple.) The formula simply replaces the dot product with the standard inner product:

$$P_{\{g(x)\}}(f(x)) := \frac{\langle f,\, g\rangle}{\|g\|^2}\ g(x).$$

For example, if we wish to find the multiple of sin(3x) which is closest to the function x on the interval $0 \le x \le \pi$, we find:

$$P_{\{\sin(3x)\}}(x) = \frac{\int_0^{\pi} \tilde{x}\,\sin(3\,\tilde{x})\,d\tilde{x}}{\int_0^{\pi} \left(\sin(3\,\tilde{x})\right)^2\,d\tilde{x}}\ \sin(3\,x) = \frac{2}{3}\,\sin(3\,x).$$

Now consider what we mean when we project a function onto the constant function 1. This should be the best approximation to f(x) consisting of a single number. What could this be but the average of f? Indeed, the projection formula gives us

$$P_{\{1\}}(f(x)) := \frac{\langle f,\, 1\rangle}{\|1\|^2}\ 1 = \frac{\int_a^b f(x)\,dx}{b-a},$$

which is familiar as the average of a function.

Model Problem II.9. Consider the set of functions on the interval $0 \le x \le \pi$. We wish to find the function in the span of the functions 1, x, and x^2, which is the closest to $f(x) = \cos(\pi x/2)$ on the interval $-1 \le x \le 1$. In other words, find the best quadratic fit to the function f in the mean-square sense. In Exercise II.5 you are asked to compare a similar global polynomial approximation to this function with the Taylor series.

Solution. The calculations can be done with Mathematica or Maple, if you prefer. Conceptually, the calculations are closely analogous to those of Model Problem II.8.

First let us ask whether the three functions are orthogonal. Actually, no. The function x is orthogonal to each of the other two, but 1 and x^2 are not orthogonal. We can see this immediately because x is odd, while 1 and x^2 are both even, and both positive a.e.

A more suitable function than x^2 would be x^2 minus its projection onto the direction of 1, that is, x^2 minus its average, which is easily found to be 1/3. The set of functions $\{1, x, \text{and } x^2 - 1/3\}$ is an orthogonal set, as you can check.

Then we can project $\cos(\pi x/2)$ onto the span of the three functions 1, x, and $x^2 - 1/3$:

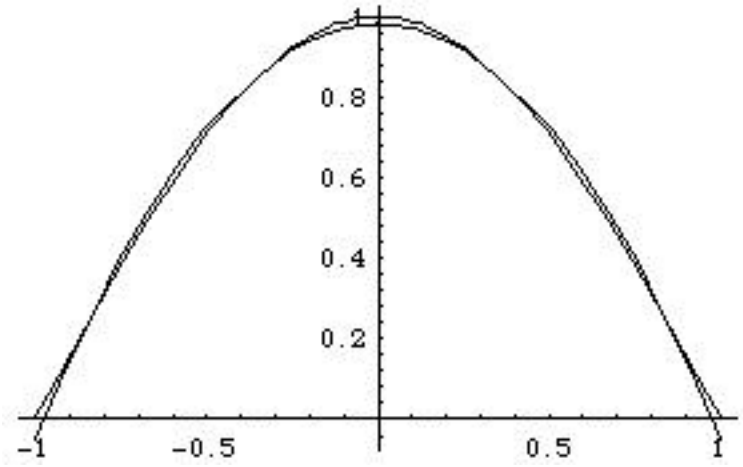
$$P_{\{1\}}\left(\cos(\pi x/2)\right) = \frac{1}{2}\int_{-1}^{1} \cos(\pi x/2)\, dx = \frac{2}{\pi}.$$

$$P_{\{x\}}\left(\cos(\pi x/2)\right) = 0 \quad \text{(the cosine is evenand x is odd).}$$

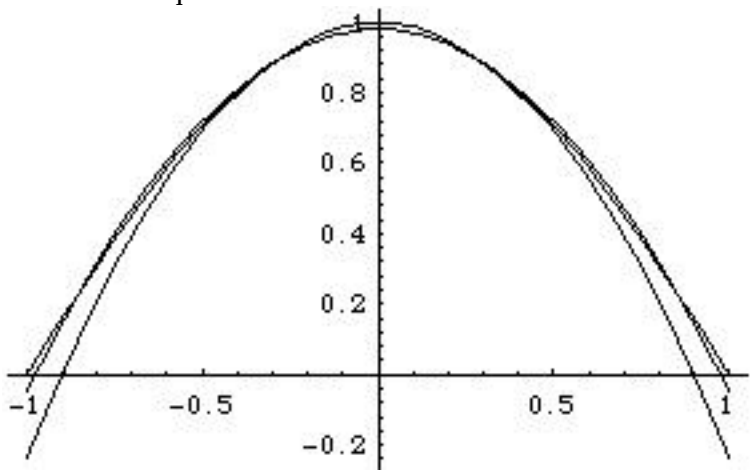
$$P_{\{x^2-1/3\}}\left(\cos(\pi x/2)\right) = \frac{\int_{-1}^{1}\left(t^2-1/3\right)\cos(\pi t/2)\, dt}{\int_{-1}^{1}\left(t^2-1/3\right)^2 dt}\left(x^2-1/3\right)$$

$$= \frac{-15\left(24-2\pi^2\right)}{2\pi^3}\left(x^2-1/3\right)$$

The best quadratic approximation to $\cos(\pi x/2)$ on the interval $-1 \le x \le 1$ is the sum of these three functions. Here is a graph showing the original function and its approximation:



For comparison, here is a plot which shows the Taylor approximation as well as the original and the best quadratic:



The general algorithm for finding the best approximation to a function is as follows. Suppose that we want to find the best approximation to f(x), $a \le x \le b$, of the form

$$\alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_n g_n(x)$$

where $g_1 \ldots g_n$ are some functions with nice properties - they may oscillate with definite frequencies, have simple shapes, etc. They could be chosen to capture the important features of f(x), while possibly simplifying its form or filtering out some noise.

Step 1. Replace $g_1 \ldots g_n$ by an orthogonal set with the same span. (A systematic way to do this, the Gram-Schmidt procedure, is described below.) Let's call the orthogonal set $h_1 \ldots h_n$.

Step 2. Project f onto each of the h_k

Step 3. Sum the projections. If P denotes the span of $g_1 \ldots g_n$, then

$$\text{Proj}_P(f) = \sum_{k=1}^{n} \frac{\langle f, h_k \rangle}{\|h_k\|^2} h_k(x) \qquad (2.12)$$

Perhaps the most important functional approximation uses the *Fourier functions* (2.8), as we shall learn to call them. The coefficients of these functions in the projection are called *Fourier coefficients*, and the approximation is as follows:

$$f(x) \quad a_0 + \sum_{m=1}^{M} a_m \cos\left(\frac{2\pi m x}{L}\right) + \sum_{n=1}^{N} b_n \sin\left(\frac{2\pi n x}{L}\right).$$

The right side should be the projection of f(x) on the span of the sines and cosines (including the constant) on the right. To get the coefficients we use the analogue of the formula
(2.10).
For example,<p>

$$P_{\cos\left(\frac{2\pi mx}{L}\right)} f = \frac{\left\langle f,\cos\left(\frac{2\pi mx}{L}\right)\right\rangle \cos\left(\frac{2\pi mx}{L}\right)}{\left\|\cos\left(\frac{2\pi mx}{L}\right)\right\|^2}$$
$$= \frac{2}{L}\left(\int_0^L \cos\left(\frac{2\pi mt}{L}\right) f(t)\, dt\right) \cos\left(\frac{2\pi mx}{L}\right).$$

In other words, the formula for the coefficients a_m must be:

$$a_m = \frac{2}{L}\int_0^L \cos\left(\frac{2\pi mx}{L}\right) f(x)\, dx,\ m = 1, 2, \ldots \tag{2.13}$$

The coefficient a_0 has a different form:

$$a_0 = \frac{1}{L}\int_0^L f(x)\, dx,\ m = 1, 2, \ldots \tag{2.14}$$

As mentioned above, the projection of f onto a constant function is its average.
Finally, the analogous calculation for the sines gives:

$$b_n = \frac{2}{L}\int_0^L \sin\left(\frac{2\pi nx}{L}\right) f(x)\, dx,\ n = 1, 2, \ldots \tag{2.15}$$

Formulae (2.13)-(2.15) will be the basis of the Fourier series in the next section.

Constructing orthonormal sets. It is often convenient to have orthonormal, or at least orthogonal sets. These are analogous to the usual basis vectors for the plane or for 3-space (denoted R^3), but you may recall that there are many choices of orthogonal bases. For instance, you may have unit vectors oriented along the x,y, and z axes, but someone else may have unit vectors oriented along a rotated set of axes. Although we shall first concentrate on the set (2.8) as a basis for a vector space of functions, the other sets of orthonormal functions (2.5)-(2.7) and (2.9) will be useful later for the same purpose. The choice among these sets is analogous to the choice of different bases for R^3.

But how do we come up with a basis in the first place? Suppose you are given several vectors, such as $v_{1,2,3}$ = (1,0,0), (1,1,0), and (1,1,1), and you want to recombine them to

get an orthonormal, or at least orthogonal set. You can do this by projecting away the parts of the vectors orthogonal to each other. The systematic way of doing this is called the Gram-Schmidt procedure, and it depends a great deal on the order in which it is done.

Model Problem II.10. Find an orthonormal set with the same span as $v_{1,2,3} = (1,0,0)$, (1,1,0), and (1,1,1), beginning with $w_1 = v_1 = (1,0,0)$. (We rename it because it is the first vector in a new set of recombined vectors.)

Solution. The next vector v_2 is not orthogonal to v_1, so we subtract off the projection of v_2 onto the direction of v_1:

$$w_2 = v_2 - P_{w_1} v_2 = \begin{bmatrix}1\\1\\0\end{bmatrix} - \begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\1\\0\end{bmatrix}.$$

For w_3, we begin with v_3 and project away the parts in the plane spanned by v_1 and v_2, which is the same as the plane spanned by w_1 and w_2. We find the standard basis vector

$$w_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Notice that the results are different if we take the same vectors in a different order:

Model Problem II.11. Find an orthonormal set with the same span as $v_{1,2,3} = (1,0,0)$, (1,1,0), and (1,1,1), beginning with v_3.

Solution. Fist let's turn v_3 into a unit vector:

$$\tilde{w}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}.$$

For the second unit vector we could take v_2 and project away the part pointing along v_3, getting

$$\begin{bmatrix}1\\1\\0\end{bmatrix} - \frac{2}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}\frac{1}{3}\\\frac{1}{3}\\-\frac{2}{3}\end{bmatrix}, \text{ which we can normalizeas}$$

$$\tilde{w}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix}.$$

Finally, taking v_1, projecting out the components in these directions, and normalizing, gives us the vector

$$\tilde{w}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Guess what's coming? The same procedure with sets of functions!

Model Problem II.12. Construction of the Legendre polynomials. Let us consider the interval $-1 \le x \le 1$, and find a set of orthonormal functions which are more mundane than the trigonometric functions, namely the polynomials. We begin with the power functions $1, x, x^2, x^3, \ldots$. Some of these are orthogonal because some are even functions and others are odd, but they are not all orthogonal to one another. For instance,

$$\langle 1, x^2 \rangle = \int_{-1}^{1} x^2 \, dx = \frac{2}{3}.$$

Let us denote the set of orthogonal polynomials we get using the Gram-Schmidt procedure on the power functions (in this order) $p_0, p_1, p_2, \ldots$. These are important in approximation theory and differential equations, and are known as the *normalized Legendre polynomials*. Beginning with the function $x^0 = 1$, after normalization we find

$$p_0(x) = \sqrt{\frac{1}{2}}$$

The next power is $x^1 = x$. Since x is already orthogonal to any constant function, all we do is normalize it:

$$p_1(x) = \sqrt{\frac{3}{2}}\, x$$

To make x^2 orthogonal to p_0 , we need to subtract a constant: $x^2 - 1/3$. Because of symmetry, it is already orthogonal to p_1, so we don't worry about p_1 yet, and just normalize:

$$p_2(x) = \sqrt{\frac{5}{2}}\left(\frac{3}{2}x^2 - \frac{1}{2}\right)$$

Similarly, when orthogonalizing x^3 we need to project out x but not 1 or x^2. We find $x^3 - 3x/5$, or, when normalized:

$$p_3(x) = \sqrt{\frac{7}{2}}\left(\frac{5}{2}x^3 - \frac{3}{2}\, x\right)$$

etc.

By the way, Legendre polynomials are traditionally not normalized as we have done, but rather are denoted $P_k(x)$ and scaled so that $P_k(1) = 1$. The normalization for Legendre polynomials of arbitrary index is such that

$$p_n(x) = (n + 1/2)^{1/2}\, P_n(x).$$

Most special functions are known to Mathematica and Maple, so calculations with them are no more difficult than calculations with sines and cosines.

Exercises.

II.1. Find the norms of and "angles" between the following functions defined for $-1 \le x \le 1$. (In the case of complex quantities, the "angle" is not so meaningful, but you can still define the cosines of the angles.) Use the standard inner product:

$1,\ x,\ x^2,\ \cos(\pi x),\ \exp(i\pi x)$

II.2. Verify the inner product of Example II.2.5 for 2×2 matrices. Find all matrices orthogonal to $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Do they form a subspace of the 2×2 matrices? If so, of what dimension?

II.3. It was remarked above that the projection operator on ordinary 3-vectors is equivalent to a matrix. What is the matrix for the projection onto (1,2,-3)? What is the matrix for the projection onto the plane through the origin determined by the vectors (1,2,-3) and (-2,0,0)?

II.4. Use the Gram-Schmidt procedure to construct other orthonormal sets from the basis vectors $v_{1,2,3}$ given in the text, taken in other orders. How many distinct basis sets are there?

II.5. Find the first 8 normalized Legendre polynomials.

II.6. a) Calculate the Taylor series for $\cos(\pi x/2)$ about the point x=0, up to the term with x^4. (Go to x^6 if you choose to use software to do this problem.)
b) Use the Legendre polynomials to find the polynomial of degree 4 which is the best approximation to $\cos(\pi x/2)$ for $-1 \le x \le 1$.
c) On a single graph, sketch $\cos(\pi x/2)$ and these two approximations.

II.7. In each case below, find a) the multiple of g which is closest in the mean-square sense to f, and b) a function of the form $f - \alpha g$ which is orthogonal to f. If for some reason this is not possible, interpret the situation geometrically.

(i) $f(x) = \sin(\pi x),\ g(x) = x,\ 0 \le x \le 1$

(ii) $f(x) = \sin(\pi x),\ g(x) = x,\ -1 \le x \le 1$

(iii) $f(x) = \cos(\pi x),\ g(x) = x,\ -1 \le x \le 1$

(iv) $f(x) = x^2 - 1,\ g(x) = x^2 + 1,\ -1 \le x \le 1$

II.8. Show with explicit examples that the formula (2.11) is definitely wrong if the vectors v_n are not orthogonal.

Link to HTML version of Chapter II
Chapter I
Chapter III
Table of Contents
Evans Harrell's home page
Jim Herod's home page

version of 23 October 1997

III. Fourier series. Introduction.†

As mentioned in the previous section, perhaps the most important set of orthonormal functions is the set of sines and cosines (2.8). These are what is known as a complete orthonormal set for the square-integrable functions on the interval [0,L]. They are also complete and orthonormal on any interval [a,b] with b-a = L.

Definition III.1. Square-integrable functions on [a,b] are functions f(x) for which $\int_a^b |f(x)|^2\,dx$ exists and isfinite.

The set of square-integrable functions is usually denoted L^2, and we shall see that it is an inner-product space. Later, we will also speak of square-integrable functions on regions in two or three dimensions, in which case we have multiple integrals over those regions.

Roughly speaking, a function on a finite interval is square integrable unless it is infinite somewhere. It can be very discontinuous, and in fact can even be slightly infinite - like the function ln(x), or even $|x|^{-1/3}$. Most familiar functions are square-integrable.

Definition III.2. An orthonormal set $\{e_n(x)\}$ is *complete* (on some fixed set of values of x) if for any square integrable function f(x) and any $\varepsilon>0$, there is a finite linear combination

$$\sum_{n=1}^{N} \alpha_n e_n(x) \text{ such that } \left\| f(x) - \sum_{n=1}^{N} \alpha_n e_n(x) \right\| < \varepsilon.$$

In other words, any reasonable function can be approximated as well as you wish (in the mean-square sense) by finite sums of the set. Indeed, we will say that it is the limit of an infinite series:

$$f(x) = \sum_{n=1}^{\infty} \alpha_n e_n(x)$$

The sense in which this infinite sum converges is, of course, in mean square. An equivalent way to describe completeness is: An orthonormal set $\{e_n(x)\}$ is complete if the statement that

$$\langle f, e_n \rangle = 0 \text{ for all } e_n$$

implies $f(x) = 0$ a.e. This is often a practical way to show that a set is incomplete, because you just have to exhibit a nonzero function which is orthogonal to the entire set.

Unfortunately, it is not as easy to prove completeness as it is to disprove it. It is a theorem that each of the sets of trigonometric functions (2.6)-(2.9) is complete and orthonormal, but the techniques for proving such a theorem are beyond the scope of this course. To understand the issue better, consider a finite linear combination of an orthonormal basis, $f(x) = \sum_{n=1}^{N} \alpha_n e_n(x)$. What is the norm of f? Because orthogonality makes the cross terms vanish, a little calculation shows us that

$$\|f(x)\|^2 = \sum_{n=1}^{N} |\alpha_n|^2 .$$

This is an extension of Pythagoras's theorem, since it says that the square of the hypotenuse is the sum of the squares of the lengths of the sides, if the sides are at right angles; only in function space, lots of things can all be at right angles. Suppose we had left out some of the components. Then we would have

$$\|f(x)\|^2 \geq \sum_{\text{someofthen}} |\alpha_n|^2 .$$

This inequality is still true if $N = \infty$, and is known as *Bessel's inequality*. The problem of completeness is that it is not easy to tell if we have included all the basis elements necessary to make both sides equal. We can leave many of them out and still have an infinite number left; in the set of Fourier functions (2.8), for example, we could leave out all the sine functions and still have all the cosine functions left.

The set (2.8) is particularly useful for *periodic* functions, that is, functions such that $f(x+L) = x$ for some fixed length L, called the *period*, and all x. Since each of the functions in the set is periodic with period L, any linear combination of them is also periodic with the same period. The completeness just alluded to means that every periodic function can be resolved into the trigonometric functions of the same period. If the independent variable is time (which you might prefer to denote t rather than x), a periodic function of the form $\sin(2n\pi t/L)$ or $\cos(2n\pi t/L)$ may be detected by your ear and perceived as a pure musical tone with frequency n/L. *Any periodic sound wave can be resolved into pure musical tones.* If you are given a sound wave f(t), which is periodic with period L

you can extract its components of frequency k/L with formulae (2.12)-(2.14), setting m or n = k. There are two such components, one with the sine function and the other with the cosine. This degree of freedom corresponds to the phase of the sound wave, because of the trig identity:

$$a\cos(\theta) + b\sin(\theta) = \sqrt{a^2+b^2}\,\cos(\theta-\phi), \tag{3.1}$$

where $\tan(\phi) = \frac{b}{a}$.

The intensity (power) carried by the component of a sound wave at pure frequency k/L is proportional to $|a_k|^2 + |b_k|^2$.

In the next chapter this resolution into pure frequencies is carried out for the square wave and the results are plotted, among other things. You may wish to glance at those plots now to get an intuitive feel for how a Fourier series can approximate a function.

The powerful theorem behind the Fourier series III.3. We shall now carefully formulate a theorem which justifies the use of Fourier series for square-integrable functions and tells us several useful facts. Historically, parts of this theorem were contributed by Fourier, Parseval, Plancherel, Riesz, Fischer, and Carleson, and in most sources it is presented as several theorems. For proofs and further details, refer to Rudin's *Real Analysis* or Rogosinski's *Fourier Series.*

If f(x) is square-integrable on an interval [a,b], then

a) All of the coefficients a_m and b_n are definite numbers uniquely given by (2.12)-(2.14).

b) All of the coefficients a_m and b_n depend linearly on f(x)

c) The series

$$f(x) \quad a_0 + \sum_{m=1}^{M} a_m\cos\left(\frac{2\pi mx}{b-a}\right) + \sum_{n=1}^{N} b_n\sin\left(\frac{2\pi nx}{b-a}\right).$$

converges to f(x) in the mean-square sense. In other words,

$$\left\| a_0 + \sum_{m=1}^{M} a_m\cos\left(\frac{2\pi mx}{b-a}\right) + \sum_{n=1}^{N} b_n\sin\left(\frac{2\pi nx}{b-a}\right) - f(x)\right\| \to 0.$$

We shall express this by writing

$$f(x) = a_0 + \sum_{m=1}^{\infty} a_m\cos\left(\frac{2\pi mx}{b-a}\right) + \sum_{n=1}^{\infty} b_n\sin\left(\frac{2\pi nx}{b-a}\right). \tag{3.2}$$

(Remember, however, that this series converges in a mean sense, and not necessarily at any given point.) It is also true that it converges a.e.

d) $$\|f(x)\|^2 = (b-a)\,|a_0|^2 + \frac{b-a}{2}\left(\sum_{m=1}^{\infty}|a_m|^2 + \sum_{n=1}^{\infty}|b_n|^2\right), \tag{3.3}$$

and the right side is guaranteed to converge.

e) If g is a second square-integrable function, with Fourier coefficients $\tilde{a}_m$ and $\tilde{b}_n$, then

$$\langle f,g\rangle = (b-a)\bar{a}_0\tilde{a}_0 + \frac{b-a}{2}\sum_{m=1}^{\infty} \bar{a}_m\tilde{a}_m + \frac{b-a}{2}\sum_{n=1}^{\infty} \bar{b}_n\tilde{b}_n \qquad (3.4)$$

This is known as the *Parseval formula*.

Conversely, given two square-summable sequences a_m and b_n, i.e., real or complex numbers such that

$$\sum_{m=0}^{\infty} |a_m|^2 + \sum_{n=1}^{\infty} |b_n|^2,$$

is finite, they determine a square integrable function f(x) uniquely a.e. such that (2.12)-(2.14) and statements b)-d) hold.

Let's stand back and think about what this big theorem tells us. Square-integrable functions are very general, so this is telling us that any reasonable function can be approximated arbitrarily well, in the r.m.s. sense, by a "trigonometric series." We have a formula to generate the coefficients, and in fact a full correspondence between the square-integrable functions and the square-summable sequences. The square-integrable functions L^2 form an inner product vector space. The set of double sequences $\{a_m,b_n\}$ is also a vector space with an inner product given by (3.4). You can think of each object in this space as a vector with an infinite number of components, some of which are denoted a_m and others b_n.

Here are some examples showing why mean-square approximation is not always good enough. At a later stage we shall discuss when Fourier series converge at individual points. We shall see that if f(x) and f'(x) are continuous near x, the Fourier series converges at x.

Examples III.4.

1. The *indicator function* $\chi_S(x)$ of a set S takes on the value 1 when x is in S and 0 otherwise. The sequence of indicator functions $f_n = \chi_S(x)$ for S = the sets $S_n = [0, 1/n]$ converges to the 0 function in the mean-square sense, but for each n there is always an interval near 0 such that $|f_n - 0| = 1$.

2. This one is more disturbing than Example 1. The function f_n is still always an indicator function $\chi_S(x)$, but the sets S_n are as follows

n=1	$0 \le x \le 1$
n=2	$0 \le x \le 1/2$
n=3	$1/2 \le x \le 1$
n=4	$0 \le x \le 1/4$
n=5	$1/4 \le x \le 1/2$
n=6	$1/2 \le x \le 3/4$
n=7	$3/4 \le x \le 1$
n=8	$0 \le x \le 1/8$

etc. This sequence also converges to the zero function in the mean-square sense. For any fixed x, $f_n(x)$ will always have either the value 0 or 1, but it will never settle down definitively to 0 as $n \to \infty$. There are an infinite number of n's for which $f_n(x) = 1$.

If you look at the various Fourier series that are plotted in the next chapter, the crazy stuff of Example 2 doesn't happen. In fact, the convergence is very good except at the ends of the intervals or at places where the function is discontinuous. If you look at the periodic extension of a function, you see that the end of an interval is a place where there is likely to be a discontinuity, and it is when this happens that the series did not converge at the end of the interval. (A good example in chapter IV is $f(x) = x$ on the basic interval [0,L].) What we observe is described by a general theorem, which we now formulate.

A function is said to be *piecewise continuous* (some say *sectionally continuous*) if it is continuous except at a discrete set of jump points, where it at least has an identifiable value on the left and a different one on the right. Here is a formal way to state this:

Definition III.5. A function f(x) is *piecewise continuous* on a finite interval $a \le x \le b$ if it is continuous except at a finite number of points $a =: x_0 < x_1 < \ldots < x_n := b$, and all the one-sided limits

$$f(x_k^-) := \lim_{x \uparrow x_k} f(x_k),$$

and

$$f(x_k^+) := \lim_{x \downarrow x_k} f(x_k)$$

exist (except that we only assume the limit from above at a and the limit from below at b).

Here the up-arrow indicates that the limit is taken for values of x tending to x_k from below, and the down-arrow indicates that the limit is taken for values of x tending to x_k from above.

What we see from the examples is that where a function has a discontinuity, the Fourier series, when truncated to a large but finite number of terms, takes on a value between the right and left limits. The theorem says that the Fourier series finds the average of the two possibilities.

Theorem III.6. Suppose that f(x) and f'(x) are piecewise continuous on a finite interval [a,b]. Then the Fourier series (3.2) converges at every value of x between a and b as follows:

$$\lim_{M,N\to\infty}\left(a_0 + \sum_{m=1}^{M} a_m \cos\left(\frac{2\pi mx}{b-a}\right) + \sum_{n=1}^{N} b_n \cos\left(\frac{2\pi nx}{b-a}\right)\right) = \frac{1}{2}\left(f(x^+) + f(x^-)\right).$$

At the end points, we have:

$$\begin{aligned}&\lim_{M,N\to\infty}\left(a_0 + \sum_{m=1}^{M} a_m \cos\left(\frac{2\pi ma}{b-a}\right) + \sum_{n=1}^{N} b_n \cos\left(\frac{2\pi na}{b-a}\right)\right)\\ &= \lim_{M,N\to\infty}\left(a_0 + \sum_{m=1}^{M} a_m \cos\left(\frac{2\pi mb}{b-a}\right) + \sum_{n=1}^{N} b_n \cos\left(\frac{2\pi nb}{b-a}\right)\right)\\ &\quad = \frac{1}{2}\left(f(a^+) + f(b^-)\right).\end{aligned}$$

The limit at the end points is reasonable, because when the function is extended periodically, they are effectively the same point. And a particular consequence is that: *At places where such a function is continuous, the Fourier series does indeed converge to the function.*

In addition, if the function is continuous on the interval [a,b], and f(a) = f(b), then we can state a bit more, namely that the Fourier series converges *uniformly* to the function. This means that the error can be estimated independently of x:

Theorem III.7. Suppose that f'(x) is piecewise continuous, f(x) itself is continuous on a finite interval [a,b], and f(a) = f(b). Then

$$\max_{a\le x\le b}\left|a_0 + \sum_{m=1}^{M} a_m \cos\left(\frac{2\pi mx}{b-a}\right) + \sum_{n=1}^{N} b_n \sin\left(\frac{2\pi nx}{b-a}\right) - f(x)\right| \to 0.$$

The condition that f(a) = f(b) is again reasonable if you think of f as a periodic function extending beyond the interval [a,b] - the extended function would be discontinuous at the end points if f(a) did not match f(b).

Exercises.

III.1. Find the entire
Fourier (sine and cosine) series
Fourier sine series (see appendix)
Fourier cosine series (see appendix)
Fourier exponential series
on the interval $0 \le x \le L$ for the function $f(x) = x$ for $0 \le x \le L/2$, otherwise 0. Discuss what happens if the series is evaluated outside the interval $0 \le x \le L$.

III.2. Numerically find the first eight terms in the
Fourier series
Fourier sine series
Fourier cosine series
Fourier exponential series
on the interval $0 < x < 2\pi$ for the function $f(x) = \cosh(\cos(x))$

III.3. Find the Fourier series of the following functions:

(i) $\cosh(2x)$, $-\pi \le x \le \pi$

(ii) $x|x|$, $-\pi \le x < \pi$. Notice that this is an odd function.

(iii) $1 - |x|$, $-1 \le x \le 1$.

(iv) $|\sin(x)|$, $0 \le x \le \pi$.

(v) $2 - 2\cos(\pi x)$, $-1 \le x \le 1$.

(vi) $f(x) = x$ for $0 \le x < L/2$, 0 for $L/2 \le x < L$

(Implicitly, the functions are extended periodically from these basic intervals.)

III.4. When we calculate the Fourier series for a square pulse and for x mod L in chapter IV, we find that $a_m = 0$ for all $m \ge 1$. Explain why, using ideas about symmetry.

III.5. a) Is it possible for a sequence of functions to converge in the r.m.s. sense for $0 \le x \le 1$, to converge at every point of the interval $0 \le x \le 1/2$, but not converge at any point of the interval $1/2 < x \le 1$? Give an example or explain why not.
b) Is it possible for a sequence of square-integrable functions to converge at every point of the interval $0 \le x \le 1$, but not converge in the r.m.s. sense? Give an example or explain why not.

III.6. Use the Parseval formula (3.4) to calculate
$\langle 3 + \cos(x) + \sin(x) + 2\cos(2x) + 2\sin(2x) + 3\cos(3x),\ 1 - \sin(2x) - \sin(4x) \rangle$
(standard inner product for $0 \le x \le 2\pi$), without doing any integrals.

Appendix on Fourier sine and cosine series.

There are three distinct versions of the Fourier series in common use, the "full" Fourier series, the Fourier sine series, and the Fourier cosine series. They are based on slightly different choices of an orthogonal basis set, whether (2.6) for the Fourier sine series, (2.7) for the Fourier cosine series, or (2.8) for the full Fourier series. Notice that while the full set (2.8) contains both sines and cosines, it does not contain as many frequencies as either of the other sets, since the trig functions depend on $2\pi nx/L$ rather than all integer multiples of $\pi x/L$. We shall also discuss an "exponential" Fourier series, which uses complex functions of the form $\exp(2\pi ix/L)$, but this is actually the same series as the full Fourier series: It has simply been written in a convenient way using exponential functions.

The three different sets do an equally good job of approximating functions in the r.m.s. sense, but their relationship to symmetries is different. For this reason, different sets of Fourier functions will arise in our study of partial differential equations, depending on the kind of "boundary conditions" which are physically appropriate for a given problem.

The full Fourier series consists of functions which are periodic, and are particularly good for representing or approximating periodic functions, that is, functions for which

$$f(x + L) = f(x).$$

Examples include sounds built on a basic frequency with overtones and any physical function depending on an angular variable such as longitude. The Fourier sine series, in contrast, is well adapted to functions which are zero at x=0 and x=L, since all the functions $\sin(n\pi x/L)$ have this property. A related but less obvious property that begs for the use of sine functions is an odd symmetry: $f(-x) = - f(x)$. If the function you wish to approximate is an odd function, then the sines are a good candidate for a basis set. Notice that the sines are not only odd under reflection about 0, but also about the other end of the interval, x=L. Fourier cosine series are better suited for even functions. A Fourier sine series represents a function as

$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi n x}{L}\right) \qquad (3.a1)$$

for $0 < x < L$; there are no cosine terms at all. A Fourier cosine series, on the other hand, represents a function as

$$f(x) = a_0 + \sum_{m=1}^{\infty} a_m \cos\left(\frac{\pi m x}{L}\right). \qquad (3.b1)$$

(No sine terms this time.) In most applications a = 0 < x < L, but this is not strictly necessary. If the lower limit for x is a > 0, then either of these series can still be used if we change variables from x to x-a, so the basis functions are $\sin(n\pi(x-a)/L)$ or, respectively, $\cos(n\pi(x-a)/L)$.

The formulae for the coefficients a_n and b_n are the same as in (2.14), except that the alternative Fourier sine series use all integer multiples of $\pi x/L$, where L=b-a, not just the even multiples: For the sine series,

$$b_n = \frac{2}{L}\int_0^L f(x)\,\sin\left(\frac{n\pi x}{L}\right)dx, \qquad (3.a2)$$

while for the cosine series,

$$a_0 = \frac{1}{L}\int_0^L f(x)\;dx.$$

while for m>0,

$$a_m = \frac{2}{L}\int_0^L f(x)\,\cos\left(\frac{m\pi x}{L}\right)dx. \qquad (3.b2)$$

Just as for the full Fourier series, the Fourier sine functions are a complete set, and they satisfy some useful identities, which I have collected in a theorem (which is almost identical to the one for the full series):

The powerful theorem behind the Fourier sine series III.3a.

If f(x) is square-integrable on an interval [a,b], then:

a) All of the coefficients b_n of (3.a2) are well defined.

b) All of the coefficients b_n depend linearly on f(x).

c) The series

$$f(x) \quad \sum_{n=1}^{N} b_n \sin\left(\frac{\pi n x}{L}\right).$$

converges to f(x) in the r.m.s. sense. In other words,

$$\left\| \sum_{n=1}^{N} b_n \sin\left(\frac{\pi n x}{L}\right) - f(x) \right\| \to 0.$$

We shall express this by writing

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

(Remember, however, that this series converges in a mean sense, and not necessarily at any given point.) It is also true that it converges a.e.

d) $$\|f(x)\|^2 = \frac{L}{2}\sum_{n=1}^{\infty} |b_n|^2,$$

and the right side is guaranteed to converge.

e) If g is a second square-integrable function, with Fourier sine coefficients $\tilde{b}$, then

$$\langle f,g\rangle = \frac{L}{2}\sum_{n=1}^{\infty} \bar{b}_n\tilde{b}_n.$$

This is known as the *Parseval formula* for the Fourier sine series.

Conversely, given a square-summable sequence b_n, i.e., real or complex numbers such that

$$\sum_{n=1}^{\infty} |b_n|^2$$

is finite, they determine a square integrable function f(x) uniquely a.e. such that all the statements of this theorem hold.

In addition, if f(x) is piecewise continuous, then Theorem III.6 likewise holds, except at the end points of the interval. If $0 < x < L$, the Fourier sine series will converge to the average of the right and left limiting values of f at x. When x=0 or L, however, the Fourier sine series will always converge to 0, since all the sine functions are 0 at those two points.

The Fourier cosine functions are also a complete set, and they satisfy similar useful relationships:

The powerful theorem behind the Fourier cosine series III.3b.

If f(x) is square-integrable on an interval [a,b], then

a) All of the coefficients a_m of (3.b2) are well defined.

b) The series

$$f(x) \quad a_0 + \sum_{m=1}^{M} a_m \cos\left(\frac{\pi m x}{L}\right)$$

converges to f(x) in the r.m.s. sense. In other words,

$$\left\| a_0 + \sum_{m=1}^{N} a_m \cos\left(\frac{\pi m x}{L}\right) - f(x)\right\| \to 0.$$

We shall express this by writing

$$f(x) = a_0 + \sum_{m=1}^{\infty} a_m \cos\left(\frac{\pi m x}{L}\right).$$

As before, it is also true that it converges a.e.

c) $$\|f(x)\|^2 = L|a_0|^2 + \frac{L}{2}\sum_{m=1}^{\infty} |a_m|^2,$$

and the right side is guaranteed to converge.

d) All of the coefficients a_m depend linearly on f(x)

e) If g is a second square-integrable function, with Fourier cosine coefficients $\tilde{a}_m$, then

$$\langle f,g\rangle = L a_0 \tilde{a}_0 + \frac{L}{2}\sum_{m=1}^{\infty} a_m \tilde{a}_m.$$

This is known as the *Parseval formula* for the Fourier cosine series.

Conversely, given a square-summable sequence a_m, i.e., real or complex numbers such that

$$\sum_{m=0}^{\infty} |a_m|^2$$

is finite, they determine a square integrable function f(x) uniquely a.e. such that all the statements of this theorem hold.

In addition, if f(x) is piecewise continuous, then Theorem III.6 likewise holds, except at the end points of the interval. If $0 < x < L$, the Fourier cosine series will converge to the average of the right and left limiting values, $\frac{1}{2}(f(x^+) + f(x^-))$. When x=0 or L, however, the Fourier cosine series will always converge to the one-sided limit of f at that point, .

$f(0^+)$ or $f(L^-)$.

When we study partial differential equations beginning in chapter VI, Fourier sine or cosine functions will arise depending on the *boundary conditions*. If we know that the functions we seek are equal to 0 at the end points 0 and L, then we can ensure this fact by writing them as sums of sines of the form $\sin(n\pi x/L)$. Similarly, if we know that the derivatives are equal to 0 at the end points, we can accomplish this with a sum of cosines of the form $\cos(m\pi x/L)$.

vversion of 24 October 1997

IV. Calculating Fourier series†

In this section we calculate several Fourier series. As we know, to find a Fourier series simply means calculating various integrals, which can often be done with software or with integral tables. Since performing integrals is not much more interesting in the modern age than long division, our goals in this section will be to get a visual and analytic impression of what to expect from Fourier series, and to understand the rôle of symmetry in the calculations.

Let's begin by evaluating the Fourier series for the functions:

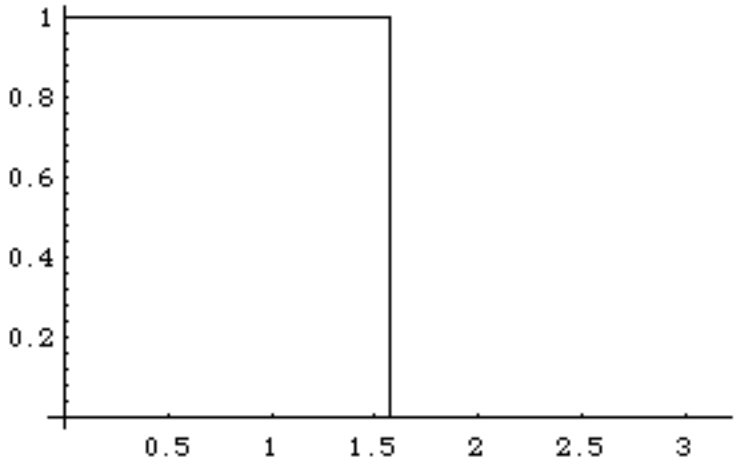
$f(x) = 1$ for $0 \le x \le L/2$, but 0 for $L/2 < x < L$

and

$g(x) = x$, $0 \le x < L$.

The functions have not been defined at the points of discontinuity, but as we know, the Fourier series will converge there to the average of the limit from the left and the limit from the right. The end point L is essentially a jump point, because the periodic extension of the functions make the values x=L and x=0 equivalent.

Here is a graph of the function f, called a "square pulse" or "square wave" (when extended periodically):



We want to represent these functions in the form

$$a_0 + \sum_{m=1}^{\infty} a_m \cos\left(\frac{2\pi mx}{b-a}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nx}{b-a}\right),$$

beginning with f(x).

Model Problem IV.1. Using software, calculate the full Fourier series for the function f(x) as defined above, and investigate its convergence.

Solution. The formulae for these coefficients were given in chapter 2:

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

(i.e., the average of f), and for the other coefficients:

$$a_m = \frac{2}{L}\int_0^L \cos\left(\frac{2\pi mx}{L}\right) f(x)\,dx,\ m = 1, 2, \ldots$$

and

$$b_n = \frac{2}{L}\int_0^L \sin\left(\frac{2\pi nx}{L}\right) f(x)\,dx,\ n = 1, 2, \ldots\ .$$

Here are the results of the calculation:

$a_0 = 1/2$;
$a_m = 0$ for $m = 1, 2, \ldots$;
$b_n = (1 - (-1)^n/(n\pi)$, $n = 1, 2, \ldots$.

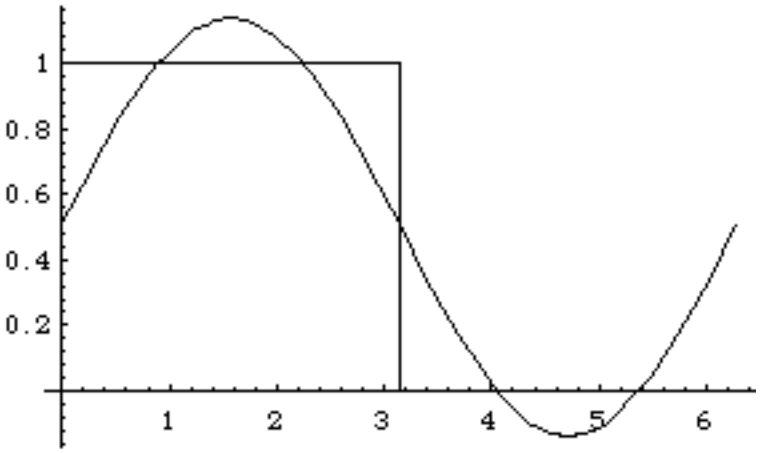
A point to ponder: why are all the $a_m = 0$? Surely this is no coincidence!

The sort of expression we got for the b_n simplifies in a way which arises often in Fourier series for elementary functions. If n=2k is even, $b_{2k} = 0$. Thus only odd terms survive, in which case

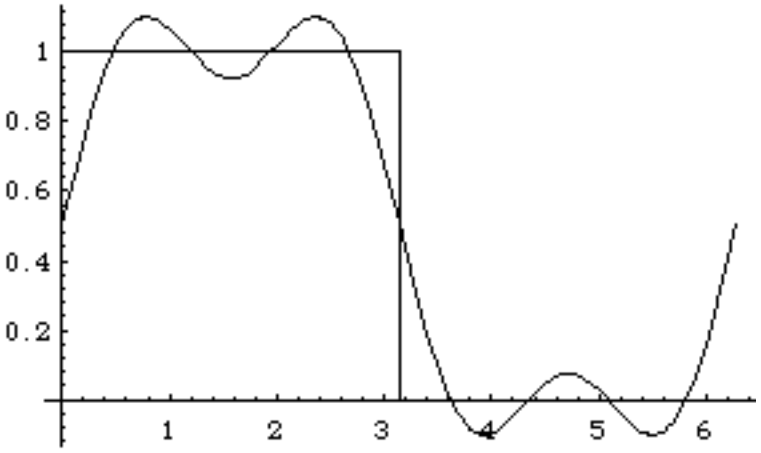
$$b_{2k+1} = 2/(\pi (2\ k + 1)).$$

We can plot the Fourier series and watch it converge to the original function as as more and more terms are included:

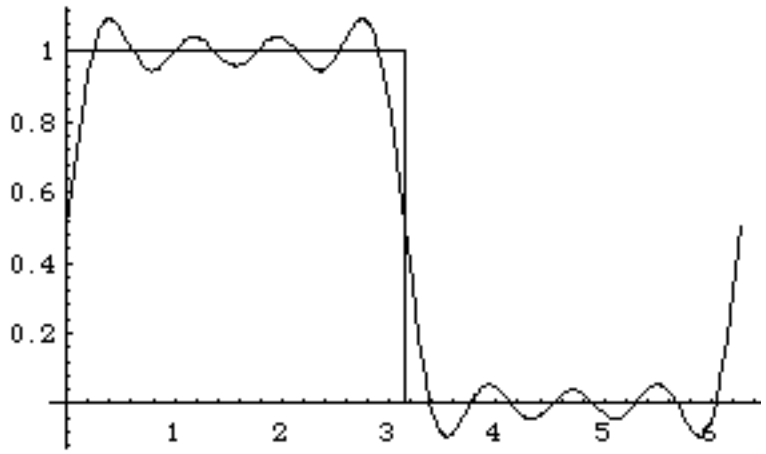
With one sine term:



With three sine terms:



With seven sine terms:



And so on. We call attention to the systematic overshoot that occurs at the edges of the jump. Curiously, the size of this last overshoot does not tend to 0 as we include more terms. The bumps next to the jump get thinner rather than shorter. This is known as the Gibbs phenomenon.

- -

Model Problem IV.2. We can now investigate some questions about the excitation of mechanical resonances. Suppose that experiments by K. Battle at the Wiener Staatsoper in Austria show that a crystal goblet will shatter if the intensity of the tone at frequency 1760 Hz (high A) exceeds .01. We use physical units in which the proportionality between the power and the square of the amplitude is 1, i.e., we define the intensity simply by

$$I = a_k^2 + b_k^2 \ .$$

At Georgia Tech, not a particularly musical spot, we can generate a square pulse with amplitude A for 1/704 sec., then amplitude 0 for the same length of time, then A, etc., periodically with period 1/352.

Question: What amplitude A will cause the glass to shatter in our laboratory?

Solution. The issue is essentially to estimate the magnitude of the Fourier coefficient for our function corresponding to the frequency 1760 Hertz.

The mapping from functions to their Fourier coefficients is linear. Hence the coefficients can be obtained from the ones we just calculated.

Step 1. Scale the independent variable by replacing x with t and L with 1/352.
Step 2. Multiply all the coefficients by A.

The harmonic with frequency 1760 is the one with n=5, so the Fourier coefficients corresponding to this frequency are

$a_5 = 0$, and

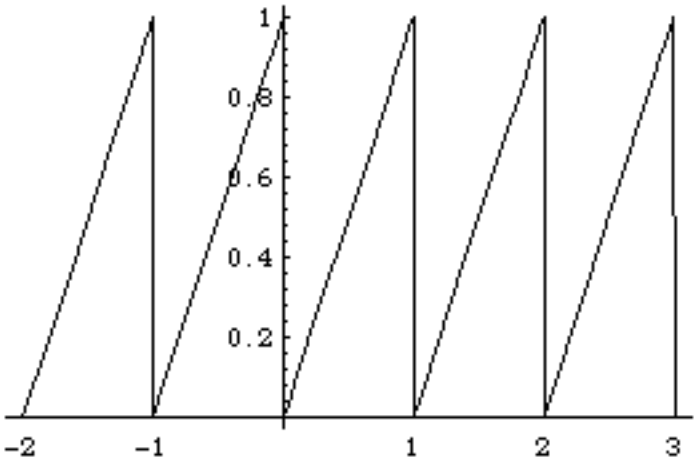
$b_5 = 2A/(5\pi)$

The intensity is $4A^2/(25\pi^2)$. The amplitude A above which the glass will shatter is $\pi/4$, numerically about 0.785398.

- -

Model Problem IV.3. For comparison, let us find another Fourier series, namely the one for the periodic extension of $g(x) = x$, $0 \le x \le 1$, sometimes designated x mod 1. Watch it converge.

Solution. For x between L and 2L, the function is (x-L), for x between 2L and 3L it is (x-2L), etc. Its graph has a sawtooth shape:



The Fourier coefficients are calculated as:

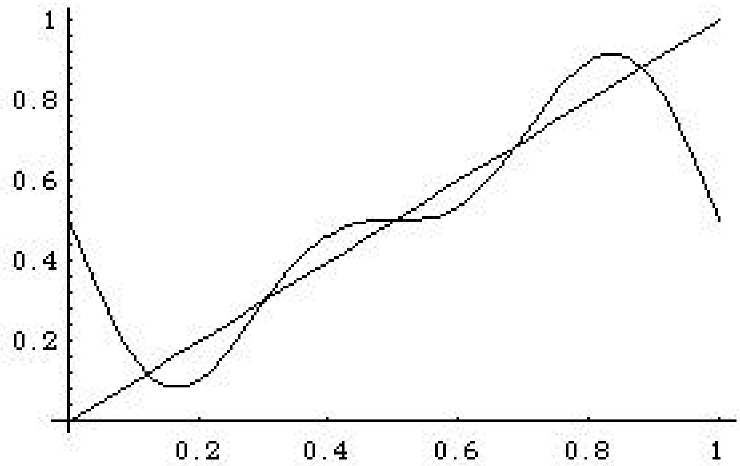
$a_0 = 1/2$;

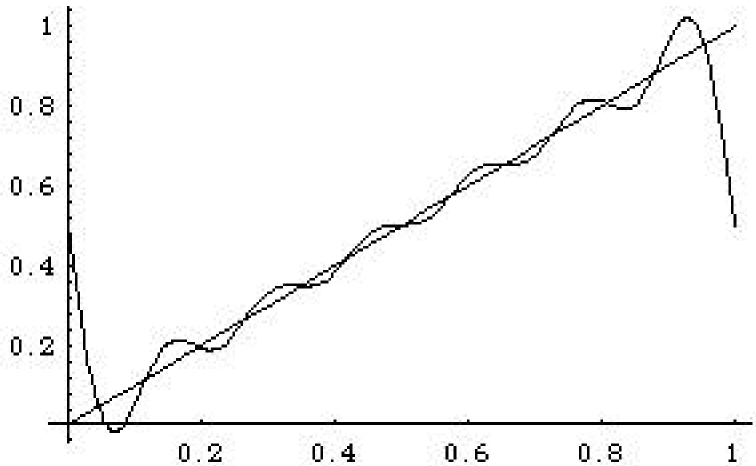
$a_m = 0$ for m = 1, 2, ...;

$b_n = (-1)/(n\pi)$ for n = 1, 2, ...

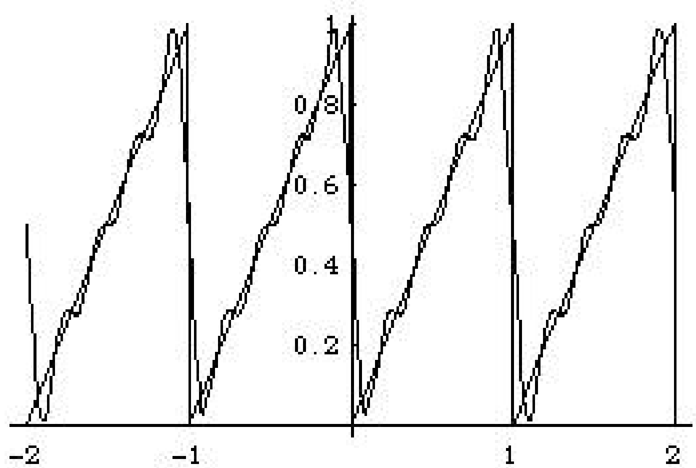
The coefficients a_m are all zero, again. Why?

Let us get an impression of how the series converges, by plotting the contributions of the first two sines, and then of six sines:





We also recall that on a longer interval, the Fourier series produces a periodic function:



- -

The Fourier series is converging nicely to the function except at the end-points of the interval, which are places where the full periodic saw-tooth function has jumps, and we saw something similar with the square pulse. In both cases we see numerical evidence for the theorem that the Fourier series converges to f(x) where f(x) is continuous, and where it has a jump, the Fourier series converges to the average of the upper and the lower value at the jump.

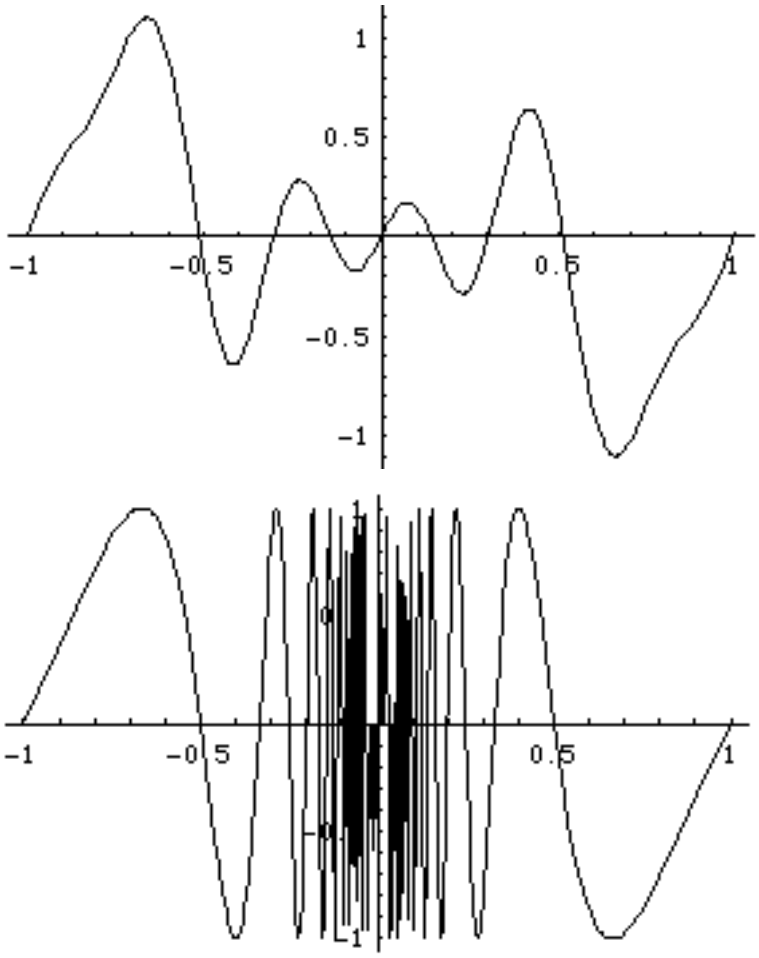
Let's try out a different sort of example, where we need to integrate numerically.

Model Problem IV.4. Find the Fourier series for the function $\sin(\pi/x)$, $-1 \le x \le 1$.

Solution. The integrals called for are not in the integral tables. Indeed, *most* integrals are not in the integral tables, even when they are nowhere near as wild as this function (ask yourself what happens near x=0). In the information age this is no barrier. We simply call on the software to do the integrals numerically. Since the function $\sin(\pi/x)$ is odd in x, the coefficients associated with even functions, a_m, will all be zero. With some complaints about the convergence, the software calculates the first six sine coefficients as:

```
-0.3071329778789654123
 0.5062509679699521912
-0.271449220623610834
-0.2561291634161970945
 0.1947511859947028479
 0.2192911372159977088
```

Let us plot the Fourier series with these six terms and compare with the function itself:



Notice that the convergence is pretty good away from the nasty singularity at x=0. In effect, truncating the Fourier series has filtered out the high-frequency oscillations.

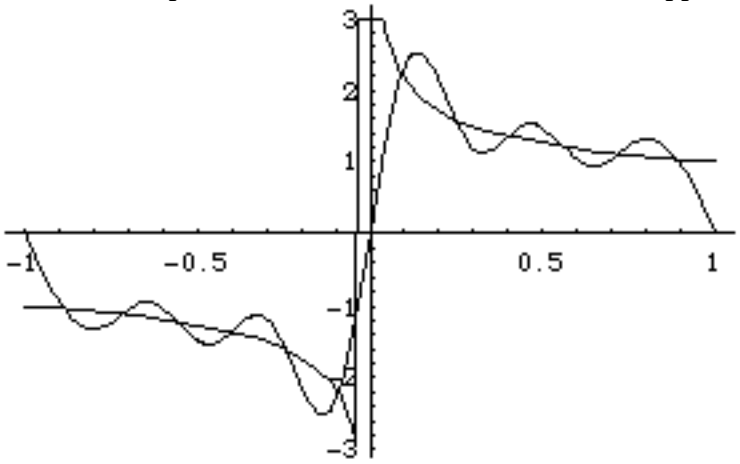
- -

Model Problem IV.5. Find the Fourier series for a function which is square-integrable but has an infinite jump.

Solution. The example we look at is $x^{1/3}$. Like the previous example, the function is odd, and there will only be sine contributions. The first several coefficients are numerically:

```
1.710689953277251171
0.3736036030974259086
0.7368949814369842807
0.2752161613201822172
0.5010701734296802681
0.2252029789384463762
```

Here is a comparison of the function and its Fourier approximation:



Think about the series calculation at x=0 and at x=±1. Did what you expected happen?

- -

As we have seen in computations, symmetries can be used to simplify the computation of Fourier series. Recall that a function is even if f(-x)=f(x) for all x, and off if f(-x) = -f(x) for all x. If f is a periodic even function, then all of the sine coefficients $b_n = 0$: b_n is the inner product of f with a sine function, but any even function and any odd function are orthogonal on the interval [-A, A]. (If f(x) is periodic with period L, then we can analyze it on any interval of length L, so even if it is originally defined on [0,L], we can just as well look at it on [-L/2, L/2]. Look at a graph of a periodic function to understand why.) For similar reasons, if f is periodic and odd, all the coefficients $a_m = 0$.

Any function can be written as the sum of an even function and an odd function, and the Fourier series picks out the two parts. If, for instance, we want to find the Fourier series of a function such as $\pi x - x^2$, $-\pi \le x \le \pi$, we can save some work by thinking about the symmetries.

Model Problem IV.6. Use symmetries to efficiently find the Fourier series for the function $\pi x - x^2$ on $-\pi \le x \le \pi$.

Solution. This function is neither even nor odd when we change x into -x, but its "odd part" is $-\pi x$, and its "even part" is $-x^2$. The sine coefficients in the Fourier series will come entirely from the odd part and the even coefficients entirely from the even part; $-x^2$ is orthogonal to all the Fourier cosine functions, and $-\pi x$ is orthogonal to all the Fourier sine

functions. Moreover, the symmetry can be used to change the integration range so that it begins at 0:

The coefficients b_n are calculated as follows:

$$b_n = \frac{2}{2\pi}\int_{-\pi}^{\pi} \pi x \sin\left(\frac{2\pi n x}{2\pi}\right) dx = \int_{-\pi}^{\pi} x \sin(nx)\, dx.$$

The product of two odd functions is even, so this integral is:

$$b_n = 2\int_{0}^{\pi} x \sin(nx)\, dx = \frac{\pi(-1)^{n+1}}{n}.$$

The final result was calculated by software. In a calculus class you would have probably calculated it by integration by parts.

The calculation of the coefficients a_n can also be simplified. First, notice that only $-x^2$ contributes:

$$a_n = \frac{2}{2\pi}\int_{-\pi}^{\pi} (-x^2) \sin\left(\frac{2\pi n x}{2\pi}\right) dx = -\frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \sin(nx)\, dx.$$

The value of this integral is

$$\frac{(-1)^{m+1}\,4}{m^2}.$$

When you combine these, you have the Fourier coefficients for $\pi x - x^2$ on this interval.

- -

Whoever was the marketing expert consulted when they invented “complex numbers” should be fired! Complex numbers almost always make things much simpler, and this is true in particular for Fourier series. The key to the simplification is Euler’s formula:

$$e^{ix} := \exp(ix) := \cos(x) + i \sin(x), \qquad (4.1)$$

It is useful to know several special cases like the above,

$$\exp(in\pi/2) = i,$$

and

$$\exp(in\pi) = (-1)^n.$$

Euler's formula allows us to discuss complex numbers not only in the Cartesian system, $z = x+iy$, but also in the polar system: $z = r \exp(i\theta)$, where θ, the argument of z, is the polar angle in the complex plans, and $r = |z|$ (the modulus of z) is the distance from the

origin. One of the truly great things about this formula is that it allows us to replace calculations with trigonometric functions with much easier calculations with exponential functions. As an example, let us think of how to calculate the Fourier series for the function f(x) = x, $0 \le x < L$ (repeated periodically, if you wish, outside this interval - notice that this is not a centered interval, so this calculation is different from the one done above.). a_0 is the average of the function, so, obviously, $a_0 = L/2$. For the other a_m, we have to calculate integrals of the form

$$\frac{2}{L}\int_0^L x \cos\left(\frac{2\pi x}{L}\right) dx \text{ and } \frac{2}{L}\int_0^L x \sin\left(\frac{2\pi x}{L}\right) dx.$$

If you don't have software or integral tables handy, you can do these integrals with integration by parts. Or you can reason as follows:

$$a_m = \frac{2}{L}\int_0^L x \cos\left(\frac{2\pi mx}{L}\right) dx = \frac{2}{L}\mathrm{Re}\left(\int_0^L x\exp\left(\frac{2\pi imx}{L}\right) dx\right).$$

Model Problem IV.7. Evaluate this integral by a method different from integration by parts.

Solution. We can use the following useful trick:

$$\int_0^L x \exp(Ax)\, dx = \frac{d}{dA}\int_0^L \exp(Ax)\, dx = \frac{d}{dA}\left(\frac{1}{A}(\exp(AL) - 1)\right). \tag{4.2}$$

After the calculation, we set $A = 2\pi im/L$. Calling the integral in (4.2) INT, we find

$$\mathrm{INT} = \frac{-1}{A^2}(\exp(AL) - 1) + \frac{L}{A}\exp(AL).$$

which starts looking nice after we substitute $A = 2\pi im/L$, which reveals that these exponential terms are both equal to 1:

$$\mathrm{INT} = \frac{L^2}{2\pi m^2 i} = -\frac{L^2}{2\pi m^2} i \quad \text{in standard form.} \tag{4.3}$$

This is purely imaginary, so all $a_m = 0$.

- -

Model Problem IV.8. Evaluate the coefficients b_n with a similar method.

Solution. STOP! Don't go back to the beginning of the calculation. We don't have to start all over, since

$$\frac{2}{L}\int_0^L x \sin\left(\frac{2\pi m x}{L}\right) dx = \frac{2}{L}\mathrm{Im}\left(\int_0^L x \exp\left(\frac{2\pi i m x}{L}\right) dx\right),$$

which is the integral we already did. From (4.3) we see that this imaginary part is just

$$b_m = -\frac{L^2}{2\pi m^2}.$$

- -

Actually, we can be much more systematic if we simply replace the complete set of functions (2.8) by (2.9):

$$\left\{\frac{1}{\sqrt{L}}\exp\left(\frac{2\pi i n x}{L}\right)\right\}_{n=-\infty}^{\infty}$$

These are simply related to the sines and cosines by Euler's formula (4.1) and its easy consequence:

$$\cos(\alpha) = \frac{\exp(i\alpha) + \exp(-i\alpha)}{2} \text{ and } \sin(\alpha) = \frac{\exp(i\alpha) - \exp(-i\alpha)}{2i}.$$

Definition IV.9. The Fourier exponential series is an expansion (for an arbitrary square-integrable function):

$$f(x) = \sum_{k=-\infty}^{\infty} c_k \exp\left(\frac{2\pi i k x}{L}\right) \qquad (4.4)$$

Since the exponential functions are an orthonormal set, a familiar kind of calculation shows us that the formula for c_k is:

$$c_k = \frac{1}{L}\left\langle \exp\left(\frac{2\pi i k x}{L}\right), f\right\rangle = \frac{1}{L}\int_a^b \exp\left(\frac{-2\pi i k x}{L}\right) f(x)\, dx. \qquad (4.5)$$

Notice the tricky minus sign - this is a place where the complex conjugate in the inner product is important.

Let's observe how much more convenient this formula is than the one without complex numbers. There is only one sum, not two sums and a constant term. There is only one formula for c_k, not separate ones for a_0, a_m, and b_n. The constant term c_0 fits in with all the others, for instance, and is not put aside as a special case. The Parseval formula (3.4) is now much simpler:

$$\langle f,g\rangle = L\sum_{k>-\infty}^{\infty} c_k\tilde{c}_k. \qquad (4.6)$$

Here, the coefficients for g have the tildes. In particular,

$$\|f\|^2 = L\sum_{k>-\infty}^{\infty} |c_k|^2. \qquad (4.7)$$

Notice that the Parseval formula is similar to the Pythagorean theorem, since, other than the normalization factor L, it states that a certain length squared is equal to the sum of the squares of its components in an orthogonal basis. We see here, however, that the space in which f lives is infinite-dimensional.

The important thing to know is that the Fourier exponential series is completely equivalent to the usual "full" Fourier series. We will later look at the Fourier sine and Fourier cosine series, which are truly different series, but the exponential series is not. If we substitute with Euler's formula, the full series becomes the exponential series or *vice versa*. We can recombine c_k with c_{-k} as follows:

$$c_k \exp\left(\frac{2\pi i k x}{L}\right) + c_{-k}\exp\left(\frac{-2\pi i k x}{L}\right) = (c_k + c_{-k})\cos\left(\frac{2\pi k x}{L}\right) + (c_k - c_{-k})\, i \sin\left(\frac{2\pi k x}{L}\right)$$

From this we get the connecting formulae that:

$$a_0 = c_0,\ a_m = c_m + c_{-m},\ \text{and}\ b_n = i\,(c_n - c_{-n}). \qquad (4.8)$$

Another useful fact to know about Fourier series is that you can normally safely differentiate them and integrate them, as long as the functions you are representing are periodic and differentiable. Actually, even if the function is not smooth, these manipulations are still often possible. This topic is explored in the next chapter.

Exercises.

Exercise IV.1. If f is a real-valued function, the coefficients a_m and b_n are real-valued. What do you conclude from (4.8) about the coefficients c_k in this case? What can you conclude about the c_k if f is an even function? if f is an odd function?

Exercise IV.2. Calculate the Fourier series for the saw-tooth function you get as the periodic extension of f(x) := 1 - |x| from the basic interval [-1,1]. Differentiate the series term by term and compare with the Fourier series for the derivative of f(x). Notice that the Fourier series is not bothered by the corners in the function at -1,0, and 1.

Exercise IV.3. Calculate the Fourier series for the saw-tooth function you get as the periodic extension of f(x) := x from the basic interval [-1,1]. Differentiate the series term by term and compare with the Fourier series for the derivative of f(x). Explain any discrepancies you find.

Exercise IV.4. Calculate the second derivatives of the Fourier series of the last two problems. Does the resulting series converge? If so, to what?

Exercise IV.5. Calculate the full and exponential Fourier series for the power functions x^2, x^3, and x^4 on a) the interval $[-\pi,\pi]$, and b) the interval [0,1].

Exercise IV.6. Calculate the exponential Fourier series forthe power functions of Exercises III.

version of 23 October 1997

V. Differentiating Fourier series†

One of the most important uses of Fourier series is as a tool in solving differential equations. Because of that we will frequently want to differentiate a Fourier series. Since the Fourier series, like any infinite series, is a limit, questions can arise about whether it is permissible to differentiate in x before summing in n. In your advanced calculus class you should have seen examples where interchanging the order of two limits leads to different answers. This can happen quite easily.

Example V.1. Let f[n,x] := 0 if $-1/n \le x \le 1/n$, and 1 otherwise - the indicator function of the complement of the interval [-1/n, 1/n]. Then f[n,0] = 0 for all n, so if we let x tend to zero before n tends to infinity,

$$\lim_{n\to\infty}\lim_{x\to 0} f[n,x] = \lim_{n\to\infty} 0 = 0,$$

whereas if n tends to infinity first,

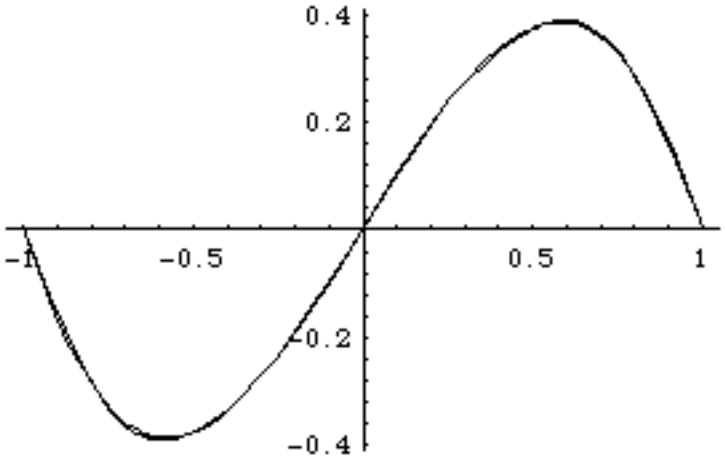
$$\lim_{x\to 0}\lim_{n\to\infty} f[n,x] = \lim_{x\to 0} 1 = 1,$$

One of the great things about Fourier series is that, despite a very reasonable worry, it is usually completely reliable to interchange limits. Indeed, one nice way to calculate Fourier series is to differentiate or integrate other Fourier series. We'll do some examples and see how reliable the answer is, as well as a couple of situations where we have to be careful.

With the formulae for the Fourier coefficients given in chapter IV it is a routine matter to calculate the Fourier series for the function $x-x^3$ defined for $-1 < x < 1$. We find that there are no cosine contributions, and

$$x - x^3 = \sum_{k=1}^{\infty} \frac{12\,(-1)^{k+1}}{(k\pi)^3} \sin(k\pi x).$$

To get a feel for this series, let us plot the sum of the first three terms and compare with $x-x^3$:



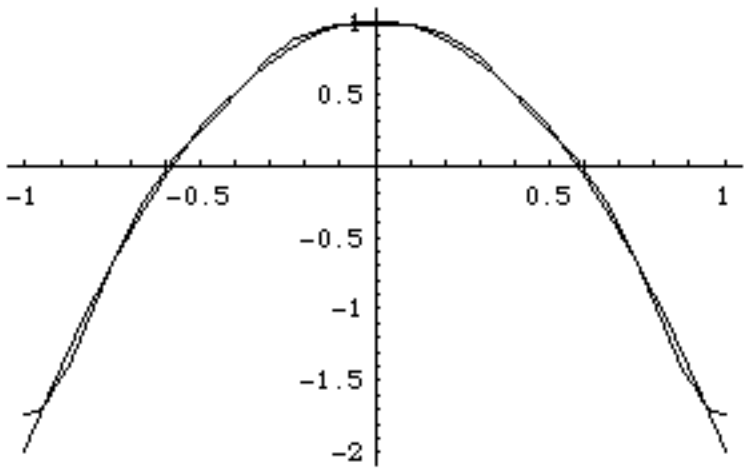
Wonderful! The function and the truncated series match rather closely.
What happens if we differentiate?

Model Problem V.1. Differentiate the terms in the Fourier series for this function, and compare with $1 - 3\,x^2$.

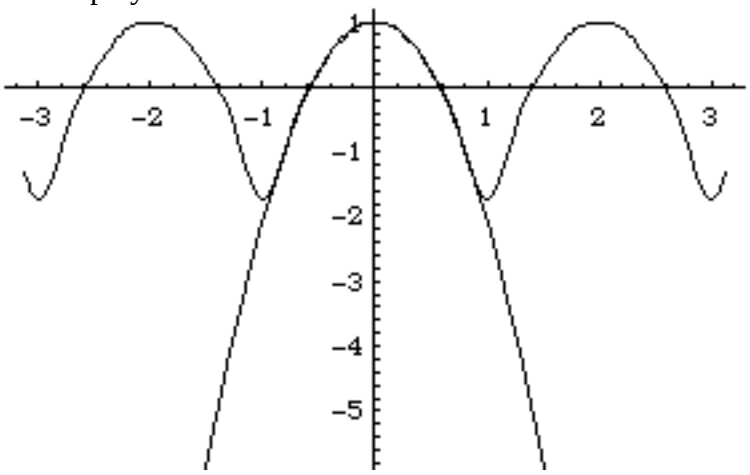
Solution.
If we differentiate the series for $x-x^3$, we get the series

$$\sum_{k=1}^{\infty} \frac{12(-1)^{k+1}}{(k\pi)^2} \cos(k\pi x),$$

supposing that it is legitimate to differentiate the infinite series term by term. (Sometimes it definitely is not.) We could calculate the Fourier series for $1-3x^2$ directly with the formulae of chapter IV, either by hand or with software, and compare, but before doing so, let's evaluate the differentiated series on its own merits. First note that the function $1-3x^2$ is even, so there will only be cosine contributions, as we have found. Next, let us plot the sum of the first four terms in the differentiated series and compare with the exact function:



A superb match! Just for fun, let's see the comparison outside the interval where we cut off the polynomials:



Remember - the Fourier series always corresponds to the periodic extension of the function on the basic interval. The formula

$$1 - 3x^2 = \sum_{n=1}^{\infty} \frac{12\,(-1)^{k+1}}{(k\pi)^2} \cos(k\pi x)$$

is not valid for the series outside [-1,1].

Finally, let us calculate the Fourier series using the integral formulae. The result, using software, is:

$a_0 = 0$,

$a_k = \frac{12\,(-1)^{k+1}}{(k\pi)^2}$, k = 1, 2, ...,

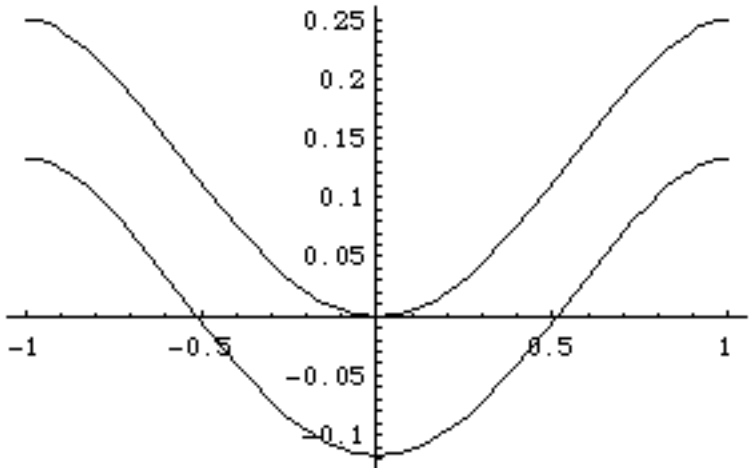
as expected. (Notice that this example has something to do with the Legendre polynomials.)

Suppose now that instead of differentiating, we integrate a Fourier series term by term. If a_0 (which = c_0) is different from 0, we get another Fourier series! If a_0 is not 0, then we would only get another Fourier series after replacing the function x with a Fourier series, but we won't consider that case now.

Does the integrated series converge to the integral of the original function? According to our experiment with differentiation, it seems so. Let us try again, by integrating the series for $x - x^3$ term by term. We find:

$$\sum_{n=1}^{\infty} \frac{12(-1)^k}{(k\pi)^4} \cos(k\pi x)$$

plus a constant of integration. Let us plot the sum of the first 4 of these functions and compare with $x^2/2 - x^4/4$:



What does this show about the constant of integration? When we integrated the series we got the integral with average average a_0, chosen here as 0. Remember that a_0 is always the average of the function. The difference in height between the two graphs is just the average of $x^2/2 - x^4/4$, which doesn't happen to be 0.

Exercises.

Exercise V.1. Use the tricks of this chapter to calculate the Fourier series for the functions x^n, n = 0, 1, ..., on the interval $-1 < x < 1$.

Exercise V.2. What series do you obtain by differentiating the Fourier series for the square pulse (cf. model problem IV.1)? If you know about the Dirac*delta function*, comment on the relationship between the series you get for the square pulse and the one for the delta function $\delta(x-a)$ (periodically extended).

Exercise V.3. Consider what happens when a Legendre series is differentiated term by term. Is the result a Legendre series? Does the result converge to the derivative of the original function?

Exercise V.4. Check the convergence of Exercise V.3 numerically in a specific example such as $f(x) = \sin(\pi x)$, $-1 < x < 1$.

version of 8 December 1997

VI. Notes (played) on the vibrating string†

There are three basic models in the subject of partial differential equations, which describe very different physics and have very different mathematical properties. They are simple in form, yet realistic for many problems in science and engineering. Moreover, they teach us what to expect when we encounter more complicated equations used to model complicated laboratory situations. They are:

The wave equation,

$$u_{tt} = c^2 u_{xx}; \qquad \text{(WE)}$$

the heat equation, also known as the diffusion equation,

$$u_t = k u_{xx}; \qquad \text{(HE)}$$

and the potential equation, also known as Poisson's equation,

$$u_{xx} + u_{yy} = \alpha \rho(x,y), \qquad \text{(PE)}$$

where ρ is a given function corresponding in electrostatics to a charge distribution, and the other quantities c, k, and α are physical constants. Often there is no charge in the region where you want to measure a potential , and (PE) reduces to Laplace's equation

$$u_{xx} + u_{yy} = 0. \qquad \text{(LE)}$$

One thing you are likely to notice if you take several advanced courses in science or engineering is that these same equations arise in many different contexts. The function u

may have a different interpretation in each incarnation, but the mathematics remains the same. For example, the function u in (PE) or (LE) might be an electric potential, a magnetic potential, the height of a taut membrane, or the potential of the velocity field in a fluid.

The reason these equations arise again and again is that they are the most symmetric second order partial differential equations that can be written down. Fundamental physical principles usually do not involve derivatives of higher than second order. If a physical model is isotropic, has a superposition principle, and is based on fundamental physical principles, it is very likely to lead to an equation of one of these forms. In the appendix, however, you will find a derivation of the wave equation for the problem of the vibrating string.

The mathematics and physics of the three types of equations differ in important ways, and lead to a classification of the equations as *hyperbolic* (WE), *parabolic* (HE), and *elliptic* (PE/LE). Other second-order equations of a given type will share the qualitative features of our three models.

We shall begin our study of the method of separation of variables with the wave equation, which was in fact the first of these equations to be analyzed in this way. The wave equation (WE) can be written with the aid of a *wave operator*

$$\Box\, f(t,x)) := \left(\frac{1}{c^2}\,\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right) f(t,x)),$$

so that (WE) is the equation for the kernel of this operator. This operator is also known as the d'Alembertian. The wave equation is a linear, homogeneous equation, so the superposition principle holds for its solutions. If we can find some solutions $u_1, \ldots, u_k$, then any linear combination of these solutions is again a solution. We also classify it as "second-order," since derivatives up to second order occur, but no higher-order derivatives. This will be the case for the great majority of PDE's in this class.

We shall solve the wave equation in two entirely different ways. The first method is known as separation of variables. Let's set c = 1 for convenience. (Equivalently, we could choose a clock with time variable t' = c t.) Let's test out the wave operator by calculating its action on some representative functions:

- $\square\ t \sin^2(x) = -2t \cos^2(x) + 2\, t \sin^2(x);$
- $\square\ \sin(x - t) = 0.$

Somewhat surprisingly, the second seemingly arbitrary function is a solution of the PDE! As we shall see, there are lots of them, and several strategies for finding them.

With the method of separation of variables, we try to find solutions of a special form, u = T(t) X(x). This is a special assumption of form - what is referred to as an *ansatz*, to give some definiteness to our search for solutions. Some other physical circumstances, called the boundary and initial conditions, will be quite important, and will constrain us in choosing T and X; a partial differential equation coupled with some boundary conditions will be called a *boundary-value problem* . For the vibrating string, the most common boundary conditions are the zero Dirichlet conditions, named after the German mathematician Gustav Lejeune-Dirichlet (1805-1859), a pioneer in the study of PDEs:

$$u(t,0) = 0, \quad u(t,L) = 0 \qquad \text{(DBC)}$$

This means that the string is fastened at two ends, x=0 and x=L. Actually, if we assigned some other fixed values to u besides 0 at the end points, we could just redefine the dependent variable:

If $u(t,0) = a$, $u(t,L) = b$,

then define $v := u(t,x) - a - (b-a)\, x/L$.

Then v will still solve the wave equation, but will have the simpler boundary conditions (DBC). When we act on a product T(t)X(x) with the wave operator and divide the result by T(t)X(x), we get

$$\frac{T''(t)}{T(t)} - \frac{X''(t)}{T(t)}.$$

If u solves the wave equation, this expression must be zero. (We could quibble about what happens when T or X = 0, but that would either mean that u is the uninteresting "trivial" solution or else the problem occurs only at isolated values of x and t.) Thus

$$T''/T = X''/X.$$

Since the left side is independent of x and the right side is independent of t, both sides must in fact a constant. We don't know its value yet, so let us call it $-\mu$. (It turns out to be negative.

Thus

$$-X''(x) = \mu X(x) \qquad (6.1)$$

This sort of equation is similar to a matrix-vector equation in linear algebra and is called an *eigenvalue equation*. The function X and the number μ are both unknown, and are called and eigenfunction and eigenvalue. We do not allow the possibility X(x) = 0 for all x. While it is always a solution for all μ, it is not useful. By definition, an eigenfunction must be a nonzero solution of an equation like (6.1). Along with (6.1) we have the boundary conditions

$$X(0) = X(L) = 0 \qquad (6.2)$$

and this will restrict the μ's for which (6.1) has nonzero solutions. Eq. (6.1) is an elementary ordinary differential equation with the familiar solution (also obtainable with software),

$$X(x,\mu) := D_1 \cos(\sqrt{\mu}\, x) + D_2 \sin(\sqrt{\mu}\, x).$$

Because of the boundary condition at x=0, we see that $D_1 = 0$. We are left with only one constant, which we can scale to 1, remembering that we may want to multiply our solution by an arbitrary constant later. Because of the BC at x=0, we now know that up to a constant multiple, $X(x) = \mathrm{Sin}(\sqrt{\mu}\, x)$. The other boundary condition in (6.2) tells us what the allowable choices for μ are, namely $\sqrt{\mu} = n\pi/L$, n = 1,2,.... Taking L=1 for convenience, the eigenfunctions and eigenvalues are:

$$X_n(x) := \sin(n \pi x)$$

$$\mu_n := (n \pi)^2$$

Now let's look at the other part of the PDE which we "separated off":

$$T''(t) = -\mu T(t) \qquad (6.3)$$

As we saw, there are many possible values of μ which are consistent with the boundary condition for the function X, and the same possibilities occur in equation (6.3) (multiply μ by c^2 if c is not set to 1). Therefore $T_n(t) = A_n \cos(n \pi t) + B_n \sin(n \pi t)$. We cannot go

further with the T function without being given the *initial condition*, that is, information about what happens at t=0.

We already have something of importance in science and engineering, however: the *normal modes* . Many physical systems, like the vibrating string, preferentially undergo simple motions with pure frequencies. Such a possible motion is exactly what we find by positing a product solution, imposing any boundary conditions, and solving the resulting eigenvalue problem.

Definition VI.1.
A *normal mode* of a boundary value problem is a product solution of the form

$$T(t)\, X(\mathbf{x}),$$

where $X(\mathbf{x})$ incorporates all boundary conditions.

Here the spatial variable **x** might be more than one-dimensional, as for the problem of the vibrating drum, which will be encountered later. I have generalized the notion of a normal mode so that any time dependence is allowed, although most commonly it will be sinusoidal, $T(t) = \sin(\omega t + \varphi)$. When it is sinusoidal, the angular frequency ω is usually a simple function of the eigenvalue.

The general solution we come up with is a linear combination of the particular solutions we get by separating the equation. It is not yet obvious that this is a completely general solution, but it is.

$$u(t,x) = \sum_{n=1}^{\infty} (A_n \cos(n \pi t) + B_n \sin(n \pi t)) \sin(n \pi x), \qquad (6.4)$$

The coefficients A and B have to be determined from the initial conditions, that is, by the condition of the string at t=0.

Suppose that

$$u(0,x) = f(x) \text{ and } \partial u(0,x)/\partial t = g(x) \qquad (IC)$$

are given. Plugging in to formula (6.4) shows that:

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

$$g(x) = \sum_{n=1}^{\infty} n\pi B_n \sin(n\pi x),$$

In other words, A_n are the Fourier sine coefficients for f(x), and the more complicated expression

$$(n\pi) B_n$$

gives the Fourier sine coefficients of g(x). Notice that the stuff in the sine is $n\pi x/L$ and not $2n\pi x/L$ (with L=1). In other words, there are more sines in this series than in the usual Fourier series, but there are no cosines. It is precisely as if we imagine the string to be defined on the interval [-1,1], twice as long as the actual interval, but on the nonphysical part from x=-1 to x=0 we pretend that the shape of the string is defined to give us an odd function: f(-x) = -f(x). This would give us only sine functions, as we have seen before. Earlier we called those coefficients B_n, but here they happen to be called A_n. The thing to remember is that since they go along with sine functions (of x) the formula for them contains sines. we would thus write

$$A_n[f] := \int_{-1}^{1} \sin(n\pi x)\, f(x)\, dx,$$

but we can simplify this a bit by using the symmetry:

$$A_n := 2\int_{0}^{1} \sin(n\pi x)\, f(x)\, dx.$$

What about the other coefficients? Well, this time it is the whole expression $(n\pi) B_n$ which is a Fourier sine coefficient for g(x), so we need to divide through by the extra factor:

$$B_n[g] := \frac{2}{n\pi}\int_{0}^{1} \sin(n\pi x)\, g(x)\, dx.$$

Here is an example to illustrate the situation.

Model Problem VI.2. Solve the problem of the vibrating string using Fourier sine series, and initial conditions $u(x,0) = 0$, $u_t(x,0) = 1$.

Solution.

We can easily calculate the formulae for the coefficients A and B, either by hand or with software:

$$A_m = 0$$

and

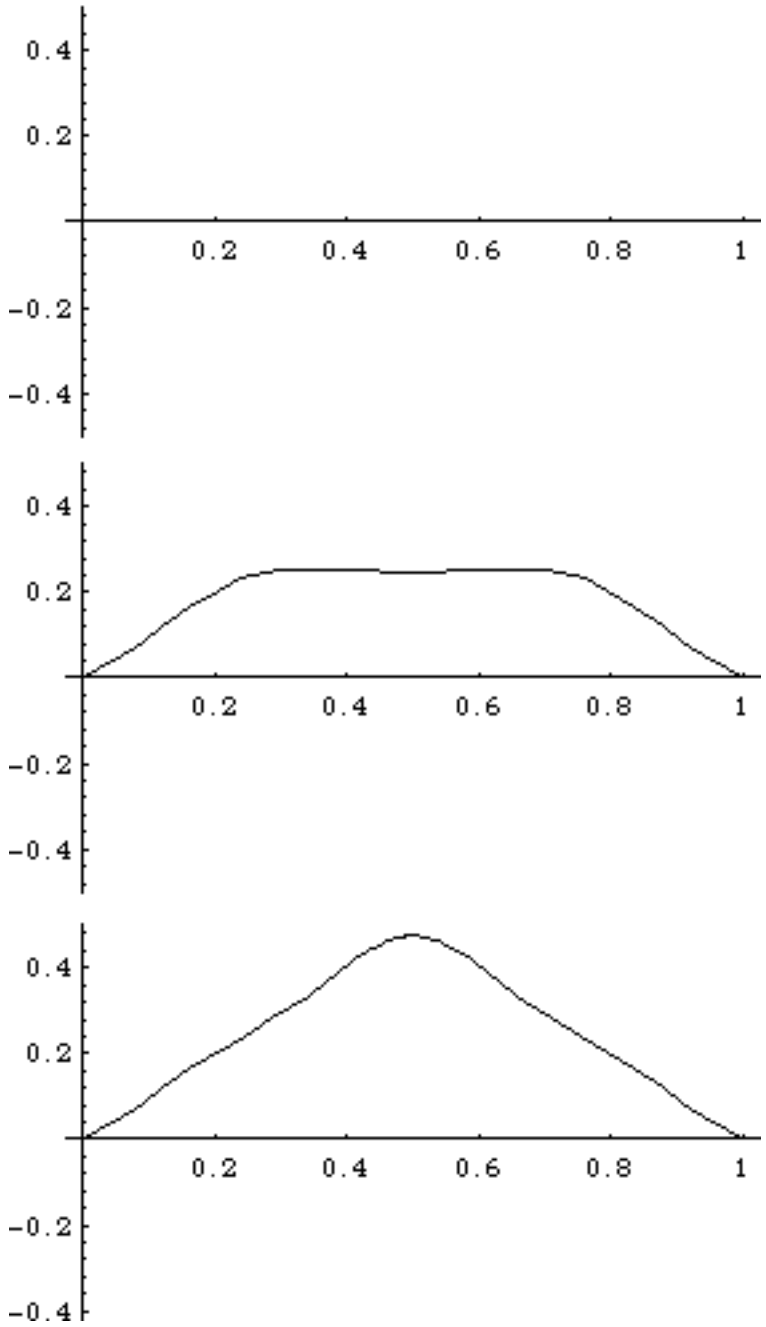
$$B_n = 2\,\frac{1-(-1)^n}{(n\pi)^2}.$$

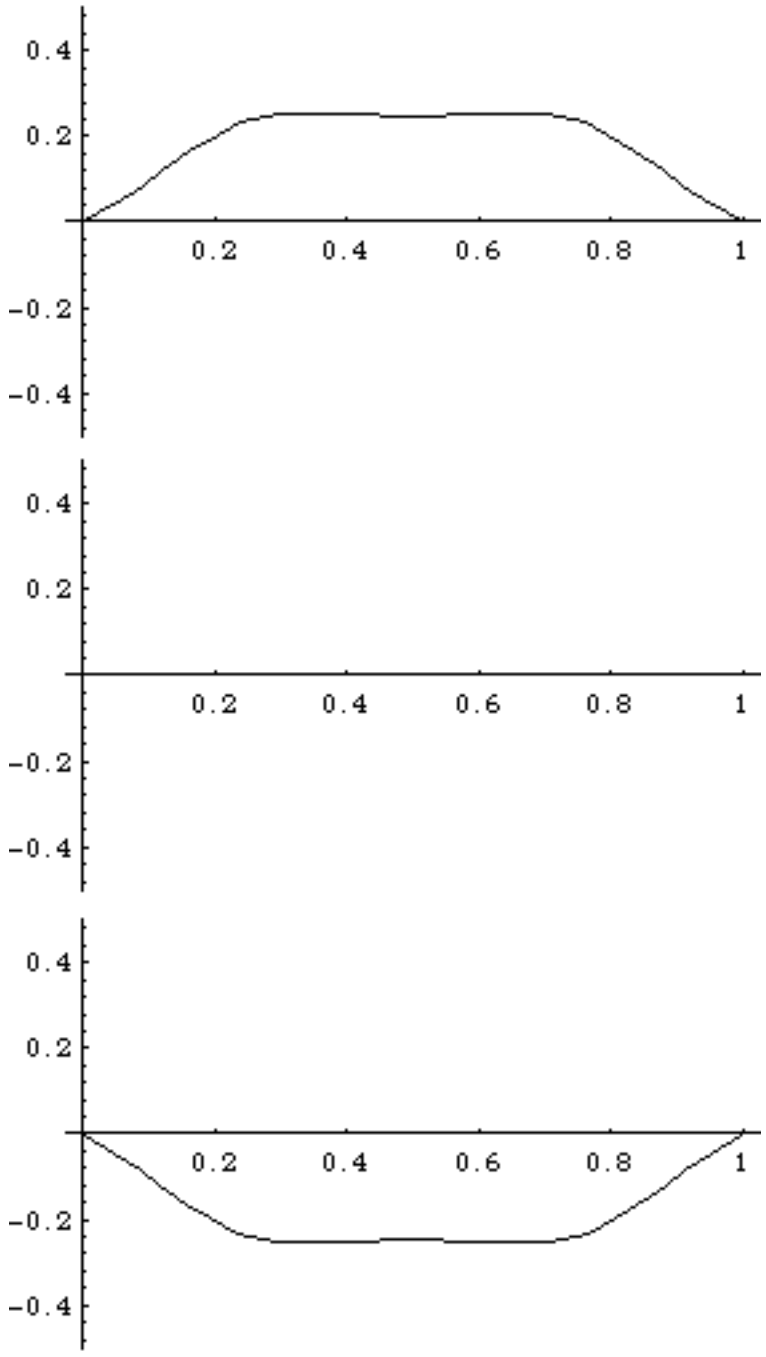
If n is even, $B_n = 0$, else $B_n = \frac{4}{n^2\,\pi^2}$.

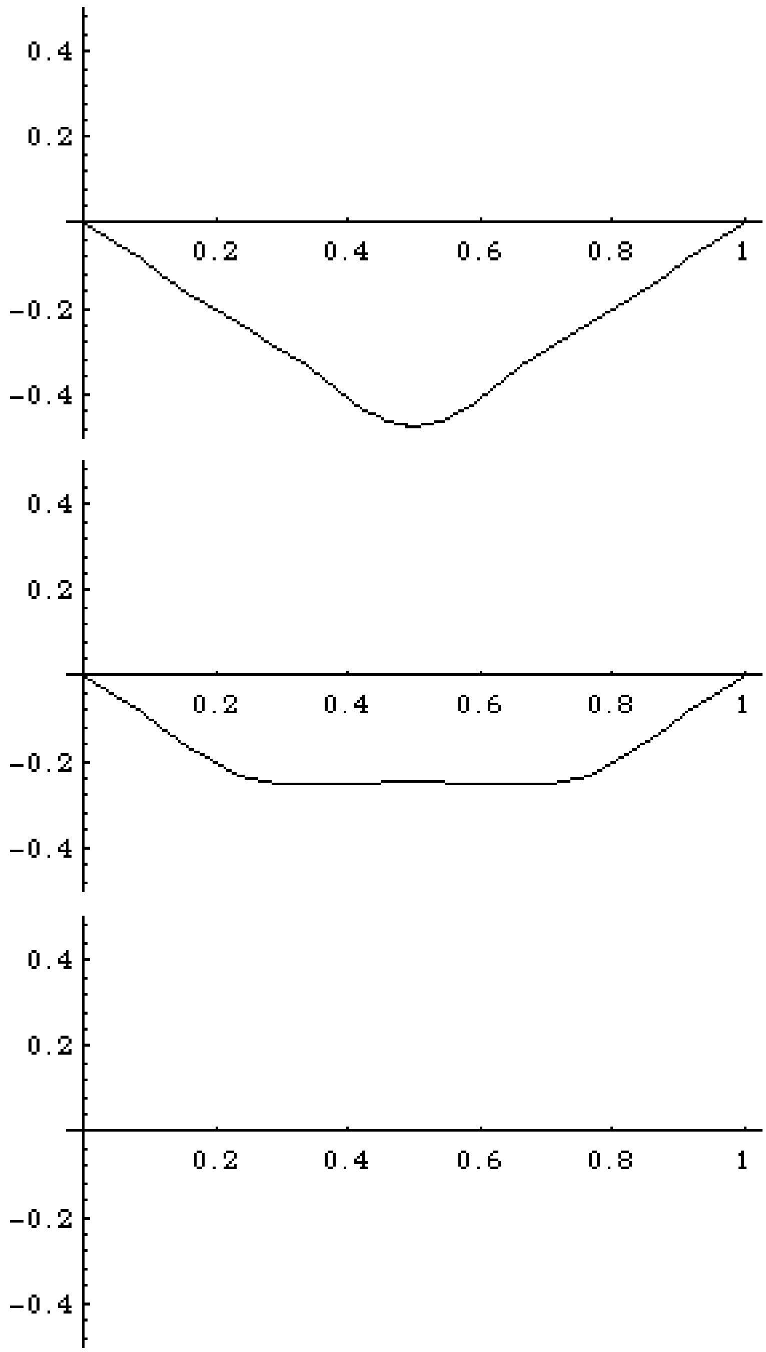
The general solution is:

$$u(t,x) = \frac{1}{4\pi^2} \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{\sin(n\pi t)\,\sin(n\pi x)}{n^2}.$$

Next we show some plots of the solution at various times. For computational purposes we keep terms only up to n=7. With software you can see the string vibrate by selecting the graphs and choosing "animate." By the way, infinity is larger than 7, but 7 is close enough to give us a reasonable picture of the solution:


0.4
0.2
0.2
0.4
0.6
0.8
1
−0.2
−0.4
0.4
0.2
0.2
0.4
0.6
0.8
1
−0.2
−0.4
0.4
0.2
0.2
0.4
0.6
0.8
1
−0.2
−0.4


0.4
0.2
0.2
0.4
0.6
0.8
1
−0.2
−0.4
0.4
0.2
0.2
0.4
0.6
0.8
1
−0.2
−0.4
0.4
0.2
0.2
0.4
0.6
0.8
1
−0.2
−0.4



- -

The boundary conditions we have been using are not the only useful ones. Suppose that our string is attached to a pair of tracks in the vertical direction perpendicular to the x-axis using a frictionless roller. Then u(t,0) and u(t,L) are not specified, but instead we have:

$$\frac{\partial u}{\partial x}(t,0) = 0,$$
$$\frac{\partial u}{\partial x}(t,L) = 0, \qquad \text{(NBC)}$$

These are called Neumann boundary conditions. after the German physicist F.E. Neumann (1798-1895), who used them in his investigations of electricity and magnetism. Obviously, as with DBC, it would be a minor change to impose them at x=a and x=b where a does not have to be 0. At this stage, we can exactly repeat the analysis of separation of variables, until the point where we first used the boundary conditions, i.e., (6.1). With NBC, we instead evaluate the x-derivative at 0:

Model Problem VI.3. Set up the general solution of the vibrating string with Neumann BC., with the simplifications c = L = 1.

Solution.
The partial differential equation separates exactly as before, so we still know that when we assume u = X(x) T(t), we get separated equations (6.1) and (6.3). The values of μ will be different, however, because they were determined by the BC. Since as before,

$$X(x,\mu) := D_1 \cos(\sqrt{\mu}\, x) + D_2 \sin(\sqrt{\mu}\, x),$$

we see that

$$X'(x,\mu) := -\sqrt{\mu}\ D_1 \sin(\sqrt{\mu}\, x) + \sqrt{\mu}\ D_2 \cos(\sqrt{\mu}\, x),$$

and for this to be 0 when x=0, there are two possibilities, namely D_1=0 or μ=0. (We ignored the latter possibility before since sin(0) = 0 is not a useful solution, but cos(0) = 1 can be one.)

Possibility 1. μ=0, which means that the output just given is not really the solution to (6.1), which has become X''(x) = 0. The solutions of this are of the form A + B x, but B has to be 0 so that X'(0) = 0. Thus with this possibility, X(x) = 1 (times any constant). Eq. (6.3) now also becomes T''(t) = 0, which leads to the product solution of (WE):

$$(A_0 + B_0\, t)\, 1 = A_0 + B_0\, t.$$

Since the boundary conditions do not apply to the t variable, we keep the general linear combination. A_0 and B_0 will be determined by the initial conditions.

Possibility 2. $\mu > 0$. The boundary condition at x=0 now forces $X(x) = \cos(\sqrt{\mu}\, x)$ (again, up to a constant multiple), and the other boundary condition forces one of the values $\mu_n = (n\,\pi)^2$. The product solutions become

$$(A_n \cos(n\,\pi\, t) + B_n \sin(n\,\pi\, t)) \cos(n\,\pi\, t)$$

and the general solution will be

$$u(t,x) = A_0 + B_0\, t + \sum_{n=1}^{\infty} \left(A_n \cos(n\,\pi\, t) + B_n \sin(n\,\pi\, t)\right)\cos(n\,\pi\, x). \qquad (6.5)$$

The coefficients are determined by the initial conditions because:

$$u(0,x) = A_0 + \sum_{n=1}^{\infty} A_n \cos(n\,\pi\, x),$$

and

$$u_t(0,x) = B_0 + \sum_{n=1}^{\infty} n\,\pi\, B_n \cos(n\,\pi\, x).$$

The numbers A_n are the Fourier cosine coefficients of $u(x,0) = f(x)$, while B_0 and $n\pi B_n$ are the Fourier cosine coefficients of $u_t(x,0) = g(x)$. (Notice the extra factors of $n\pi$.)

- -

A realistic string or optical fiber may not be uniform, so some of the simplifying assumptions in our derivation fo the wave equation are not valid. For example, the constant c^2 = (spring constant) / (mass density) for the spring may be different at different positions, leading to an equation of the form

$$\frac{\partial^2 u}{\partial t^2} = \left(c(x)\right)^2 \frac{\partial^2 u}{\partial t^2}. \qquad \text{(ModWE1)}$$

or, as a thorough examination of the derivation of the wave equation reveals, we could more generally have:

$$\frac{\partial^2 u}{\partial t^2} = p(x)\, \frac{\partial}{\partial x}\left(s(x)\frac{\partial u}{\partial x}\right). \qquad \text{(ModWE2)}$$

for some potentially complicated positive functions p(x) and s(x). How well does the method of separation of variables do for problems like this? Rather well, actually, although we may have to encounter some new functions.<P>

Model Problem VI.4. Suppose that the wave speed depends on position, so that

$c^2 = 1/(1 + x)$, $0 < x < 1$, with DBC at 0 and 1.

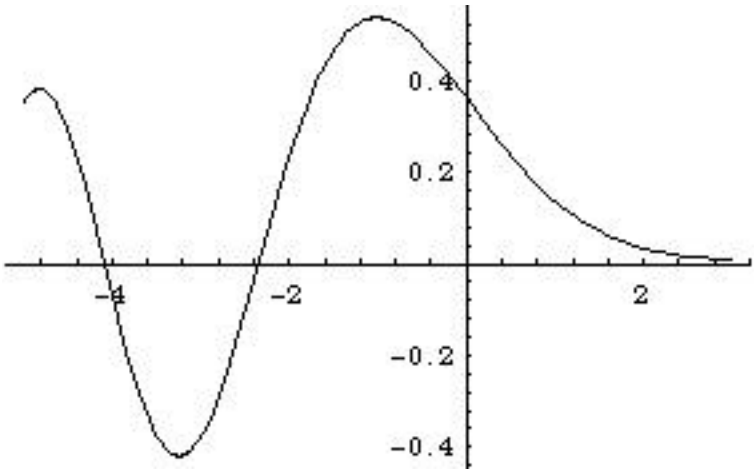
Find the normal modes of vibration

Solution (using software).

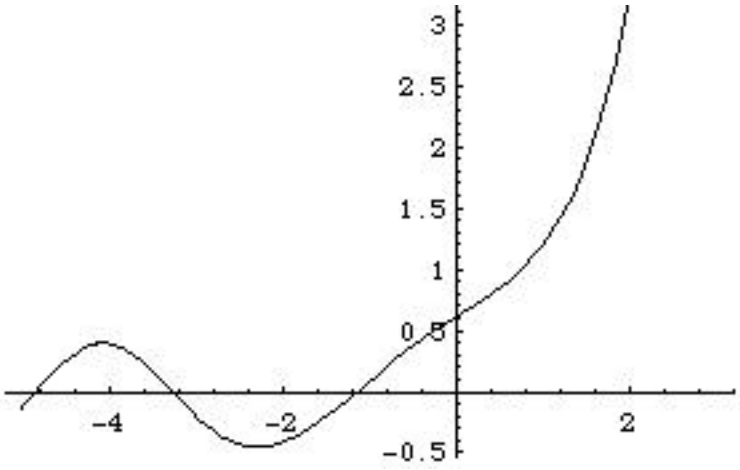
When we attempt to solve the equation with the ansatz $u(t,x) = T(t)\, X(x)$, we find the following eigenvalue problem.

$$- X''(x) = (1 + x)\, \mu\, X(x)$$

There are actually some special functions, called Airy functions, which solve the ODE $y''(x) = x\, y$. Two independent solutions are called Ai(x) and Bi(x):



The Airy function Ai(x).



The Airy function Bi(x).

The function Bi explodes exponentially to the right, while Ai decays exponentially. They both oscillate to the left (why?). By changing variables we can get these functions to solve our eigenvalue equation: If

$$y(x) := Ai(-\mu^{1/3}(x+1)),$$

then

$$y''(x) = -\mu(1+x)\,y(x),$$

and similarly for $Bi(-\mu^{1/3}(x+1))$.

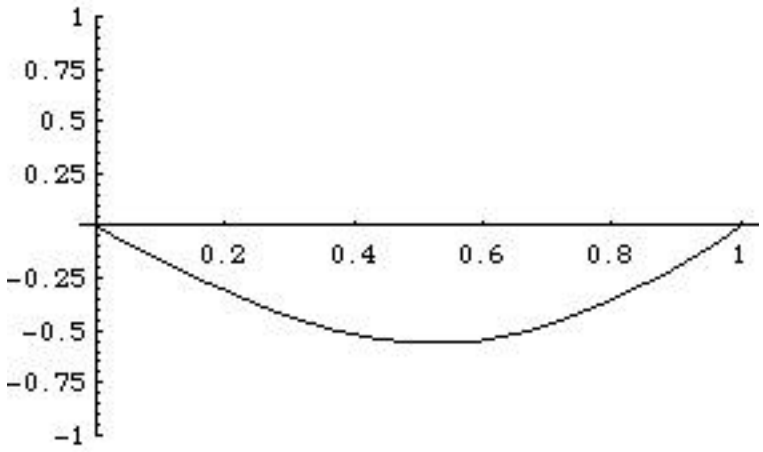
We need a linear combination

$$u(x) = Ai(-\mu^{1/3}(1+x)) + C\,Bi(-\mu^{1/3}(1+x))$$ which is 0

at x=0 and 0 at x=1. Mathematica or Maple can solve for these conditions numerically, and we find:

$$\mu^{1/3} = 1.87088,\ C = 0.819688.$$

A plot of the function u shows that it has no nodes between 0 and 1:



It thus resembles $\sin(\pi x)$, indicating that this is the spatial part of the fundamental (lowest-frequency) mode. The eigenfunction and eigenvalue are numerically:

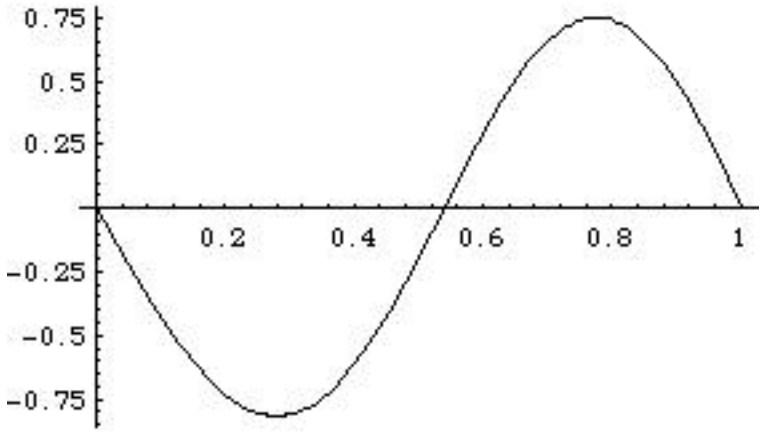
```
Ai(-1.87088 (1 + x)) + 0.819688 Bi(-1.87088 (1 + x)),
```
$\mu = 6.54844$

The time-dependence is a combination

$A_0 \cos(6.54844\, t) + B_0 \sin(6.54844\, t)$.

The second eigenfunction and eigenvalue correspond to a solution which changes sign once between 0 and 1:

```
Ai(-2.98005 (1+x)) -1.75913 Bi(-2.98005 (1+x))
```
$\mu = 26.4649$



The second eigenfunction

Exercises VI

VI.1. Solve the wave equation with NBC at 0 and 1 and:

$$f(x) = |x-1/2|,\ g(x) = x^2$$

Do you encounter any difficulties because the IC are not consistent with the BC?

VI.2. Put DBC at x=0 and NBC at x=L. Separate variables, solve the eigenvalue problem for X(x), and write down a general solution.

VI.3. Solve the wave equation $u_{tt} = u_{xx}$ with IC:

$$f(x) = \sin(\pi x),\ g(x) = 0$$

and

a) DBC at x=0 and x=1
b) NBC at x=0 and x=1
c) DBC at x=0 and NBC at x=1

Plot the solutions and compare.

VI.4. Solve the wave equation with IC:

$$f(x) = 0,\ g(x) = 1$$

and

a) DBC at x=0 and x=1
b) NBC at x=0 and x=1
c) DBC at x=0 and NBC at x=1

Plot the solutions and compare.

VI.5. Solve the wave equation with IC:

$$f(x) = \sin(\pi x),\ g(x) = 1$$

and

a) DBC at x=0 and x=1
b) NBC at x=0 and x=1
c) DBC at x=0 and NBC at x=1

Plot the solutions and compare.

VI.6. Solve the wave equation $u_{tt} = \mathbf{4}\, u_{xx}$ for $0 < x < \pi$, $0 < t < \infty$, initial conditions:

$$u(0,x) = 0, \quad u_t(0,x) = x(\pi - x) = \frac{8}{\pi} \sum_{n\ \mathrm{odd}} \frac{1}{n^3} \sin(nx),$$

and

Dirichlet boundary conditions at x=0 and $x=\pi$, i.e., $u(t,0) = u(t,\pi) = 0$, for $0 < t$.

VI.7. Find all normal modes (product solutions) for the modified wave equation

$$u_{tt} + u = u_{xx}$$

with Neumann boundary conditions:

$u_x(t,0) = 0$

$u_x(t,\pi) = 0$

VI.8. (Refer to the appendix for the derivation of the wave equation for this problem.) Derive a wave equation for the vibrating string if the x-axis is horizontal and we take gravity into account. The gravitational potential energy of the string is

$$\int_a^b g\, u(t,x)\, dx,$$

where g, the gravitational constant, is approximately 980 in the cgs system.

Appendix. Derivation of the wave equation.

The physics and mathematics of the vibrating string were studied by Jean le Rond d'Alembert and later by Joseph Louis Lagrange, Leonhard Euler and Daniel Bernoulli, who gave the first satisfactory discussion of the physics of the vibrating string. I have not been able to locate a detailed discussion of Bernoulli's derivation of the wave equation, but it is likely that he based it on an energy principle, somewhat as follows.

First let us recall that in classical mechanics, Newton's equations for the motion of a particle are equivalent to the vanishing of the first variation of the Lagrangian action integral,

$$L := \int_C (\mathrm{KE} - \mathrm{PE})\,dt.$$

The kinetic energy is usually of the form

$$\mathrm{KE} = \frac{m\,|\mathbf{v}|^2}{2},$$

while the potential energy PE may be a function of position, U(**q**). (I call the position **q** to avoid confusion with an x used below). The action integral is calculated along a trajectory C = (**q**(t)) beginning at position and time coordinates ($\mathbf{q}_0$, t_0) and ending at coordinates ($\mathbf{q}_1$, t_1). The trajectory chosen by physics is one which is stationary with respect to variations of the path. In other words, if we replace C with

$$C_\delta = (\mathbf{q}(t) + \delta\,\mathbf{h}(t)),$$

where **h**(t) is a smooth function with $\mathbf{h}(t_0) = \mathbf{h}(t_1) = 0$, and we calculate $L(\delta)$, then

$$L'(0) = 0.$$

(You may hear that the physical trajectory *minimizes* the action, but this is only a necessary condition for minimum, and the physical trajectory is not always an actual minimum.) If we expand L in powers of δ and retain only the first-order term, the stationary condition becomes

$$\int_{t_0}^{t_1} (m\,\dot{\mathbf{q}}(t)\;\dot{\mathbf{h}}(t) - \nabla U(\mathbf{q}(t))\;\mathbf{h}(t))\,dt = 0,$$

and if we integrate by parts, it is:

$$\int_{t_0}^{t_1} \mathbf{h}(t)\;(-m\,\ddot{\mathbf{q}}(t) - \nabla U(\mathbf{q}(t)))\,dt = 0.$$

(The boundary terms vanish because of the conditions $\mathbf{h}(t_0) = \mathbf{h}(t_1) = 0$, which served to fix the beginning and end of the trajectory.) If, now, we let the three components of **h**

range over a complete set such as $\sin(n\pi t/L)$, with $L = t_1 - t_0$, we see that the only possibility is for

$$m\ddot{\mathbf{q}}(t) = \mathbf{F} = -\nabla U(\mathbf{q}),$$

which is Newton's law.

The big advantage of Lagrangian mechanics is that it allows us relatively easily to find the equations of motion of an extended body, such as a string. Suppose now that we have a taut string along the x-axis between positions a and b, but displaced laterally by an amount u(t,x). If the density is ρ, then the kinetic energy of a small bit of string at position x is

$$dKE = \frac{\rho |u_t(t,x)|^2}{2}\, dx,$$

so the total kinetic energy is

$$KE = \int_a^b \frac{\rho |u_t(t,x)|^2}{2}\, dx.$$

It is plausible that the microscopic force transmitted by the string to a point x is proportional to the amount by which the string is stretched at x, i.e., it depends on the arc length element ds at x. As we know,

$$ds = \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2}\, dx.$$

Thus we may assume that the differential potential energy depends on u only through $\partial u/\partial x$. The total potential energy would be of the form

$$PE = \int_a^b F(u_x^2, x)\, dx.$$

The Lagrangian action integral will be the integral of KE - PE with respect to time. Notice that it is u and not x which corresponds to the **q** used above; the "trajectory" of the string is specified by the function u(t,x), and a variation would entail replacing u(t,x) with

$$u(t,x) + \delta h(t,x),$$

where $h(t_0,x) = h(t_1,x) = 0$. If the ends of the string are fixed, we would also have $h(t,a) = h(t,b) = 0$. The action is a double integral,

$$L = \int_{t_0}^{t_1} \int_a^b \frac{\rho u_t^2}{2} - F(u_x^2, x)\, dx\, dt,$$

and if we use Taylor's theorem to keep only the first-order term in δ, Lagrange's condition reads:

$$0 = \int_{t_0}^{t_1} \int_a^b \rho\, u_t h_t - F'(u_x^2, x)\, 2u_x h_x\, dx\, dt.$$

This general formula may be useful in deriving realistic wave equations for non-homogeneous strings, but let us simplify at this stage by assuming that ρ and $T := 2\,F'$ are constants. If we integrate the first integrand by parts in the t variable and the second by parts in the x variable, we now find that:

$$0 = \int_{t_0}^{t_1} \int_a^b h\left(-\rho\, u_{tt} + T\, u_{xx}\, dx\right)\, dt.$$

If h is arbitrary enough to run through a complete set, we must conclude that

$$u_{tt} = c^2\, u_{xx}, \qquad \text{(WE)}$$

with $c^2 = T/\rho$ a positive constant with dimensions of velocity^2.

Judging from guitar strings, for which a 1 meter taut string gives a musical note in the mid range of the musical scale, typical values of c for thin metal strings are on the order of 1000 m./sec. The wave equation (WE) also describes one-dimensional acoustic waves ($c \approx 344$ m/sec. in air at room temperature or 330 m./sec. at 0 C) and light waves ($c \approx 3 \times 10^8$ m./sec.), although the derivation is not the same in these cases.

version of 10 November 1996

VII. Traveling waves and the method of d'Alembert.†

This is an interlude from our study of wave equations by the method of separation of variables. For the standard wave equation

$$\frac{\partial^2 u(t,x)}{\partial t^2} = c^2 \frac{\partial^2 u(t,x)}{\partial x^2},$$

where c is a constant, there is a completely different-looking method of solution, due to the French mathematical physicist Jean le Rond d'Alembert (1717-1783). Suppose φ is a function of one variable z, but we set z = x + ct. Then $\varphi(x+ct)$ represents a fixed wave-form which moves to the left at speed x as time goes on. Provided that φ can be differentiated twice, $\varphi(x+ct)$ is automatically a solution of the wave equation, which we can rewrite as in chapter VI in terms of a wave operator as $\Box u = 0$. (Recall that by definition,

$$\Box f(t,x)) := \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right) f(t,x)),$$

for any t,x and suitable f.) It is not hard to verify this fact with the chain rule, and you may recall that we saw a particular example in chapter VI.

Since the wave equation is linear, given any two twice differentiable functions of a single variable, the expression $\varphi(x+ct)+ \psi(x-ct)$ is also a solution. Suppose now that we have an infinitely long string, and initial conditions of the usual form:

$$u(0,x) = f(x)$$
$$\partial u(0,x)/\partial t = g(x).$$

We do not need boundary conditions for this problem, although sometimes it is physically important to have "boundary conditions at infinity" that as $x \to \pm\infty$, $u \to 0$, so that the wave energy is finite.

Can we always find a solution of the form

$$u(t,x) = \varphi(x+ct) + \psi(x-ct), \qquad (7.1)$$

which solves this initial-value problem? Yes indeed! As an example, suppose g(x) := 0. Then, substituting t=0 shows that

$$\varphi(x)+ \psi(x) = f(x)$$

while, since $\partial u(0,x)/\partial t = c\, f'(x+ct)- c\, y'(x-ct)$,

$$c\, \varphi'(x)- c\, \psi'(x) = 0.$$

This last equation can be integrated to give $\psi(x) = \varphi(x)+ C$, but we take this constant to be zero (if $f(x) \to 0$ at infinity, this is forced on us by the BC at ∞). Thus:

$$\varphi(x) = \psi(x) = f(x) /2.$$

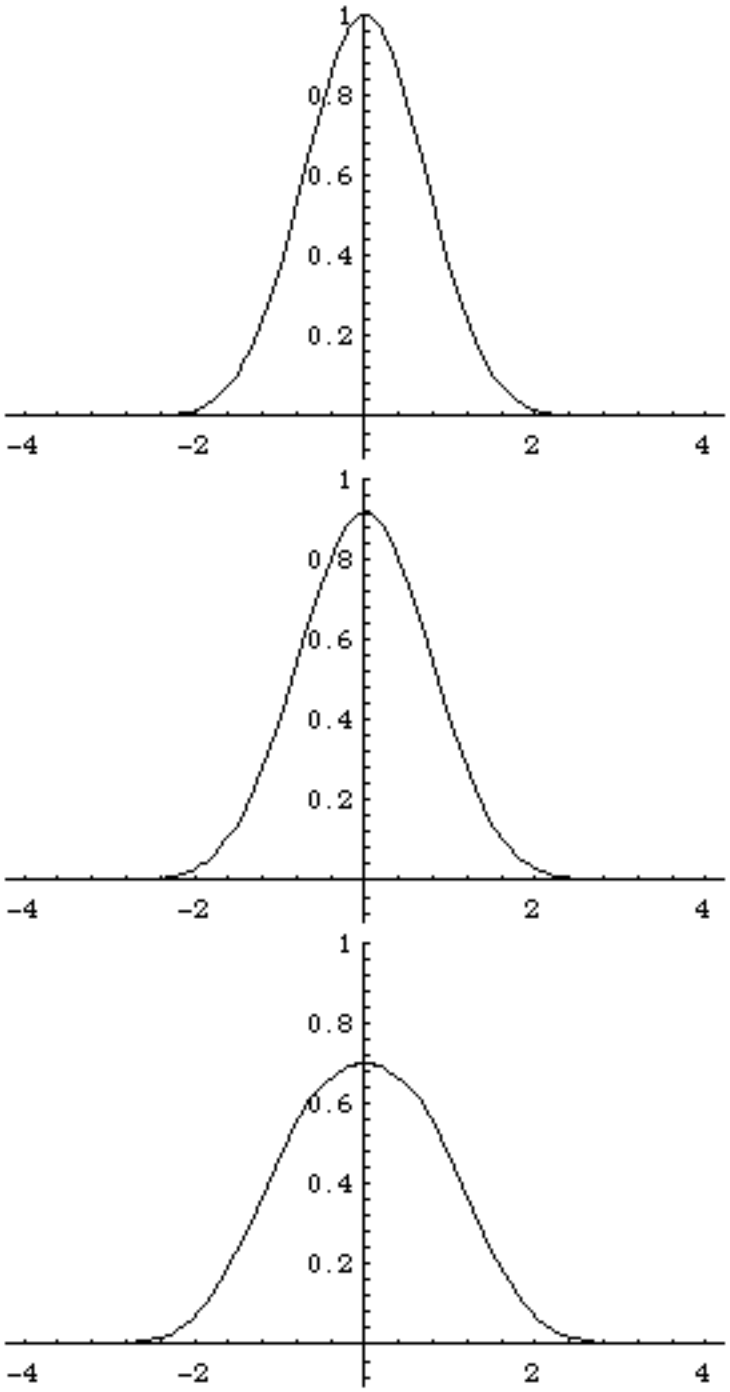
and

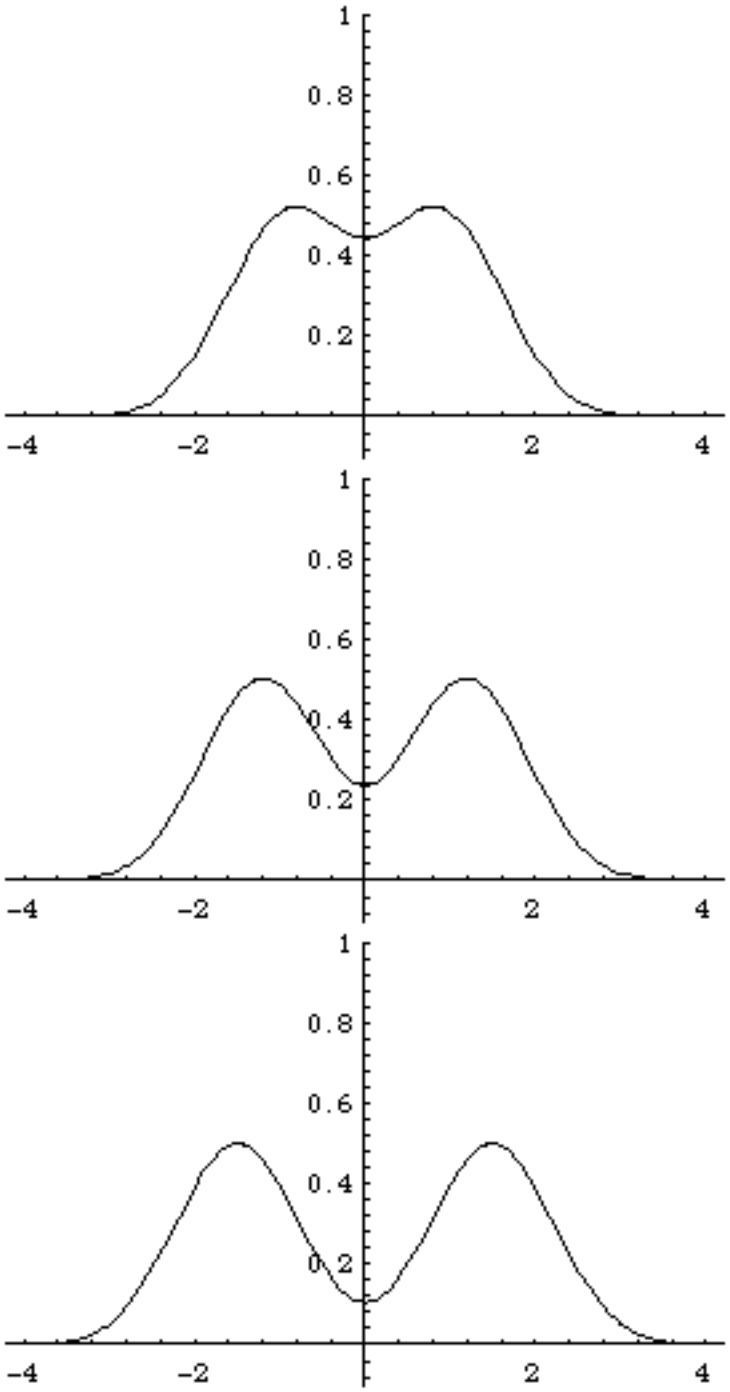
$$u(t,x) = f(x+ct) /2 + f(x-ct) /2.$$

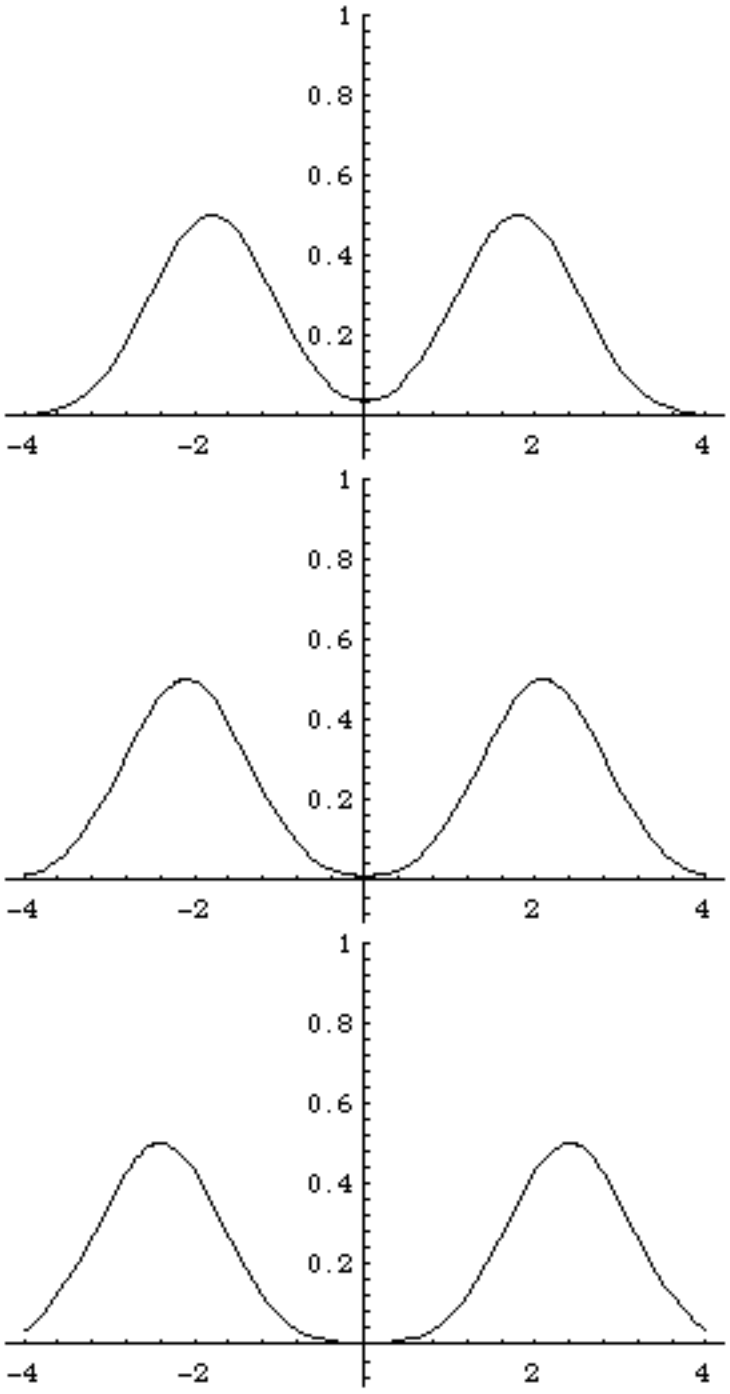
Model Problem VII.1. Let the initial displacement f(x) be $\mathrm{Exp}(-x^2)$, and let the initial velocity g(x) = 0. Then the solution in the future is

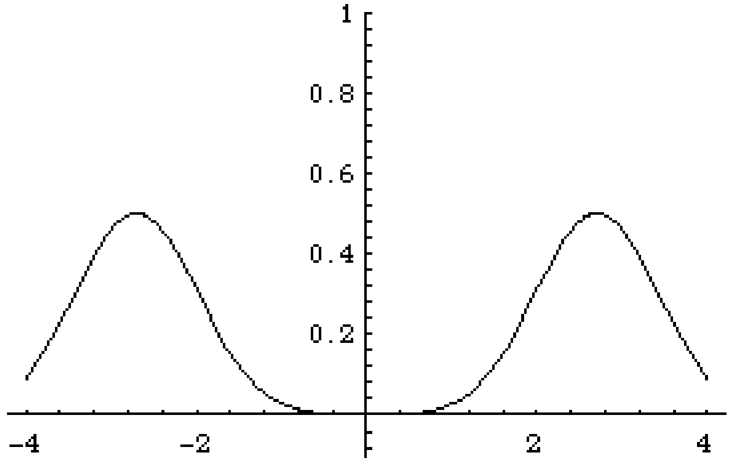
$$u(t,x) = \frac{1}{2}\left(\mathrm{Exp}\left(-(x+ct)^2\right) + \mathrm{Exp}\left(-(x-ct)^2\right)\right)$$

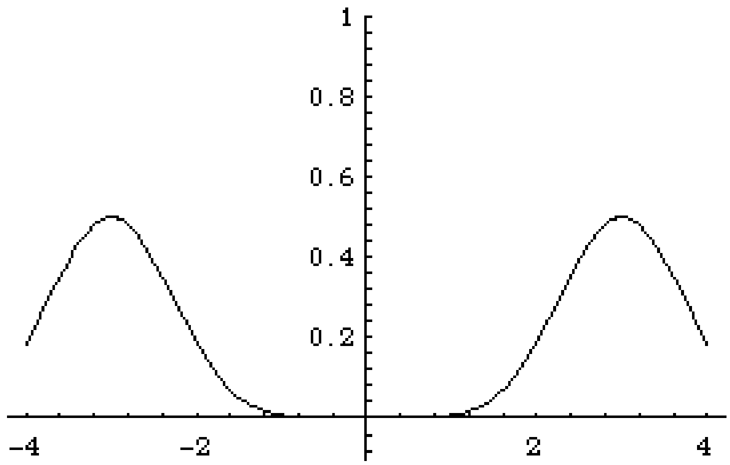
We should see the initial bump break into the superposition of two bumps of half the height, one moving to the right and one to the left. Here are some snapshots at different times of the displacement as a function of x:


1
0.8
0.6
0.4
0.2
-4
-2
2
4
1
0.8
0.6
0.4
0.2
-4
-2
2
4
1
0.8
0.6
0.4
0.2
-4
-2
2
4


1
0.8
0.6
0.4
0.2
-4
-2
2
4
1
0.8
0.6
0.4
0.2
-4
-2
2
4
1
0.8
0.6
0.4
0.2
-4
-2
2
4


1
0.8
0.6
0.4
0.2
-4
-2
2
4
1
0.8
0.6
0.4
0.2
-4
-2
2
4
1
0.8
0.6
0.4
0.2
-4
-2
2
4





- -

One thing we can learn from the form of this solution and can visualize in these graphs is that the integral of u(t,x) with respect to x is independent of time.

Next, let's turn the tables and suppose that f(x) = 0 but g(x) ≠ 0. Putting in t=0, we see that

$$\varphi(x) = -\psi(x),$$

while

$$c\,\varphi'(x) - c\,\psi'(x) = 2\,c\,\varphi'(x) = g(x).$$

Clearly, then,

$$\varphi(x) = \frac{1}{2c}\int^{x} g(x)dx \quad .$$

The constant of integration is not specified, but since the full solution is of the form

$f(x+ct) - f(x-ct)$,

it will cancel out.

The solution to (WE) with these IC is:

$$u(t,x) = \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\, ds$$

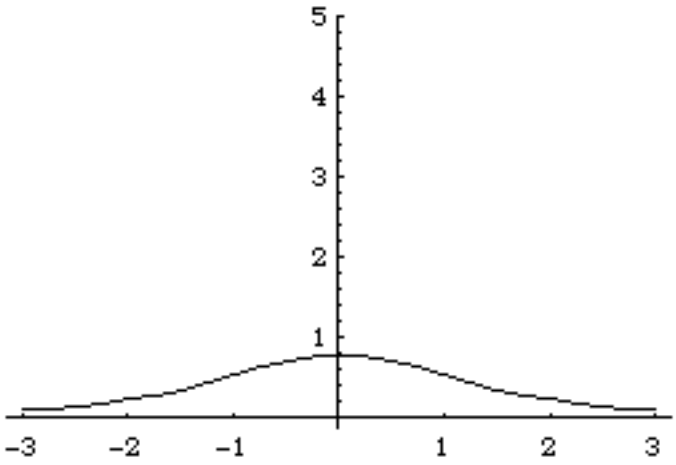
If, for example, c=1 and

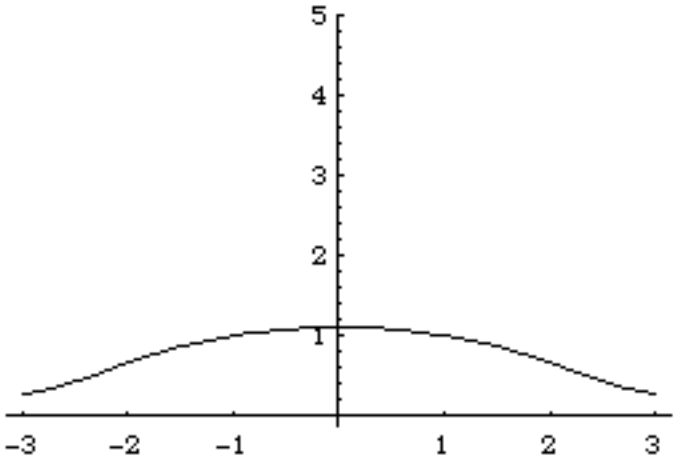
$$g(x) = \frac{1}{1+x^2},$$

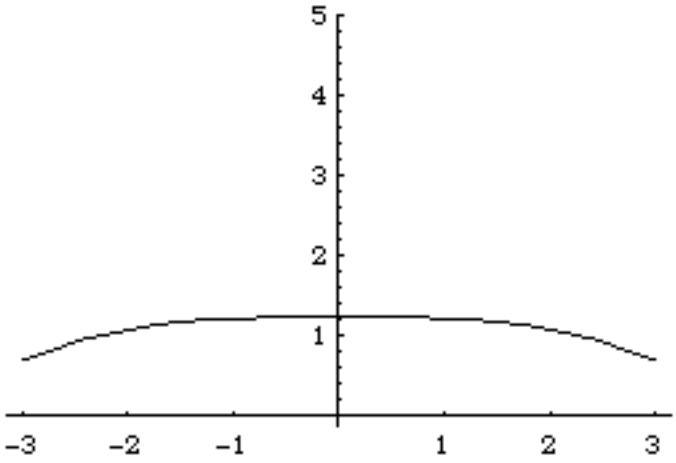
Then a calculation shows that

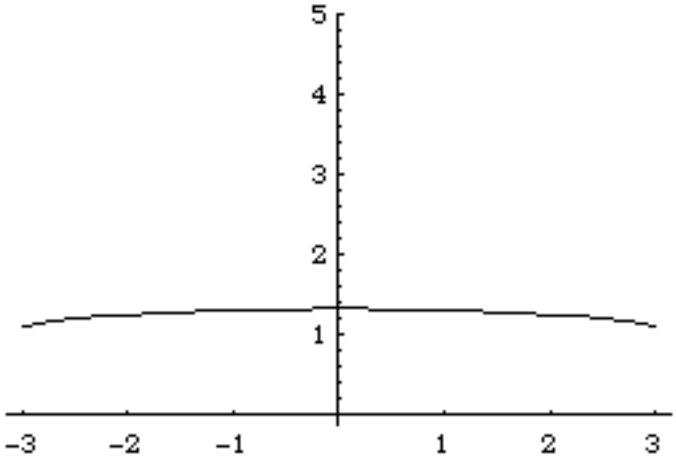
$$u(t,x) = \frac{\arctan(t-x) + \arctan(t+x)}{2}.$$

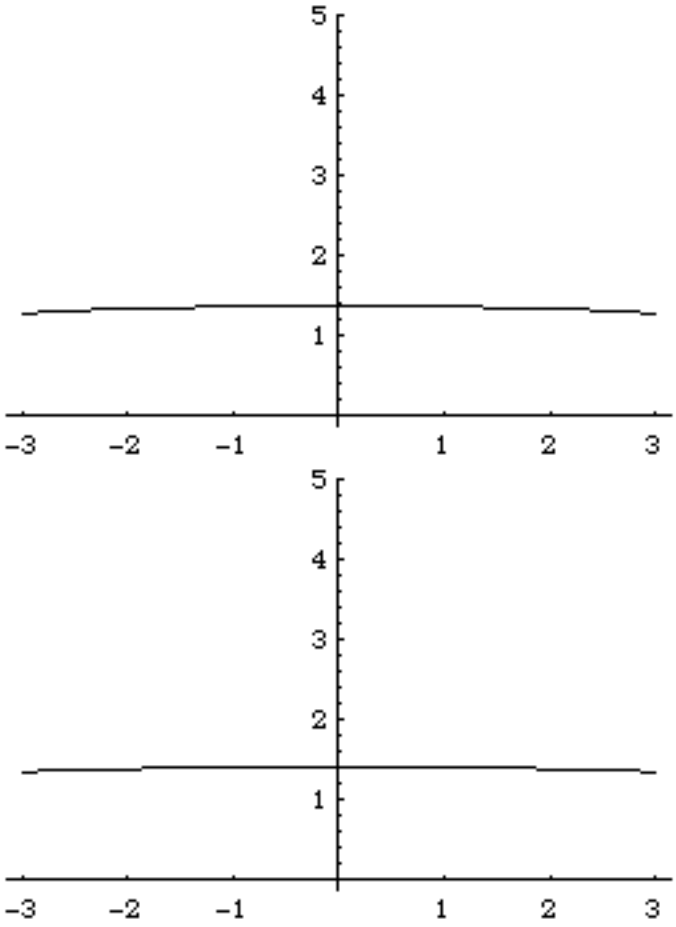
Here are some snapshots of the solution:




5
4
3
2
1
-3
-2
-1
1
2
3


5
4
3
2
1
-3
-2
-1
1
2
3


5
4
3
2
1
-3
-2
-1
1
2
3



The solution with both initial conditions

$$u(0,x) = f(x)$$
$$u_t(0,x) = g(x)$$

is just the sum of these two solutions, because of the superposition principle:

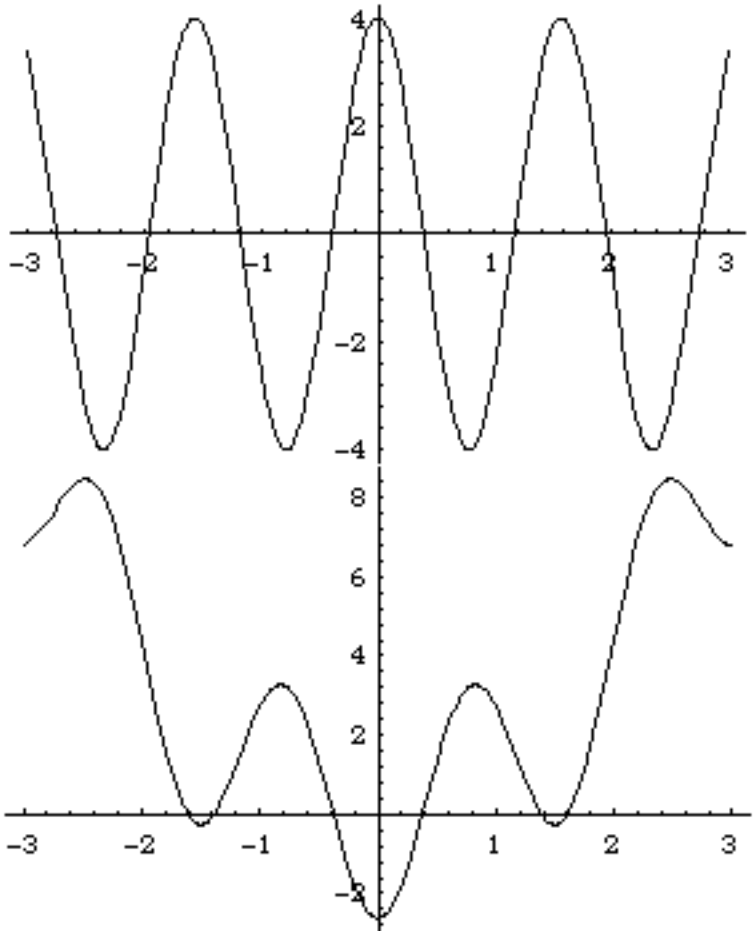
$$u(t,x) = \frac{1}{2}\big(f(x+ct) + f(x-ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)ds\ . \qquad (7.2)$$

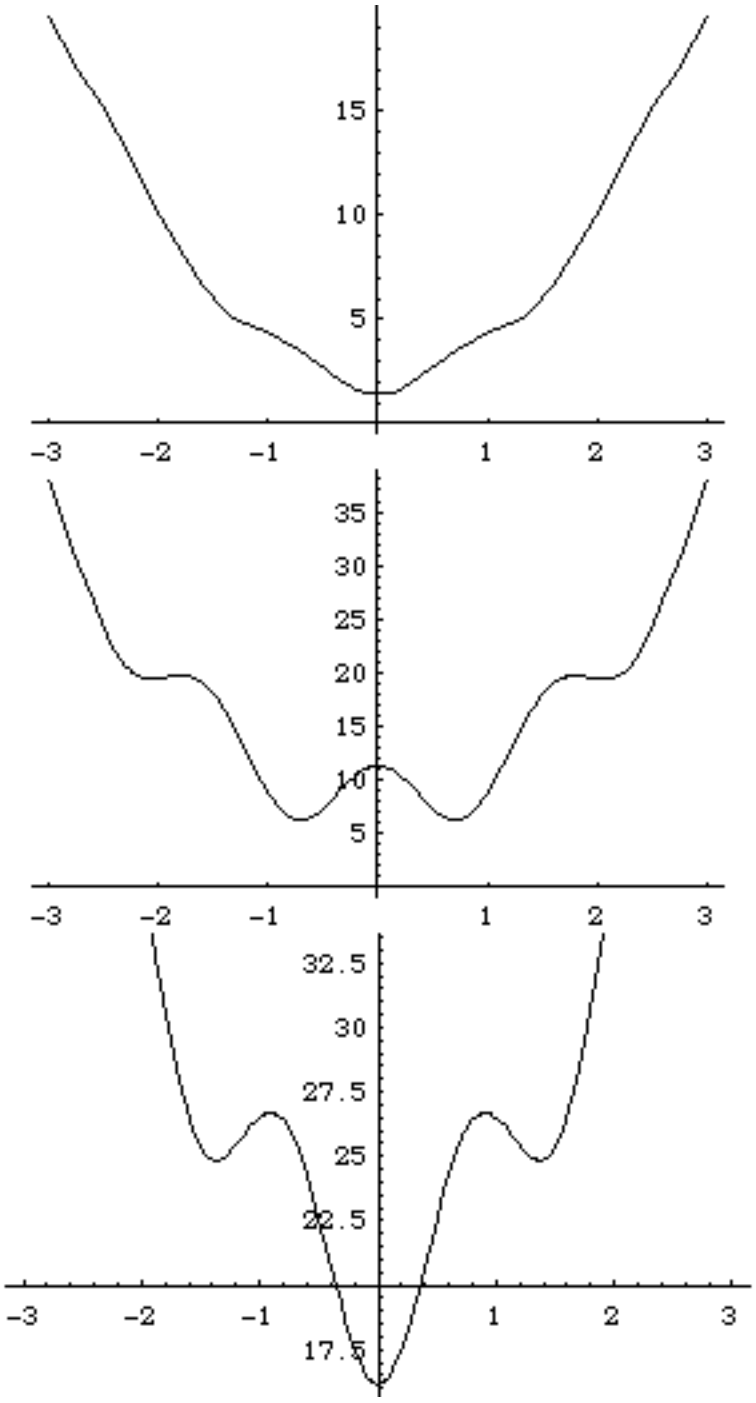
Model Problem VII.2. Find the general solution of the wave equation with c=1 and initial conditions $f(x) = \cos(4x)$, $g(x) = x^2 - 1/3$.

Solution. Use Mathematica to calculate with formula (7.2). The answer (see notebook) is:

$$u(t,x) = \frac{2}{3}\left(t^3 + 3tx^2 - t\right).$$

Although the formula for the solution is fairly simple, the wave forms it describes can be complicated. For example, at unit time steps beginning at t=0, this solution looks as follows:





- -

The method of d'Alembert is great if there is no boundary condition, but you may think it fails to be of use when there are boundary conditions. Not so!

Here is the trick. Suppose that we have (say) DBC at 0, and we watch the vibrating string through a window where we can see only what is going on at x>0. We can't tell whether

the string is actually clamped down at 0, but the height of the string just happens to be 0 at 0 (at all times). The way we can arrange this is to exploit symmetries.

If f(x) and g(x) are defined on $-\infty < x < \infty$ and are even about the point 0 (i.e., f(-x) = f(x) and g(-x) = g(x)), then the solution by d'Alembert's method will also be even, since

$$\begin{aligned} u(t,-x) &= \tfrac{1}{2}\big(f(-x+ct) + f(-x-ct)\big) + \frac{1}{2c}\int_{-x-ct}^{-x+ct} g(s)\,ds \\ &= \tfrac{1}{2}\big(f(x-ct) + f(x+ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(r)\,dr \qquad (7.3) \\ &= u(t,x) \end{aligned}$$

The substitution in the middle line is r=-s.

It works the same way if f and g are even about some point other than 0. Likewise if f and g are odd, then so is the solution (see Exercise 1).

Model Problem VII.3. Solve the wave equation (with c=1) for $0 < x < \infty$ with the following initial and boundary conditions. Use the superposition principle and the uniqueness theorem for this problem to simplify your analysis.

a) $u(0,x) = 0$, $(\partial u/\partial t)(0,x) = 0$, $u(t,0) = t^2$.

b) $u(0,x) = \sin(x)$, $(\partial u/\partial t)(0,x) = 0$, $u(t,0) = t^2$.

c) $u(0,x) = 0$, $(\partial u/\partial t)(0,x) = 0$, $(\partial u/\partial x)(t,0) = t\ \exp(-t)$.

Solutions.

a) $u(0,x) = 0$, $(\partial u/\partial t)(0,x) = 0$, $u(t,0) = t^2$.

The uniqueness theorem for this problem tells us that it has only one solution. The point of uniqueness theorems is often lost on students, but in fact they can be extremely practical: If you can solve a different problem from the one which was posed, but it happens to satisfy all the conditions, then it must be the solution of the original problem, too.

In this case, we know how to solve the wave equation if there is no boundary and we are given initial conditions for all x. D'Alembert discovered that the general solution in that case is a superposition of a wave moving to the right and a wave moving to the left. The shapes of the moving waves can be arbitrary until they are fixed by the boundary conditions.

With this in mind, imagine a string with $-\infty < x < \infty$ carrying with a rightward moving wave, which happens to be of zero amplitude when $t = 0$ and $x > 0$. In other words, $u(t,x) = \psi(x-t)$, where $\psi(x-0) = \psi(x) = 0$ when $x > 0$. This satisfies the initial conditions, so it must be the boundary condition at x=0 which determines the unknown function y: Putting x=0, we see that

$\psi(0-t) = t^2$ for all $t > 0$

We now know $\psi(x) =$ for all values of x, because given any negative x , we can substitute $x = -t$ and use this last observation to find that $\psi(x) = (-x)^2 = x^2$. In the traveling wave, we replace x by x-t, so the solution is:

$u(t,x) = \psi(x-t) = (x-t)^2$ when $0 < x \leq t$, and otherwise 0.

b) $u(0,x) = \sin(x)$, $(\partial u/\partial t)(0,x) = 0$, $u(t,0) = t^2$.

Let us think about the superpositioon principle here to simplify things, rather than beginning the whole analysis at the beginning. A linear combination of solutions of the wave equation is again a solution of the wave equation, but the boundary condition is non-homogeneous. This means that we can't just add two solutions such that $u(t,0) = t^2$; if we did the boundary condition satisfied by the sum would be $2t^2$, not t^2. If, however, we added a function satisfying our boundary condition to one having a zero Dirichlet condition, the sum would equal t^2 at the boundary. This suggests the following strategy:

Step 1, Solve problem a), with $u(0,x) = 0$, which we did above.
Step 2, Solve a simpler problem, the wave equation with
b') $u(0,x) = \sin(x)$, $(\partial u/\partial t)(0,x) = 0$, $u(t,0) = 0$.
Step 3. Add the two solutions.

Since the sum solves the wave equation and the initial and boundary conditions, it is the one and only correct answer to problem b).

Here is how to solve b'). If we extend x to $-\infty$ and u(0,x) = sin(x) there, then out initial condition would be odd, and we might expect that by symmetry, u(t,0)= 0 when t > 0. The solution of this problem, by d'Alembert's method, is:

(1/2) (sin(x+t) + sin (x-t)),

and when x=0, this is (1/2) (sin(t) + sin (-t)), which is indeed zero, since the sine is an odd function. It may also help to notice that by the sine-sum formula,
(1/2) (sin(x+t) + sin(x-t)) = sin(x) cos(t)

The solution to problem b) can be written:

$$u(t,x) = (x-t)^2 + \sin(x)\cos(t), \text{ when } 0 < x \le t, \text{ and}$$
$$\sin(x)\cos(t), \text{ otherwise.}$$

c) $u(0,x) = 0$, $(\partial u/\partial t)(0,x) = 0$, $(\partial u/\partial x)(t,0) = t \exp(-t)$.

We can solve this problem just as for problem a), by imagining that the solution is a traveling wave coming from the left of 0, but which hasn't disturbed the string at x>0 before it arrives at t=0. Thus $u(t,x) = \psi(x-t)$, where $\psi(x-0) = \psi(x) = 0$ for $x > 0$.

The difference is the boundary condition. This time we know that

$\partial u/\partial t = -\psi'(x-t)$, and at time 0, we have

$-\psi'(0-t) = t \exp(-t)$.

Perhaps it is best here to change variables so that x = -t:

$-\psi'(x) = -x \exp(x)$, so, according to Mathematica:

```
In:=
psi[x_] = Integrate[x Exp[x], x]
Out=
 x
E  (-1 + x)
```

Mathematica fails to write the constant of integration, which we know here should be +1 in order that the solution will be 0 when x = 0. The solution is thus

```
In:=
% +1 /. x -> x-t
```

```
Out=
     -t + x
1 + E       (-1 - t + x)
```
for $x-t \le 0$ (and 0 for $x-t > 0$).

Let's check the boundary condition:

```
In:=
Simplify[D[%,t] /. x -> 0]
Out=
Simplify[%]
t
--
 t
E
```

If we are given initial conditions f and g only between 0 and L, we are free to imagine that they are defined for $x > L$ as well as for negative x, in any way we wish. If we extend the definition of f and g so as to be odd functions about the two points 0 and L, and imagine solving the problem of an infinite string with d'Alembert's method, then the boundary conditions will continue to be satisfied when $t>0$. That has to be the same as our solution for the finite problem, because of uniqueness.

Model Problem VII.4. Suppose that $g(x) = 0$, but $f(x) = x(1-x)$, $0 < x < 1$. Extend these functions to all of x and thereby solve the initial-boundary-value problem with these initial conditions and zero Dirichlet boundary conditions at 0 and 1.

Solution. We need to extend f and g so as to be odd around both 0 and 1. Obviously, we extend g(x) as 0 for all x, but f(x) is a little trickier For x from -1 to 0, we change f(x) to -f(-x). Mathematica can be invoked to show what is going on (see notebook for the calculations):

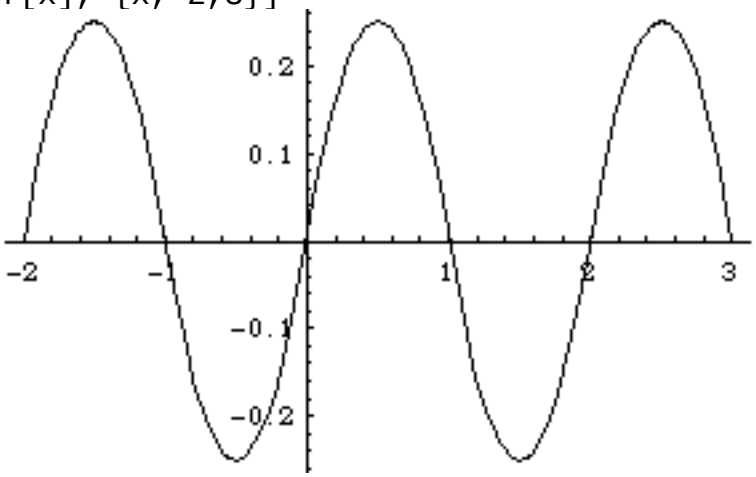
```
f[x_] := x (1 - x) /; 0 <= x <= 1
f[x_] := x (1 + x) /; -1 <= x < 0
```

(Change x to -x and multiply the whole by -1 to get this.) For 1<x<2, we have to reflect about the point x=1, so we replace x by 2-x. To get the odd extension beyond 2, we multiply the result by -1 again:

```
f[x_] := - (2 - x) (1 - (2 - x)) /; 1 < x <= 2
```

This new territory can then be flipped over to the interval [-2,-1], etc:

```
f[x_] := x (1 - x) /; 0 <= x <= 1
f[x_] := x (1 + x) /; -1 <= x < 0
f[x_] := - (2 - x) (1 - (2 - x)) /; 1 < x <= 2
f[x_] := + (2 + x) (1 - (2 + x)) /; -2 <= x < -1
f[x_] := + (x - 2) (1 - (x - 2)) /; 2 < x <= 3
Plot[f[x], {x,-2,3}]
```



It looks much like a sine function, but numerically it isn't. We could go on like this all day, or we could use modular arithmetic (the int and frac commands in programming languages, or, in Mathematica, Floor[x] and x - Floor[x]) to automate it better, as follows. (For more details see the Mathematica notebook.) Int and Frac give the integer and fractional parts of a number, for example:

```
In:=
{Int[3.89], Frac[3.89], Int[-3.89], Frac[-3.89]}
Out=
{3, 0.89, -4, 0.11}
```

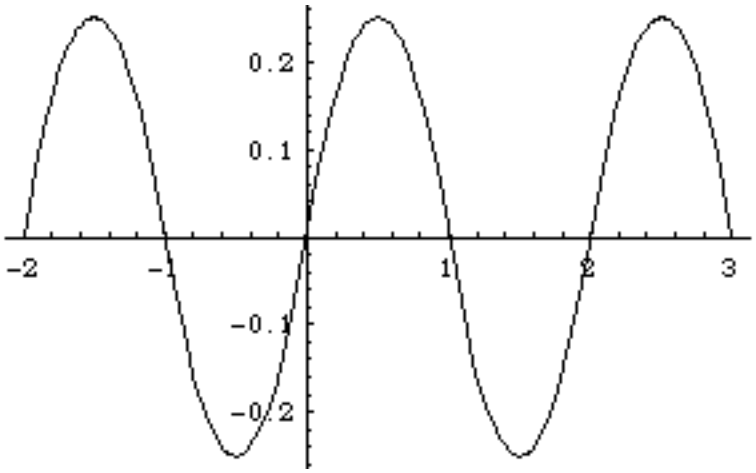
Notice that when we piece the function together, it is periodic with period 2. Here is an alternative way to write the extended function f(x):

```
In:=
f[x_] := If[0 <= Frac[x/2] < 1/2, Frac[x] (1 - Frac[x]), \
```

```
Frac[x] (Frac[x] - 1)]
```

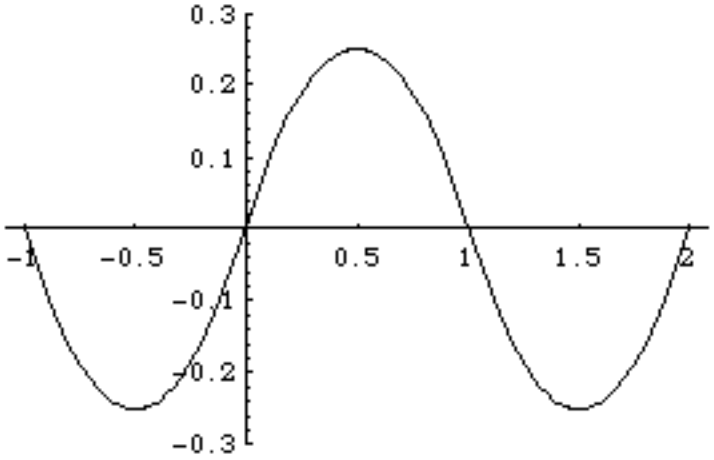
```
In:=  Plot[f[x], {x,-2,3}]
```

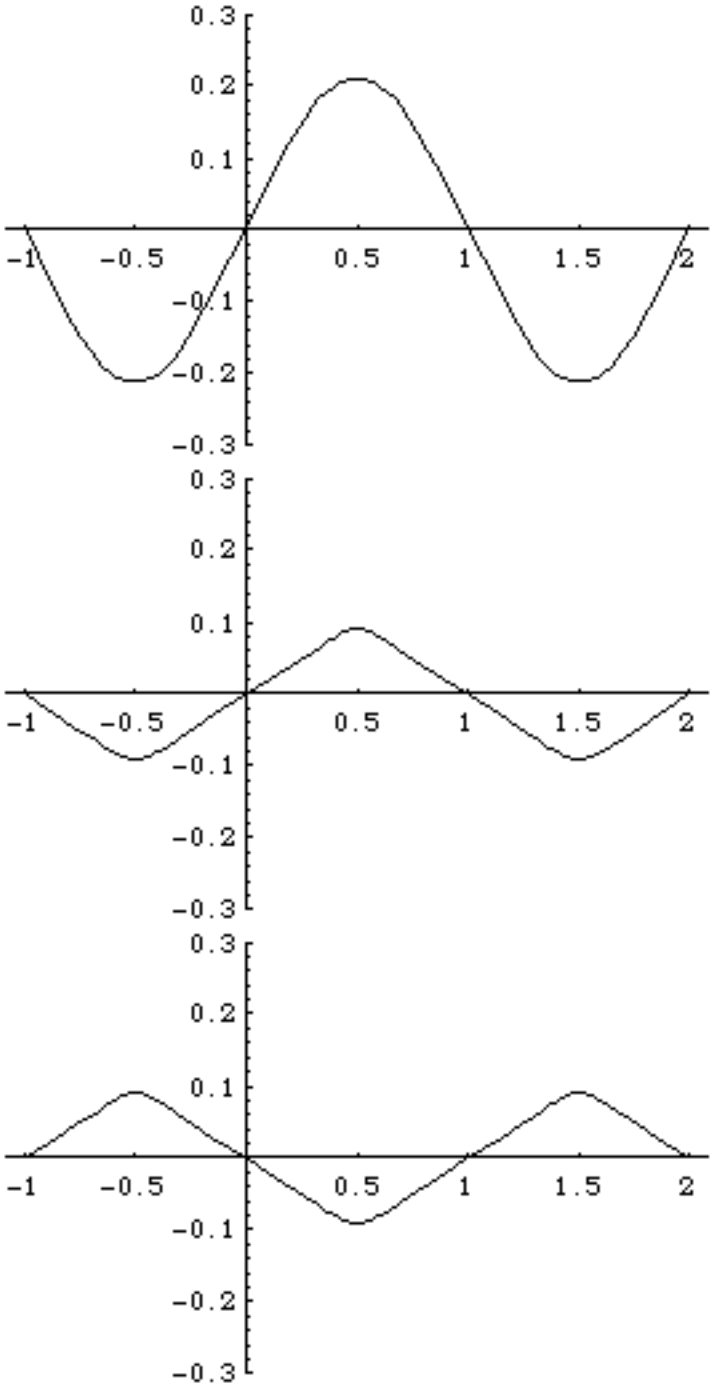
Out=

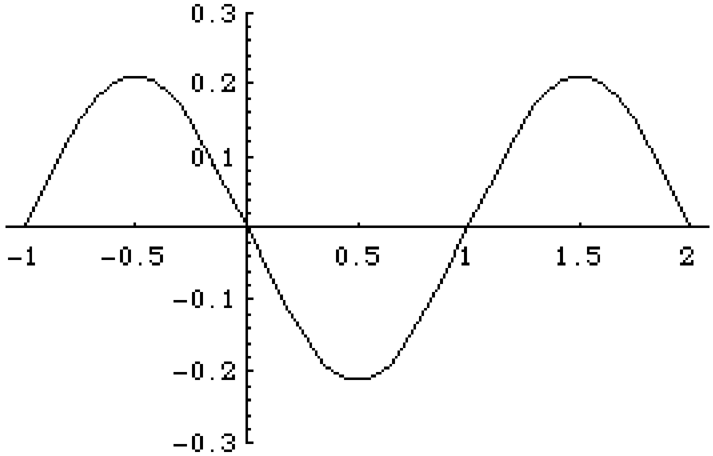


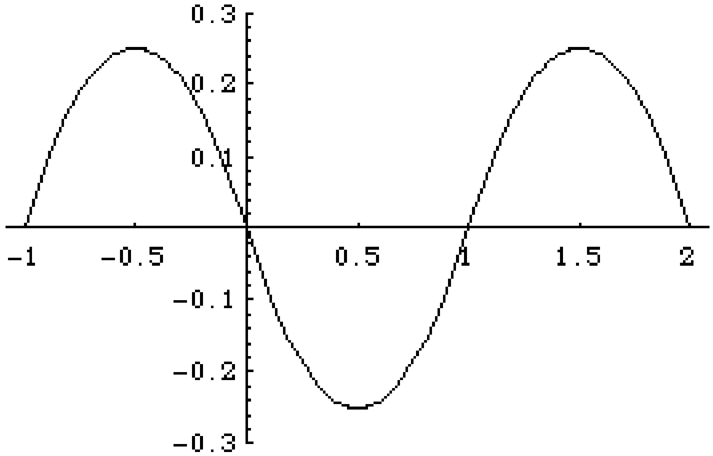
Finally, we solve the IBVP with formula (7.2) as evaluated with Mathematica (see notebook):

```
In:=  Table[Plot[u[f,g,t,x,1],
{x, -1, 2}, PlotRange -> {-.3,.3}],{t,0,1.6,.2}]
```




0.3
0.2
0.1
-1
-0.5
0.5
1
1.5
2
-0.1
-0.2
-0.3
0.3
0.2
0.1
-1
-0.5
0.5
1
1.5
2
-0.1
-0.2
-0.3
0.3
0.2
0.1
-1
-0.5
0.5
1
1.5
2
-0.1
-0.2
-0.3





With software it is possible to animate this and see the standing wave motion. Notice that the wave does not simply jiggle up and down; its shape is changing, too.

The wave equation is the prototype for a class of partial differential equations, the *hyperbolic* equations. Other ones exhibit wave motion, although the speed may vary and dispersion and dissipation may occur. (In more ordinary language, the waves may spread and die down.) Another feature of the hyperbolic equations is that they are quite tolerant of irregularities. The general solution (7.1) makes no reference to the differentiability of φ and ψ, and indeed, even if φ and ψ are discontinuous, that formula has a reasonable interpretation as a superposition of traveling waves. Such an expression is referred to as a *weak* or *generalized* solution of the wave equation, and is important in applications because it describes shock waves, sonis booms, etc.

A weak solution can be defined formally by integration by parts. Suppose that a function of the form (7.1) is put into an inner product with another function $\zeta(t,x)$ which

a) is differentiable at all orders, and
b) equals 0 for |x| sufficiently large or for |t| sufficiently large. The standard inner product here is defined as in chapter II, but by integrating over all x and t, $-\infty < x,t \leq \infty$.

I will refer to the region on which $\zeta \neq 0$, together with the boundary of this region, as its support. If u(t,x) is differentiable twice with respect to t and x (and the derivatives are continuous functions), and A denotes the wave operator defined above, then integration by parts shows that

$$<A\,\zeta,u> = <\zeta,\, A\,u> = 0 \qquad (7.4)$$

for all such ζ. The functions ζ are referred to as *test functions*, and you can think of (7.4) as a way of testing whether u is a solution of the wave equation without differentiating it directly.

Definition. A function u which satisfies (7.4) for all test functions is a*weak solution* of the wave equation.

The method of d'Alembert works in a modified form for other hyperbolic, and even parabolic, partial differential equations, and is more generally known as the method of characteristics. Instead of simple, traveling waves, the solutions consist of waves which travel along special curves, and which can have variable speeds and can change their shapes. If the equations are not linear, the waves will not satisfy the superposition principle, but the method of characteristics is less closely tied to linearity than is the method of separation of variables.

Exercises.

VII.1. As in eq. (7.2), verify that

a) If f and g are even about a general point x=a, i.e., f(2a - x) = f(x) and g(2a - x) = g(x), then so is the d'Alembert solution with these initial conditions.

b) If f and g are odd about a general point x=a, i.e., f(2a - x) = -f(x) and g(2a - x) = -g(x), then so is the d'Alembert solution with these initial conditions.

VII.2. Use the int and frac functions to extend the function f(x) = x (1-x), $0 \le x \le 1$, to all $-\infty < x < \infty$ such that

a) f(x) is even about x = 0 and x = 1

b) f(x) is odd about x = 0 and even about x = 1.

VII.3. Use the int and frac functions to extend the function f(x) = x (1-x), $0 \le x \le 2$, to all $-\infty < x < \infty$ such that

a) f(x) is even about x = 0 and x = 2

b) f(x) is odd about x = 0 and even about x = 2.

VII.4. Show that the solution obtained with d'Alembert's method with boundary conditions is equivalent to the solution we got by separation of variables, even though they look quite different. You will need to use trigonometric identities.

VII.5. Consider the wave equation on the whole line $-\infty < x < \infty$, but change variables so that z = x+ct and w = x-ct. Show that the wave equation becomes

$$\partial^2 u / \partial z \, \partial w = 0.$$

Derive d'Alembert's solution from this equation, by integrating it first with respect to z and then with respect to w. (Hint: Since the partial derivative with respect to z means that w is being held temporarily fixed, if you integrate something with respect to z, the "constant" of integration may be different for different values of w. In other words, instead of a constant of integration you get a function of w. Similarly for integrating partial derivatives with respect to z.

VII.6. Derive (7.4)

VII.7. Show that if $u(t,x) = \varphi(x-ct)$ for a function φ which is merely assumed piecewise continuous, then u is a weak solution of the wave equation.

VII.8. Solve the modified wave equation

$$u_{tt} = u_{xx} - (1/4)\, u_t - 32$$

with IC:

$$u(0,x) = \sin(2x)$$
$$u_t(0,x) = 0$$

and BC:

$$u(t,0) = 0$$
$$u(t,2\pi) = 1.$$

(You will want to use ideas from both this chapter and chapter VI to solve this nonhomogeneous problem.)

version of 11 November 1997

VIII. The mathematics of hot rods.†

In many ways the simplest of our three model partial differential equations is the heat equation, also known as the diffusion equation,

$$u_t = k \quad ^2u \qquad \text{(HE-3D)}$$

or, specializing for now to one dimension,

$$u_t = k\, u_{xx}. \qquad \text{(HE)}$$

A physical derivation of this equation is given in an appendix; the constant

$$k = \sigma/\rho\, \kappa_s,$$

where σ is the thermal conductivity, ρ is the density, and κ_s is the specific heat capacity of a homogeneous substance. Ordinary substances have values of k ranging from about 5 to 9000 cm^2/hr (see the table at the end of the chapter).

We shall solve the heat equation with the method of separation of variables, as we did for the wave equation. The solutions will be quite similar, and you may find it instructive to compare with what we do in chapter VI. We begin with the one-space-dimensional heat equation (HE), which can be written in terms of a heat operator

$$H_k[f] := \left(\frac{1}{k}\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}\right) f(t,x),$$

so that (HE) is the equation for the kernel of this operator. This equation describes the temperature u(t,x) in a long, thin rod along the x-axis, if the rod is wrapped along its length with a perfect insulator. The heat equation is a linear, homogeneous equation, so, just as

---

for the wave equation, the superposition principle holds; if we can find some solutions u_1, …, u_k, then any linear combination of these solutions is again a solution.

With the method of separation of variables, we try to find solutions of a special form,

$$u = T(t)\, X(x).$$

As before, boundary and initial conditions will be important, and arise from the physical situation we wish to study. The end points of the rod may be in thermal contact with ice water, which will ensure that we have homogeneous Dirichlet boundary conditions (DBC)

$$u(t,a) = u(t,b) = 0.$$

(Actually, stirred wet crushed ice would guarantee these boundary conditions in the laboratory better than ice water.)

The fact that these boundary conditions are homogeneous (something equals 0) is going to be useful in our analysis, so you may wonder if something goes wrong if we use a different temperature scale or hold the ends at a different fixed temperature T different from 0. Fortunately, we only need to redefine u by a constant to have these same boundary conditions. Since $H_k\,[T] = 0$, if

$$v(t,x) := u(t,x) - T, \qquad (8.1)$$

then

$$H_k[v] = H_k[u] = 0.$$

By this simple change, the problem with some other fixed temperature at the ends is converted into out standard problem with homogeneous boundary conditions of the Dirichlet type.

Neumann boundary conditions (NBC),

$$u_x(t,a) = u_x(t,b) = 0,$$

correspond physically to the situation when the ends of the rod are also insulated, since the heat flow through either end of the rod is proportional to the temperature gradient at that end. If there is no heat flow, the gradient must be 0.

We begin by considering the case of Dirichlet boundary conditions, and suppose that the rod has length 1, with a = 0 and b = 1. We shall solve this problem with the method of separation of variables, by looking for special solutions analogous to normal modes for the wave equation. Ignoring the initial conditions for the time being, we begin by seeking product solutions,

$u(t,x) = T(t)\, X(x)$

to the heat equation (HE), which also satisfy the homogeneous boundary conditions.

The efficient way to proceed is to divide the wave equation by the product solution:

$$\frac{1}{T(t)\, X(x)} H_k [T(t)\, X(x)] = \frac{T'}{k\, T} - \frac{X''}{X} .$$

If u solves the heat equation, this expression must be zero. Thus

$T'/(k\, T) = X''/X.$

Since the left side is independent of x and the right side is independent of t, both sides must in fact a constant. We don't know its value yet, so let us call it $-\mu$. (It will turn out to be negative.

Thus

$$-X''(x) = \mu\, X(x) \qquad (8.2)$$

This is precisely the same eigenvalue equation as in chapter VI, for the wave equation with (DBC). We know that the eigenfunctions and eigenvalues are:

$X(x) := \sin(n\, \pi\, x),$

$\mu_n = (n\pi)^2.$

The other part of the partial differential equation which has been separated differs from what we got with the wave equation:

$$T'(t) = -\ \mu_n\, k\, T(t) \qquad (8.3))$$

This is a simple first-order equation, the solutions of which are exponential functions,

$T(t) = C \exp(-\mu_n\, k\, t).$

The eigenvalue problem (8.2) fixed the possibilities for μ_n, leaving us with normal modes of the form

$c_n \exp(-\, n^2\pi^2\, k\, t) \sin(n\pi x),\ n = 1, \ldots$

With DBC at a and b, b-a = L, we would get:

$$c_n \exp(- n^2\pi^2 k\, t) \sin(n\pi(x-a)/L),\ n = 1, \ldots;$$

the easiest way to see this is to change variables to x' = (x-a)/L. In passing, we note that we could have set k=1 throughout the derivation of the normal modes, and then changed variables to t' = k t at the end.

Since the heat equation is linear, we can make linear combinations of the solutions we have found and thereby produce new solutions. The most general solution to the heat equation that we get in this way is:

$$\sum_{n=1} c_n \exp(- n^2\pi^2 t) \sin(n\pi x). \qquad (8.4)$$

At this stage it may not be completely clear that this is the absolutely most general solution, but the completeness theorem for the Fourier sine series will guarantee that this can indeed describe the solution to the heat equation with an arbitrary square-integrable initial condition.

The initial conditions appropriate to the heat equation are somewhat different from what we had for the wave equation, again because the time variable enters differently. The heat equation is first-order in time, so we need only one initial condition, the initial temperature distribution in the rod,

$$u(t,x) = f(x). \qquad (IC)$$

There is no need to look at u_t at t=0. Physically, this would be inappropriate, because our experience teaches us that the temperature changes only in response to the initial temperature, not to some unmeasurable effect of the rate of change of the initial temperature. If this were not the case, cookbooks would be nearly impossible. Fortunately, the initial conditions to bake a cake are, typically, batter at room temperature, and if we have identical batter in identical pans at room temperature, we need to bake the cake for identical lengths of time. No other physical initial condition is necessary.

Model Problem VIII.1. Find the future temperature within a rod of length 1, k=1, insulated along its length, and with ends in thermal contact with ice water, if at time 0, the rod is at a uniform temperature u(0,x) = 50.

Solution (using software for calculations). At time t=0, the general solution becomes a Fourier sine series,

$$\sum_{n=1} c_n \sin(n\pi x),$$

so the coefficients c_n are determined as the Fourier *sine* series coefficients of the initial function, f(x) = 50. We know the formula for the Fourier sine coefficients, (2.14), so we simply calculate the integral to get:

$$c_n = \frac{100}{n\pi}(1 - (-1)^n).$$

This is 0 when n is even and otherwise 200/(n π). Our solution to the heat equation is thus:

$$u(t,x) = \sum_{n \text{ odd}} (200/n\pi) \exp(- n^2\pi^2 t) \sin(n\pi x).$$

- -

Suppose now that the rod is insulated not only along its length but also on its ends. Then u(t,0) and u(t,L) are not specified, but instead we have:

$$\partial u[t,0]/\partial x = 0, \qquad \partial u[t,L]/\partial x = 0 \qquad \text{(NBC)}$$

Model Problem VIII.2. Set up the general solution of the insulated rod with Neumann BC., with the simplifications k = L = 1.

Solution.
This is very similar to the analysis in chapter VI, and the eigenfunctions are the basis functions of the Fourier *cosine* series,

$$\{1, \cos(n\pi x)\}.$$

The normal modes are

$$c_n \exp(- n^2\pi^2 t) \cos(n\pi x), \ n = \mathbf{0}, 1, \ldots$$

and the general solution is:

$$\sum_{n=0} c_n \exp(- n^2\pi^2 t) \cos(n\pi x), \tag{8.5}$$

- -

Notice that after a long time, all of the terms of this series will become negligible except the one corresponding to n=0. Remembering that c_0 is nothing other than the average

temperature, we see that if the rod is entirely insulated, then after a long time the temperature has settled down to equilibrium at the one temperature which is consistent with conservation of energy.

Another way to see the same effect is as follows. I will do the calculation in three dimensions, since it is not much harder than in one dimension. In the derivation of the heat equation in the appendix, we reason that the rate of change of heat content is

$$\frac{dQ}{dt} = \iint_{\partial\Omega} \sigma \, \mathbf{n} \quad u(t,\mathbf{x}) \, d^2x.$$

If the object is insulated, then $\mathbf{n}$ $u = 0$. (This higher-dimensional Neumann condition is what simplifies to $u_x = 0$ in one space dimension.) It follows that the total heat content

$$Q = \iiint_{\Omega} \rho \kappa_s \, u(t,\mathbf{x}) \, d^3x$$

remains constant. For a homogeneous material, this is just a multiple of the average temperature.

Something more subtle happens in the case of DBC, where energy is lost (or gained) through the ends of the rod. In that case, the general solution (8.4) ensures that not only will the temperature settle down to an equilibrium $u(t,x) \to 0$ as $x \to \infty$, but it usually will do so with a particular temperature profile, for assuming that $c_1 \neq 0$, we find that

$$\exp(\pi^2 t) \, u(t,x) = c_1 + r(t,x),$$

where the remainder terms in

$$r(t,x) = \sum_{n=2} c_n \exp(-(n^2-1)\pi^2 t) \sin(n\pi x)$$

all vanish rapidly as $t \to \infty$. As can be seen from the calculations in the Mathematica notebook, the temperature profile of the rod rapidly approaches that of a simple sine function, times a time-varying factor tending to 0.

Model Problem VIII.3. Suppose now that we have non-homogeneous boundary conditions, such as what we would get if the end of the rod at x=0 is held at temperature 100 and the end at x=1 at temperature 0. Find the general solution to the heat equation with k=1.

Solution.

We can no longer convert this problem to the usual one by the simple trick (8.1). A cleverer idea is to think about what should happen at equilibrium in this problem. At equilibrium, $u_t = 0$, so the temperature will depend on space alone according to

$$u^{(eq)}_{xx} = 0. \tag{8.6}$$

In this one-dimensional situation, this means that $u^{(eq)}(x) = c_1 x + c_2$ - the general solution to (8.6). With the given boundary conditions, we can easily solve for the constants, and find that

$$u^{(eq)}(x) = 100 - 100\, x.$$

Since the heat equation is linear we can adjust the solution to obtain our familiar boundary conditions by subtracting the equilibrium, letting

$$v(t,x) := u(t,x) - u^{(eq)}(x).$$

It is easy to see that

$$H_k[u - u^{(eq)}] = 0,$$

and the general solution is thus

$$u(t,x) = 100 - 100\, x + \sum_{n=1} c_n \exp(- n^2\pi^2\, t) \sin(n\pi x)$$

- -

The properties of a solution to the heat equation are quite different from those of the wave equation. Instead of wave motion and a tolerance for discontinuity, we will find

1. Equilibration - after a long time, the solution becomes independent of time.
2. Intolerance for extremes and discontinuity. Even if the solution starts off as quite a wild function f(x), it will instantaneously become a smooth function, which can be differentiated arbitrarily often.

We investigated equilibration above, but have not yet seen how the heat equation enforces smoothness. One aspect of this phenomenon can be seen from the general solution, say with DBC, (8.4). We learned in chapters II-III that a function can behave rather terribly and still be the sum of a Fourier series, where the squares of the coefficients must tend to 0 as $n \to \infty$, owing to the Parseval formula (3.4). Because of this, suppose that all we know about the coefficients is that they are bounded by some tremendous number B. Now, if we could say that we are safe in neglecting the terms in a Fourier series beyond some finite point, $n > N$, then we would be left with a finite sum of sines or cosines, and that would clearly be differentiable as often as we would like. We also learned in chapter V that we can safely differentiate Fourier series, provided that the differentiated series makes sense.

With this in mind, let's differentiate our general solution, say, 900 times with respect to x. It is not hard to see from (8.4) that we get

$$\sum_{n=1} c_n \exp(- n^2\pi^2 t) (n\pi)^{900} \sin(n\pi x), \qquad (8.7)$$

where all we know about the c_n is that they are bounded by B. (We would have some minuses or cosines if we differentiated a number of times not divisible by 4.) If t=0, we have amplified all the coefficients, but if $t > 0$, then for large n, the exponential goes to zero faster than any power of n, and the series (8.7) converges uniformly. This happens no matter how small t is, and shows that the 900-th x-derivative of u exists as a continuous function (a theorem from advanced calculus states that a uniform limit of continuous functions is continuous).

A careful analysis suggested by this argument proves that even if the initial condition is extremely irregular, a solution of the heat equation instantly becomes smooth. A precise definition of what we mean by "smooth" is as follows.

Definition VIII.4. A function f C^∞ if it can be differentiated arbitrarily often.

Theorem VIII.5. Let u solve the heat equation on an interval $a < x < b$, and $t > 0$, with initial condition $u(0,x) = f(x)$ square integrable and either Dirichlet or Neumann boundary conditions. Then for all $t > 0$, $u(t,x)$ C^∞.

Later we shall consider the heat equation on higher dimensional regions Ω. The same theorem holds for all $\mathbf{x}$ in the interior of Ω. Also, other standard kinds of boundary conditions can be allowed in this theorem.

Another important property of the heat equation is the *maximum principle*.

Theorem VIII.6. Let u solve the heat equation. If we consider any finite space-time region, such as a rectangle $a \le x \le b$, $c \le t \le d$, then the greatest value of the temperature occurs on the boundary of the region, in this case when $x = a$ or b or when $t = c$ or d. In other words, a hot spot cannot be spontaneously generated in the interior of the material. The same is true of the minimum, a potential cold spot - the minimum always occurs on the boundary.

As with the smoothing theorem, the maximum principle also applies to higher-dimensional regions.

To prove this, let us first consider the heat equation with a *heat sink*, which absorbs energy. The equation describing this would be

$$u_t = u_{xx} - Q(t,x), \qquad (8.8)$$

where $Q \geq 0$. For now, suppose that Q is strictly positive, so $Q \geq \varepsilon > 0$ throughout the rod. Suppose that there is a hot spot, i.e., a local maximum, at (t_0,x_0), in the interior of the space-time region. By basic calculus, the space-time gradient of u must be 0 there (we know that the solution is differentiable), that is, $u_t(t_0,x_0) = u_x(t_0,x_0) = 0$. If the point is really a maximum, then $u_{xx}(t_0,x_0) \leq 0$. But this is a contradiction, since

$$u_{xx} = u_t + Q \geq \varepsilon > 0.$$

This shows that there can be no local maximum in the interior of the region, so the only possibility is for the maximum to be on the boundary.

Now suppose that there is no sink, so we have a solution of the heat equation with no sinks or sources,

$u_t = u_{xx}$, and that there is a point (t_0,x_0) where the temperature exceeds that on the boundary of a finite region. To be specific, suppose that we replace u with the function

$$v(t,x) := u(t,x) + \delta \exp(x),$$

where δ is a small but positive quantity, so small that the maximum value of v is also attained in the interior - this is possible by the intermediate value theorem of calculus, which applies to continuous functions. What equation does v solve?

$$v_t = u_t,$$

while

$$v_{xx} = u_{xx} + \delta \exp(x) = u_t + \delta \exp(x) = v_t + \delta \exp(x).$$

Solving for v_t, we have

$$v_t = v_{xx} - Q,$$

where $Q = \delta \exp(x)$ is strictly positive. We already considered this case, and found that it cannot happen. The transformation shows that it cannot happen for the basic heat equation, either.

QED

The statement about the minimum follows because -u solves the heat equation whenever u does.

Exercises VIII

VIII.1. Solve the heat equation with DBC at 0 and 1 and:

$f(x) = |x-1/2|$.

VIII.2. Solve the heat equation with NBC at 0 and 1 and:

$f(x) = \sin(\pi x)$

VIII.3. Put DBC at x=0 and NBC at x=L. Separate variables, solve the eigenvalue problem for X[x], and write down a general solution.

VIII.4. Solve the heat equation with IC:

$f(x) = 1$

and

a) DBC at x=0 and x=1
b) NBC at x=0 and x=1
c) DBC at x=0 and NBC at x=1

Plot the solutions and compare.

VIII.5. Find some initial conditions for which a unit rod with DBC will not have an asymptotic temperature profile with spatial dependence $\sin(\pi x)$, and use the computer to study the approach to equilibrium in this case.

VIII.6. Solve the initial-boundary value problem

$u_t = u_{xx}$
$u(t,0) = 50,\ u(t,2) = 100$
$u(0,x) = 20\ x$

VIII.7. Solve the initial-boundary value problem

$u_t = u_{xx} + 1$
$u(t,0) = 100,\ u(t,2) = 100$
$u(0,x) = 20\ x$

Hint: subtract something off to eliminate the non-homogeneity in the PDE as well as in the BC.

VIII.8. Use the ratio test and l'Hôpital's rule of advanced calculus to show that the series (8.7) converges uniformly.

VIII.9. It is possible for a rod to be in equilibrium at a constant temperature, and in that case it attains its maximum, and only, temperature in the interior. Explain why this state of affairs is consistent with the maximum principle and with its proof. Prove the maximum principle with 2 or 3 spatial variables, using the same argument.

VIII.10. Does the maximum principle hold for $u_t = u_{xx} + Q(t,x)$, $Q > 0$? Either give a proof or a counterexample. Interpret physically.

Appendix. Derivation of the heat equation.

Newton was the first to articulate some principles of heat flow, but it was Fourier who created the correct systematic theory of how heat flows through solids. Inside a solid there is no convective transfer of heat energy and little radiative transfer, so temperature changes only by conduction, as the energy we now recognize as molecular kinetic energy flows from hotter regions to cooler regions. The first basic principle of heat is

1. the heat energy contained in a material is proportional to the temperature, the density of the material, and a physical characteristic of the material called the *specific heat capacity*. In mathematical terms,

$$Q = \iiint_{\Omega} \rho \kappa_s \, u(t,\mathbf{x}) \, d^3x \qquad (8.a1)$$

For the other principles of heat transfer, let us do some experiments with the following materials: a hot stove, some iron rods of different, relatively short lengths and various widths, and various ceramic rods of different lengths and widths. Since these will be thought experiments only, it will be safe to use a finger as the probe. Putting your finger right on the stove will convince you that the energy transfer is proportional to the difference in temperature between your finger and the stove. Using, if necessary, a different, undamaged finger, you will also find that the rate of heat transfer is inversely proportional to the length of an iron rod intervening between your finger to the stove (fixing the cross-sectional area). In other words, the rate of heat flow from one region to another is proportional to the temperature gradient between the two regions. You will probably also agree that the rate of heat flow will be proportional to the area of the contact; for example, a short pin with one end on a hot stove and the other touching your hand is preferable to putting the palm of your hand on a frying pan. Finally, a ceramic material on the stove being usually more pleasant to the touch than hot iron, we see that the rate of heat transfer depends on the material, as measured with a physical constant known as the *heat conductivity* σ. The second basic principle is thus:

2. The heat transfer through the boundary of a region is proportional to the heat conductivity, to the gradient of the temperature across the region, and to the area of contact, or, in mathematical terms, if the boundary of the region is $\partial\Omega$, with outward normal vector $\mathbf{n}$, then

$$\frac{dQ}{dt} = \iint_{\partial\Omega} \sigma \, \mathbf{n} \quad u(t,\mathbf{x}) \, d^2x.$$

If we differentiate eq. (8.a1) with respect to time and apply Gauss's divergence theorem to (8.a2), we find that dQ/dt can be expressed in two ways as an integral:

$$\iiint_\Omega \rho \kappa_s u_t(t,\mathbf{x})\, d^3x = \iiint_\Omega \sigma\, u(t,\mathbf{x})\, d^3x$$

Since the region Ω can be an arbitrary piece of the material under study, the integrands must be equal at almost every point. If the material under study is a slab of a homogeneous substance, then ρ, κ_s, and σ are independent of the position $\mathbf{x}$, and we obtain the heat equation

$$u_t = k \cdot \nabla^2 u$$

where $k = \sigma/\rho\,\kappa_s$. The one-dimensional heat equation

$$u_t = k\, u_{xx}$$

would apply, for instance, to the case of a long, thin metal rod wrapped with insulation, since the temperature of any cross-section will be constant, due to the rapid equilibration to be expected over short distances.

Here are some approximate physical data, from sources including the CRC *Handbook of Chemistry and Physics*, 48th ed., Cleveland: Chemical Rubber Company, 1967, E.S. Dana and W.E. Ford, *Textbook of Mineralogy*, New York: Wiley, 1964, and C.J. Giankoplis, *Transport Processes and Unit Operations*, Englewood Cliffs, NJ: Prentice-Hall, 1993. These characteristics actually depend on such factors as temperature and pressure. According to the Dulong-Petit law, the specific heat capacity of an element is roughly inversely proportional to its atomic weight, but there is no simple rule for solid compounds. Thus these are not necessarily accurate.

	σ	ρ	κ_s	k
units	cal/(hr-cm-deg)	gm/cm^3	cal/(gm-deg)	cm^2/hr
glass	80	2.8	.2 (guess)	140
iron	6800	7.9	.1	8600
granite	300	2.7	.19 (orthoclase)	585
pork	75	1.1	0.6	113
water	5	1	1 (or .5 for ice)	5
wood	350	.7	.2 (guess)	2500
air	.002	.0011	.25 (N_2)	7

version of 21 Nov 1997

IX. PDEs in space†

Until now we have studied partial differential equations in one space dimension, x. Now we are ready to consider some problems in two or three space dimensions. Fortunately, the technique of separation of variables in more dimensions presents only a few new conceptual issues, and the technical complications are quite manageable..

In three space dimensions the wave equation has the form:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \qquad u := c^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right).$$

It describes many sorts of waves that move through space, such as acoustic waves in air or another fluid medium, where c is the speed of sound in that medium. It also describes any component of an electromagnetic field in free space if c is the speed of light (see the derivation from Clerk Maxwell's equations in the appendix). The expression on the right is c^2 times the *Laplacean* of u, which is abbreviated $\nabla^2 u$ (some texts prefer the notation Δu). It is sometimes convenient to think of $\nabla^2 u$ as the *Laplace operator* ∇^2 applied to u.

The heat equation in more than one space dimension likewise involves the Laplace operator:

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

Both these equations describe how a quantity will change in time - but it might be static. If an electromagnetic field or a temperature is at equilibrium, the time derivatives will be 0, and we are led to the potential equation of Laplace,

---

$$\nabla^2 u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{9.1}$$

or, in three dimensions,

$$\nabla^2 u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0. \tag{9.2}$$

Laplace's equation arises in many other applications as well. It is the equation which describes the

- displacement of an elastic membrane
- electrostatic potential function
- equilibrium concentration of a suspensate in a still fluid

and still further physical quantities.

Solutions of Laplace's equation are called *harmonic functions*.

Let's begin with the two-dimensional Laplace equation on a rectangle

$$0 < x < a$$
$$0 < y < b.$$

This could arise from the wave equation for a long, rectangular wave-guide, assuming that the electric field is independent of z and static, that is, independent of t; or it might also arise from heat flow in a rectangular solid at equilibrium. In either case we are left with an equation in only two independent variables. It should thus be no more complicated than the earlier equations in the variables t and x.

Laplace's equation is the prototype of an elliptic equation, with different qualitative properties from either hyperbolic equations (like the wave equation) or parabolic equations (like the heat equation). The auxiliary conditions that are imposed are quite different from those for the wave equation. Specifically, we are normally given boundary conditions on the entire boundary, and are not specially concerned with an "initial" value of either variable x or y. Dirichlet boundary conditions for this problem are typically nonhomogeneous, of the form:

$$u(0,y) = f_1(y) \tag{9.3}$$

$$u(a,y) = f_2(y) \tag{9.4}$$
$$u(x,0) = f_3(x) \tag{9.5}$$
$$u(x,b) = f_4(x) \tag{9.6}$$

where the four boundary functions are supposed given. In electromagnetic theory, for example, they are determined if the charge is measured on the boundary. In thermodynamics, they are the temperature determined by the temperature at the edge of the rectangle.

Separation of variables works much as it did for the heat equation and the wave equation. As with our earlier equations, we begin with the *ansatz* that u is a product solution

$$u(x,y) = X(x)\,Y(y)$$

and substitute into Laplace's equation; it is again convenient to divide the result by u. We use first information from the partial differential equation, then the information from homogeneous boundary conditions, and lastly information from inhomogeneous boundary conditions, which are treated much like the initial conditions for the earlier PDE's.

Rather than tackling four non-homogeneous boundary conditions (9.3)-(9.6) all at once, we begin by setting three of the four boundary functions to 0. For definiteness, we also make a specific choice of the fourth while developing the ideas:

Model Problem IX.1. Let us solve Laplace's equation with boundary conditions which are homogeneous on three of the four sides of a rectangle:

$$u(0,y) = 0 \tag{9.7}$$
$$u(1,y) = 1 \tag{9.8}$$
$$u(x,0) = 0 \tag{9.9}$$
$$u(x,2) = 0. \tag{9.10}$$

Solution.

Separation begins much as before, but note that boundary condition (9.8) is not homogeneous, so it is not consistent with the superposition principle. If we add two functions satisfying (9.8), for example, we will get a function with the boundary value 2, not 1, when x = a. As in our earlier solutions by separating variables, let us guess that the solution is a product, u(x,y) = X(x) Y(y) and subsitute into Laplace's equation (9.1). Dividing through by X(x) Y(y), we find that

$$\frac{X''(x)}{X(x)} + \frac{Y''(x)}{Y(x)} = 0.$$

Evidently, if the y-independent quantity X"/X equals the x-independent quantity -Y"/Y, both must be a constant, and we have the familiar ordinary differential equations,

$$X'' = \mu X$$
$$-Y'' = \mu Y$$

(notice the difference of sign).

A good rule of thumb in this subject is to deal with the most homogeneous boundaries first. It is the variable y which has two homogeneous boundary conditions in this case. It satisfies essentially the same eigenvalue problem as we have seen in previous chapters:

$$-Y'' = \mu Y$$
$$Y(0) = 0,\ Y(2) = 0$$

At this stage we can probably recognize the eigenfunctions and eigenvalues as

$$Y_n(y) = \sin\left(\frac{n\pi}{2} y\right), \quad \mu_n = \left(\frac{n\pi}{2}\right)^2.$$

This same constant μ_n enters in to the X equation, but with the other sign. The general solution for X is therefore a linear combination of $\exp(n \pi x/2)$ and $\exp(- n \pi x/2)$, rather than sines and cosines. There is a better choice for the basis of this solution space, though, namely

$$X_n(x) = C_1 \cosh(n \pi x/2) + C_2 \sinh(n \pi x/2).$$

One homogeneous boundary condition applies to this function, forcing

$$X_n(0) = 0.$$

The Dirichlet boundary condition thus eliminates the hyperbolic cosine, allowing us to normalize the function X_n as

$$X_n(x) = \sinh(n \pi x/2).$$

For the general solution we get:

$$u(x, y) = \sum_{k=1}^{\infty} c_k X_k(x)\, Y_k(y) = \sum_{n=1}^{\infty} c_n \sin(n \pi y / 2) \sinh(n \pi x / 2).$$

There is only one boundary condition left, the nonhomogeneous (9.8) at x = 1. Before incorporating it, we make a general linear combination of all the product solutions which are consistent with what we know so far:

$$u(x, y) = \sum_{n=1}^{\infty} c_n \sin(n \pi y / 2) \sinh(n \pi x / 2)$$

This is a Fourier sine series in the variable y with the complication that the coefficients have an extra factor,

$$b_n = c_n \sinh(\pi n x/2).$$

The series must be matched to the series for the function $f_2(y) = 1$, the coefficients in which can be calculated with the familiar formula. If

$$\sum_{n=1}^{\infty} b_n \sin(n \pi y / 2) = 1,$$

then the Fourier coefficients are easily calculated as:

$$b_n = \frac{4}{n\pi} \text{ if n is odd, and otherwise 0.}$$

To find c_n, now divide by $\sinh(n \pi/2)$. The solution becomes

$$u(x, y) = \sum_{\substack{n=1 \\ n\,\text{odd}}}^{\infty} \frac{4}{n\pi} \sin(n \pi y / 2) \frac{\sinh(n \pi x / 2)}{\sinh(n \pi/ 2)}.$$

- -

It wasn't critical that the function at x=1 was just the constant function 1; given any other function of y at x=1 as a Fourier sine series,

$$u(1, y)= f_2(y) = \sum_{n = 1}^{\infty} b_{2,n} \sin(n \pi y / 2),$$

the solution we get would still be of the form:

$$u(x, y) = \sum_{n= 1}^{\infty} b_{2,n} \sin(n \pi y / 2) \frac{\sinh(n \pi x / 2)}{\sinh(n \pi/ 2)}.$$

Now let us return to the full problem with boundary conditions (9.3) -(9.6). Don't start over and rederive everything! Use the solution we already obtained as a springboard, as follows.

Let us rename the solution we just obtained for the simplified boundary conditions

$u(0,y) = 0$
$u(a,y) = f_2(y)$
$u(x,0) = 0$
$u(x,b) = 0,$

with only the function f_2 different from 0; for future bookkeeping u_2 will be better than u:

$$u_2(x, y) = \sum_{n=1}^{\infty} b_{2,n} \sin(n \pi y/ 2) \frac{\sinh(n \pi x / 2)}{\sinh(n a \pi/ 2)}.$$

(Notice the a in the denominator, which was fixed as 1 in the model problem.)

Next, suppose that we had the boundary conditions:

$u(0,y) = 0$
$u(a,y) = 0$
$u(x,0) = 0$
$u(x,b) = f_4(x).$

The only differences here are that f_2 becomes f_4, a becomes b, and the variables x and y are switched. Thus:

$$u_4(x, y) = \sum_{n=1}^{\infty} b_{4,n} \sin(n\pi x/2) \frac{\sinh(n\pi y/2)}{\sinh(n b \pi/2)}. \qquad (9.11)$$

As the next piece of the puzzle, suppose that

$u(0,y) = f_1(y)$
$u(a,y) = 0$
$u(x,0) = 0$
$u(x,b) = 0.$

We can get these boundary conditions from our original ones with f_2 by interchanging x and a-x. Notice that this does not affect the potential equation, because under this change of variable there are two compensating changes of sign in $\partial^2 u/\partial x^2$; each time you use the chain rule there is a factor of -1. The answer has to be:

$$u_1(x, y) = \sum_{n=1}^{\infty} b_{1,n} \sin(n\pi y/2) \frac{\sinh(n\pi (a-x)/2)}{\sinh(n a \pi/2)}. \qquad (9.12)$$

Similarly, if

$u(0,y) = 0$
$u(a,y) = 0$
$u(x,0) = f_3(x)$
$u(x,b) = 0,$

then

$$u_3(x, y) = \sum_{n=1}^{\infty} b_{3,n} \sin(n\pi x/2) \frac{\sinh(n\pi(b-y)/2)}{\sinh(n b \pi/2)}. \qquad (9.13)$$

Because of the principle of superposition, the entire solution of the boundary-value problem with boundary conditions (9.3)- (9.6) is the sum:

$u(x,y) = u_1(x,y) + u_2(x,y) + u_3(x,y) + u_4(x,y).$

- -

Let us turn our attention now to a fully multidimensional problem. There will be few new concepts, though as we shall see the additional dimensions require some extra book-keeping with several indices to label the pieces of the solution correlating with the various dimensions. We illustrate the topic in a specific example.

Model Problem IX.2. How a mathematician cooks a cube steak. Consider how to cook a steak which is one meter on a side (it is a whale steak) under the following conditions.
BC: The steak will be put into a preheated oven at temperature 200 C at time t=0.
IC: At time t=0, the steak is pulled directly from the freezer (u(0, x,y,z) = 0) and put into the oven for four hours.

We wish to find the temperature throughout the interior of the steak at that time. Since the mathematician shops at a terrible meat market (but doesn't really notice), we will assume that the steak has the thermal properties of wood, so that in the heat equation,
$k = 2500$ cm^2/hr.

Solution.

Instead of directly solving for the temperature, let u(t,x,y,z) be the temperature minus 200 C, in order to have homogeneous Dirichlet boundary conditions. With this change, for all $t > 0$ the conditions on the six faces of the cube are

$$\begin{aligned} &u(t,0,y,z) = 0 \\ &u(t,x,0,z) = 0 \\ &u(t,x,y,0) = 0 \\ &u(t,100,y,z) = 0 \\ &u(t,x,100,z) = 0 \\ &u(t,x,y,100) = 0, \end{aligned} \tag{9.14}$$

while the PDE is the usual heat equation (HE),

$$u_t := \frac{\partial u}{\partial t} = k \quad ^2u, \tag{HE}$$

and the initial condition becomes

$$u(0,x,y,z) = -200$$

Solution.

Step 1. Construct a general solution by separation of variables.

Suppose as usual that a solution of the heat equation is a product of the form

$$u(t,x) = T(t)\, Q(x);$$

a new feature here is the vector variable x. Plugging into the heat equation and dividing by T Q, we find:

$$\frac{T'(t)}{T(t)} = k\,\frac{\nabla^2 Q(\mathbf{x})}{Q(\mathbf{x})}.$$

We have a multidimensional eigenvalue problem to solve, viz.,

$$-\nabla^2 Q(\mathbf{x}) = \mu\, Q(\mathbf{x}),$$

subject to the Dirichlet boundary conditions (9.14). I inserted the minus sign, as before, because experience teaches me to expect μ to be positive this way. (It is a matter of book-keeping and not essential.)

How do we solve this equation? Why, by separating variables again, of course! Let

$$Q(x) = X(x)\, Y(y)\, Z(z)$$

and evaluate

$$\mu = \frac{\nabla^2 Q(\mathbf{x})}{Q(\mathbf{x})}.$$

The result, after cancelling common factors is

$$\mu = \frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} + \frac{Z''(z)}{Z(z)}. \qquad (9.15)$$

which is a very curious equation, since any one of the three terms on the right could be isolated on one side of the equation. For instance, if we solve for X"/X, we find

$$X''(x)/X(x) = (\text{something independent of } x).$$

Similarly for Y"/Y and Z"/Z. The conclusion is that we have three one-dimensional eigenvalueproblems, all precisely like the familiar eigenvalue problem from Chapter 6:

$$-X'' = \mu_1 X$$
$$-Y'' = \mu_2 Y$$
$$-Z'' = \mu_3 Z$$

If we compare with (9.15), we see that

$$\mu = \mu_1 + \mu_2 + \mu_3,$$

and by this stage we can immediately see that μ_1 and X are of the form

$$\mu_1 = (m_1 \pi/100)^2,$$
$$X(x) = \sin(m_1 \pi x /100),\ m_1 = 1, 2, \ldots.$$

Likewise,

$$\mu_2 = (m_2 \pi/100)^2,\ Y(y) = \sin(m_2 \pi y /100),\ m_2 = 1, 2, \ldots, \text{ and}$$
$$\mu_3 = (m_3 \pi/100)^2,\ Z(z) = \sin(m_3 \pi z /100),\ m_3 = 1, 2, \ldots.$$

The product solutions thus look like:

$$u_{m_1 m_2 m_3}(t,x,y,z) = \exp\left(-\frac{\pi^2}{10000}(m_1^2+m_2^2+m_3^2)kt\right) \sin\left(\frac{m_1\pi x}{100}\right) \sin\left(\frac{m_2\pi y}{100}\right) \sin\left(\frac{m_3\pi z}{100}\right).$$

The general solution is a triple sum:

$$u(t,x,y,z) = \sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} \sum_{m_3=1}^{\infty} c_{m_1 m_2 m_3} u_{m_1 m_2 m_3}(t,x,y,z),$$

where the multi-indexed constants $c_{m1,m2,m3}$ can have arbitrary values, subject only to convergence of the series.

Step 2.

The constants will be determined by the initial conditions when we set t=0, obtaining a triple Fourier sine series. In this model problem the coefficients need to satisfy

$$u(0,x,y,z) = -200 = \sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} \sum_{m_3=1}^{\infty} c_{m_1 m_2 m_3} u_{m_1 m_2 m_3}(0,x,y,z),$$
$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} c_{m_1 m_2 m_3} \sin\left(\frac{m_1\pi x}{100}\right) \sin\left(\frac{m_2\pi y}{100}\right) \sin\left(\frac{m_3\pi z}{100}\right).$$

There are two ways to proceed here. It turns out that the products of three sines compose a complete orthogonal set for the cube, so we could directly evaluate the coefficients by the use of projections in function space. In this case the inner product uses a thre-dimensional integral over the cube.

You may find it more congenial, however, to rely on the usual, one-variable Fourier sine series, as follows. Imagine for the moment that y and z are fixed. What remains is a function of x, and we could use the orthogonality of the functions $\sin(m_1 \pi x/L)$ to remove the sum over m_1:

$$\int_0^{100} u(0,x,y,z) \sin\left(\frac{n_1\pi x}{100}\right) dx = \sum_{m_1}\sum_{m_2}\sum_{m_3} c_{m_1m_2m_3} \int_0^{100} \sin\left(\frac{n_1\pi x}{100}\right) \sin\left(\frac{m_1\pi x}{100}\right) dx \times \sin\left(\frac{m_2\pi y}{100}\right) \sin\left(\frac{m_3\pi z}{100}\right)$$

$$= \sum_{m_1}\sum_{m_2}\sum_{m_3} c_{m_1m_2m_3} 50\, \delta_{n_1m} \sin\left(\frac{m_2\pi y}{100}\right) \sin\left(\frac{m_3\pi z}{100}\right)$$

$$= \sum_{m_2}\sum_{m_3} c_{\mathbf{n}_1m_2m_3} 50 \sin\left(\frac{m_2\pi y}{100}\right) \sin\left(\frac{m_3\pi z}{100}\right).$$

The point here is that $\sin(m_1 \pi x /100)$ and $\sin(n_1 \pi x /100)$ are orthogonal unless $m_1=n_1$, so only one term survives from the sum over m_1.

If we now multiply by $\sin(n_2 \pi y/100)$ and integrate from 0 to 100, which will eliminate all but one term in the sum over m_2:

$$\int_0^{100}\int_0^{100} u(0,x,y,z) \sin\left(\frac{n_1\pi x}{100}\right) \sin\left(\frac{n_2\pi y}{100}\right) dx\,dy = 50^2 \sum_{m_3} c_{\mathbf{n}_1\mathbf{n}_2m_3} \sin\left(\frac{m_3\pi z}{100}\right).$$

Finally, let us multiply the result by $\sin(n_3 \pi z /100)$ and integrate over z:

$$\int_0^{100}\int_0^{100}\int_0^{100} u(0,x,y,z) \sin\left(\frac{n_1\pi x}{100}\right) \sin\left(\frac{n_2\pi y}{100}\right) \sin\left(\frac{n_3\pi z}{100}\right) dx\,dy\,dz = 50^3 c_{n_1n_2n_3}.$$

The result of the calculation is

$$c_{n_1 n_2 n_3} = \frac{8\left(1-(-1)^{n_1}\right)\left(1-(-1)^{n_2}\right)\left(1-(-1)^{n_3}\right)}{n_1 n_2 n_3 \pi^3}.$$

Notice that this is 0 unless all three integers n_k are odd, in which case it is $64/(n_1\, n_2\, n_3)\, \pi^3$.

- -

Exercises IX

IX.1. Derive solutions (9.11)-(9.13) carefully.

IX.2. Find the solution of Laplace's equation with mixed Dirichlet and Neumann boundary conditions :

$u(0,y) = f_1(y)$
$u(a,y) = f_2(y)$
$\partial\, u(x,0) / \partial\, y = f_3(x)$
$\partial\, u(x,b) / \partial\, y = f_4(x)$

Hint: Solve four simpler problems, each of which has three of these functions equal to 0. Then sum the result.

IX.3. Find the solution of Laplace's equation with the more specific mixed Dirichlet and Neumann boundary conditions :

$u(0,y) = y$
$u(1,y) = 0$
$\partial\, u(x,0) / \partial\, y = 0$
$\partial\, u(x,2) / \partial\, y = 1$

IX.4. Find the specific solution of Model Problem IX.2 by regarding the functions

$$\sin(n_1\, \pi\, x /100)\, \sin(n_2\, \pi\, y /100)\, \sin(n_3\, \pi\, z /100)$$

directly as an orthonormal set on the cube.

IX.5. Change the boundary conditions of Model Problem IX.2 so that two of the faces of the cube are insulated, i.e., Neumann boundary conditions. Discuss the differences depending on whether the two insulated faces are adjacent or opposite.

IX.6. Consider a rubber cube of side 1 firmly encased in a metal box of the same dimensions. The vertical displacement of the cube satisfies the wave equation with c=1000, and the box enforces zero Dirichlet boundary conditions analogout to (9.14) on all six faces (but at coordinates 0 and 1, not 0 and 100).

a) Find the normal modes of vibration.

b) Solve the initial-boundary-value problem with initial conditions:

$$u(0,x,y,z) = xyz \sin(\pi x)\sin(\pi y)\sin(2\pi z)$$
$$u_t(0,x,y,z) = \sin(2\pi x).$$

Appendix

Clerk Maxwell codified the laws of electricity and magnetism in four partial differential equations, which in in rationalized mks units read:

$$\nabla\times\mathbf{H}\ (\equiv \mathrm{Curl}\ \mathbf{H}) = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \qquad \text{(extension of Ampère's law)}$$

$$\nabla\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{(Faraday's law of induction)}$$

$$\nabla\cdot\mathbf{B}\ (\equiv \mathrm{Div}\ \mathbf{B}) = 0 \qquad \text{(Gauß's law for magnetism, in the absence of monopoles)}$$

$$\nabla\cdot\mathbf{D} = \rho \qquad \text{(Gauß-Coulomb law)}$$

Here, $\mathbf{H}$ is the magnetic force and $\mathbf{B} = \mu\,\mathbf{H}$ is the magnetic induction; $\mathbf{E}$ is the electric intensity and $\mathbf{D} = \varepsilon\,\mathbf{E}$ is the electric induction; $\mathbf{J}$ is the current density and ρ is the charge density. In free space, $\rho = 0$ and $J = 0$, while the value of $\mu\,\varepsilon$ turns out to equal $1/c^2$, where c is the speed of light.

Let us calculate the second time derivative of $\mathbf{B}$:

$$\frac{\partial^2 \mathbf{B}}{\partial t^2} = -\nabla\times\left(\frac{\partial \mathbf{E}}{\partial t}\right) = -\frac{1}{\varepsilon}\nabla\times\left(\frac{\partial \mathbf{D}}{\partial t}\right) = -\frac{1}{\varepsilon\mu}\nabla\times(\nabla\times\mathbf{B}).$$

There happens to be a vector identity for the curl of a curl:

$$\nabla\times(\nabla\times\mathbf{v}) = -\nabla^2\mathbf{v} + \nabla(\nabla\cdot\mathbf{v}),$$

where the Laplacian of a vector $v = (v_1, v_2, v_3)$ is defined by its action on each component:

$$\nabla^2\mathbf{v} := (\nabla^2 v_1, \quad \nabla^2 v_2, \quad \nabla^2 v_3).$$

Because of the absence of magnetic monopoles, the vector identity as applied to **B** becomes

$$\nabla\times(\nabla\times(\mathbf{B})) = -\nabla^2\mathbf{B},$$

which means that

$$\frac{\partial^2\mathbf{B}}{\partial t^2} = \frac{1}{\varepsilon\mu}\nabla^2\mathbf{B}.$$

Clerk Maxwell recognized that this equation was the wave equation in three dimensions, for each component of the magnetic field B. A very similar analysis shows that each component of E satisfies the same wave equation. Of course, Clerk Maxwell's equations imply that the electric and magnetic fields in a wave are coupled; if we assume a traveling wave, the fields **E** and **B** turn out to oscillate exactly out of phase.

After deducing the possibility of these electromagnetic waves, Clerk Maxwell was interested in studying the properties of these waves to see whether they could be experimentally verified. Imagine his excitement in 1854(?) when he calculated the characteristic speed of the waves with the best known values of μ and ε for a vacuum and discovered that it was quite close to the observed value of the speed of light. This was surely one of the crowning moments in the history of physics. In Clerk Maxwell's own words,

> ...we can scarcely avoid the inference that *light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena*.

(W.D. Niven, ed., The Scientific Papers of J. Clerk Maxwell, Cambridge, 1890, vol. I, p. 500).

version of 31 December 1995

X. PDEs on a disk[†]

Let us look again at potential equation of Laplace, but this time in polar coordinates. This equation would apply to the equilibrium temperature distribution in a thin disk insulated on its two flat faces, or to a long cylinder along the z-axis, provided that the temperature does not depend on z. A temperature could be prescribed on the outside edge as a function of the angular variable θ, producing a boundary-value problem of the form

$$u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0. \qquad (10.1)$$

$$\text{with } u(a,\theta) = f(\theta) \qquad (10.2)$$

The PDE (10.1) is what you get if you change variables using the 2-variable chain rule. I shall leave this to the exercises, but notice at least that (10.1) is dimensionally consistent: r has units of length and $\partial/\partial r$ of length^{-1}, whereas θ is dimensionless. Hence each term in (10.1) has dimensions length^{-2}.

If we try to find a product solution for (10.1) of the form

$$u(r, \theta) = R(r)\, Q(\theta)\ ,$$

we readily find

$$0 = \frac{\Delta u}{u} = \frac{R''(r)}{R(r)} + \frac{1}{r}\frac{R'(r)}{R(r)} + \frac{1}{r^2}\frac{Q''(\theta)}{Q(\theta)}.$$

The equation for Q is the familiar one that gives sines and cosines, but what are the boundary conditions? Answer: periodic BC. Since there is no physical difference between θ and $\theta + 2\pi$, we must have:

$$Q(\theta + 2\pi) = Q(\theta). \qquad \text{(PBC)}$$

In order to arrange periodicity, we take solutions in the form

$$Q(\theta,m) = a(m)\,\mathrm{Cos}(m\,\theta) + b(m)\,\mathrm{Sin}(m\,\theta), \qquad (10.3)$$

where m is an integer, or, equivalently,

$$Q(\theta,m) = c(m)\,\mathrm{Exp}(i\,m\,\theta) + c(-m)\,\mathrm{Exp}(-\,i\,m\,\theta). \qquad (10.4)$$

The function R satisfies the *equidimensional*, or *Euler*, equation:

$$R'' + (1/r)\,R' - (m^2/r^2)\,R = 0 \qquad (10.5)$$

Equidimensional equations can be solved by making the guess $R = r^a$ and figuring out what the constant a has to be. In this case, $a = \pm m$, forced by the periodicity of the solutions (10.3), so the general solution is of the form:

$$a_0 + \sum_{m=1}^{\infty} r^m\left(a_m \cos(m\,\theta) + b_m \sin(m\,\theta)\right) + \sum_{m=1}^{\infty} r^{-m}\left(a_{-m} \cos(m\,\theta) + b_{-m} \sin(m\,\theta)\right). \qquad (10.6)$$

If we are studying the equilibrium temperature (or potential, or equilibrium displacement of a membrane) on a disk, the solutions containing r^{-m}, $m>0$, are unphysical, so in this case the general solution will be of the form:

$$a_0 + \sum_{m=1}^{\infty} r^m\left(a_m \cos(m\,\theta) + b_m \sin(m\,\theta)\right) \qquad (10.7)$$

where the sum now runs from m=1 to infinity. I have written a_0 separately for the same reasons as with Fourier series.

Model Problem X.1. Solve (10.1) with the boundary condition

$$u(1,\theta) = (\cos[\theta])^2.$$

Solution.

To find a solution in the interior, we need to determine the coefficients in (10.7). This is done by substituting r = 1 and comparing:

$$a[0] + \text{Sum}(a[m]\cos[m\,\theta] + b[m]\sin[m\,\theta]) = (\cos[\theta])^2.$$

In other words, the coefficients a[m] and b[m] are none other than the usual Fourier coefficients for the function f, here $(\cos[\theta])^2$. Rather than doing integrals here, the easiest way to find these Fourier coefficients is to use the double-angle formula in the form:

$$\cos[2\theta] = 2\,(\cos[\theta])^2 - 1$$

Remember - if you can find Fourier coefficients by some trick, the answer is perfectly correct, since the Fourier series is uniquely determined. This is great thing about uniqueness theorems in mathematics - they allow you find an answer by the easiest method available, without worrying that it might be a different answer from the one you would get from a harder but more standard procedure.

We conclude that

$$(\cos(\theta))^2 = \frac{1}{2} + \frac{1}{2}\cos(2\theta),$$

or in other words that $a_0 = 1/2$, $a_2 = 1/2$, and all other coefficients are 0.

That is the solution on the boundary of the disk (r=1); inside the disk the temperature is:

$$u(r,\theta) = \frac{1}{2} + \frac{1}{2}r^2\cos(2\theta).$$

- -

Next, let us look again at a time-dependent problem, such as the heat equation on a disk, in polar coordinates.

Setting k=1 for convenience, the heat equation in takes on the form:

$$u_t = \quad ^2\,u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}. \qquad (10.8)$$

As before, we begin by searching for product solutions. Recall from Chapter IX that it is convenient to separate vector variables one at a time in the product solution. If

$$u = T[t]\, V[r,\theta],$$

then we find that

$$T'/T = (\ ^2 V)/V \qquad (10.9)$$

and we can make the usual argument that the left side is independent of space while the right side is independent of time, hence both equal some constant, called $-\lambda$. The solutions for T are of the form

$$T[t] = \text{const.}\ \ \text{Exp}[-\lambda\, t],$$

where λ should be determined from the boundary conditions in space.

Model Problem X.2. Suppose that a thin disk of radius A is insulated on its faces while the round edge is held at temperature 0. Find the eigenfunctions and eigenvalues for the spatial part of the separated equation (10.9).

Solution.

On the edge we have Dirichlet boundary conditions of the form

$$u[t,A,\theta] = 0 \qquad \text{(DBC)}$$

This two-dimensional problem, resembles (10.1) except that in the eigenvalue equation,

$$-\ ^2 V(r,\theta) = \lambda\ V(r,\theta),$$

the eigenvalue λ need not be zero. The constant λ may be zero, in which case the disk is at equilibrium and because of the boundary conditions at r=A, the equilibrium temperature is 0 throughout. This trivial solution is not considered an eigenfunction, since it is not useful in building a series for a general solution.

Separating variables again by writing

$$V = R[r]\, Q[\theta],$$

leads to the same eigenvalue equation for Q[θ] as before. The boundary conditions are also the same, periodic, so the solutions are the same sines and cosines as in (10.3), indexed by the integer m.

Something new occurs for the other part of the solution, R. Evaluating

$$\lambda = -\frac{\partial^2 V}{V}$$

with the knowledge that Q(θ) = sin(2 m θ) or cos(2 m θ), we get an ordinary differential equation for R. In place of the Euler, equation 10.5), which was solved with power functions, we get Bessel's equation,

$$R'' + (1/r)\, R' + (\lambda - m^2/r^2)\, R = 0, \qquad (10.10)$$

which no longer has elementary solutions. Just as with the equidimensional equation, near r=0 the solutions can behave like either r^m or r^{-m}, but the latter are unphysical because the temperature can't diverge at r=0. The regular solutions, behaving roughly like r^m near r=0, are called regular Bessel functions and, with a standard choice of the overall constant, denoted $J_m(\lambda^{1/2} r)$. Mathematica uses the notation BesselJ. They are normalized so that $J_m(x) \sim x^m/(2\, m!)$ as $x \to 0$:

```
In:=   Plot[BesselJ[0,x],{x,0,25}]
Out=
```