II. Partial Differential Equations and Fourier Methods

Introductory Example: The Heat Equation

The *heat equation* or *diffusion equation* in one space dimension is

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}\,. \qquad (*)$$

It's a *partial differential equation* (PDE) because partial derivatives of the unknown function with respect to two (or more) variables appear in it.

I refer to our textbooks for more thorough discussions of the following:

1. *Physical interpretation and derivation:* In the most usual application, t is time, and x is a spatial variable (say along a thin wire, a homogeneous bar, or an imaginary one-dimensional world studied in the (justified) hope that the solutions of the more difficult three-dimensional heat equation

$$\nabla^2 u = \frac{\partial u}{\partial t}$$

will be qualitatively similar). $u(t, x)$ is the temperature in the bar (possibly with something subtracted off, as we'll see). The equation follows quickly from algebraic formulations of the physical principles that

(1) the amount of heat energy in any small region of the bar is proportional to the temperature there,

(2) the rate of heat flow is proportional to the derivative of the temperature, since it's driven by temperature differences between regions.

In fact, the same equation describes many other *diffusion* processes. It — or some modification of it — arises whenever one studies the large-scale, averaged effects of the random motion of many particles. (Think of a cloud of mosquitos released from a cage in one corner of a large room.)

2. *Scaling to remove irrelevant constants:* We are free to redefine the units in which u, t, and x are measured. In general, the equation will first be presented to us as

$$K\,\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}\,,$$

where K is a constant depending on the physical properties of the material (specifically, its specific heat and thermal conductivity, which are the coefficients in the

two "algebraic formulations" mentioned above). By rescaling x or t (or both), we can change K to 1. So there is no loss of generality in ignoring K henceforth. This uses up only one of the three degrees of freedom in the units. The other two can be used in other ways.

Typically, our bar will have a finite length, say L. We can rescale x to make L have any convenient value; the most popular choices are 1 (not surprisingly) and π (for reasons that will become obvious later). After that, we can rescale t so as to keep K equal to 1. We can also add a constant to x so that the left endpoint of the bar is at $x = 0$.

Scaling u will not change the form of the equation, since it is linear (see below). However, this scaling freedom can be used to simplify a boundary condition or initial condition.

3. *Initial and boundary conditions:* To make a PDE into a well-defined problem, we have to state over what domain of the independent variables we hope to solve it, and we need to have enough information about the behavior of u on the boundary of that domain to make the solution of the problem unique. For physical and mathematical reasons, time and space enter the heat problem in different ways. One finds:

(1) If we know the temperature distribution at one time (say $t = 0$), we can hope to predict the temperature at *later* times, but not necessarily at *earlier* times. (If we observe a room full of mosquitos, it is hard to tell by looking which corner they flew out of.) Thus we will be solving $(*)$ in the region

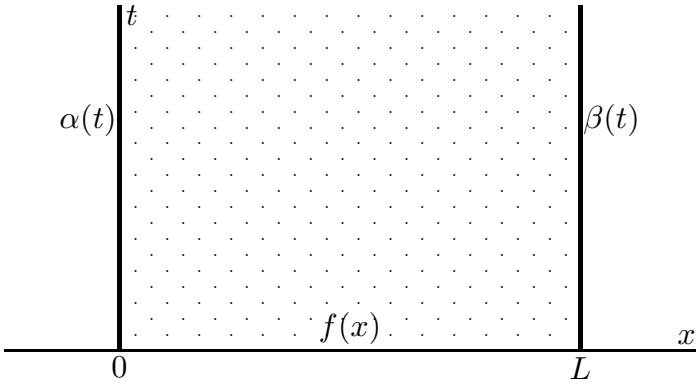
$$0 < x < L, \qquad t > 0$$

given *initial data*

$$u(0, x) = f(x) \quad \text{for } 0 < x < L.$$

(2) We need to know what happens to the heat when it reaches the end of the bar. Obviously it will make a big difference to the temperature distribution whether the end is insulated or in contact with some other material which can conduct heat away. There are four standard types of *boundary conditions* that can be considered. Each type is worthy of consideration for its own sake as a mathematical possibility, but it happens that each one has a real physical interpretation in the heat problem:

(A) *Dirichlet* condition: $u(t, 0) = \alpha(t)$ for some given function α. This says that the temperature at the end of the bar is controlled (say by contact with a "heat bath").



(B) *Neumann* condition: $\dfrac{\partial u}{\partial x}(t,0) = \alpha(t)$. This says that the heat flow through the end is controlled. This is hard to do in practice, except in the special case $\alpha = 0$, which says that the end is *insulated*.

(C) A generalization of the first two is a *Robin* condition:

$$c_1 u(t,0) + c_2 \frac{\partial u}{\partial x}(t,0) = \alpha(t),$$

where the c's are constants characteristic of the situation. Such a condition arises in *convective cooling*, when the bar is in contact with a less dense medium (such as air) which can carry away heat, but not fast enough to lower the bar temperature immediately to the medium's temperature.

In all these cases of conditions at $x = 0$, one would need another condition (not necessarily the same kind) at $x = L$ to complete the specification of the problem.

(D) *Periodic* boundary conditions: These deal with both endpoints at once.

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

The usual physical interpretation of this is that our "bar" is actually a *ring*, and x is an angle. (Thus $L = 2\pi$ when x is measured in radians.)

One tends to think of the boundary conditions as part of the definition of the *physical system* under study, while the initial conditions label the various *possible solutions* of the equations of motion of that given system. In other words, in our discussions the boundary conditions are usually "more constant", the initial conditions "more variable". Imposing the initial conditions is usually the last step in finding a solution, as it is usually is for ODEs, too.

Fundamental Concepts: Linearity and Homogeneity

This is probably the most abstract lecture of the course, and also the most important, since the procedures followed in solving PDEs will be simply a bewildering welter of magic tricks to you unless you learn the general principles behind them.

Linear equations and linear operators

I think that you already know how to recognize linear and nonlinear equations, so let's look at some examples before I give the official definition of "linear" and discuss its usefulness.

Algebraic equations:

Linear	*Nonlinear*
$x + 2y = 0,$ $x - 3y = 1$	$\tan x = 2x$

Ordinary differential equations:

Linear	*Nonlinear*
$\dfrac{dy}{dt} + t^3\, y = \cos 3t$	$\dfrac{dy}{dt} = t^2 + e^y$

Partial differential equations:

Linear	*Nonlinear*
$\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2}$	$\dfrac{\partial u}{\partial t} = \left(\dfrac{\partial u}{\partial x}\right)^2$

What distinguishes the linear equations from the nonlinear ones? The most visible feature of the linear equations is that they involve the unknown quantity (the dependent variable, in the differential cases) *only to the first power*. The unknown does not appear inside transcendental functions (such as sin and ln), or in a denominator, or squared, cubed, etc. This is how a linear equation is usually recognized by eye. Notice that there may be terms (like $\cos 3t$ in one example)

which don't involve the unknown at all. Also, as the same example term shows, there's no rule against nonlinear functions of the *independent* variable.

The formal definition of "linear" stresses not what a linear equation looks like, but the properties that make it easy to describe all its solutions. For concreteness let's assume that the unknown in our problem is a (real-valued) function of one or more (real) variables, $u(x)$ or $u(x, y)$. The fundamental concept is not "linear equation" but "linear operator":

Definition: An operation, L, on functions is *linear* if it satisfies

$$L(u + v) = L(u) + L(v) \qquad \text{and} \qquad L(\lambda u) = \lambda L(u) \tag{$*$}$$

for all functions u and v and all numbers λ.

Examples of linear operations are

differentiation of u: $\quad L(u) \equiv \dfrac{du}{dx}$,

multiplication of u by a given function of x: $\quad L(u) \equiv x^2 u(x)$,

evaluation of u at a particular value of x: $\quad L(u) \equiv u(2)$,

integration of u $\quad L(u) \equiv \int_0^1 u(x)\, dx$.

In each example it's easy to check that $(*)$ is satisfied, and we also see the characteristic *first-power* structure of the formulas (*without* u-independent terms this time). In each case L is a *function on functions*, a mapping which takes a function as input and gives as output either another function (as in the first two examples) or a number (as in the last two). Such a superfunction, considered as a mathematical object in its own right, is called an *operator*.

Now we can return to equations:

Definition: A *linear equation* is an equation of the form

$$L(u) = g,$$

where L is a linear operator, g is a "given" or "known" function (or number, as the case may be), and u is the unknown to be solved for.

So the possible u-independent terms enter the picture in the role of g. This leads to an absolutely crucial distinction:

Homogeneous vs. nonhomogeneous equations

Definition: A linear equation, $L(u) = g$, is *homogeneous* if $g = 0$ (i.e., *all* terms in the equation are *exactly* of the first degree in u); it is *nonhomogeneous* if $g \neq 0$ (i.e., "constant" terms also appear).

In the second parenthetical clause, "constant" means independent of u. The "constant" term g may be a nontrivial function of the *independent* variable(s) of the problem.

Among our original examples, the linear ODE example was nonhomogeneous (because of the $\cos 3t$) and the PDE example was homogeneous. The algebraic example is nonhomogeneous because of the 1. Here we are thinking of the system of simultaneous equations as a *single* linear equation in which the unknown quantity is a two-component vector,

$$\vec{u} \equiv \begin{pmatrix} x \\ y \end{pmatrix}.$$

The linear operator L maps $\vec{u}$ onto another vector,

$$\vec{g} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

As many of you know, the system of equations can be rewritten in matrix notation as

$$\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The linear operator is described by the square matrix:

$$L = \begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}.$$

In solving a differential equation we usually need to deal with initial or boundary conditions in addition to the equation itself. The main reason is that initial or boundary data need to be specified to give the problem a unique answer. Usually these conditions are themselves linear equations — for example, a standard initial condition for the heat equation:

$$u(0, x) = f(x).$$

Often the differential equation will be homogeneous but at least one of the boundary conditions will be nonhomogeneous. (The reverse situation also occurs.) Therefore, I think it's helpful to introduce one more bit of jargon:

Definitions: A *linear problem* consists of one or more linear conditions (equations) to be satisfied by the unknown, u. A linear problem is *homogeneous* if all of its conditions are homogeneous, *nonhomogeneous* if one or more of the conditions are nonhomogeneous.

Example 1: The ODE problem

$$u'' + 4u = 0, \qquad u(0) = 1, \quad u'(0) = 0$$

is an nonhomogeneous linear problem.

Example 2: The PDE problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + j(x), \qquad u(0, x) = 0, \quad u(t, 0) = 0, \quad u(t, 1) = 0$$

is an nonhomogeneous linear problem. The boundary conditions *and* the initial condition are homogeneous, but the heat equation itself is nonhomogeneous in this case; the function j represents generation of heat inside the bar (perhaps by combustion or radioactivity), a possibility not considered in our earlier discussion of the heat-conduction problem.

Remark: It is easy to see that every homogeneous linear equation has $u = 0$ as a solution. (One proof: $L(0) = L(u - u)$ (for any u) $= L(u) - L(u) = 0$, QED.) Therefore, any homogeneous linear problem has 0 as a solution. Therefore, *if a linear problem has a unique solution and that solution is nontrivial (not just the 0 function), then that linear problem must be nonhomogeneous.* That is, an interesting, well-posed problem always has at least one nonhomogeneous condition.

Solving linear problems

The importance of linear problems is that solving them is made easy by the *superposition principles* (which don't apply to nonlinear problems):

Principles of Superposition:

1. A linear combination of solutions of a homogeneous problem is a new solution of that problem. That is, if $L(u_1) = 0$ and $L(u_2) = 0$, then $L(c_1 u_1 + c_2 u_2) = 0$ for any numbers c_1 and c_2 (and similarly for more than two solutions, and for more than one homogeneous linear equation defining the problem).

   Example: Let **Problem 1** be the homogeneous ODE $u'' + 4u = 0$. Two solutions of this problem are

   $$u_1 \equiv \cos 2x, \qquad u_2 \equiv \sin 2x.$$

Then $u = u_1 + 3u_2$, for example, is also a solution. (In fact, we know that the *most general* solution is $c_1u_1 + c_2u_2$ where the c's are arbitrary constants. But for this we need a deeper existence-and-uniqueness theorem for second-order ODEs; it doesn't just follow from linearity.)

2. The sum of a solution of an nonhomogeneous problem and a solution of the *corresponding homogeneous problem* is a new solution of the original nonhomogeneous problem. ("Corresponding homogeneous problem" means the one with the same L's, but with all g's replaced by 0.)

   Example: Let **Problem 2** be the nonhomogeneous equation $u'' + 4u = e^x$. One solution is $u_{\rm p} \equiv \frac{1}{5}e^x$. (This has to be found by the method of undetermined coefficients, or by luck. Again, general principles of linearity by themselves can't solve the whole problem.) Now if we add a solution of Problem 1 we get a new solution of Problem 2: $u_3 \equiv \frac{1}{5}e^x + \cos 2x$.

3. The difference of two solutions of an nonhomogeneous problem is a solution of the corresponding homogeneous problem. Therefore, *every* solution of an nonhomogeneous problem can be obtained from *one particular* solution of that problem by adding some solution of the homogeneous problem.

   Example: The general solution of Problem 2 is

   $$u = \tfrac{1}{5}e^x + c_1 \cos 2x + c_2 \sin 2x.$$

4. The sum of solutions to two nonhomogeneous problems with the same L's is a solution of a *new* nonhomogeneous problem, for which the g's are the *sums* of the corresponding g's of the two original problems. (Similarly for more than two nonhomogeneous problems.)

   Example 1: The sum of two solutions of Problem 2, $u_{\rm p}$ and u_3, is $z \equiv \frac{2}{5}e^x + \cos 2x$, which is a solution of $z'' + 4z = 2e^x$. The important lesson to be learned from this example is that the right-hand side of this new equation is *not* e^x, the nonhomogeneous term of the two old equations. *Do not superpose solutions of an* NONHOMOGENEOUS *problem in the hope of getting a solution of that* SAME *problem.*

   Example 2: Note that $u_{\rm p}$ is the unique solution of **Problem 3**:

   $$u'' + 4u = e^x, \qquad u(0) = \tfrac{1}{5}, \quad u'(0) = \tfrac{1}{5}\,.$$

   Suppose that we really want to solve **Problem 4**:

   $$u'' + 4u = e^x, \qquad u(0) = 0, \quad u'(0) = 0.$$

Recalling Principles 2 and 3 as applied to the differential equation alone (not the initial conditions), we see that $u = u_p + y$, where y is some solution of $y'' + 4y = 0$. A moment's further thought shows that the correct y is the solution of **Problem 5**:

$$y'' + 4y = 0, \qquad y(0) = -\tfrac{1}{5}, \quad y'(0) = -\tfrac{1}{5}.$$

A standard calculation shows that $y = -\frac{1}{5}\cos 2x - \frac{1}{10}\sin 2x$, and from this and u_p we can get the solution of Problem 4. (Of course, in solving such problems we usually don't write out Problem 5 as an intermediate step; the standard procedure is to impose the initial data of Problem 4 on the general solution found earlier. That is just a different way of organizing the same algebra. However, consciously splitting an nonhomogeneous problem into two nonhomogeneous problems, as I've demonstrated here for an ODE, is a common technique for solving PDEs.)

In summary, these principles provide the basic strategies for solving linear problems. If the problem is nonhomogeneous and complicated, you split it into simpler nonhomogeneous problems and add the solutions. If the solution is not unique, the nonuniqueness resides precisely in the possibility of adding a solution of the corresponding homogeneous problem. (In particular, if the original problem is homogeneous, then you seek the general solution as a linear combination of some list of basic solutions.) If the problem statement contains enough initial and boundary conditions, the solution will be unique; in that case, the only solution of the homogeneous problem is the zero function.

An important example application of this strategy is the solution of the heat-conduction problem in a bar with fixed end temperatures.

PDE: $$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},$$

IC: $$u(0, x) = f(x),$$

BC: $$u(t, 0) = T_1\,, \qquad u(t, 1) = T_2\,.$$

(The end temperatures are *constants*, not functions of t as they could be in principle.) Here we have a homogeneous PDE, an nonhomogeneous initial condition, and two nonhomogeneous boundary conditions. The trick is to treat the two types of nonhomogeneity separately. We write $u = v + w$, where

(1) v is to be a solution of the problem consisting of the PDE and the nonhomogeneous BC, with *no* particular IC assumed. It is possible to find a solution

of this problem which is *independent of* t: $v(t,x) = V(x)$. (Let's return to the details later.)

(2) w is to be a solution of the problem consisting of the PDE, the *homogeneous* Dirichlet boundary conditions

$$w(t,0) = 0, \qquad w(t,1) = 0,$$

and the initial condition which is needed to make u satisfy the original IC. Namely,

$$w(0,x) = f(x) - V(x).$$

It is very important that the only nonhomogeneity in the problem is the IC. This makes it possible to solve for w by the method of *separation of variables* and then add the solutions without falling into the trap I warned you against earlier (Example 1). The solution is completed by finding the *Fourier series* of the function $f - V$. The processes of separating variables and calculating Fourier coefficients will be studied in depth by us in the near future. The important point for today is that for them to work here, it was absolutely crucial to make the boundary conditions homogeneous first. In the calculation of normal modes, *no* nonhomogeneous conditions at all are imposed. The appropriate nonhomogeneous IC is imposed on a *superposition* (w) of normal modes. Then still another term, v, is added to satisfy the nonhomogeneous BC.

The sermon, one more time

Our primary moral lessons are these:

- **Impose only HOMOGENEOUS conditions on normal modes (separated solutions).**
- **Impose nonhomogeneous conditions only on a SUPERPOSITION (sum or integral) of normal modes.**

A related principle is

- **Handle only one nonhomogeneity at a time!**

This principle is handled in practice by different strategies in different problems. Let's consider a doubly nonhomogeneous problem with the structure

$$L_1(u) = f_1\,, \qquad L_2(u) = f_2\,.$$

The two principal strategies are these:

1. **"Zero out" the other condition.** Solve

$$L_1(u_1) = f_1\,, \qquad L_2(u_1) = 0,$$
$$L_1(u_2) = 0, \qquad L_2(u_2) = f_2\,.$$

Then $u = u_1 + u_2$.

Examples where this strategy will be used (later in this course) include

(a) treatment of the initial data u and $\frac{\partial u}{\partial t}$ in the wave equation;

(b) Laplace's equation in a rectangle with boundary values given on two perpendicular sides.

2. **Temporarily ignore the other condition.** Solve $L_1(u_1) = f_1$ and let $L_2(u_1)$ be whatever it turns out to be, say $L_2(u_1) \equiv h$. Next solve

$$L_1(u_2) = 0, \qquad L_2(u_2) = f_2 - h.$$

Then $u = u_1 + u_2$.

Examples where this strategy is used include

(a) the method of undetermined coefficients for an ordinary differential equation with initial conditions;

(b) finding a *steady-state solution* for the wave or heat equation with nonzero, but time-independent, boundary conditions.

Unfinished business: the **steady-state solution** for our heat problem. Return to step (1) and assume that $v(t, x) = V(x)$. Then the equation becomes $0 = V''$, and the boundary conditions become $V(0) = T_1$, $V(1) = T_2$. We see that $V = C_1x+C_2$ and thus

$$T_1 = C_2\,, \qquad T_2 = C_1 + C_2\,.$$

Therefore,

$$\begin{aligned} V(x) &= (T_2 - T_1)x + T_1 \\ &= T_1(1 - x) + T_2x. \end{aligned}$$

Notice that $V(x)$ is a linear superposition of two terms, one of which vanishes when $x = 1$ and completely handles the nonhomogeneous data at $x = 0$, and the other — vice versa. This reconfirms that every nonhomogeneous term makes a separate contribution to the solution, even in a case where it was not technically necessary to make the separation at the beginning of the process of constructing the solution.

Separation of Variables

Now return to the second half of the problem, the initial-value problem for the heat equation with homogenized boundary conditions:

PDE: $$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2}\,,$$

BC: $$w(t,0) = 0, \qquad w(t,1) = 0,$$

IC: $$w(0,x) = g(x) \quad [= f(x) - V(x)].$$

Our strategy will be to look first for functions of the form

$$w_{\text{sep}}(t,x) = T(t)X(x)$$

which satisfy all the *homogeneous* equations of the problem (namely, the PDE and BC) — but *not* (usually) the nonhomogeneous equations (the IC, in this case). Then we will try to satisfy the nonhomogeneous conditions by a "superposition" (or infinite linear combination) of these separated solutions: It will look something like

$$w(t,x) = \sum_{n=1}^{\infty} c_n\, T_n(t) X_n(x).$$

At risk of tedium, let me emphasize again that

(1) since the separated (product) solutions satisfy the homogeneous conditions, the sum will also;

(2) attempting to impose the nonhomogeneous conditions on the individual w_{sep}'s will lead to catastrophe, since nonhomogeneous conditions are not preserved under summation. If we found an infinite string of functions that *each* satisfied the nonhomogeneous condition $u(t,0) = T_1$, then the sum of their boundary values would be an infinite series of equal constants, which would not converge — certainly not to T_1.

Substitute $w = TX$ into the PDE:

$$T'(t)X(x) = T(t)X''(x).$$

Now we *separate* the variables: Divide by $T(t)X(x)$, getting

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}\,.$$

In this equation the left side depends only on t and the right side depends only on x. The only way the equation can then hold for all t and all x is that both quantities are *constant*:

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

(I have advance information that the most interesting values of this constant will be negative, so I call it $-\lambda$. However, we are not yet ready to make any commitment as to whether λ is positive, negative, zero, or even complex. All possibilities must be considered.)

We have split the equation into two equations,

$$X'' + \lambda X = 0, \tag{1}$$

$$T' + \lambda T = 0. \tag{2}$$

Now look at the boundary conditions, which are

$$0 = w_{\text{sep}}(t, 0) = T(t)X(0), \qquad 0 = w_{\text{sep}}(t, 1) = T(t)X(1).$$

These impose restrictions on X, not T. (If we were to satisfy either of them for all t by setting $T(t) = 0$, we would make the entire solution w_{sep} identically equal to 0, a trivial and uninteresting solution.) Our next task is to find the values of λ that allow X to vanish at both 0 and 1.

Suppose first that λ is positive, and write $\lambda = \omega^2$ (where ω is positive). Then (1) with its BC is

$$X'' + \omega^2 X = 0, \qquad X(0) = 0, \quad X(1) = 0. \tag{\#}$$

The general solution of the ODE is

$$X = c_1 \cos \omega x + c_2 \sin \omega x, \tag{†}$$

and the first boundary condition forces $c_1 = 0$. We can choose $c_2 = 1$ without loss of generality (since what we are looking for is a linearly independent set of separated solutions w_{sep}). So $X = \sin \omega x$. Then the second boundary condition is

$$\sin \omega = 0.$$

The positive solutions of this equation are

$$\omega_n \equiv n\pi, \quad n = 1, 2, \ldots .$$

[Notice that the root $\omega = 0$ is irrelevant, since solutions of the ODE with $\lambda = 0$ do not have the form (†). Negative ω's give nothing new, which is why we restricted ω

to be positive when we introduced it.] Note, incidentally, that if we were working on the interval $0 < x < \pi$ instead of $0 < x < 1$, we would get just $\omega_n = n$, without the π.

The numbers ω_n that "work" in (#) are called *eigenvalues*, and the associated solutions X are called *eigenfunctions*, or *eigenvectors*, in analogy with the situation for matrices in linear algebra.

We can now solve the time equation, (2):

$$T(t) = e^{-{\omega_n}^2 t} = e^{-n^2\pi^2 t}.$$

The full separated solution for each n is thus

$$w_{\text{sep}}(t, x) = \sin(n\pi x)\, e^{-n^2\pi^2 t}.$$

Now consider the possibility that $\lambda = 0$. In place of (†) we have the general solution

$$X = c_1 + c_2 x.$$

Applying the two BC, we swiftly get $c_1 = 0$, $c_2 = 0$. So there is no nontrivial solution with $\lambda = 0$.

Similar arguments show that negative and complex λ's give only trivial solutions. In the negative case, write $\lambda = -\kappa^2$; then

$$X = c_1 \cosh \kappa x + c_2 \sinh \kappa x,$$

and the result follows (since $\cosh 0 \neq 0$ and $\sinh z \neq 0$ unless $z = 0$). If λ is complex, it has two complex square roots, which are negatives (*not* complex conjugates!) of each other. Thus

$$X = c_1 e^{(\kappa + i\omega)x} + c_2 e^{-(\kappa + i\omega)x},$$

where $(\kappa + i\omega)^2 = -\lambda$ and $\kappa \neq 0$ (else we would be back in the case of positive λ). $X(0) = 0$ implies that $c_2 = -c_1$, and then $X(1) = 0$ implies that

$$e^{(\kappa + i\omega)} = e^{-(\kappa + i\omega)}.$$

Since $\kappa \neq 0$, these two complex numbers have different moduli (absolute values), so this conclusion is a contradiction.

There is a more modern, less grubby way to see that λ has to be positive. Using the ODE ($X'' = -\lambda X$) and the BC (which allow us to discard all endpoint terms

which arise in integration by parts), we see that

$$\begin{aligned}\lambda \int_0^1 |X(x)|^2\,dx &= -\int_0^1 X^*\,X''\,dx \\ &= +\int_0^1 |X'|^2\,dx \\ &= -\int_0^1 (X'')^*\,X\,dx \\ &= +\lambda^* \int_0^1 |X|^2\,dx.\end{aligned}$$

Comparing the first and last members of this chain of equalities, we see that $\lambda = \lambda^*$ — that is, λ must be real. Comparing either of the extreme members with the one in the middle, we see that λ is positive, since two integrals are positive.

This argument suggests a general method for handling such questions when the second-derivative operator is replaced by a more general linear differential operator $L[X]$. If the L can be moved by integration by parts from one side of the integral to the other,

$$-\int_a^b X^*\,L[X]\,dx = -\int_a^b (L[X])^*\,X\,dx,$$

then all the allowed *eigenvalues* λ must be real. (Here it is understood that $X(x)$ satisfies the boundary conditions of the problem, though not necessarily the differential equation. An operator with this integration-by-parts symmetry is called *self-adjoint*.) If, in addition, an intermediate step in the integration by parts is a manifestly positive (or nonnegative) integral, then the λ's must be positive (or nonnegative, respectively).

To summarize, in the one-dimensional heat problem with Dirichlet boundary conditions we have found the eigenvalues

$$\lambda_n \equiv {\omega_n}^2 = (n\pi)^2$$

and the corresponding solutions

$$w_{\text{sep}}(t,x) = \sin(\omega_n x)\,e^{-{\omega_n}^2 t}.$$

We still need to investigate how to superpose such solutions to obtain a solution with the arbitrary initial data $w(0,x) = g(x)$. So, let us assume that such a solution exists, and see if that assumption leads us either to useful information (good), or to a contradiction (bad):

$$w(t,x) = \sum_{n=1}^{\infty} b_n \sin(\omega_n x)\,e^{-{\omega_n}^2 t}$$

for some (not yet known) coefficients b_n . Then

$$g(x) = \sum_{n=1}^{\infty} b_n \sin(\omega_n x). \qquad (*)$$

This is supposed to hold on the interval $0 < x < 1$.

More generally, if the spatial interval is $0 < x < L$, then we would like $(*)$ to be true for the appropriate choice of the ω_n's — namely,

$$\omega_n = \frac{n\pi}{L}$$

(the positive solutions of $0 = X(L) = \sin(\omega L)$). In particular, if $L = \pi$, then $\omega_n = n$. I shall develop the theory of $(*)$ for the case $L = \pi$, rather than the case $L = 1$ that I've been discussing heretofore. There are two reasons for this: (1) It requires less writing. (2) Most of the homework problems for this week have $L = \pi$. Exercises with other choices of L have been postponed to next week.

To find the b_n in $(*)$ we multiply that equation by $\sin mx$ and integrate from 0 to π. We assume that the integral of the infinite series exists and is equal to the sum of the integrals of the individual terms:

$$\int_0^\pi g(x) \sin mx \, dx = \sum_{n=1}^{\infty} b_n \int_0^\pi \sin nx \sin mx \, dx.$$

(In the general case, of course, the integral would be from 0 to L. See the textbooks for details and a later segment of notes for summary.) Now

$$\sin nx \sin mx = \tfrac{1}{2} \cos{(nx - mx)} - \tfrac{1}{2} \cos(nx + mx),$$

so

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

— provided that $n \neq m$. If $n = m$ we have

$$\begin{aligned} \int_0^\pi \sin^2 mx \, dx &= \left[\frac{1}{2} \, x - \frac{1}{4m} \sin(2mx) \right]_0^\pi \\ &= \frac{\pi}{2} \, . \end{aligned}$$

Thus only the $n = m$ term in the sum survives, and

$$\int_0^\pi g(x) \sin mx \, dx = \frac{\pi}{2} \, b_m \, .$$

Conclusion: If $(*)$ is true,

$$\boxed{g(x) = \sum_{n=1}^{\infty} b_n \sin nx,}$$

then

$$\boxed{b_n = \frac{2}{\pi} \int_0^{\pi} g(x) \sin nx \, dx.}$$

$(*)$ is called the *Fourier sine series* of the function g, and the b_n are its *Fourier coefficients*.

Convergence Theorems

So far we've seen that we can solve the heat equation with homogenized Dirichlet boundary conditions and arbitrary initial data (on the interval $[0, \pi]$), provided that we can express an arbitrary function g (on that interval) as an infinite linear combination of the eigenfunctions $\sin(nx)$:

$$g(x) = \sum_{n=1}^{\infty} b_n \sin nx.$$

Furthermore, we saw that if such a series exists, its coefficients must be given by the formula

$$b_n = \frac{2}{\pi} \int_0^{\pi} g(x) \sin nx \, dx.$$

So the burning question of the hour is: Does this Fourier sine series really converge to $g(x)$?

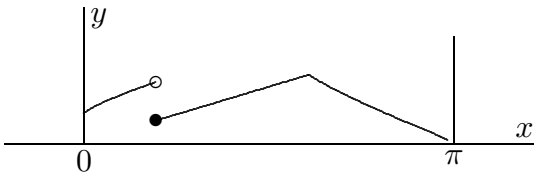
No mathematician can answer this question without first asking, "What kind of convergence are you talking about? And what technical conditions does g satisfy?" There are three standard convergence theorems, each of which states that certain technical conditions are sufficient to guarantee a certain kind of convergence. Generally speaking,

more smoothness in g

$\Longleftrightarrow$ more rapid decrease in b_n as $n \to \infty$

$\Longleftrightarrow$ better convergence of the series.

Definition: g is *piecewise smooth* if its derivative is piecewise continuous. That is, $g'(x)$ is defined and continuous at all but a finite number of points (in the domain $[0, \pi]$, or whatever finite interval is relevant to the problem), and at those bad points g' has finite one-sided limits. (At such a point g itself is allowed to be discontinuous, but only the "finite jump" type of discontinuity is allowed.)



This class of functions is singled out, not only because one can rather easily prove convergence of their Fourier series (see next theorem), but also because they are a natural type of function to consider in engineering problems. (Think of

electrical voltages under the control of a switch, or applied forces in a mechanical problem.)

Pointwise Convergence Theorem: If g is continuous and piecewise smooth, then its Fourier sine series converges at each x in $(0, \pi)$ to $g(x)$. If g is piecewise smooth but not necessarily continuous, then the series converges to

$$\tfrac{1}{2}[g(x^-) + g(x^+)]$$

(which is just $g(x)$ if g is continuous at x). [Note that at the endpoints the series obviously converges to 0, regardless of the values of $g(0)$ and $g(\pi)$.]

Uniform Convergence Theorem: If g is both continuous and piecewise smooth, and $g(0) = g(\pi) = 0$, then its Fourier sine series converges *uniformly* to g throughout the interval $[0, \pi]$.

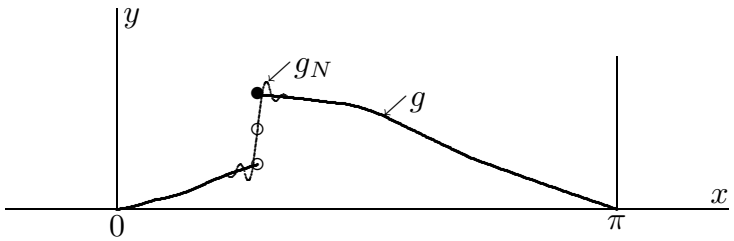
Remarks:

1. Recall what uniform convergence means: For every ϵ we can find an N so big that the partial sum

$$g_N(x) \equiv \sum_{n=1}^{N} b_n \sin(nx)$$

approximates $g(x)$ to within an error ϵ everywhere in $[0, \pi]$.

2. In contrast, if the convergence is nonuniform (merely pointwise), then for each x we can take enough terms to get the error $|g(x) - g_N(x)|$ smaller than ϵ, but the N may depend on x as well as ϵ. It is easy to see that if g is discontinuous, then uniform convergence is impossible, because the approximating functions g_N need a finite "time" to jump across the gap. There will always be points near the jump point where the approximation is bad.



It turns out that g_N develops "ears" or "overshoots" right next to the jump. This is called the *Gibbs phenomenon*.

3. For the same reason, the sine series can't converge uniformly near an endpoint where g doesn't vanish. An initial-value function which violated the condition

$g(0) = g(\pi) = 0$ would be rather strange from the point of view of our heat problem, since we want $w(t, 0) = w(t, \pi) = 0$ and also $w(0, x) = g(x)$! We'll soon see another way of understanding the significance of this condition.

4. If g is piecewise continuous, it can be proved that $b_n \to 0$ as $n \to \infty$. In other words, $b_n = o(n^0)$. (This is one form of the *Riemann–Lebesgue theorem*.) This is a key step in proving the pointwise convergence theorem.

   If g satisfies the conditions of the uniform convergence theorem, then integration by parts shows that

$$b_n = \frac{2}{n\pi} \int_0^\pi g'(x) \cos(nx)\, dx,$$

   and by another version of the Riemann–Lebesgue theorem this integral also approaches 0 when n is large, so that $b_n = o(n^{-1})$. This additional falloff is "responsible" for the uniform convergence of the series.

   (This remark is as close as we'll come in this course to proofs of the convergence theorems.)

5. There are continuous (but not piecewise smooth) functions whose Fourier series do *not* converge, but it is hard to construct an example!

   Now, the third kind of convergence.

Parseval's Equation: $$\int_0^\pi |g(x)|^2\, dx = \frac{\pi}{2} \sum_{n=1}^{\infty} |b_n|^2.$$

(In particular, the integral converges if and only if the sum does.)

"Proof": Taking convergence for granted, let's calculate the integral. (I'll assume that $g(x)$ and b_n are real, although I've written the theorem so that it applies also when things are complex.)

$$\begin{aligned}
\int_0^\pi |g(x)|^2\, dx &= \int_0^\pi \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_n \sin(nx)\, b_m \sin(mx)\, dx \\
&= \int_0^\pi \sum_{n=1}^{\infty} {b_n}^2 \sin^2 nx\, dx \\
&= \frac{\pi}{2} \sum_{n=1}^{\infty} {b_n}^2.
\end{aligned}$$

(Only terms with $m = n$ contribute, because when $m \neq n$

$$\int_0^\pi \sin(mx)\sin(nx)\,dx = 0.$$

This property is called *orthogonality*, and we will see later that it always arises (and plays a very valuable role) in eigenfunction expansions. The integral with $m = n$ can be evaluated by a well known rule of thumb: The integral of $\sin^2 \omega x$ over any integral number of quarter-cycles of the trig function is half of the integral of $\sin^2 \omega x + \cos^2 \omega x$ — namely, the length of the interval, which is π in this case.)

Definition: g is *square-integrable* on $[0, \pi]$ if the integral in Parseval's equation converges:

$$\int_0^\pi |g(x)|^2\,dx < \infty.$$

L^2 (or Mean) Convergence Theorem: If g is square-integrable, then the series converges *in the mean*:

$$\int_0^\pi |g(x) - g_N(x)|^2\,dx \to 0 \quad \text{as } N \to \infty.$$

Remarks:

1. Recalling the formulas for the length and distance of vectors in 3-dimensional space,

   $$|\vec{x}|^2 \equiv \sum_{n=1}^{3} {x_n}^2, \qquad |\vec{x} - \vec{y}|^2 \equiv \sum_{n=1}^{3} (x_n - y_n)^2,$$

   we can think of the Parseval integral as a measure of the "length" of g, and the integral in the theorem as a measure of the "distance" between g and g_N. (More about this when we cover general orthogonal basis functions, Constanda Chap. 3.)

2. A function can be square-integrable without being piecewise smooth, or even bounded. Example:

   $$g(x) \equiv \left(x - \tfrac{1}{2}\right)^{-\frac{1}{3}}.$$

   Also, a series can converge in the mean without converging pointwise (not to mention uniformly). This means that the equation

   $$g(x) = \sum_{n=1}^{\infty} b_n \sin nx$$

   must not be taken too literally in such a case — such as by writing a computer program to add up the terms for a fixed value of x. (The series will converge (pointwise) for "almost" all x, but there may be special values where it doesn't.)

More General Kinds of Fourier Series

(1) We have been discussing the Fourier sine series for functions defined on the interval $(0, 1)$,

$$g(x) = \sum_{n=1}^{\infty} b_n \sin n\pi x, \qquad b_n = 2\int_0^1 g(x)\, \sin n\pi x\, dx,$$

or on the interval $(0, \pi)$,

$$g(x) = \sum_{n=1}^{\infty} b_n \sin nx, \qquad b_n = \frac{2}{\pi}\int_0^{\pi} g(x)\, \sin nx\, dx.$$

By rescaling the independent variable, we can reduce an arbitrary interval (a, b) to either of these.

Often it is useful in applications to have available the formulas for the Fourier series on the original interval, so that you don't have to stop to repeat the scaling argument every time. If $L \equiv b - a$, those formulas are

$$g(x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi(x-a)}{L}\,,$$

$$b_n = \frac{2}{L}\int_a^b g(x)\, \sin \frac{n\pi(x-a)}{L}\, dx.$$

The important point is that the coefficients are calculated by integrating g over the interval concerned! The numbers $n\pi/L$ are the roots of $\sin(\omega L) = 0$, and the numerical coefficient in the integral depends on L because of the action of the chain rule on the differential.

In theoretical discussions, on the other hand, it is convenient to assume that $a = 0$ and $L = \pi$ or 1, thereby freeing the formulas of irrelevant clutter.

(2) There is a corresponding series in terms of cosines:

$$g(x) = \sum_{n=0}^{\infty} a_n \cos nx$$

(if the interval is $(0, \pi)$). Note that $n = 0$ corresponds to a nonzero function in this case ($\cos(0x) = 1$), and it must be included. Unfortunately, the formula for the coefficient of that term is different from the others by a factor of 2:

$$a_n = \frac{2}{\pi}\int_0^{\pi} g(x)\, \cos nx\, dx \quad \text{for } n > 0,$$

$$a_0 = \frac{1}{\pi}\int_0^\pi g(x)\,\cos 0x\,dx$$
$$= \frac{1}{\pi}\int_0^\pi g(x)\,dx.$$

It is helpful to remember that a_0 is the *average value* of g on the interval $[0,\pi]$. (This is reasonable, since the cosine terms with $n \neq 0$ all average to 0.)

Unfortunately, the notation for cosine series is not standardized. Many books define a_0 to be twice my a_0 so that the a_n formula is valid for $n = 0$, but the formula for the series itself is messed up:

$$g(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx.$$

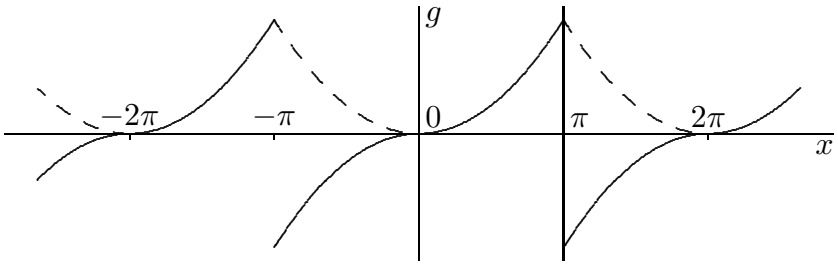
The formal derivation of the cosine formulas is just like the derivation of the sine series, being based on the orthogonality of the functions $\{\cos nx\}$ on an interval of length π.

(3) What happens to the series when x is *not* in the interval $[0,\pi]$?

Recall that $\sin nx$ is an *odd* function [satisfies $f(-x) = -f(x)$], but $\cos nx$ is an *even* function [satisfies $f(-x) = f(x)$]. Thus the partial sum $g_N(x)$, a linear combination of sines or cosines, is an odd function in the sine case, an even function in the cosine case. Therefore, the same is true of the limit; the Fourier sine series converges to an odd function, the Fourier cosine series to an even one.

Also, each g_N is periodic with period 2π (since each term of it is), and so the sum of the series is such a periodic function.

If the original function g is given by a formula, such as $g(x) = x^2$, the sum of the series will generally *not* agree with that formula anywhere except in the interval $[0,\pi]$. (This was a source of great confusion to the mathematicians of 200 years ago, who didn't clearly distinguish between a function and an algebraic expression.) The limit of the sine series is the odd periodic extension of g (solid line), while that of the cosine series is the even periodic extension (dashed line), which happens to coincide with x^2 on $[-\pi,0]$ in our example just because x^2 is already an even function. That the series should disagree with g off the original interval is really no surprise, since the coefficients in the series are calculated from integrals of g over that interval and therefore carry absolutely no information about the behavior of g elsewhere.



(4) Any function can be written as the sum of an even function and an odd function:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}\,. \qquad (\dagger)$$

If f is defined on the interval $[-\pi, \pi]$ (and is smooth enough to satisfy the hypotheses of a convergence theorem), then the even term in (†) is correctly represented by a cosine series *everywhere in* $[-\pi, \pi]$. Similarly, the odd term is represented by a sine series. Consequently, an arbitrary function on $[-\pi, \pi]$ has a series containing both sines and cosines:

$$f(x) = \sum_{n=0}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx.$$

This is what is known as **the** Fourier series of f, or the "full" Fourier series (as distinct from one containing only one type of trig function).

The formulas for the coefficients are

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \sin nx\, dx,$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \cos nx\, dx \qquad (n \neq 0),$$

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

To check that these are correct, consider the special cases where f is even or odd. If f is even, the integral for b_n equals 0 since its integrand is odd; the integrand of a_n is even, so the contribution of the negative half of the integral is equal to the positive half, and the formula (for $n \neq 0$) can be rewritten

$$a_n = \frac{2}{\pi} \int_{0}^{\pi} f(x)\, \cos nx\, dx.$$

Thus the sine terms of the series are absent, and the cosine terms have the coefficients that give the correct Fourier cosine series for f on $[0, \pi]$; this is exactly as it should be. Similarly, an odd f has a Fourier series consisting entirely of sines.

The sum of a full Fourier series is a periodic function without any special symmetry (even or odd) or any special boundary behavior. It is appropriate physically to the problem of heat flow in a ring or cylinder.

Again, many books write the Fourier series as

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

This is a *different definition* of a_0, which makes the formula

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \cos nx\, dx$$

correct for $n = 0$ as well as for the other n's.

(5) The trigonometric form of the Fourier series is quite cumbersome, because three separate formulas for the coefficients are needed. It can be greatly simplified by converting to the complex-exponential function,

$$e^{inx} = \cos nx + i \sin nx.$$

Note that

$$e^{-inx} = \cos nx - i \sin nx$$

and hence

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2}\,, \qquad \sin nx = \frac{e^{inx} - e^{-inx}}{2i}\,.$$

Also, $\cos 0x = 1 = e^{0x}$. Therefore, the cosine and sine functions appearing in the Fourier series can be expressed in terms of the functions e^{ikx}, where k ranges over all integers, positive, negative, and zero.

Let us work out this transformation. Substituting into the Fourier series we get

$$\begin{aligned} f(x) &= a_0 + \sum_{n=1}^{\infty} a_n \frac{e^{inx} + e^{-inx}}{2} + \sum_{n=1}^{\infty} b_n \frac{e^{inx} - e^{-inx}}{2i} \\ &= \sum_{k=-\infty}^{-1} \left(\frac{a_{-k}}{2} - \frac{b_{-k}}{2i}\right) e^{ikx} + a_0 e^{i0x} + \sum_{k=1}^{\infty} \left(\frac{a_k}{2} + \frac{b_k}{2i}\right) e^{ikx}. \end{aligned}$$

(The terms with negative k come from setting $k \equiv -n$ in the terms with factors e^{-inx}; the terms with positive k come from setting $k \equiv n$ in the terms with factors e^{inx}.) Therefore, we will have

$$\boxed{f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}}$$

if we define

$$c_k = \begin{cases} \frac{1}{2}(a_k - ib_k) & \text{if } k > 0, \\ a_0 & \text{if } k = 0, \\ \frac{1}{2}(a_{-k} + ib_{-k}) & \text{if } k < 0. \end{cases}$$

It is easy to see that in all three cases,

$$\boxed{c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx} f(x)\, dx.}$$

For instance, if $k > 0$,

$$\frac{1}{2}(a_k - ib_k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)[\cos kx - i \sin kx]\, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx} f(x)\, dx.$$

In terms of the original, nonnegative index n,

$$c_n = \tfrac{1}{2}(a_n - ib_n), \qquad c_{-n} = \tfrac{1}{2}(a_n + ib_n).$$

The two equations in boxes are clearly a vast improvement over the original Fourier series from the point of view of simplicity. This is the way the Fourier series is usually written and used in modern work.

Unfortunately, in order to evaluate the integral c_k for a particular function f, using the standard integral tables, you will usually need to break the integrand up into sine and cosine parts, since the tables list only real integrals, not complex ones. However, sometimes the integral is easy enough to be evaluated by elementary means. In such a case the usual identities for the exponential function may be used, even though the argument is an imaginary number. As an example (which will be used later) I'll calculate the orthogonality and normalization integrals for the functions e^{ikx}.

Suppose first that $k \neq l$. Then

$$\begin{aligned} \int_{-\pi}^{\pi} e^{ikx}\, e^{-ilx}\, dx &= \int_{-\pi}^{\pi} e^{i(k-l)x}\, dx \\ &= \frac{1}{i(k-l)}\, e^{i(k-l)x} \Big|_{-\pi}^{\pi} \\ &= \frac{1}{i(k-l)} \left[e^{i(k-l)\pi} - e^{-i(k-l)\pi} \right] \\ &= \frac{2}{k-l} \sin[(k-l)\pi] \\ &= 0 \quad \text{(since } k - l \text{ is an integer).} \end{aligned}$$

On the other hand, if $k = l$ we have

$$\begin{aligned}\int_{-\pi}^{\pi} e^{ikx}\, e^{-ilx}\, dx &= \int_{-\pi}^{\pi} e^{ikx}\, e^{-ikx}\, dx \\ &= \int_{-\pi}^{\pi} 1\, dx \\ &= 2\pi.\end{aligned}$$

(6) Like the sine series, the full Fourier series can be written for functions defined on intervals of any length, not just 2π. Naturally, it is often convenient to use the physical length involved in a given problem, even though this clutters the formulas with extra constants.

For a full Fourier series, it is natural to let L stand for *half* the length of the interval, which is chosen to be $[-L, L]$. By a scaling transformation one can easily see that the proper formulas are

$$f(x) = \sum_{k=-\infty}^{\infty} c_k\, e^{i\pi kx/L},$$

$$c_k = \frac{1}{2L} \int_{-L}^{L} e^{-i\pi kx/L}\, f(x)\, dx.$$

Using the periodicity of the trigonometric functions, it is possible to write Fourier series on an arbitrary interval of length $2L$, say $[c, c + 2L]$, and Fourier sine or cosine series on an arbitrary interval of length L.

(7) For each type of Fourier series there is a corresponding Parseval equation and three standard convergence theorems (pointwise, uniform, and mean). Obviously, I can't write these out for all the possible cases. I'll discuss one example of each, in enough detail to make it possible in principle to reconstruct the others.

The Parseval equation for the full Fourier series on the standard interval is

$$\int_{-\pi}^{\pi} |f(x)|^2\, dx = 2\pi \sum_{k=-\infty}^{\infty} |c_k|^2.$$

It is now necessary to allow the quantities to be complex; $|c_k|^2$, for instance, means the complex modulus squared (c_k times its complex conjugate). To derive this equation, substitute the boxed Fourier series and its complex conjugate

$$f(x)^* = \sum_{l=-\infty}^{\infty} c_l{}^*\, e^{-ilx}$$

into the left-hand side, and integrate term-by-term using the orthogonality integrals worked out in (5).

The Parseval equation for the Fourier cosine series is complicated, because the a_0 term needs to be treated separately. In fact, there are two different versions of the formula, depending on where you put that annoying factor of 2 in the definition of the series. Perhaps it is best to work out this Parseval formula from first principles (the orthogonality integrals) every time you need it.

To guarantee that the full Fourier series converges *uniformly*, we must have f continuous and piecewise smooth, and also

$$f(-L) = f(L).$$

The reason for the last condition is that the function which really must be continuous and piecewise smooth is the *periodic extension* of f to the whole real line. If the endpoint values of f do not match, the extension will have a discontinuity at the endpoints. Another way of saying this is that f must be a continuous function when regarded as a function defined on a *circle* or *ring*; that is, when x is an angle, so that $-L$ and L label the same physical point.

We can now understand from a more fundamental point of view why the uniform convergence theorem for the sine series requires that the function vanish at the endpoints. In that case, the relevant function is the *odd periodic extension*, which will have a discontinuity at 0 if the original function's graph does not pass through the origin (and a discontinuity at L if the function doesn't vanish there). The cosine series needs no such supplementary condition, because the *even* extension of a function is continuous at the endpoints. (The *derivative* of the extension is likely to be discontinuous, but that will not disrupt the uniform convergence.)

A Big Example

This will break up into many small examples, which will demonstrate many of the principles we've talked about — often in a slightly new context.

Problem statement

We will consider heat conduction in a two-dimensional region, a rectangle. The ranges of the variables, therefore, will be

$$0 < x < a, \qquad 0 < y < b, \qquad t > 0.$$

Without loss of generality, we can assume that the variables have been scaled so that $a = \pi$.

The heat equation is

PDE: $$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\,.$$

Let us assume that the boundary conditions are

BC_1: $$\frac{\partial u}{\partial x}(t, 0, y) = 0 = \frac{\partial u}{\partial x}(t, \pi, y)\,,$$

BC_2: $$u(t, x, 0) = p(x), \qquad u(t, x, b) = q(x).$$

That is, the plate is insulated on the sides, and the temperature on the top and bottom edges is known and given by the functions p and q. Finally, there will be some initial temperature distribution

IC: $$u(0, x, y) = f(x, y).$$

Steady-state solution

From our experience with the one-dimensional problem, we know that we must eliminate the nonhomogeneous boundary condition (BC_2) before we can solve the initial-value problem by separation of variables! Fortunately, p and q are independent of t, so we can do this by the same technique we used in one dimension: hunt for a time-independent solution of (PDE) and (BC), $v(t, x, y) = V(x, y)$, then consider the initial-value problem with homogeneous boundary conditions satisfied by $u - v$.

So, we first want to solve

PDE: $$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

BC_1: $$\frac{\partial V}{\partial x}(0, y) = 0 = \frac{\partial V}{\partial x}(\pi, y)\,,$$

BC_2: $$V(x, 0) = p(x), \qquad V(x, b) = q(x).$$

This is still a *partial* differential equation (namely, the two-dimensional *Laplace equation*). Furthermore, it still contains two nonhomogeneous conditions. Therefore, we split the problem again:

$$V = V_1 + V_2\,,$$

$$\begin{aligned} V_1(x, 0) &= p(x), & V_2(x, 0) &= 0, \\ V_1(x, b) &= 0, & V_2(x, b) &= q(x). \end{aligned}$$

Each V_j is supposed to satisfy Laplace's equation and (BC_1).

Remark: This splitting is slightly different from the one involving the steady-state solution. In each subproblem here we have replaced every nonhomogeneous condition except one by its *corresponding homogeneous condition*. In contrast, for the steady-state solution we simply *discarded* the inconvenient nonhomogeneous condition, and later will modify the corresponding nonhomogeneous condition in the other subproblem to account for the failure of the steady-state solution to vanish on that boundary. Which of these techniques is best varies with the problem, but the basic principle is the same: Work with only one nonhomogeneous condition at a time, so that you can exploit the superposition principle correctly.

Let us solve for V_2 by separation of variables:

$$V_{2\mathrm{sep}}(x, y) = X(x)Y(y).$$

$$0 = X''Y + XY'' \;\Rightarrow\; -\frac{X''}{X} = \lambda = \frac{Y''}{Y}\,.$$

The boundary condition (BC_1) implies that

$$X'(0) = 0 = X'(\pi).$$

Therefore, up to a constant,

$$X(x) = \cos nx, \qquad \lambda = n^2.$$

Now Y must be a solution of $Y'' = n^2 Y$ that vanishes at $y = 0$; that is, up to a constant,

$$Y(y) = \sinh ny \quad \text{if } n \neq 0.$$

The case 0 must be treated separately: $Y(y) = y$. We have now taken care of three of the four boundaries. The remaining boundary condition is nonhomogeneous, and thus we cannot apply it to the individual separated solutions XY; first we must adding up the separated solutions with arbitrary coefficients:

$$V_2(x, y) = a_0 y + \sum_{n=1}^{\infty} a_n \cos nx \sinh ny.$$

Now we must have

$$q(x) = a_0 b + \sum_{n=0}^{\infty} a_n \cos nx \sinh nb.$$

This is a Fourier cosine series, so we solve for the coefficients by the usual formula:

$$a_n \sinh nb = \frac{2}{\pi} \int_0^{\pi} \cos nx \, q(x) \, dx \qquad (n > 0).$$

Divide by $\sinh nb$ to get a formula for a_n. For $n = 0$ the Fourier formula lacks the factor 2, and we end up with

$$a_0 = \frac{1}{\pi b} \int_0^{\pi} q(x) \, dx.$$

This completes the solution for V_2.

Solving for V_1 is exactly the same except that we need $Y(b) = 0$ instead of $Y(0) = 0$. The appropriate solution of $Y'' = n^2 Y$ can be written as a linear combination of $\sinh ny$ and $\cosh ny$, or of e^{ny} and e^{-ny}, but it is neater to write it as

$$Y(y) = \sinh\bigl(n(y - b)\bigr),$$

which manifestly satisfies the initial condition at b as well as the ODE. (Recall that hyperbolic functions satisfy trig-like identities, in this case

$$\begin{aligned} \sinh\bigl(n(y - b)\bigr) &= \cosh nb \sinh ny - \sinh nb \cosh ny \\ &= \tfrac{1}{2} e^{-nb} \, e^{ny} - \tfrac{1}{2} e^{nb} \, e^{-ny}, \end{aligned}$$

so the three forms are consistent.) Again the case $n = 0$ is special: $Y(y) = y - b$. We now have

$$V_1(x, y) = A_0(y - b) + \sum_{n=1}^{\infty} A_n \cos nx \sinh n(y - b).$$

At $y = 0$ this becomes

$$p(x) = -A_0 b - \sum_{n=1}^{\infty} A_n \cos nx \sinh nb.$$

Thus

$$A_n = -\frac{2}{\pi \sinh nb} \int_0^\pi \cos nx\, p(x)\, dx \qquad (n > 0),$$

$$A_0 = -\frac{1}{\pi b} \int_0^\pi p(x)\, dx.$$

This completes the solution for V_1 and hence for $v(t, x, y)$.

Remark: Since the boundary conditions at $y = 0$ and $y = b$ refer to the *same variable*, it was not really necessary to treat them separately. We could have separated variables in the problem [(Laplace PDE) + (BC_1)] satisfied by the function V, getting

$$V_\tau sep(x, y) = \cos nx\, Y(y), \qquad Y'' = n^2 Y.$$

Then we could find the *general solution* of this last equation,

$$Y(y) = a_n \sinh ny + b_n \cosh ny$$

— or, better,

$$Y(y) = a_n \sinh ny + A_n \sinh n(y - b);$$

write the general superposition as a sum of these over n; and then use the two nonhomogeneous boundary conditions (BC_2) to determine the constants a_n and A_n in the summation.

This works because the nonhomogeneous conditions refer to parallel parts of the boundary. It definitely will not work for perpendicular edges! When in doubt, follow the injunction to deal with just one nonhomogeneity at a time.

Homogeneous problem

Next we're supposed to solve for $w \equiv u - v$, which must satisfy

PDE:
$$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\,,$$

BC_1:
$$\frac{\partial w}{\partial x}(t, 0, y) = 0 = \frac{\partial w}{\partial x}(t, \pi, y)\,,$$

BC_2: $$w(t,x,0)=0, \qquad w(t,x,b)=0,$$

IC: $$w(0,x,y)=f(x,y)-V(x,y)\equiv g(x,y).$$

Since there is only one nonhomogeneous condition, we can separate variables immediately:

$$w_{\text{sep}}(t,x,y)=T(t)X(x)Y(y).$$

$$T'XY=TX''Y+TXY''.$$

$$\frac{T'}{T}=\frac{X''}{X}+\frac{Y''}{Y}=-\lambda.$$

(We know that λ is a constant, because the left side of the equation depends only on t and the right side does not depend on t at all. By analogy with the one-dimensional case we can predict that λ will be positive.) Since X''/X depends only on x and Y''/Y depends only on y, we can introduce *another* separation constant:

$$\frac{X''}{X}=-\mu, \qquad \frac{Y''}{Y}=-\lambda+\mu.$$

The boundary conditions translate to

$$X'(0)=0=X'(\pi), \qquad Y(0)=0=Y(b).$$

Thus for X we have the familiar solution

$$X(x)=\cos mx, \qquad \mu=m^2.$$

Similarly, we must have

$$Y(y)=\sin\frac{n\pi y}{b}\,, \qquad -\lambda+\mu=-\frac{n^2\pi^2}{b^2}$$

$$\Rightarrow\ \lambda=m^2+\frac{n^2\pi^2}{b^2}\equiv\lambda_{mn}\,.$$

Then

$$T(t)=e^{-\lambda t}.$$

(As usual in separation of variables, we have left out all the arbitrary constants multiplying these solutions. They will all be absorbed into the coefficients in the final Fourier series.)

We are now ready to superpose solutions and match the initial data. The most general solution of the homogeneous problem is a double infinite series,

$$w(t,x,y)=\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}c_{mn}\cos mx\,\sin\frac{n\pi y}{b}\,e^{-\lambda_{mn}t}.$$

The initial condition is

$$g(x,y) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} c_{mn} \cos mx \, \sin \frac{n\pi y}{b} \, .$$

To solve for c_{mn} we have to apply Fourier formulas twice:

$$\sum_{m=0}^{\infty} c_{mn} \cos mx = \frac{2}{b} \int_0^b \sin \frac{n\pi y}{b} \, g(x,y) \, dy;$$

$$c_{mn} = \frac{2}{\pi} \frac{2}{b} \int_0^{\pi} dx \int_0^b dy \, \cos mx \, \sin \frac{n\pi y}{b} \, g(x,y) \qquad (m > 0),$$

$$c_{0n} = \frac{2}{\pi b} \int_0^{\pi} dx \int_0^b dy \, \sin \frac{n\pi y}{b} \, g(x,y).$$

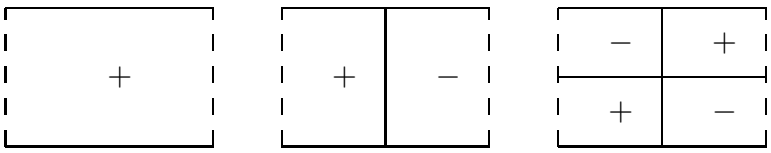
This completes the solution for w. Now we have the full solution to the original problem:

$$u(t,x,y) = w(t,x,y) + V(x,y).$$

Furthermore, along the way we have constructed a very interesting family of functions defined on the rectangle:

$$\phi_{mn}(x,y) \equiv \cos mx \sin \frac{n\pi y}{b} \, .$$

A few early members of the family look like this:



(Recall that $\cos(0x) = 1$.) The function is positive or negative in each region according to the sign shown. The function is zero on the solid lines and its normal derivative is zero along the dashed boundaries. The functions have these key properties for our purpose:

- They are *eigenvectors* of the Laplacian operator:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \phi_{mn} = -\lambda_{mn} \phi_{mn} \, .$$

- *Completeness:* Any function (reasonably well-behaved) can be expanded as an infinite linear combination of them (the double Fourier series $(*)$).

- *Orthogonality:* Each expansion coefficient c_{mn} can be calculated by a relatively simple integral formula, involving the corresponding eigenfunction ϕ_{mn} only.

These functions form an *orthogonal basis* for the vector space of functions whose domain is the rectangle (more precisely, for the space L^2 of square-integrable functions on the rectangle), precisely analogous to the orthogonal basis of eigenvectors for a symmetric matrix that students learn to construct in linear-algebra or ODE courses.

Remark: The double Fourier series raises complicated convergence questions, which it is not feasible to study here. We can say that if $g(x, y)$ is very smooth, then the coefficients go to 0 fast as m or $n \to \infty$, and everything is OK. (More precisely, what needs to be smooth is the extension of g which is even and periodic in x and odd periodic in y. This places additional conditions on the behavior of g at the boundaries.) Also, if g is merely square-integrable, then the series converges in the mean, but not necessarily pointwise. (In that case the series for g can be used for certain theoretical purposes — e.g., inside the integrand of certain integrals — but an attempt to add it up on a computer is likely to lead to disaster.) However, when $t > 0$ the series for w will converge nicely, even if g is rough, because the exponential factors make the terms decrease rapidly with m and n. This is a special feature of the heat equation: Because it describes a diffusive process, it drastically smooths out whatever initial data is fed into it.

Fourier Transforms

So far, all our variable-separation problems have involved finite intervals. The boundary conditions at the two ends of the intervals restricted the eigenvalues to some discrete set, such as $\{n^2\pi^2/L^2\}$. Fourier series are sums over such index sets. When the interval is infinitely long, the boundary restriction will disappear, and the Fourier sum will turn into an integral.

For a function on a finite interval of length $2L$, we have

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

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

Let's write

$$\omega_n \equiv \frac{n\pi}{L} \, .$$

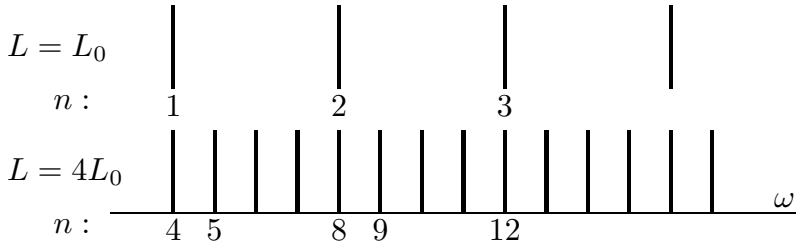
Then

$$f(x) = \sum_{\omega_n} c_n \, e^{i\omega_n x}.$$

The numbers ω_n are called "frequencies" or "wave numbers". As L increases, the frequencies become more closely spaced:

$$\Delta\omega_n \equiv \frac{(n+1)\pi}{L} - \frac{n\pi}{L} = \frac{\pi}{L} \, .$$

This suggests that for f defined on the whole real line, $\infty < x < \infty$, *all* values of ω should appear.



To make sense of the limit $L \to \infty$, we have to make a change of variable from n to ω. Let

$$\hat{f}(\omega_n) \equiv L\sqrt{\frac{2}{\pi}} \, c_n \, .$$

Then

$$f(x) = \sqrt{\frac{\pi}{2}} \sum_{\omega_n} \frac{1}{L} \hat{f}(\omega_n)\, e^{i\omega_n x}$$
$$= \frac{1}{\sqrt{2\pi}} \sum_{\omega_n} \hat{f}(\omega_n)\, e^{i\omega_n x}\, \Delta\omega_n\,,$$

$$\hat{f}(\omega_n) = \frac{1}{\sqrt{2\pi}} \int_{-L}^{L} f(x)\, e^{-i\omega_n x}\, dx.$$

As $L \to \infty$ the first formula looks like a Riemann sum. In the limit we therefore expect

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega x}\, d\omega,$$
$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-i\omega x}\, dx.$$

Note the surprising symmetry between these two formulas! $\hat{f}$ is called the *Fourier transform* of f, and f is the *inverse Fourier transform* of $\hat{f}$. The foregoing is not a proof that applying the two formulas in succession will take you back to the function f from which you started; all the convergence theorems for Fourier series need to be reformulated and reproved for this new situation. In fact, since the integrals are improper, the function f needs to satisfy some technical conditions before the integral $\hat{f}$ will converge at all. I'll discuss such questions briefly later, but first let's look at an example of how Fourier transforms are used in solving boundary-value problems.

Laplace's equation in the upper half-plane

Let the ranges of the variables be

$$-\infty < x < \infty, \qquad 0 < y < \infty.$$

Consider the equation

PDE: $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

with the boundary data

BC: $$u(x, 0) = f(x).$$

This equation might arise as the steady-state problem for heat conduction in a *large* plate, where we know the temperature along one edge and want to simplify

the problem by ignoring the effects of the other, distant edges. It could also arise in electrical or fluid-dynamical problems.

It turns out that to get a unique solution we must place one more condition on u: it must remain bounded as x or y or both go to infinity. (In fact, it will turn out that usually the solutions go to 0 at ∞.) Excluding solutions that grow at infinity seems to yield the solutions that are most relevant to real physical situations, where the region is actually finite. (You may sniff a hint of "boundary-layer theory" here.) But it is the mathematics of the partial differential equation which tells us that to make the problem well-posed we do *not* need to *prescribe some arbitrary function* as the limit of u at infinity, as we needed to do in the case of finite boundaries.

Separating variables for this problem at first gives one a feeling of *déjà vu*:

$$u_{\text{sep}}(x,y) = X(x)Y(y) \;\Rightarrow\; 0 = X''Y + XY'';$$
$$-\frac{X''}{X} = \lambda = \frac{Y''}{Y};$$

write λ as ω^2. The remaining steps, however, are significantly different from the case of the finite rectangle, which we treated earlier.

If $\lambda \neq 0$, the solution of the x equation can be

$$X(x) = e^{i\omega x},$$

where any ω and its negative give the same λ. The condition of boundedness requires that ω be *real* but does not further restrict it! Taking $\omega = 0$ yields the only bounded solution with $\lambda = 0$. Therefore, we take the X in each separated solution to be $e^{i\omega x}$ for some real ω. The corresponding λ will be positive or zero.

Turning now to the y equation, we see that Y is some linear combination of $e^{\omega y}$ and $e^{-\omega y}$. For boundedness we need the exponent to be negative, so we write

$$Y(y) = e^{-|\omega| y} \,(= e^{-\sqrt{\lambda}\, y})$$

to get an expression that's valid regardless of whether ω is positive or negative.

We are now finished with the homogeneous conditions, so we're ready to superpose the separated solutions. Since ω is a continuous variable, "superpose" in this case means "integrate", not "sum":

$$u(x,y) = \int_{-\infty}^{\infty} d\omega\, c(\omega)\, e^{i\omega x} e^{-|\omega| y}.$$

Here $c(\omega)$ is an arbitrary function, which plays the same role as the arbitrary coefficients in previous variable separations. The initial condition is

$$f(x) = \int_{-\infty}^{\infty} d\omega\, c(\omega)\, e^{i\omega x}.$$

Comparing with the formula for the inverse Fourier transform, we see that $c(\omega) = \frac{1}{\sqrt{2\pi}}\,\hat{f}(\omega)$. That is,

$$c(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\,e^{-i\omega x}\,dx.$$

In other words, the solution can be written

$$u(x,y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} d\omega\,\hat{f}(\omega)\,e^{i\omega x}e^{-|\omega|y}.$$

A Green function

We can get a simpler expression for u in terms of f by substituting the formula for $\hat{f}$ into the one for u. But to avoid using the letter x to stand for two different things in the same equation, we must first rewrite the definition of the Fourier transform using a different variable:

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dz\,e^{-i\omega z}\,f(z)\,.$$

Then

$$u(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\int_{-\infty}^{\infty} dz\,e^{i\omega(x-z)}e^{-|\omega|y}\,f(z).$$

We'll evaluate this multiple integral with the ω integral on the inside. (This step requires some technical justification, but that is not part of our syllabus.) The inner integral is

$$\begin{aligned}\int_{-\infty}^{\infty} d\omega\,e^{i\omega(x-z)}e^{-|\omega|y} &= \int_{-\infty}^{0} d\omega\,e^{i\omega(x-z)}e^{\omega y} + \int_{0}^{\infty} d\omega\,e^{i\omega(x-z)}e^{-\omega y}\\ &= \frac{e^{i\omega(x-z-iy)}}{i(x-z-iy)}\,\Big|_{-\infty}^{0} + \frac{e^{i\omega(x-z+iy)}}{i(x-z+iy)}\,\Big|_{0}^{\infty}\\ &= \frac{1}{i(x-z-iy)} - \frac{1}{i(x-z+iy)}\\ &= \frac{2y}{(x-z)^2+y^2}\,.\end{aligned}$$

Thus

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} dz\,\frac{y}{(x-z)^2+y^2}\,f(z). \qquad (*)$$

The function

$$G(x-z,y) \equiv \frac{1}{\pi}\,\frac{y}{(x-z)^2+y^2}$$

is called a *Green function* for the boundary-value problem we started from. It is also called the *kernel* of the *integral operator*

$$u = G(f)$$

defined by $(*)$. The point of $(*)$ is that it gives the solution, u, as a function of the boundary data, f. Some properties of the Green function will be pointed out in a later homework assignment.

In principle, Green functions exist for the boundary-value problems on finite regions which we have solved earlier. However, in those cases the G is given by an infinite *sum* arising from the Fourier *series*, rather than the *integral* which expresses G in a Fourier-transform problem.

Gaussian integrals

The Green function for the *heat* equation on an infinite interval can be derived from the Fourier-transform solution. This, and also several of the homework problems from the Schaum's Outline book, make use of a basic integral formula, which I'll now derive.

The integral in question is

$$H(x) \equiv \int_{-\infty}^{\infty} e^{i\omega x}\, e^{-\omega^2 t}\, d\omega,$$

where t is positive.

Note first that

$$\frac{d}{d\omega} e^{-\omega^2 t} = -2\omega t\, e^{-\omega^2 t}.$$

This will allow us to find a differential equation satisfied by H: From the definition we calculate

$$\begin{aligned} H'(x) &= \int_{-\infty}^{\infty} i\omega\, e^{i\omega x}\, e^{-\omega^2 t}\, d\omega \\ &= \frac{-i}{2t} \int_{-\infty}^{\infty} e^{i\omega x} \left(\frac{d}{d\omega} e^{-\omega^2 t} \right) d\omega \\ &= \frac{+i}{2t} \int_{-\infty}^{\infty} \left(\frac{d}{d\omega} e^{i\omega x} \right) e^{-\omega^2 t}\, d\omega \\ &= \frac{-x}{2t} \int_{-\infty}^{\infty} e^{i\omega x}\, e^{-\omega^2 t}\, d\omega \\ &= -\frac{x}{2t}\, H(x). \end{aligned}$$

Thus

$$\frac{H'}{H} = -\frac{x}{2t};$$
$$\ln H = -\frac{x^2}{4t} + \text{const.};$$
$$H = C\, e^{-x^2/4t}.$$

To find the constant we evaluate the integral for $x = 0$:

$$\begin{aligned} C &= H(0) \\ &= \int_{-\infty}^{\infty} e^{-\omega^2 t}\, d\omega \\ &= \frac{1}{\sqrt{t}} \int_{-\infty}^{\infty} e^{-q^2}\, dq, \end{aligned}$$

by the substitution $q = \omega\sqrt{t}$. But it is well known that

$$\int_{-\infty}^{\infty} e^{-q^2}\, dq = \sqrt{\pi},$$

because its square is

$$\begin{aligned} \iint_{\mathbf{R}^2} e^{-x^2} e^{-y^2}\, dx\, dy &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2}\, r\, dr\, d\theta \\ &= 2\pi \int_0^{\infty} e^{-u}\, \frac{1}{2}\, du \\ &= \pi. \end{aligned}$$

So

$$C = \sqrt{\frac{\pi}{t}}\,.$$

Therefore, we have shown that $H(x)$ is

$$\boxed{\int_{-\infty}^{\infty} e^{i\omega x}\, e^{-\omega^2 t}\, d\omega = \sqrt{\frac{\pi}{t}}\, e^{-x^2/4t}.}$$

In particular,

$$G(t, x - z) \equiv \frac{1}{2\pi}\, H(x - z) = \frac{1}{\sqrt{4\pi t}}\, e^{-(x-z)^2/4t}$$

is the Green function for the one-dimensional infinite-space heat equation, which we have now derived by Fourier-transform methods. Note also that the formula in the box is also useful for evaluating similar integrals with the roles of x and ω interchanged. (Taking the complex conjugate of the formula, we note that the sign of the i in the exponent doesn't matter at all.)

More about Fourier Transforms

(1) The Fourier transform can be written in terms of the basic functions $\sin\omega x$ and $\cos\omega x$ $(0 \le \omega < \infty)$ in place of $e^{i\omega x}$ $(-\infty < \omega < \infty)$. In the normalization convention of the Schaum's Outline book, the formulas are

$$f(x) = \int_0^\infty [A(\omega)\,\cos\omega x + B(\omega)\,\sin\omega x]\,d\omega,$$

$$A(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \cos\omega x\, f(x)\,dx,$$
$$B(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \sin\omega x\, f(x)\,dx.$$

This is seldom done in practical calculations with functions defined on $-\infty < x < \infty$, except by people with a strong hatred for complex numbers.

However, the trigonometric functions do become very useful in calculations on a half-line (semi-infinite interval) with a boundary condition at the end. The basic argument is the same as for Fourier series: An arbitrary function on $0 \le x < \infty$ can be identified with its *even extension* to the whole real line. An *even* function has a Fourier transform consisting entirely of cosines (rather than sines), and the formula for the coefficient function can be written as an integral over just the positive half of the line:

$$f(x) = \int_0^\infty A(\omega)\,\cos\omega x\,d\omega,$$

$$A(\omega) = \frac{2}{\pi}\int_0^\infty \cos\omega x\, f(x)\,dx.$$

An equally common normalization convention splits the constant factor symmetrically between the two formulas:

$$f(x) = \sqrt{\frac{2}{\pi}}\int_0^\infty A(\omega)\,\cos\omega x\,dx,$$

$$A(\omega) = \sqrt{\frac{2}{\pi}}\int_0^\infty f(x)\,\cos\omega x\,dx.$$

Still other people put the entire factor $\frac{2}{\pi}$ into the $A \mapsto F$ equation. In any case, A is called the *Fourier cosine transform* of f, and it's often given a notation such as $\hat{f}_{\mathrm{c}}(\omega)$ or $F_{\mathrm{C}}(\omega)$.

Correspondingly, there is a *Fourier sine transform* related to *odd* extensions of functions. The sine transform arises naturally in problems where the functions vanish at the boundary ($x = 0$), and the cosine transform is appropriate when the derivative vanishes there.

For example, let's solve the heat equation for a semi-infinite rod with a constant temperature at the end:

$$0 < x < \infty, \qquad 0 < t < \infty,$$

PDE:
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},$$

BC:
$$u(t, 0) = T,$$

IC:
$$u(0, x) = f(x).$$

A steady-state solution in this case is obviously $u(t, x) = T$, so let's assume that that constant has already been subtracted off, and take $T = 0$. (So f is the true initial temperature distribution minus T.)

The usual separation of variables obviously would lead to

$$X'' = -\lambda X, \qquad T' = -\lambda T,$$

and the BC together with the usual boundedness assumption gives

$$X(x) = \sin \omega x, \qquad \omega > 0, \qquad \lambda = \omega^2.$$

(Negative ω's give the same functions and hence are redundant.) Thus

$$u(t, x) = \int_0^\infty d\omega \, B(\omega) \, \sin \omega x \, e^{-\omega^2 t}.$$

And so

$$f(x) = \int_0^\infty d\omega \, B(\omega) \, \sin \omega x,$$

or

$$B(\omega) = \frac{2}{\pi} \int_0^\infty dx \, f(x) \, \sin \omega x,$$

according to the sine transform formulas.

From these results one can easily calculate a Green function for u in terms of f:

$$G(t, x, z) = \frac{1}{\sqrt{4\pi t}} \, e^{-(x-z)^2/4t} - \frac{1}{\sqrt{4\pi t}} \, e^{-(x+z)^2/4t}.$$

It turns out to be a combination of the heat Green function for the whole real line with the heat Green function for a fictitious "image" source in the unphysical region,

$x < 0$. The terms are subtracted so that the total will satisfy the homogeneous Dirichlet boundary condition at $x = 0$. This "method of images" is a very efficient method of solving boundary value problems, when it applies.

(2) The shortest possible summary of the subject so far:

Data on finite intervals $\Rightarrow$ Fourier series (sums);

Data on infinite intervals $\Rightarrow$ Fourier transforms (integrals).

To elaborate further,

Dirichlet BC at 0 (the function vanishes) $\Rightarrow$ sines;

Neumann BC at 0 (the derivative vanishes) $\Rightarrow$ cosines;

No restriction at 0 (e.g., periodic case) $\Rightarrow$ both sines and cosines, or complex exponentials.

All this assumes that the basic ODE which arises from the separation of variables is $X'' = -\lambda X$. More complicated equations have more complicated eigenfunctions, with related new types of sum or integral expansions. (More about this later.)

(3) Convergence theorems: First, let's state the generalization to Fourier transforms of the pointwise convergence theorem for Fourier series. To get a true theorem, we have to make a seemingly fussy, but actually quite natural, technical condition on the function: Let's define a function with domain $(-\infty, \infty)$ to be *piecewise smooth* if its restriction to every finite interval is piecewise smooth. (Thus f is allowed to have infinitely many jumps or corners, but they must not pile up in one region of the line.) The Fourier transform is defined by

$$\hat{f}(\omega) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-i\omega x}\, dx.$$

Pointwise convergence theorem: If $f(x)$ is piecewise smooth, and

$$\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$$

(f is *absolutely integrable*), then:

a) $\hat{f}(\omega)$ is continuous.

b) $\hat{f}(\omega) \to 0$ as $|\omega| \to \infty$ (but $\hat{f}$ is not necessarily absolutely integrable itself). (This is a new version of the Riemann–Lebesgue theorem.)

c) The inverse Fourier transform

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega x}\, d\omega$$

converges pointwise to $\frac{1}{2}[f(x^+) + f(x^-)]$ (just like Fourier series).

The next theorem treats the variables x and ω on a completely symmetrical basis.

Mean convergence theorem: If $f(x)$ is sufficiently smooth to be integrated, and

$$\int_{-\infty}^{\infty} |f(x)|^2\, dx < \infty$$

(f is *square-integrable*), then:

a) $\hat{f}(\omega)$ is also square-integrable. (The integral defining $\hat{f}(\omega)$ may not converge at every point ω, but it will converge "in the mean", just like the inversion integral discussed below.)

b) A Parseval equation holds:

$$\int_{-\infty}^{\infty} |f(x)|^2\, dx = \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\, d\omega.$$

(If you define $\hat{f}$ so that the 2π is kept all in one place, then this formula will not be so symmetrical.)

c) The inversion formula converges in the mean:

$$\lim_{\Lambda\to\infty} \int_{-\infty}^{\infty} dx\, |f(x) - f_\Lambda(x)|^2 = 0$$

where

$$f_\Lambda(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\Lambda}^{\Lambda} \hat{f}(\omega)\, e^{i\omega x}\, d\omega.$$

(4) *The Fourier transform of f' is $i\omega$ times that of f* (assuming, of course, that f' is "nice" enough for its transform to be defined). This can be seen either by differentiating

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega x}\, d\omega$$

with respect to x, or by integrating by parts in

$$\widehat{f'}\,(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(x)\, e^{-i\omega x}\, dx.$$

Similarly, differentiation with respect to ω corresponds to multiplication by $-ix$.

(5) This differentiation property can be used to streamline the process of solving a PDE by Fourier transforms. So far I have taught you to find separated solutions $e^{i\omega x}\cdots$, superpose them, and then calculate the coefficient function as the Fourier transform of the initial or boundary data, f. The method preferred by many people is to start by taking the Fourier transform of the entire solution, u:

$$\hat{u}(t,\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, e^{-i\omega x}\, u(t,x).$$

The result is an ordinary differential equation in t (or whatever the other variable is), which can be easily solved.

To demonstrate this, let's return to the example of Laplace's equation in the half-plane:

$$-\infty < x < \infty, \qquad 0 < y < \infty,$$

PDE:
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

BC:
$$u(x,0) = f(x),$$

$u(x,y)$ bounded. Apply the Fourier transform to both equations. By property (4), (PDE) becomes

$$-\omega^2 \hat{u}(\omega,y) + \frac{d^2}{dy^2}\hat{u}(\omega,y) = 0.$$

(BC) transforms to

$$\hat{u}(\omega,0) = \hat{f}(\omega).$$

Treating ω as a constant, we find the solution to the DE

$$\hat{u}(\omega,y) = c_1 e^{\omega y} + c_2 e^{-\omega y},$$

but the requirement of boundedness forces $c_1 = 0$ if $\omega > 0$ and $c_2 = 0$ if $\omega < 0$. This can be written in a unified equation as

$$\hat{u}(\omega,y) = c\, e^{-|\omega| y}.$$

Now we remember that c may be a function of ω; and in fact the boundary condition implies that $c(\omega) = \hat{f}(\omega)$. Thus

$$\hat{u}(\omega,y) = \hat{f}(\omega) e^{-|\omega| y},$$

and hence

$$u(x,y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega\, e^{i\omega x}\, \hat{f}(\omega)\, e^{-|\omega| y}.$$

Of course, this is the same answer we found before.

This technique can be applied in this simple way *only to problems where the domain of x is the whole real line.* Although we aren't explicitly writing down a separation of variables, we are using our prior knowledge that separation of variables *would* lead to a full Fourier transform, to which the differentiation property (4) applies. It *is* possible to treat problems on finite or semi-infinite domains by variations of this method. (In this context the list of Fourier coefficients $\{c_n\}$ is called a "finite Fourier transform".) But the differentiation property must then be *modified* to take account of the boundary data, and also the difference between sines and cosines. Being careless about these points can easily lead to mistakes — especially if the appropriate periodic or even/odd extension of the data function is discontinuous at the endpoint! — and since we do not have time to study the procedure carefully in this course, it seems safer to avoid it entirely.

(6) Relation to the Laplace transform: Suppose $f(x) = 0$ for $x < 0$. Then

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(x)\, e^{-i\omega x}\, dx.$$

Recall that the *Laplace transform* of f is

$$F(s) = \int_0^\infty f(x)\, e^{-sx}\, dx.$$

Allow s to be complex:

$$s = \sigma + i\omega, \qquad \sigma \text{ and } \omega \text{ real.}$$

Then

$$\begin{aligned} F(s) &= \int_0^\infty f(x)\, e^{-\sigma x}\, e^{-i\omega x}\, dx \\ &= \sqrt{2\pi} \times \text{Fourier transform of } f(x)e^{-\sigma x} \qquad (\sigma \text{ fixed}). \end{aligned}$$

For "most" f's, $f(x)e^{-\sigma x}$ will be square-integrable if σ is sufficiently large, even if f itself is not square-integrable (e.g., f = polynomial for $x > 0$). To attain this result it was crucial that we cut f off below $x = 0$; when we multiply by $e^{-\sigma x}$, $\sigma > 0$, what we gain at $x = +\infty$ we lose at $-\infty$. The Laplace transform (with *time* in the role of x) is useful for solving initial-value problems, where the data and solution functions may not fall off to 0 as the time approaches $+\infty$, but negative values of time are not of interest. (In particular, the Laplace transform with respect to time can be applied to nonhomogeneous boundary data that depend on time, so that the steady-state solution technique does not apply.)

The Fourier inversion formula for $fe^{-\sigma x}$ says

$$f(x)e^{-\sigma x} = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\sigma + i\omega)\, e^{i\omega x}\, d\omega,$$

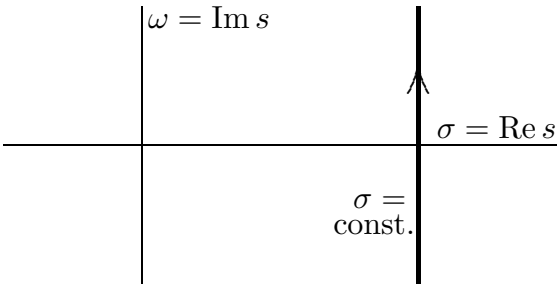
or

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\sigma + i\omega)\, e^{(\sigma+i\omega)x}\, d\omega.$$

In the exponent we recognize the complex variable $s \equiv \sigma + i\omega$. If we do a formal integration by substitution, taking $ds = i\, d\omega$, we get

$$\boxed{f(x) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} F(s)\, e^{sx}\, ds.}$$

In courses on complex analysis (such as Math. 407 and 601), it is shown that this integral makes sense as a line integral in the complex plane. It provides an *inversion formula for Laplace transforms*. Remember that in Math. 308 no such formula was available; the only way to invert a Laplace transform was to "find it in the right-hand column of the table" — that is, to know beforehand that that function can be obtained as the direct Laplace transform of something else. The complex analysis courses also provide techniques for evaluating such integrals, so the number of problems that can be solved exactly by Laplace transforms is significantly extended.



In short, the Laplace transform is really the Fourier transform, extended to complex values of ω and then rewritten in a notation that avoids complex numbers — until you want a formula to calculate the inverse transform, whereupon the complex numbers come back with a vengeance.

(7) Convolutions, autocorrelation function, and power spectrum: In this course we emphasize the use of the Fourier transform in solving partial differential equations. The Fourier transform also has important applications in signal processing and the analysis of data given as a function of a time variable. The subject is presented with that "EE flavor" in many books, such as the Harcourt Brace Jovanovich College Outline Series competition to our Schaum's Outline book:

Applied Fourier Analysis by H. P. Hsu. Here we take a quick look at some of the tools of that trade.

The Fourier transform of a product of functions is *not* the product of their Fourier transforms! Instead, it is easy to show that that transform is a certain integral involving the transforms of the two factors. This fact is most often used in the inverse direction, so that is how I'll state the formula:

Convolution Theorem: The inverse Fourier transform of $\hat{f}_1(\omega)\hat{f}_2(\omega)$ is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_1(u) f_2(x-u)\, du.$$

This integral is called the *convolution* of f_1 and f_2. Note that

$$f_1 * f_2 = f_2 * f_1\,,$$

although that is not immediately visible from the integral formula.

This theorem is proved on p. 86 of the Schaum book. However, the proof contains a misprint:

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \qquad \text{should be} \qquad \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix}.$$

By manipulating the formulas defining the Fourier transform and its inverse, it is easy to show the following:

Theorem:

(a) If $f(x)$ is real-valued, then $\hat{f}(\omega)^* = \hat{f}(-\omega)$.

(b) If $g(x) \equiv f(-x)$, then $\hat{g}(\omega) = \hat{f}(-\omega)$.

(c) If $f(x)$ is real-valued and $g(x) \equiv f(-x)$, then $\hat{g}(\omega) = \hat{f}(\omega)^*$.

Now take $\hat{f}_1 = \hat{f}$ and $\hat{f}_2 = \hat{f}^*$ in the convolution theorem and apply the theorem just stated:

Corollary: If $f(x)$ is real-valued, then the Fourier transform of $|\hat{f}(\omega)|^2$ is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) f(u-x)\, du.$$

This integral is called the *autocorrelation function* of f, because it measures to what extent values of f at arguments displaced a distance x tend to coincide. The function $|\hat{f}(\omega)|^2$ is called the *power spectrum* of f; it measures the extent to which the signal in f is concentrated at frequency ω.

