Partial Differential Equations

Notes from the texts by R. Haberman and W. Strauss

Jeffrey A. Oregero
Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA
August 26, 2024

Contents

1 Introduction

These notes provide an introduction to the subject of partial differential equations.

1.1 Preliminary definitions

- **Ordinary differential equation (ODE)** – The key defining property is that there is one independent variable.
- **Partial differential equation (PDE)** – The key defining property is that there is more than one independent variable.

Definition 1.1. A PDE is a mathematical relation between an unknown function and its partial derivatives.

Definition 1.2. The order of a PDE is the value of the highest derivative appearing in the equation.

For example, the most general form of a first-order PDE in two independent variables is given by

$$F(x, y, u, u_x, u_y) = 0, \tag{1.1}$$

where $u = u(x, y)$ is the unknown.

Definition 1.3. A solution of a PDE is a function $u(x, y, \dots)$ that satisfies the equation indentically, at least in some region of the $x, y, \dots$ variables.

Definition 1.4. A linear PDE is one that can be written in the form:

- $Lu = 0$, (homogeneous),
- $Lu = f$, (non-homogeneous),

where L is a linear operator, i.e., for functions u, v, and scalar c, L satisfies the following two properties:

(i) $L(u+v) = L(u) + L(v)$,

(ii) $L(cu) = cL(u)$.

The key advantage of the linear homogeneous equation $Lu = 0$ is that if u, v are solutions so is $c_1u + c_2v$, i.e., linear combinations of solutions is again a solution. This is referred to as the "superposition principle".

In these notes we wish to discuss the solution of elementary problems involving partial differential equations. Here are a few of the equations that we will study:

- $u_t + \mathbf{c} \cdot \nabla u = 0$ (transport equatin - N spatial dimensions)
- $u_t = \kappa u_{xx}$ (heat equation - one spatial dimension)
- $u_t = \kappa(u_{xx} + u_{yy})$ (heat equation - two spatial dimensions)
- $u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta}$ (Laplace's equation - polar coordinates)
- $u_{tt} = c^2\Delta u$ (wave equation - three spatial dimensions)

Recall that

$$\Delta := \nabla \cdot \nabla = (\partial_x, \partial_y, \partial_z) \cdot (\partial_x, \partial_y, \partial_z) = \partial_x^2 + \partial_y^2 + \partial_z^2 \tag{1.2}$$

is a linear operator called the "Laplacian" (in Cartesian coordinates), and moreover,

$$\Delta u = 0\,, \quad u = u(x, y, z) \tag{1.3}$$

is called "Laplace's equation".

Why are PDEs useful? Most of the natural laws of physics are stated in terms of PDEs. Applications include acoustics, aerodynamics, fluid dynamics, wave propagation, optics, quantum mechanics, thermodynamics, etc. Note PDEs also arise in the social sciences. For example, the famous Black-Scholes equation is a PDE governing the fair-market value of a call option in finance. Finally, PDEs also arise in geometry.

Goals:

- How to formulate the PDE
- How to solve the PDE

Solution techniques:

- Method of characteristics (first-order PDEs)
- Separation of variables (Fourier series)
- Integral transforms (Fourier transform)
- Change of coordinates
- Eigenfunction expansions
- Impulse-response (Green's function)

- Similarity solution method

Types:

$$a(x,y)u_x + b(x,y)u_y + c(x,y)u = d(x,y), \quad \text{linear first-order},$$

$$a(x,y,u)u_x + b(x,y,u)u_y = c(x,y,u), \quad \text{quasi-linear first-order},$$

$$a(x,y)u_{xx} + b(x,y)u_{xy} + c(x,y)u_{yy} + d(x,y)u_x + e(x,y)u_y + f(x,y)u = g(x,y), \quad \text{linear second-order}$$

Phenomena:

- Transport (linear transport processes)
- Heat–parabolic (diffusion processes)
- Wave–hyperbolic (vibrating systems)
- Laplace–elliptic (steady-state phenomena)

Mathematics:

- solution techniques
- well-posedness
  - (i) existence of solutions
  - (ii) uniqueness of solutions
  - (iii) continuous dependence on the data
- asymptotic and qualitative analysis of solutions
- stability

At the same time many diverse subject areas in engineering and the physical sciences are dominated by the study of PDEs. For example, the following areas are highly dependent on the study of PDEs: acoustics, aerodynamics, elasticity, electrodynamics, fluid dynamics, geophysics (seismic wave propagation), meteorology, oceanography, optics, plasma physics, quantum mechanics, and thermodynamics.

1.2 Initial-value problems (IVPs) versus boundary-value problems (BVPs)

Recall from your introduction to ODEs course that the general solution to the linear second-order homogeneous equation

$$y'' + p(t)y' + q(t)y = 0 \tag{1.4}$$

is a linear combination of two linearly independent solutions. Moreover, we have the important theorem:

Theorem 1.5. There exists a unique solution to the IVP

$$y'' + p(t)y' + q(t)y = 0, \quad y(t_o) = y_0, \quad y'(t_o) = y_1, \tag{1.5}$$

on an interval $I(t_o)$ containing the initial point t_o so long as $p(t)$ and $q(t)$ are continuous on the interval $I(t_o)$.

Distinct from the IVP is the so-called BVP. The answer to the question of existence/uniqueness of solutions to BVPs is more subtle as shown by the following example.

Example 1.6. Solve the following BVP:

$$y'' - y = 0, \quad y(0) = y(\pi) = 0.$$

The characteristic equation for this problem is $r^2 - 1 = 0 \iff r = \pm 1$. Thus, the general solution is given by

$$y(t) = c_1 e^t + c_2 e^{-t}.$$

Then applying the first BC we see that $y(0) = c_1 + c_2 = 0 \iff c_2 = -c_1$. Thus, the solution must be of the form $y(t) = c_1(e^t - e^{-t})$, or equivalently, $y(t) = 2c_1 \sinh(t)$. Applying the second BC gives $y(\pi) = 2c_1 \sinh(\pi) = 0 \iff c_1 = 0$ (since $\sinh(\pi) \neq 0$). Thus, we get the unique trivial solution $y(t) \equiv 0$ since $c_1 = c_2 = 0$ is the only way to satisfy the BCs.

Example 1.7. Solve the following BVP:

$$y'' + y = 0, \quad y(0) = y(\pi) = 0.$$

The characteristic equation for this problem is $r^2 + 1 = 0 \iff r = \pm i$. Thus, the general solution is given by

$$y(t) = c_1 \cos(t) + c_2 \sin(t).$$

Then applying the first BC we see that $y(0) = c_1 = 0$. Thus, the solution must be of the form $y(t) = c_2 \sin(t)$. Applying the second BC gives $y(\pi) = c_2 \sin(\pi) = 0$. Since $\sin(\pi) = 0$ we see that the BCs are satisfied for any value of c_2. Thus, in this case we have infinitely many solutions.

Recall from ODEs that the linear homoegeneous second-order equation with constant coefficients, i.e.,

$$ay'' + by' + cy = 0 \tag{1.6}$$

has solution $y(t) = e^{rt}$ if and only if r is a root of the "characteristic equation"

$$ar^2 + br + c = 0. \tag{1.7}$$

Moreover, it can be shown that the general solution is given by the following:

- distinct real roots r_1, r_2 – $y(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}$
- repeated real root $r_1 = r_2 = r$ – $y(t) = (c_1 + c_2 t)e^{rt}$
- complex conjugate roots $r_{1,2} = \alpha \pm i\beta$ – $y(t) = e^{\alpha t}(c_1 \cos(\beta t) + c_2 \sin(\beta t))$

Finally, by general solution we mean that all solutions live within the given family of solutions. By particular solution we mean a solution selected from the family that satisfies some initial or boundary conditions.

We prefer to start by investigating a physical problem. There are two reasons for this (i) the mathematical techniques will be of greater interest when it becomes clear that these models analyze 'real-world' problems; and (ii) we will find that physical considertations will motivate many of our mathematical developments.

2 Transport equation

We start by studying the physical phemonena of linear transport. We make the following assumptions:

(i) Let $u = u(x,t)$ denote the density (mass per unit volume) of some substance at the point x and time t.

(ii) The geometry is a line with coordinate x (e.g. pollutant in a river or cars on the road).

(iii) The total amount of the substance remains fixed (no sources or sinks).

Thus, in an arbitrary interval (x_1, x_2) it follows that

$$M = \int_{x_1}^{x_2} u(x,t)\,\mathrm{d}x, \tag{2.1}$$

is the total amount of the substance in the interval (x_1, x_2). Thus, the time-rate of change of the amount is

$$\frac{dM}{dt} = \frac{d}{dt}\int_{x_1}^{x_2} u(x,t)\,\mathrm{d}x = \int_{x_1}^{x_2} \frac{\partial}{\partial t}u(x,t)\,\mathrm{d}x = \int_{x_1}^{x_2} u_t(x,t)\,\mathrm{d}x$$

where we assume that u is smooth enough to differentiate under the integral. Next, let $q(x,t)$ denote the flux, i.e., change of the substance per unit time at point x. Thus,

- $q(x_1,t)$ = amount of substance that enters,
- $q(x_2,t)$ = amount of substance that leaves

Hence, since there are no sources or sinks by conservation of mass

$$\begin{aligned}\int_{x_1}^{x_2} u_t(x,t)\,\mathrm{d}x &= q(x_1,t) - q(x_2,t)\\ &= -(q(x_2,t) - q(x_1,t))\\ &= -\int_{x_1}^{x_2} q_x(x,t)\,\mathrm{d}x\end{aligned}$$

where the last line follows from the Fundamental Theorem of Calculus. Thus, since (x_1, x_2) is an arbitrary interval it follows that

$$\int_{x_1}^{x_2} u_t(x,t) + q_x(x,t)\,\mathrm{d}x = 0 \iff q_t + q_x = 0.$$

Hence, we've derived the following scalar conservation law:

$$u_t(x,t) + q_x(x,t) = 0. \tag{2.2}$$

Finally, letting $q = cu$, where c is the velocity of the substance we arrive at the transport equation, namely,

$$u_t + cu_x = 0. \tag{2.3}$$

The transport equation (2.3) is a particularly simple example of a first-order linear PDE. Next, we will derive the general solution using three different techniques.

Geometric method: Notice that $u_t + cu_x$ is the directional derivative of u in the direction $\mathbf{v} = (c, 1)$. Recall that the directional derivative in the direction $\mathbf{v}$ is given by

$$(\nabla u) \cdot \mathbf{v}, \tag{2.4}$$

where $\nabla u = (\partial_x u, \partial_t u)$ is the gradient of u. Thus, the transport equation can be written as

$$(\nabla u) \cdot \mathbf{v} = 0, \tag{2.5}$$

which says that u is constant in the direction $\mathbf{v} = (c, 1)$. Thus, u is constant along the line $t = x/c$, or $x - ct = 0$. More generally, the solution is constant along lines $x - ct = x_o$ for some constant x_o. Thus, along each line $x - ct = x_o$ the solution is constant, i.e., $u = f(x_o)$. Therefore, the general solution of the transport equation is

$$u(x, t) = f(x - ct), \tag{2.6}$$

where f is an arbitrary function.

Coordinate method: Let's make the invertible change-of-coordinates

$$\tau = cx + t \qquad \eta = x - ct,$$

so that $u = u(\eta(x, t), \tau(x, t))$. Thus, by the chain rule

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial t} + \frac{\partial u}{\partial \tau}\frac{\partial \tau}{\partial t} = -cu_\eta + u_\tau, \\
\frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial \tau}\frac{\partial \tau}{\partial x} = u_\eta + cu_\tau.
\end{aligned}$$

Plugging into the transport equation gives

$$\begin{aligned}
u_t + cu_x &= -cu_\eta + u_\tau + c(u_\eta + cu_\tau) \\
&= (1 + c^2)u_\tau \\
&= 0.
\end{aligned}$$

Thus, since $1 + c^2 \neq 0$ it follows that in the new coordinates the transport equation is

$$u_\tau = 0, \tag{2.7}$$

which has general solution $u(\eta, \tau) = f(\eta)$. Thus,

$$u(x, t) = f(x - ct). \tag{2.8}$$

Method of characteristics Assume we have a moving reference frame $x = x(t)$. Then since $u = u(x(t), t)$ by the chain rule:

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial t}.$$

Thus, if the observer moves with velocity c, i.e., if $dx/dt = c$, then

$$\frac{du}{dt} = cu_x + u_t = 0,$$

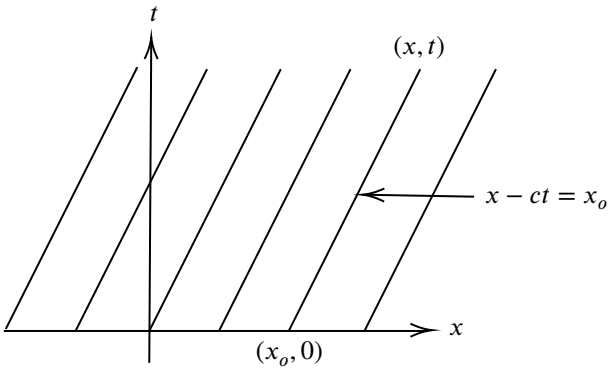

Figure 1: Characteristic curves given by–$dx/dt = c$.

which implies that u remains constant. Thus, an observer moving with speed $|c|$ would measure no change in u. Thus, the PDE has been replaced by a set of ODEs, namely,

$$\frac{dx}{dt} = c, \tag{2.9a}$$

$$\frac{du}{dt} = 0, \tag{2.9b}$$

which are called the "characteristic ODEs".

Definition 2.1. A characteristic is a curve along which a PDE reduces to an ODE.

Note: u is constant along the lines $x = ct + x_o$ not necessarily constant everywhere (see Fig. 1). Thus, u propagates as a wave with speed $|c|$. This is known as advection.

Example 2.2. Let's consider the following example:

$$u_t + 2u_x = 0, \tag{2.10}$$

subject to the initial condition:

$$u(x,0) = \begin{cases} 4x, & 0 < x < 1, \\ 0, & x < 0,\ x > 1. \end{cases} \tag{2.11}$$

We know that u is constant along the characteristic curves $x - 2t = \text{const.}$, keeping its same shape with velocity two (moving to the right). The important characteristics (where the initial data changes form) are

$$x = 2t + 0,$$
$$x = 2t + 1.$$

Thus, $u = 0$ if $x > 2t + 1$, or if $x < 2t$. Otherwise, by shifting

$$u(x,t) = 4(x - 2t), \qquad 2t < x < 2t + 1.$$

Thus, the solution is given by

$$u(x,t) = \begin{cases} 4(x-2t) & 0 < x - 2t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

3 The method of characteristics

The method of characteristics provides a systematic way to reduce the first-order quasi-linear PDE

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) \tag{3.1}$$

to a system of ODEs called the characteristic ODEs.

Suppose we can find a solution $z = u(x, y)$. The graph of the solution is then given by

$$\mathcal{S} := \{(x, y, z = u(x, y))\}. \tag{3.2}$$

Thus, if $u(x, y)$ solves (3.1), we know that at each point $(x, y, z) \in \mathcal{S}$

$$(a(x, y, u), b(x, y, u), c(x, y, u)) \cdot (u_x, u_y, -1) = 0, \tag{3.3}$$

where $N(x, y) = (u_x, u_y, -1)$ is the normal vector to the level surface $u(x, y) - z = 0$.

Recall that a level surface is the set of points (x, y, z) such that $F(x, y, z) = c$ for some constant c. Thus we parametrize a curve $(x(s), y(s), z(s))$ on the level surface $F(x, y, z) = c$. Then it follows that

$$\begin{aligned}\frac{dF}{ds} &= \frac{\partial F}{\partial x}\frac{dx}{ds} + \frac{\partial F}{\partial y}\frac{dy}{ds} + \frac{\partial F}{\partial z}\frac{dz}{ds} \\ &= (\nabla F) \cdot (x'(s), y'(s), z'(s)) \\ &= 0\end{aligned}$$

where $(x'(s), y'(s), z'(s))$ is the tangent vector to the parametrized curve. Thus, the gradient vector is normal to the level surface. In our case $\mathcal{S}$ is equivalent to the level surface $u(x, y) - z = 0$. This is precisely why the vector $(u_x, u_y, -1)$ is normal to the surface $\mathcal{S}$. Hence, the vector $(a(x, y, u), b(x, y, u), c(x, y, u))$ lies in the tangent plance. (Since it is orthogonal to the surface.)

Thus, the objective is to construct a surface $\mathcal{S} = \{(x, y, z = u(x, y))\}$ such that at each point (x, y, z) on $\mathcal{S}$, the vector $(a(x, y, u), b(x, y, u), c(x, y, u))$ lies in the tangent plane. Let's parameterize a curve $\mathcal{C} = \{(x(s), y(s), z(s))\}$ on the surface. Thus, at any point on the surface we must have $u(x(s), y(s)) - z(s) = 0$. Thus, the tangent vector at this point is

$$\frac{dz}{ds} = \frac{\partial u}{\partial x}\frac{dx}{ds} + \frac{\partial u}{\partial y}\frac{dy}{ds}.$$

Thus, from the PDE (3.1) if we impose $dx/ds = a(x, y, u)$ and $dy/ds = b(x, y, u)$, then $dz/ds = c(x, y, u)$.

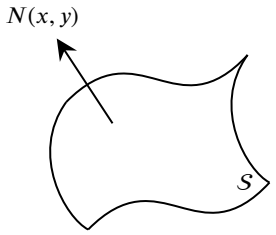

Figure 2: Normal vector to the surface $\mathcal{S}$.

Characteristics ODEs – First-order quasi-linear PDE:

$$\frac{dx}{ds} = a(x, y, u), \tag{3.4a}$$
$$\frac{dy}{ds} = b(x, y, u), \tag{3.4b}$$
$$\frac{dz}{ds} = c(x, y, u). \tag{3.4c}$$

The logic is the following: Suppose we know the surface $z = u(x, y)$. Then the vector $N(x, y) = (u_x, u_y, -1)$ is normal to the surface at each point. Since the PDE (3.1) can be written as

$$(a(x, y, u), b(x, y, u), c(x, y, u)) \cdot (u_x, u_y, -1) = 0,$$

this implies that the vector field $(a(x, y, u), b(x, y, u), c(x, y, u))$ is tangent to the surface at every point $(x, y, z = u(x, y))$. Thus, the surface is a union of integral curves of this vector field, i.e., curves that satisfy the characteristics ODEs.

Example 3.1. Find the general solution to the PDE

$$u_t + e^{x+t} u_x = 0. \tag{3.5}$$

Then solve IVP for (3.5) with initial condition $u(x, 0) = \phi(x)$. We apply the method of characteristics:

$$\frac{dt}{ds} = 1, \tag{3.6a}$$
$$\frac{dx}{ds} = e^{x+t}, \tag{3.6b}$$
$$\frac{dz}{ds} = 0, \tag{3.6c}$$

or equivalently,

$$\frac{dx}{dt} = e^{x+t}, \tag{3.7a}$$
$$\frac{du}{dt} = 0. \tag{3.7b}$$

Thus, $e^{-x}dx = e^t dt \iff -e^{-x} = e^t + C \iff x(t) = -\ln(C - e^t)$. Also, $u = \text{const.}$ along the characteristics. Thus, the general solution is given by

$$u(x, t) = f(e^{-x} + e^t). \tag{3.8}$$

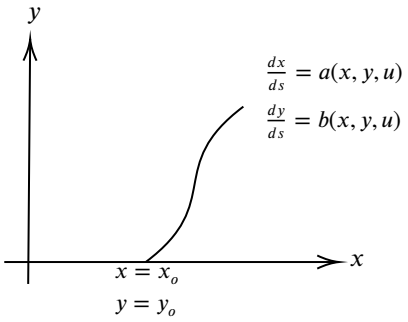

Figure 3: Characteristic curve given by–$dy/dx = b/a$.

Next, since u is constant along characteristics it follows that $u(x,t) = u(x_o, 0)$, or more precisely $u(x(t), t) = u(x_o, 0)$ (see Fig. 3). Thus, since $x(0) = -\ln(C-1) = x_o \iff C = 1+e^{-x_o}$. Thus, $x(t) = -\ln(1+e^{-x_o}-e^t)$. Finally, solving for x_o gives $x_o = -\ln(e^t + e^{-x} - 1)$. Hence,

$$u(x,t) = \phi(x_o) = \phi(-\ln(e^t + e^{-x} - 1)). \tag{3.9}$$

3.1 Method of characteristics for quasi-linear PDEs (special case)

Consider the special form of a quasi-linear PDE:

$$\frac{\partial \rho}{\partial t} + c(\rho)\frac{\partial \rho}{\partial x} = 0 \tag{3.10}$$

The equations for the characteristics for (3.10) reduce to

$$\frac{dx}{dt} = c(\rho), \tag{3.11a}$$

$$\frac{d\rho}{dt} = 0. \tag{3.11b}$$

The characteristic velocity is no longer constant, but instead depends on the density ρ. It is known as the **density wave velocity**. From (3.11a), it follows that the density ρ remains constant along each as yet undetermined characteristic. The velocity of each characteristic, $c(\rho)$, will be constant, since ρ is constant. Each characteristic is thus a straight line. However, different characteristics will move at different constant velocities because they may start with different densities. The characteristics, though each is straight, are not parallel to one another.

Consider the characteristic that is initially at the position $x = x_o$. Along the curve $dx/dt = c(\rho)$, it follows $d\rho/dt = 0$, or ρ is constant. Initially, ρ equals the value at $x = x_o$ (i.e., at $t = 0$). Thus, along this one characteristic

$$\rho(x,t) = \rho(x_o, 0) = f(x_o). \tag{3.12}$$

which is a known constant. The local wave velocity that determines the characteristic is a constant, $dx/dt = c(f(x_o))$. Consequently, this characteristic is a straight line,

$$x(t) = c(f(x_o))t + x_o, \tag{3.13}$$

since $x = x_o$ at $t = 0$. Different values of x_o yield different straight-line characteristics. Along each characteristic, the density is a constant. To determine the density at some later time, the characteristic with parameter x_o that goes through that space-time point must be obtained from (3.13).

3.2 Traffic flow

As an approximation it is possible to model a conjested one-directional highway by a quasi-linear PDE. We introduce the **traffic density** $\rho(x,t)$, the number of cars per mile at time t located at position x. An easily observed and measured quantity is the **traffic flow** $q(x,t)$, the number of cars per hour passing a fixed place x (at time t).

Conservation of cars: Consider an arbitrary section of roadway, between $x = a$ and $x = b$. If there are neither entrances nor exits on this segment of the road, then the number of cars between $x = a$ and $x = b$ ($N = \int_a^b \rho(x,t)\,dx$) might still change in time. The rate of change of the number of cars dN/dt, equals the

number per unit time entering $x = a$ (the traffic flow $q(a,t)$ there) minus the number of cars per unit time leaving at $x = b$ (the traffic flow $q(b,t)$ there):

$$\frac{d}{dt}\int_a^b \rho(x,t) = q(a,t) - q(b,t), \tag{3.14}$$

where (3.14) is the integral form of the conservation of cars. Now let's note that the boundary conditions may be represented as an integral over the region, i.e.,

$$q(a,t) - q(bt) = -\int_a^b \frac{\partial q}{\partial x}\, dx. \tag{3.15}$$

Thus, we get a scalar conservation law for our traffic flow problem:

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{3.16}$$

Car velocity: The number of cars per hour passing a place equals the density of cars times the velocity of cars. By introducing $v(x,t)$ as the **car velocity**, we have

$$q = \rho v \tag{3.17}$$

Let's make a simplifying but physically realistic assumption: car velocity depends only on the density, $v = v(\rho)$, with cars slowing down as the traffic density increases–i.e., $\frac{dv}{d\rho} \leq 0$. In the case that velocity depends only on density our conservation law (3.16) becomes:

$$\frac{\partial \rho}{\partial t} + c(\rho)\frac{\partial \rho}{\partial x} = 0, \tag{3.18}$$

where $c(\rho) = q'(\rho)$, a quasi-linear PDE with $Q \equiv 0$.

3.3 Inviscid Burgers' equation–shock waves

For quasi-linear PDEs it is quite usual for characteristics to intersect. The resolution of this scenario will require the introduction of moving discontinuities called *shock waves*. In order to make the mathematical presentation relatively simple, we restrict attention to a quasi-linear PDE known as the inviscid Burger's equation:

$$\begin{cases} u_t + uu_x = 0, \quad -\infty < x < \infty, \quad t > 0, \\ u(x,0) = f(x). \end{cases} \tag{3.19}$$

Thus, the ODEs which define the characteristics are given by:

$$\frac{du}{dt} = 0, \tag{3.20a}$$

$$\frac{dx}{dt} = u. \tag{3.20b}$$

Integrating (3.20a) tells us that u is constant along the characteristic. Integration of (3.20b) shows that the characteristics are straight lines. Hence,

$$u = c, \qquad x = ut + x_o,$$

where x_o is the point on the x-axis of the xt-plane from which the characteristic curve is drawn. Recall that u is known along the x-axis from the initial condition. Moreover, u does not change as we move along the characteristic curve emanating from each point x_o along the x-axis.

Hence along each characteristic we have

$$u = f(x_o), \tag{3.21}$$

and moreover,

$$x - ut = x_o. \tag{3.22}$$

Thus, the solution is given by

$$u(x,t) = f(x - f(x_o)t). \tag{3.23}$$

This is an implicit relation that determines the solution to the inviscid Burgers' equation provided characteristic curves do not intersect. If the characteristic curves do intersect, then a classical solution to the PDE does not exist and leads to the formation of a shock wave. Whether characteristics can intersect or not depends on the initial data. In fact, the *breaking time*–time at which the shock occurs–is given by

$$t_b := \frac{-1}{\inf_x(f'(x))}. \tag{3.24}$$

4 Heat equation

4.1 Derivation and conservation of heat energy

We begin our study by formulating the equations of heat flow describing the transfer of thermal energy. Heat energy is caused by the agitation of molecular matter. Two basic processes take place in order for thermal energy to move: conduction and convection. **Conduction** results from the collisions of neighboring molecules in which the kinetic energy of vibration of one molecule is transferred to its nearest neighbor. Thermal energy is thus spread by conduction even if the molecules themselves do not move their location appreciably. Moreover, if a vibrating molecule moves from one region to another, it takes its thermal energy with it. This type of movement of thermal energy is called **convection**. In order to begin with relatively simple problems, we will study heat flow only in the cases in which the conduction of heat energy is much more significant than its convection. We will thus think of heat flow primarily in the case of solids, although heat transfer in fluids (liquids and gases) is also primarily by conduction if the fluid velocity is sufficiently small.

Thermal energy density. We begin by considering a rod of constant cross-sectional area A oriented in the x-direction (from $x = 0$ to $x = L$) as illustrated in Fig. 4. We temporarily introduce the amount of thermal energy per unit volume as an unknown variable and call it the **thermal energy density**:

$$e(x,t) := \text{thermal energy density}. \tag{4.1}$$

We assume that all thermal quantities are constant across a section; the rod is one-dimensional. The simplest way this may be accomplished is to insulate perfectly lateral surface area of the rod. Then no thermal energy can pass through the lateral surface. The dependence on x and t corresponds to a situation in which the rod is not uniformly heated; the thermal energy density varies from one cross section to another.

Heat energy. We consider a thin slice of the rod contained between x and $x+\Delta x$ as shown in Fig. 4. If the thermal energy density is constant throughout the volume, then the total energy in the slice is the product of the thermal energy density and the volume. In general, the energy density is not constant. However, if Δx is exceedingly small, then $e(x,t)$ may be approximated as a constant throughout the volume so that

$$\text{heat energy} = e(x,t)A\Delta x \tag{4.2}$$

Conservation of heat energy. The heat energy between x and $x+\Delta x$ changes in time due only to heat energy flowing across the edges (x and $x+\Delta x$) and heat energy generated inside (due to positive or negative sources of heat energy). No heat energy changes are due to flow across the lateral surface, since we have assumed that the lateral surface is insulated. The fundamental heat flow process is described by the "word equation": **rate of change of heat energy in time = heat energy flowing across boundaries per unit time + heat energy generated inside per unit time**. This is called **conservation of heat energy**. For the small slice, the rate of change of heat energy is

$$\frac{\partial}{\partial t}\big(e(x,t)A\Delta x\big) \tag{4.3}$$

where the partial derivative $\frac{\partial}{\partial t}$ is used since x is being held fixed.

Heat flux. Thermal energy flows to the right or left in a one-dimensional rod. We introduce the **heat flux**:

$$\phi(x,t) = \text{amount of thermal energy } \textit{per unit time} \text{ flowing to the right } \textit{per unit surface area} \tag{4.4}$$

If $\phi(x,t) < 0$, it means that heat energy is flowing to the left. Heat energy flowing per unit time across the boundaries of the slice is $\phi(x,t)A - \phi(x+\Delta x,t)A$, since the heat flux is the flow per unit surface area and it must be multiplied by the surface area. If $\phi(x,t) > 0$ and $\phi(x+\Delta x,t) > 0$, as illustrated in Fig. 4, then the heat energy flowing per unit time at x contributes to an increase of the heat energy in the slice, whereas the heat flow at $x+\Delta x$ decreases the heat energy.

Heat sources. We also allow for internal sources of thermal energy:

$$Q(x,t) = \text{heat energy } \textit{per unit volume} \text{ generated } \textit{per unit time}\,, \tag{4.5}$$

perhaps due to chemical reactions or electrical heating. Moreover, $Q(x,t)$ is approximately constant in space for a thin slice, and thus the total thermal energy generated per unit time in the thin slice is approximately $Q(x,t)A\Delta x$.

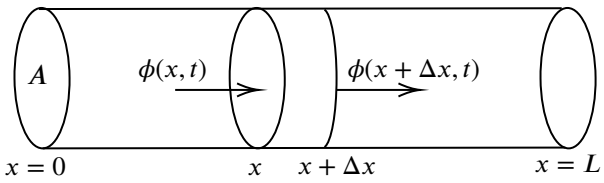

Figure 4: One-dimensional rod with heat energy flowing into and out of a thin slice.

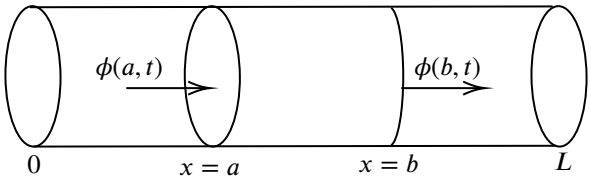

Figure 5: Heat energy flowing into and out of a finite segment of a rod

Conservation of heat energy (thin slice) The rate of change of heat energy is due to thermal energy flowing across the boundaries and internal sources:

$$\frac{\partial}{\partial t}\big(e(x,t)A\Delta x\big) \approx \phi(x,t)A - \phi(x+\Delta x,t)A + Q(x,t)A\Delta x\,. \tag{4.6}$$

Equation (4.6) is not precise because various quantities were assumed approximately constant for the small cross-sectional slice. We claim that (4.6) becomes increasingly accurate as $\Delta x \to 0$. Before giving a careful (and mathematically rigorous) derivation, we will just attempt to explain the basic ideas of the limit process, $\Delta x \to 0$. In the limit as $\Delta x \to 0$, (4.6) gives no interesting information, namely, $0 = 0$. However, if we first divide by Δx and then take the limit $\Delta x \to 0$, we obtain

$$\frac{\partial e}{\partial t} = \lim_{\Delta x \to 0} \frac{\phi(x,t) - \phi(x+\Delta x,t)}{\Delta x} + Q(x,t)\,, \tag{4.7}$$

where the constant cross-sectional area has been cancelled. We claim that this result is exact (with no small errors), and hence we replace the $\approx$ in (4.6) by $=$ in (4.7). In this limiting process, $\Delta x \to 0$, t is being held fixed. Consequently, from the definition of a partial derivative,

$$\frac{\partial e}{\partial t} = -\frac{\partial \phi}{\partial x} + Q\,. \tag{4.8}$$

Conservation of heat energy (exact). An alternative derivation of conservation of heat energy has the advantage of not being restricted to small slices. The resulting approximate calculation of the limiting process ($\Delta x \to 0$) is avoided. We consider any *finite* segment (from $x = a$ to $x = b$) of the original one-dimensional rod (see Fig. 5). We will investigate the conservation of heat energy in this region. The total heat energy is $\int_a^b e(x,t)A\,\mathrm{d}x$, the sum of the contributions of the infinitesimal slices. Again it changes only due to heat energy flowing through the side edges ($x = a$ and $x = b$) and heat energy generated inside the region, and thus (after canceling the constant A)

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b e(x,t)\,\mathrm{d}x = \phi(a,t) - \phi(b,t) + \int_a^b Q(x,t)\,\mathrm{d}x\,. \tag{4.9}$$

Technically, an ordinary derivative $\mathrm{d}/\mathrm{d}t$ appears in (4.9), since $\int_a^b e\,\mathrm{d}x$ depends only on t, not also on x. However,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b e\,\mathrm{d}x = \int_a^b \frac{\partial e}{\partial t}\,\mathrm{d}x \tag{4.10}$$

if a and b are constants (and if e is continuous). This holds since inside the integral the ordinary derivative now is taken keeping x fixed, and hence it must be replaced by a partial derivative. Every term in (4.9) is now an ordinary integral if we notice that

$$\phi(a,t) - \phi(b,t) = -\int_a^b \frac{\partial \phi}{\partial x}\,\mathrm{d}x \tag{4.11}$$

(this being valid if ϕ is continuously differentiable). Consequently,

$$\int_a^b \left(\frac{\partial e}{\partial t} + \frac{\partial \phi}{\partial x} - Q \right) \mathrm{d}x = 0 \,. \tag{4.12}$$

This integral must be zero *for arbitrary a and b*; the area under the curve must be zero for arbitrary limits. This is possible only if the integrand is identically zero. Thus, we rederive (4.8) as

$$\frac{\partial e}{\partial t} = -\frac{\partial \phi}{\partial x} + Q \,. \tag{4.13}$$

Equation (4.9), the **integral conservation law**, is more fundamental than the differential form (4.13). Equation (4.13) is valid in the usual case in which the physical variables are continuous.

A further explanation of the minus sign preceding $\partial\phi/\partial x$ is in order. For example, if $\partial\phi/\partial x > 0$ for $a \le x \le b$, then the heat flux ϕ is an increasing function of x. The heat is flowing greater to the right at $x = b$ than at $x = a$ (assuming that $b > a$). Thus (neglecting any effects of sources Q), the heat energy must decrease between $x = a$ and $x = b$, resulting in the minus sign in (4.13).

Temperature and specific heat. We usually describe materials by there temperature,

$$u(x,t) = \text{temperature} \,, \tag{4.14}$$

not their thermal energy density. Distinguishing between the concepts of temperature and thermal energy is not necessarily trivial. Only in the mid-1700's did the existence of accurate experminental apparatus enable physicists to recognize that it may take different amounts of thermal energy to raise two different materials from one temperature to another larger temperature. This necessitates the introduction of the **specific heat** (or heat capacity):

$$c = \text{heat energy supplied to a unit mass of a substance to raise its temperature one unit} \,. \tag{4.15}$$

In general, the specific heat c if a material depends on the temperature u. This significantly complicates the mathematical problem (nonlinear). Often for restricted temperature intervals, the specific heat is approximately independent of the temperature. On the other hand in order to allow for non-uniform materials we can assume that the specific heat depends on the spatial variable, i.e., $c = c(x)$.

Thermal energy. The thermal energy in a thin slice is $e(x,t)A\Delta x$. However, it is also defined as the energy it takes to raise the temperature from a reference temperature 0° to its actual temperature $u(x,t)$. Since the specific heat is independent of temperature, the heat energy per unit mass is just $c(x)u(x,t)$. We need to introduce the **mass density**:

$$\rho(x) = \text{mass per unit volume} \,, \tag{4.16}$$

allowing it to vary with x, possibly due to the rod being composed of nonuniform material. The total mass of the thin slice is $\rho A\Delta x$. The total thermal energy in any thin slice is thus $c(x)u(x,t)\rho A\Delta x$, so that

$$e(x,t)A\Delta x = c(x)u(x,t)\rho A\Delta x \,.$$

In this way we have explained the basic relationship between thermal energy and temperature:

$$e(x,t) = c(x)\rho(x)u(x,t) \tag{4.17}$$

Thus, the thermal energy per unit volume equals the thermal energy per unit mass per unit degree times the temperature times the mass density. Replacing the thermal energy density in (4.13) gives

$$c(x)\rho(x)\frac{\partial u}{\partial t} = -\frac{\partial \phi}{\partial t} + Q \,. \tag{4.18}$$

Fourier's law of heat conduction. It follows that (4.18) is one equation with two unknowns, the temperature u and the heat flux ϕ (flow per unit surface area per unit time). How and why does heat energy flow? In other words, we need an expression for the dependence of the flow of heat energy on the temperature field. First we summarize certain qualitative properties of heat flow:

1. If the temperature is constant in a region, no heat energy flows (thermal equilibrium).
2. If there are temperature differences, the heat energy flows from the hotter to the colder regions.
3. The greather the temperature differences (for the same material), the greater is the flow of heat energy.
4. The flow of heat energy will vary for different materials, even with the same temperature differences.

Fourier (1768–1830) summarized the above properties by the formula:

$$\phi = -K_o \frac{\partial u}{\partial x}\,, \tag{4.19}$$

which is known as Fourier's law of heat conduction. Here $\partial u/\partial x$ is the derivative of the temperature; it represents temperature differences (per unit length). Equation (4.19) states that the heat flux is proportional to the temperature differences (per unit length). For example, if the temperature u increases as x increases (i.e., temperature is hotter to the right), $\partial u/\partial x > 0$, then we know (property 2) that heat energy flows to the left. This explains the minus sign in (4.19).

We designate the proportionality constant K_o. It measure the ability of the material to conduct heat and is called the **thermal conductivity**. Experiments indicate that different materials conduct heat differently; K_o depends on the particular material. The larger K_o is, the greater the flow of heat energy with the same temperature differences. A material with a low K_o would be a poor conductor of heat energy (and ideally suited for home insulation). For a rod composed of different materials, K_o will be a function of x. (Experiments show that K_o can also depend on temperature u.) Throughout these notes we will assume that the thermal conductivity K_o depends only on x, $K_o(x)$. Usually, we will discuss uniform rods in which K_o is a constant.

Heat equation. If Fourier's law, (4.19), is substituted into the conservation of heat energy equation, (4.18), a PDE results:

$$c\rho \frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(K_o \frac{\partial u}{\partial x}\right) + Q \tag{4.20}$$

We usually think of Q as being given and thus the only unknown is $u(x,t)$. The thermal coefficients c, ρ, K_o all depend on the material and hence may be functions of x. In the special case of a uniform rod, in which c, ρ, K_o are all constants, the PDE (4.20) becomes

$$c\rho \frac{\partial u}{\partial t} = K_o \frac{\partial^2 u}{\partial x^2} + Q\,.$$

If, in addition, there are no sources, $Q = 0$, then after dividing by $c\rho$, the PDE becomes

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2} \tag{4.21}$$

where the constant κ,

$$\kappa = \frac{K_o}{c\rho}$$

is called the **thermal diffusivity**. Equation (4.21) is often called the **heat equation**; it corresponds to constant thermal properties and no sources. If heat energy is initially concentrated in one place, (4.21) will describe

how the heat energy spreads out, a physical process known as **diffusion**. Other physical quantities besides temperature smooth out in much the same manner, satisfying the same PDE (4.21). For this reason (4.21) is also known as the **diffusion equation**. For example the concentration $u(x,t)$ of chemicals (such as perfumes and pollutants) satisfies the diffusion equation (4.21) in certain one-dimensional settings.

Initial conditions. The PDEs describing the flow of heat energy, (4.20) or (4.21), have one time derivative. When an ordinary differential equation (ODE) has one derivative, the initial value problem consists of solving the ODE with one initial condition. Newton's law of motion for the position x of a particle yields a second-order ODE, $mx'' = F(t, x, x')$, where prime denotes differentiation with respect to the independent variable t. It involves second derivatives. The initial value problem consists of solving the ODE with two initial conditions, the initial position x and the initial velocity x'. From these pieces of information (including the knowledge of forces), by solving the ODE with the initial conditions, we can predict the future motion of a particle in the x-direction. We wish to do the same process for our PDE, that is, predict the future temperature. Since the heat equation has one time derivative, we must be given one **initial condition** (IC) (usually $t = 0$), the initial temperature. It is possible that the initial temperature is not constant, but depends on x. Thus, we must be given the initial temperature distribution,

$$u(x, 0) \equiv u_o(x).$$

Is this enough information to predict the future temperature? We know the initial temperature distribution and know that the temperature changes according to the PDEs (4.20) or (4.21). However, we need to know what happens at the two boundaries $x = 0$ and $x = L$. Without knowing this information, we cannot predict the future. Two conditions are needed corresponding to the second spatial derivatives present in (4.20) or (4.21), usually one condition at each end. We will discuss **boundary conditions** (BCs) throughout these notes.

Diffusion of a chemical pollutant. Let $u(x,t)$ be the **density** or **concentration** of the chemical per unit volume. Consider a one-dimensional region (Fig. 5) between $x = a$ and $x = b$ with a constant cross-sectional area A. The total amount of the chemical in the region is $\int_a^b u(x,t)A\,\mathrm{d}x$. We introduce the **flux** $\phi(x,t)$ of the chemical, the amount of the chemical per unit surface area flowing to the right per unit time. The rate of change with respect to time of the total amount of chemical in the region equals the amount of chemical flowing in per unit time minus the amount of chemical flowing out per unit time. Thus, after canceling the constant cross-sectional area A, we obtain the **integral conservation law** for the chemical concentration:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b u(x,t)\,\mathrm{d}x = \phi(a,t) - \phi(b,t)\,. \tag{4.22}$$

Since $\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b u(x,t)\,\mathrm{d}x = \int_a^b \frac{\partial u}{\partial t}\,\mathrm{d}x$ and $\phi(a,t) - \phi(b,t) = -\int_a^b \frac{\partial \phi}{\partial x}\,\mathrm{d}x$, it follows that $\int_a^b (\frac{\partial u}{\partial t} + \frac{\partial \phi}{\partial x})\,\mathrm{d}x = 0$. Since the integral is zero for arbitrary regions, the integrand must be zero, and in this way we derive the **differential conservation law** for the chemical concentation:

$$\frac{\partial u}{\partial t} + \frac{\partial \phi}{\partial x} = 0\,. \tag{4.23}$$

In solids, chemicals spread out from regions of high concentration to regions of low concetration. According to **Fick's law of diffusion**, the flux is proportional to $\frac{\partial u}{\partial x}$, the spatial derivative of the chemical concetration:

$$\phi = -k\frac{\partial u}{\partial x}\,. \tag{4.24}$$

If the concentration $u(x,t)$ is constant in space, there is no flow of the chemical. If the chemical concentation is increasing to the right ($\frac{\partial u}{\partial x} > 0$), then atoms of chemicals migrate to the left, and vice-versa. The proportionality constant k is called the chemical **diffusivity**, and it can be measured experimentally. When Fick's law (4.24) is used in the basic conservation law (4.23), we see that the chemical concentration satisfies the **diffusion equation**:

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}\,, \tag{4.25}$$

since we are assuming as an approximation that the diffusivity is constant. Fick's law of diffusion for chemical concetration is analogous to Fourier's law for heat diffusion.

4.2 Boundary conditions

In solving the heat equation, either (4.20) or (4.21), one **boundary condition** (BC) is needed at each end of the rod. The appropriate condition depends on the physical mechanism in effect at each end.

Prescribed temperature. In this case the temperature at the end of the rod is prescribed, for example,

$$u(0,t) = u_B(t)\,, \quad \text{(DirichletBC)} \tag{4.26}$$

where $u_B(t)$ is the temperature of a fluid bath (or resevoir) with which the rod is in contact.

Prescribed heat flux. In other situations, it is possible to prescribe the heat flow rather than the temperature,

$$-K_o(0)\frac{\partial u}{\partial x}(0,t) = \phi(t)\,, \quad \text{(Neumann BC)} \tag{4.27}$$

where $\phi(t)$ is given. Equation (4.27) cannot be integrated in x because the slope is known only at one value of x. If the boundary is "perfectly" insulated then

$$\frac{\partial u}{\partial x}(0,t) \equiv 0\,, \tag{4.28}$$

which means there is no flow of heat across the boundary.

Newton's law of cooling. When a one-dimensional rod is in contact at the boundary with a moving fluid (e.g. air), then neither prescribed temperature nor presribed heat flow may be appropriate. Experiments show, as a good approximation, the heat flow leaving the rod is proportional to the temperature difference between the bar and the prescribed external temperature, i.e., Newton's law of cooling:

$$-K_o(0)\frac{\partial u}{\partial x}(0,t) = -H[u(0,t) - u_B(t)]\,, \quad \text{(Robin BC)} \tag{4.29a}$$

$$-K_o(0)\frac{\partial u}{\partial x}(L,t) = H[u(L,t) - u_B(t)]\,, \quad \text{(Robin BC)} \tag{4.29b}$$

where the proportionality constant $H > 0$ is called the **heat transfer coefficient**.

4.3 Equilibrium temperature distribution

Let us formulate a simple, but typical, problem of heat flow. For simplicity assume constant thermal coefficients and no sources of thermal energy. Then the temperature $u(x,t)$ in a one-dimensional rod $0 \le x \le L$ satisfies

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2}\,. \tag{4.30}$$

The solution of this PDE must satisfy the initial condition

$$u(x,0) = f(x)\,, \tag{4.31}$$

and one boundary condition at each end. For example, assume that the temperature is presribed at each end

$$u(0,t) = g_1(t)\,, \tag{4.32a}$$
$$u(L,t) = g_2(t)\,. \tag{4.32b}$$

This is an example of an initial boundary-value problem (IBVP). Moving forward we will write IBVPs as follows:

$$\begin{cases} u_t = \kappa u_{xx}\,, \quad 0 < x < L\,, \quad t > 0 \\ u(x,0) = f(x) \\ u(0,t) = g_1(t)\,, \quad u(L,t) = g_2(t) \end{cases} \tag{4.33}$$

Before we begin to attack problems such as (4.33), we discuss a physically related question for ordinary differential equations. Suppose that the boundary conditions at $x = 0$ and $x = L$ were **steady** (i.e., independent of time),

$$u(0,t) \equiv A\,, \qquad u(L,t) \equiv B\,, \tag{4.34}$$

where A and B are given constants.

Definition 4.1. An **equilibrium** or **steady-state** solution to a given equation of evolution type (e.g., a PDE with a time derivative) is a solution that does not depend on time.

Thus, in this case $u(x,t) = u(x)$ and so $\partial_t u(x) = 0$. Thus the heat equation in one spatial dimension reduces to a second-order ODE, namely,

$$u'' = 0\,, \tag{4.35}$$

and thus the steady-state solution to the heat equation in one spatial dimension is given by

$$u(x) = c_1 x + c_2\,. \tag{4.36}$$

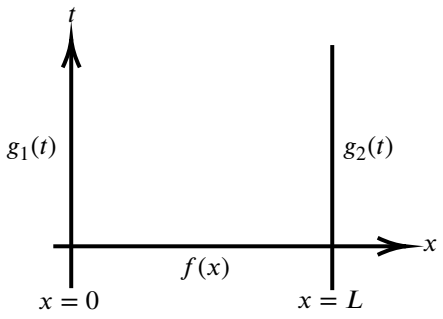

Figure 6: Depiction of the initial boundary-value problem defined by (4.33).

Applying the BCs (4.34) gives

$$u(x) = \left(\frac{B-A}{L}\right)x + A\,. \tag{4.37}$$

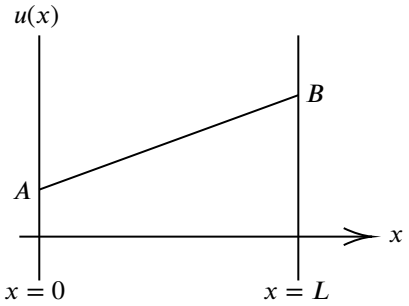

Figure 7: Equilibrium temperature distribution.

Approach to equilibrium. For the time-dependent problem (4.30) with IC (4.31) and with steady BCs (4.34), we expect the temperature distribution $u(x,t)$ to change in time; it will not remain equal to its initial distribution $f(x)$. If we wait a very, very long time, we would expect that the influence of the two ends should dominate. The initial conditions are usually forgotten. Thus, eventually the temperature distribution is expected to approach the equilibrium temperature distribution, since the BCs are independent of time:

$$\lim_{t\to\infty} u(x,t) = u(x) = \left(\frac{B-A}{L}\right)x + A\,.$$

Next, we study the steady-state solution for the IBVP:

$$\begin{cases} u_t = \kappa u_{xx} \\ u(x,0) = f(x) \\ u_x(0,t) = 0\,, \quad u_x(L,t) = 0 \end{cases} \tag{4.38}$$

Physically, this means that there is zero thermal flux at the endpoints $x = 0$ and $x = L$. Further, solving the ODE (since $u_t = 0$) and applying the BCs gives

$$u(x) \equiv c_2\,, \tag{4.39}$$

for any constant c_2. However, it does not make sense that the solution approaches an arbitrary constant; we should know what constant the solution approaches in the limit $t \to \infty$. The particular constant is obtained by considering the IC in the time-dependent problem. Since both ends are insulated (no heat flux) and there are no sources, the total thermal energy is constant (by conservation of energy). Thus, integrating from $x = 0$ to $x = L$ gives

$$\frac{d}{dt}\int_0^L u\,dx = \int_0^L \frac{\partial u}{\partial t}\,dx = \kappa\int_0^L \frac{\partial^2 u}{\partial x^2}\,dx = \kappa\frac{\partial u}{\partial x}(L,t) - \kappa\frac{\partial u}{\partial x}(0,t) = 0\,.$$

Thus,

$$\int_0^L u\,dx = \text{constant} \tag{4.40}$$

Thus, the initial amount of thermal energy must equal the final ($t \to \infty$) amount of thermal energy. Thus,

$$\int_0^L f(x)\,dx = \int_0^L c_2\,dx \iff c_2 = \frac{1}{L}\int_0^L f(x)\,dx\,,$$

which means that the steady-state solution is the average of the initial temperature distribution.

4.4 Heat equation in higher dimensions.

Consider an arbitrary sub-region R of some three-dimensional object. Again by conservation of energy we get the balance equation: *Rate of change of heat energy = heat energy flowing across the boundary per unit time + heat energy generated insider per unit time*.

$$\iiint_R c\rho u \, \mathrm{d}V \,, \tag{4.41}$$

where $\mathrm{d}V$ is the volume element.

Heat flux vector and normal vectors. We need an expression for the flow of heat energy. In a three-dimensional problem, heat flows in some direction, and hence the heat flux is a vector $\boldsymbol{\phi}$. The magnitude of $\boldsymbol{\phi}$ is the amount of heat energy flowing per unit time per unit surface area. Importantly, by conservation of heat energy we are only concerned with the amount of heat flowing across the boundaries per unit time. In fact, it is only the normal component of the heat flux vector that is important. (For example, if the heat flux vector is parallel to the surface to heat if flowing across the boundary.) Thus, we let $\hat{\mathbf{n}}$ be the unit outward normal vector. Consequently, conservation of heat energy for an arbitrary three-dimensional region R beomes

$$\frac{d}{dt}\iiint_R c\rho u \, \mathrm{d}V = -\oiint_{\partial R} \boldsymbol{\phi}\cdot\hat{\mathbf{n}} \, \mathrm{d}S + \iiint_R Q \, \mathrm{d}V \,., \tag{4.42}$$

where ∂R denotes the "boundary" of the region R and $\oiint \mathrm{d}S$ is a "closed" surface integral.

Divergence theorem. The divergence theorem deal with a vector $\mathbf{v} = [v_1, v_2, v_2]^{\mathrm{T}}$ and its divergence defined as

$$\nabla\cdot\mathbf{v} := \partial_x v_1 + \partial_y v_2 + \partial_z v_3 \,. \tag{4.43}$$

Thus, the divergence of a vector is a scalar quantity. The divergence theorem states that the volume integral of the divergence of any continuously differentiable vector field $\mathbf{v}$ is equal to the closed surface integral of the outward normal component of $\mathbf{v}$. Thus,

$$\iiint_R \nabla\cdot\mathbf{v} = \oiint_{\partial R} \mathbf{v}\cdot\hat{\mathbf{n}} \, \mathrm{d}S \,. \tag{4.44}$$

Thus, applying the Divergence theorem to (4.42) gives

$$\frac{d}{dt}\iiint_R c\rho u \, \mathrm{d}V = -\iiint_R \nabla\cdot\boldsymbol{\phi} \, \mathrm{d}V + \iiint_R Q \, \mathrm{d}V \,. \tag{4.45}$$

This gives the following conservation law:

$$c\rho\frac{\partial u}{\partial t} = -\nabla\cdot\boldsymbol{\phi} + Q \tag{4.46}$$

Next, as in the one-dimensional case we use Fourier's law of heat conduction, i.e., the heat flux is proportional to the temperature gradient:

$$\boldsymbol{\phi} = -K_o\nabla u \,, \tag{4.47}$$

where $\nabla u := (\partial_x u, \partial_y u, \partial_z u)$ is the gradient. Substitution into (4.46) gives

$$c\rho\frac{\partial u}{\partial t} = K_o\nabla\cdot\nabla u + Q \,. \tag{4.48}$$

Recalling that $\nabla \cdot \nabla u = \Delta u$, where $\Delta u := \partial_x^2 u + \partial_y^2 u + \partial_z^2 u$ is the Laplacian one gets the heat equation in three spatial dimensions

$$c\rho \frac{\partial u}{\partial t} = K_o \Delta u + Q\,. \tag{4.49}$$

Finally, assuming that c and ρ are constant, and there are no sources gives

$$\frac{\partial u}{\partial t} = \kappa \Delta u\,. \tag{4.50}$$

Important objects:

- Divergence – $\nabla \cdot \mathbf{v} = \partial_x v_1 + \partial_y v_2 + \partial_z v_3$
- Divergence theorem – $\iiint_R \nabla \cdot \mathbf{v}\,\mathrm{d}V = \oiint_{\partial R} \mathbf{v} \cdot \hat{\mathbf{n}}\,\mathrm{d}S$
- Gradient – $\nabla u = (\partial_x u, \partial_y u, \partial_z u)$
- Laplacian – $\nabla \cdot \nabla u = \Delta u = \partial_x^2 u + \partial_y^2 u + \partial_z^2 u$

5 Laplace's equation

If the boundary conditions and any sources of thermal energy are independent of time, it is possible that there exist steady-state solutions to the heat equation satisfying the given steady-state boundary condition. Note that in higher dimensions the equilibrium temperature satisfies a PDE that is independent of time:

$$\Delta u = -\frac{Q}{K_o} \tag{5.1}$$

which is known as **Poisson's equation**.

If, in addition, there are no sources ($Q \equiv 0$), then

$$\Delta u = 0, \tag{5.2}$$

the Laplacian of the temperature distribution is zero. Equation (5.2) is known as **Laplace's equation**. It is also known as the potential equation, since the gravitational and electrostatic potentials satisfy (5.2) if there are no sources. Since there is no time-dependence, Laplace's equation is a boundary-value problem (as opposed to an initial boundary-value problem). Solutions of Laplace's equation are called **Harmonic functions**.

- In 1D– $u_{xx} = 0$
- In 2D– $u_{xx} + u_{yy} = 0$
- In 3D– $u_{xx} + u_{yy} + u_{zz} = 0$

5.1 Applications

- **Steady-state:** Steady-state solutions of heat/wave flow problems.
- **Electrostatics:** Consider Maxwell's equations $\nabla \times \mathbf{E} = \mathbf{0}$ and $\nabla \cdot \mathbf{E} = 4\pi\rho$. From the first equation $\mathbf{E} = -\nabla\phi$ for some electrostatic potential ϕ. From the second equation $\Delta\phi = -4\pi\rho$ which is Poisson's equation.
- **Steady fluid flow:** Assume irrotational so that $\nabla \times \mathbf{v} = \mathbf{0}$ where $\mathbf{v}(x, y, z)$ is the fluid velocity. Assume incompressible and no sources/sinks so that $\nabla \cdot \mathbf{v} = 0$. Hence $-\nabla\phi = \mathbf{v}$ implies $\Delta\phi = 0$.
- **Analytic function theory:** Let $z = x + iy$ and $f(z) = u(z) + iv(z) = u(x + iy) + iv(x + iy)$, where u and v are real-valued. An analytic function is one that can be written as a power series in z, i.e., $f(z) = \sum_{n=0}^{\infty} c_n z^n$ or $u(x+iy)+iv(x, y) = \sum_{n=0}^{\infty} c_n(x+iy)^n$. Formal differentiation of the series yields the **Cauchy-Riemann equations**:

  $$u_x = v_y \qquad u_y = -v_x.$$

  Differentiate to get $u_{xx} = -u_{yy}$ and $v_{xx} = -v_{yy}$. Thus the real and imaginary parts of an analytic function are harmonic.

5.2 Coordinate systems

The geometry of a boundary-value problem may be complicated. In certain cases using well-known coordinate systems may be useful. Thus, we define the Laplacian in Cartesian, cylindrical, and spherical coordinates.

Cartesian coordinates:

$$\begin{aligned} x &= x \\ y &= y \\ z &= z \end{aligned}$$

$$\Delta u = \partial_x^2 u + \partial_y^2 u + \partial_z^2 u \tag{5.3}$$

Cylindrical coordinates:

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}$$

$$\Delta u = \frac{1}{r}\partial_r\big(r\partial_r u\big) + \frac{1}{r^2}\big(\partial_\theta^2 u\big) + \partial_z^2 u \tag{5.4}$$

Spherical cooridnates:

$$x = \rho \sin\varphi \cos\theta$$
$$y = \rho \sin\varphi \sin\theta$$
$$z = \rho \cos\varphi$$

$$\Delta u = \frac{1}{\rho^2}\partial_\rho\big(\rho^2 \partial_\rho u\big) + \frac{1}{\rho^2 \sin\varphi}\partial_\varphi\big(\sin\varphi \partial_\varphi u\big) + \frac{1}{\rho^2 \sin^2\varphi}\partial_\theta^2 u \tag{5.5}$$

Note that Laplace's equation is invariant under all rigid motions, i.e., (i) translations, and (ii) rotations. Let's consider this idea in 2D. A **translation** in the plane is a transformation

$$x' = x + a \qquad y' = y + b.$$

It is easy to check that $u_{x'x'} + u_{y'y'} = 0 \iff u_{xx} + u_{yy} = 0$. Also, a **rotation** in the plane through an angle α is given by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Again, it is easy to check that $u_{x'x'} + u_{y'y'} = 0 \iff u_{xx} + u_{yy} = 0$.

The above invariants of Laplace's equations leads us to consider solutions of Laplace's equation that are rotationally symmetric, i.e., $u_\theta = 0$. Thus looking at Laplace's equation in polar coordinates we have

$$\begin{aligned} \frac{1}{r}\frac{\partial}{\partial r}\big(r u_r\big) = 0 &\iff \frac{\partial}{\partial r}\big(r u_r\big) = 0 \\ &\iff r u_r = c_1 \\ &\iff u_r = \frac{c_1}{r} \\ &\iff u(r) = c_1 \ln(r) + c_2 \end{aligned}$$

The radial solution $u(r) = \ln(r)$ is known as the fundamental solution of Laplace's equation in 2D. It is special and will be discussed more later in these notes.

6 Separation of variables

In this section we study a technique useful for solving IBVPs and BVPs corresponding to linear homogeneous (no sources) PDEs such as the heat and Laplace equations.

6.1 Linearity

As in the study of ODEs, the concept of linearity will be very important to us.

Definition 6.1. An operator L is **linear** if it satisfies

$$L[c_1 u_1 + c_2 u_2] = c_1 L[u_1] + c_2 L[u_2]\,. \tag{6.1}$$

Note that we can define the *heat operator*, i.e.,

$$L := \partial_t - \kappa\Delta\,, \tag{6.2}$$

which is just the heat equation in operator notation. It follows easily that

$$\begin{aligned}
L[c_1u_1 + c_2u_2] &= \partial_t(c_1u_1 + c_2u_2) - \kappa\Delta(c_1u_1 + c_2u_2) \\
&= \partial_t(c_1u_1 + c_2u_2) - \kappa(\partial_x^2 + \partial_y^2 + \partial_z^2)(c_1u_1 + c_2u_2) \\
&= c_1\partial_t u_1 + c_2\partial_t u_2 - \kappa(c_1\partial_x^2 u_1 + c_2\partial_x^2 u_2 + c_1\partial_y^2 u_1 + c_2\partial_y^2 u_2 + c_1\partial_z^2 u_1 + c_2\partial_z^2 u_2) \\
&= c_1\big(\partial_t - \kappa(\partial_x^2 + \partial_y^2 + \partial_z^2)\big)u_1 + c_2\big(\partial_t - \kappa(\partial_x^2 + \partial_y^2 + \partial_z^2)\big)u_2 \\
&= c_1(\partial_t u_1 - \kappa\Delta u_1) + c_2(\partial_t u_2 - \kappa\Delta u_2) \\
&= c_1L[u_1] + c_2L[u_2]
\end{aligned}$$

Thus, the heat equation (with constant thermal properties) is linear. Consequently, from the above calculation it follows that Laplace's equation is also linear.

Suppose that L is a linear operator, then

- $L[u] = 0$ – linear homogeneous
- $L[u] = f$ – linear non-homogeneous

Principle of superposition: If u_1 and u_2 satisfy a *linear homogeneous* equation, then any linear combination of them $c_1u_1 + c_2u_2$ also satisfies the same linear homogeneous equation.

Note that the concepts of linearity and homogeneity also apply to boundary conditions.

6.2 Heat equation with Dirichlet BCs

Now that we have been introduced to the subject let's begin studying our first technique for solving IBVPs, namely, "separation of variables". The technique is used in particular to solve IBVPs corresponding to linear homogeneous PDEs with linear homogeneous BCs.

We begin by solving the following IBVP for the temperature of a rod with constant thermal properties and a prescribed temperature at the endpoints:

$$\begin{cases} u_t = \kappa u_{xx}\,, \quad 0 < x < L\,, \quad t > 0 \quad \text{[PDE]} \\ u(x,0) = f(x) \quad \text{[IC]} \\ u(0,t) = 0\,,\ u(L,t) = 0 \quad \text{[BCs]} \end{cases} \tag{6.3}$$

First, we make the ansatz:

$$u(x,t) = \psi(x)g(t). \tag{6.4}$$

Next, substitute into the PDE in (6.3).

$$\psi\frac{dg}{dt} = \kappa\psi''g \iff \frac{1}{kg}\frac{dg}{dt} = \frac{\psi''}{\psi} = -\lambda$$

where λ is the separation constant to be determined.

Remark 6.2. The reason $\frac{1}{\kappa g}\frac{dg}{dt} = \frac{\psi''}{\psi} = \text{const.}$ can be seen as follows: Let $h(t) = \frac{1}{\kappa g}\frac{dg}{dt}$ (as it is only a function of t). Then taking a derivative with respect to t gives $h'(t) = 0$ (as ψ''/ψ is only a function of x). Thus since its derivative is zero $\frac{1}{\kappa g}\frac{dg}{dt}$ must equal a constant.

Next, consider BCs:

$$u(0,t) = \psi(0)g(t) = 0 \iff \psi(0) = 0 \quad \text{and} \quad u(L,t) = \psi(L)g(t) = 0 \iff \psi(L) = 0.$$

Notice, if $g(t)$ were zero then the solution would be $u(x,t) \equiv 0$, i.e., the trivial solution. This would only makes sense if the initial condition $f(x)$ was also zero which we did not assume. Thus, we have two ODEs:

(i) $\frac{dg}{dt} = -\kappa\lambda g$

(ii) $\psi'' + \lambda\psi = 0\,, \quad \psi(0) = 0\,, \quad \psi(L) = 0$

Note that (i) is first-order linear (also separable) and can be solved easily to get

$$g(t) = ce^{-\kappa\lambda t}\,, \tag{6.5}$$

where c is an arbitrary constant of integration. Next, let's solve (ii) which is second-order linear homogeneous with constant coefficients. The difficulty is that (ii) has BCs that must be satisfied. Recall that the characteristic equation $r^2 + \lambda = 0$ implies the following:

- $\lambda > 0 \implies$ two purely imaginary complex conjugate roots $r = \pm i\sqrt{\lambda}$.
- $\lambda = 0 \implies$ real root $r = 0$ of multiplicity two.
- $\lambda < 0 \implies$ two real distinct roots $r = \pm\sqrt{|\lambda|}$.

As we know, the general solution to the ODE in (ii) will take a different form in each of the above three cases. Further, we will not need to consider $\lambda \in \mathbb{C}$ (λ complex) as will be shown later in the course. Let's consider each case individually.

($\lambda > 0$) In this case since the roots are purely imaginary and complex conjugate it follows that

$$\psi(x) = c_1 \sin(\sqrt{\lambda}x) + c_2 \cos(\sqrt{\lambda}x)$$

Next, apply the BCs: $\psi(0) = c_1 \sin(0) + c_2 \cos(0) = c_2 = 0$. Moreover, $\psi(L) = c_1 \sin(\sqrt{\lambda}L) = 0 \iff \sqrt{\lambda}L = n\pi \iff \lambda = (n\pi/L)^2$. Thus,

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad n = 1, 2, \ldots \quad \text{(eigenvalues)} \tag{6.6}$$

$$\psi_n(x) = \sin(n\pi x/L)\,, \quad n = 1, 2, \ldots \quad \text{(eigenfunctions)} \tag{6.7}$$

(see Fig. 8) and c_1 is an arbitrary constant we take to be one.

($\lambda = 0$) In this case since the root has algebraic multiplicity two it follows that

$$\psi(x) = c_1 x + c_2$$

Next, apply the BCs: $\psi(0) = c_2 = 0$. Moreover, $\psi(L) = c_1 L = 0 \iff c_1 = 0$. Thus, $\psi(x) \equiv 0$ which is the trivial solution. Thus, $\lambda = 0$ is not an eigenvalue.

($\lambda < 0$) In this case since the roots are real and distinct it follows that

$$\psi(x) = c_1 e^{\sqrt{|\lambda|}x} + c_2 e^{\sqrt{|\lambda|}x}$$

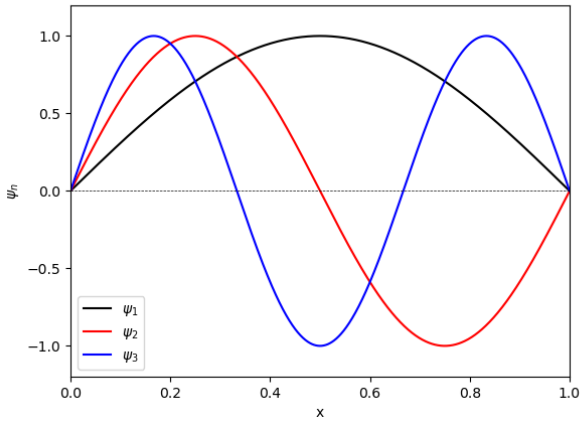

Figure 8: Eigenfunctions $\psi_n(x)$ corresponding to the IBVP (6.3).

Next, apply the BCs: $\psi(0) = c_1 + c_2 = 0 \iff c_2 = -c_1$. Thus, $\psi(x) = c_1\left(e^{\sqrt{|\lambda|}x} - e^{-\sqrt{|\lambda|}x}\right) = 2c_1 \sinh(\sqrt{|\lambda|}x)$. Thus, $\psi(L) = 2c_1 \sinh(\sqrt{|\lambda|}L) = 0$. Since $c_1 = 0$ gives the trivial solution we need $\sinh(\sqrt{|\lambda|}L) = 0$ which has no solutions for $\lambda < 0$. Thus there are no eigenvalues for $\lambda < 0$.

Putting everything together we have found for $n = 1, 2, \ldots$

$$u(x,t) = c \sin(n\pi x/L) e^{-\kappa(n\pi/L)^2 t} \tag{6.8}$$

solves the heat equation with homogeneous Dirichlet BCs. What about the IC $u(x,0) = f(x)$? Suppose that $u(x,0) = 2\sin(5\pi x/L)$. Then using (6.8) gives

$$u(x,0) = c \sin(n\pi x/L) = 2\sin(5\pi x/L) \implies c = 2 \quad \text{and} \quad n = 5.$$

Thus, the solution is $u(x,t) = 2\sin(5\pi x/L)e^{-\kappa(5\pi/L)^2 t}$. Moreover, since c is an arbitrary constant by the superposition principle it follows that a linear combination of solutions is also a solution. Thus,

$$u(x,t) = \sum_{n=1}^{N} c_n \sin(n\pi x/L) e^{-\kappa(n\pi/L)^2 t}\,. \tag{6.9}$$

also solves the heat equation with homogeneous Dirichlet BCs. What if the initial condition $u(x,0) = f(x)$ is not a sum of sine functions? In this case we consider the infinite linear combination, namely,

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin(n\pi x/L) e^{-\kappa(n\pi/L)^2 t}\,, \tag{6.10}$$

with

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} c_n \sin(n\pi x/L)\,. \tag{6.11}$$

This leads to the concept of Fourier series discussed in the next section.

6.3 Fourier series

Using separation of variables we found that the initial condition $u(x,0) = f(x)$ could be satisfied only if $f(x) =$ infinite linear combination of eigenfunctions. This leads to the notion of Fourier series. A Fourier series is given by

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L) + \sum_{n=1}^{\infty} b_n \cos(n\pi x/L)\,. \tag{6.12}$$

Remark 6.3. Each function in the Fourier series is $2L$-periodic. Thus, the Fourier series approximation of $f(x)$ on the interval $-L \le x \le L$ is $2L$-periodic. Thus, we need the periodic extension of $f(x)$.

Moreover, by using certain orthogonality relations we can find the constant a_n, b_n:

$$a_0 = \frac{1}{2L}\int_{-L}^{L} f(x)dx \tag{6.13}$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(n\pi x/L)dx \tag{6.14}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(n\pi x/L)dx \tag{6.15}$$

These relations can be found by integrating $\int_{-L}^{L}\cos(n\pi x/L)\cos(m\pi x/L)dx$, $\int_{-L}^{L}\cos(n\pi x/L)\sin(m\pi x/L)dx$, and $\int_{-L}^{L}\sin(n\pi x/L)\sin(m\pi x/L)dx$ for $n \ne m$ and $n = m$.

Fourier sine series. Suppose that $f(-x) = -f(x)$ (odd) on $-L \le x \le L$. Then it follows easily that $a_0 = a_n = 0$ for all n. In this case we get the Fourier sine series, namely,

$$F_s(x) = \sum_{n=1}^{\infty} b_n \sin(n\pi x/L)\,. \tag{6.16}$$

Thus, for $f(x)$ defined on $0 \le x \le L$ we first make an odd extension of $f(x)$ to the interval $-L \le x \le L$. Then we make a $2L$-periodic extension. This gives the Fourier sine series (6.16) for the initial datum.

Fourier cosine series. Suppose that $f(-x) = f(x)$ (even) on $-L \le x \le L$. Then it follows easily that $b_n = 0$ for all n. In this case we get the Fourier cosine series, namely,

$$F_c(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L)\,. \tag{6.17}$$

Thus, for $f(x)$ defined on $0 \le x \le L$ we first make an even extension of $f(x)$ to the interval $-L \le x \le L$. Then we make a $2L$-periodic extension. This gives the Fourier cosine series (6.17) for the initial datum.

Next, consider the following arbitrary boundary conditions for which Dirichlet, Neumann, and Robin are special cases:

$$\alpha_1\psi(a) + \beta_1\psi(b) + \gamma_1\psi'(a) + \delta_1\psi'(b) = 0 \tag{6.18a}$$

$$\alpha_2\psi(a) + \beta_2\psi(b) + \gamma_2\psi'(a) + \delta_2\psi'(b) = 0 \tag{6.18b}$$

Definition 6.4. The BCs (6.18) are **symmetric** if

$$f'(x)g(x) - f(x)g'(x)\Big|_{x=a}^{x=b} = 0\,, \tag{6.19}$$

for any pair of functions f, g satisfying the boundary conditions.

We have the following two theorems for the eigenvalue problem $\psi'' + \lambda\psi = 0$ with symmetric BCs:

Theorem 6.5. If the BCs are symmetric, then any two eigenfunctions that correspond to distinct eigenvalues are orthogonal.

Theorem 6.6. If the BCs are symmetric, then all eigenvalues are real.

Remark 6.7. Recall that two vectors in $\mathbb{R}^n$ are "orthogonal" if there dot product is zero, i.e., $\mathbf{u}\cdot\mathbf{v} = 0$. Thus, we need to generalize this concept to vectors given by functions. In this case the inner product (or dot product) is defined as

$$(f,g)_{L^2(a,b)} := \int_a^b f(x)\overline{g(x)}dx\,, \tag{6.20}$$

where the overbar denotes complex conjugation. Thus, if the function is real-valued we do not need the overbar. Moreover, the norm (abstract notion of length) is then defined as

$$\|f\|_{L^2(a,b)} := \sqrt{(f,f)} = \Big(\int_a^b |f(x)|^2 dx\Big)^{1/2}. \tag{6.21}$$

Convergence. We consider three modes of convergence for the infinite series $\sum_{n=1}^{\infty} f_n(x)$:

- (Pointwise) We say that the infinite series converges to $f(x)$ pointwise in (a,b) if it converges to $f(x)$ for each $a < x < b$. That is for each $a < x < b$ it follows

$$|f(x) - \sum_{n=1}^{N} f_n(x)| \to 0 \quad \text{as} \quad N \to \infty\,. \tag{6.22}$$

- (Uniform) We say that the infinite series converges to $f(x)$ uniformly in $[a,b]$ if

$$\sup_{a\le x\le b} |f(x) - \sum_{n=1}^{N} f_n(x)| \to 0 \quad \text{as} \quad N \to \infty\,. \tag{6.23}$$

- (Mean square–L^2) We say that the infinite series converges to $f(x)$ in $L^2(a,b)$ if

$$\|f - \sum_{n} f_n\|^2_{L^2(a,b)} = \int_a^b |f(x) - \sum_{n=1}^{N} f_n(x)|^2 dx \to 0 \quad \text{as} \quad N \to \infty\,. \tag{6.24}$$

Thus, we have the following important theorems regarding the convergence of Fourier series:

Theorem 6.8. If $f(x)$ is piecewise smooth on $-L \le x \le L$, then the Fourier series converges pointwise

(i) to the periodic extension of $f(x)$, where the periodic extension is continuous,

(ii) to the average of the jump discontinuity $\frac{1}{2}[f(x^+) + f(x^-)]$.

Recall that piecewise smooth means that f and f' are continuous on sub-intervals of $-L \le x \le L$, and only jump discontinuities are allowed, i.e., $\lim_{x\to x_o^+} f(x) = f(x^+) \neq f(x^-) = \lim_{x\to x_o^-} f(x)$.

Theorem 6.9. The Fourier series converges to $f(x)$ uniformly on the interval $[a, b]$ provided that

(i) f, f', and f'' exist and are continuous for $a \le x \le b$,

(ii) f satisfies the given BCs .

Theorem 6.10. The Fourier series converges to $f(x)$ on the interval $[a, b]$ in the L^2 sense provided that

(i) $\|f\|^2_{L^2(a,b)} = \int_a^b |f(x)|^2 dx < \infty$.

Remark 6.11. The L^2 (mean square) is the most interesting case mathematically. Recall that the norm (length) of a vector $\mathbf{v} \in \mathbb{R}^n$ is

$$\|\mathbf{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

which follows from the Pythagorean theorem. Notice that

$$\|f\|_{L^2(a,b)} = (\int_a^b |f(x)|^2 dx)^{1/2} < \infty$$

is just the generalization of the vector norm to functions. Thus, we have an abstraction of length to functions. Thus, $\text{dist}(f, g) := \|f - g\|_{L^2(a,b)}$ gives a measure of distance between two functions. In this case

$$\{1, \cos(n\pi x/L), \sin(n\pi x/L)\}, \quad n = 1, 2 \ldots$$

is a countable basis for the infinite dimensional vector space of functions $f \in L^2(a, b)$. Where a function $f \in L^2(a, b)$ simply means $\|f\|_{L^2(a,b)} = \sqrt{\int_a^b |f(x)|^2 dx} < \infty$. Moreover, the constants c_n in the Fourier series are just projections onto the space spanned by the basis of trigonometric functions, i.e.,

$$c_n = \frac{(f, \psi_n)}{(\psi_n, \psi_n)}\,.$$

Differentiation. A **Fourier series** that is continuous can be differentiated term by term if $f'(x)$ is piecewise smooth. If $f'(x)$ is piecewise smooth, then the **Fourier cosine series** of a continuous function $f(x)$ can be differentiated term by term. If $f'(x)$ is piecewise smooth, then the **Fourier sine series** of a continuous function $f(x)$ can only be differentiated term by term if $f(0) = 0$ and $f(L) = 0$.

Integration. A Fourier series of a piecewise-smooth $f(x)$ can always be integrated term by term and the result is a convergent infinite series that always converges to the integral of $f(x)$ for $-L \le x \le L$ (even if the original Fourier series has jump discontinuities).

6.4 Periodic boundary conditions.

Consider a circular rod of length $2L$ that is in perfect thermal contact at the endpoints, i.e.,

$$u(-L, t) = u(L, t)\,, \quad u_x(-L, t) = u_x(L, t) \tag{6.25}$$

where x is the arc length along the circular rod. Thus we wish to solve the IBVP with **periodic boundary conditions**:

$$\begin{cases} \frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2}, \quad -L < x < L, \quad t > 0 \\ u(x,0) = f(x) \\ u(-L,t) = u(L,t), \quad u_x(-L,t) = u_x(L,t) \end{cases} \tag{6.26}$$

First, we make the ansatz:

$$u(x,t) = \psi(x)g(t) \tag{6.27}$$

Next, substitute into the PDE in (6.26):

$$\psi \frac{dg}{dt} = \kappa \psi'' g \iff \frac{1}{\kappa g}\frac{dg}{dt} = \frac{\psi''}{\psi} = -\lambda$$

where λ is the separation constant to be determined. Next, consider the BCs:

$$u(-L,t) = \psi(-L)g(t) = \psi(L)g(t) = u(L,t) \quad u_x(-L,t) = \psi'(-L)g(t) = \psi'(L)g(t) = u_x(L,t)\,.$$

Thus, we have two ODEs:

(i) $\frac{dg}{dt} = -\kappa\lambda g$

(ii) $\psi'' + \lambda\psi = 0\,, \quad \psi(-L) = \psi(L)\,, \quad \psi'(-L) = \psi;(L)$

Note that (i) is first-order linear (also separable) and can be solved easily to get

$$g(t) = ce^{-\kappa\lambda t}\,, \tag{6.28}$$

where c is an arbitrary constant of integration. Next, let's solve (ii) which is second-order linear homogeneous with constant coefficients. The difficulty is that (ii) has periodic BCs that must be satisfied. Recall that we consider three cases (i) $\lambda > 0$, (ii) $\lambda = 0$, and (iii) $\lambda < 0$.

($\lambda > 0$) In this case since the roots are purely imaginary and complex conjugate it follows that

$$\psi(x) = c_1 \sin(\sqrt{\lambda}x) + c_2 \cos(\sqrt{\lambda}x)\,,$$

and

$$\psi'(x) = c_1\sqrt{\lambda}\cos(\sqrt{\lambda}x) - c_2\sqrt{\lambda}\sin(\sqrt{\lambda}x)$$

Next, apply the BCs:

$$-c_1\sin(\sqrt{\lambda}L) + c_2\cos(\sqrt{\lambda}L) = c_1\sin(\sqrt{\lambda}L) + c_2\cos(\sqrt{\lambda}L) \iff 2c_1\sin(\sqrt{\lambda}L) = 0 \iff \lambda = \left(\frac{n\pi}{L}\right)^2,$$

$$c_1\sqrt{\lambda}\cos(\sqrt{\lambda}L) + c_2\sqrt{\lambda}\sin(\sqrt{\lambda}L) = c_1\sqrt{\lambda}\cos(\sqrt{\lambda}L) - c_2\sqrt{\lambda}\sin(\sqrt{\lambda}L) \iff 2c_2\sqrt{\lambda}\sin(\sqrt{\lambda}L) = 0$$

Thus,

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad n = 1,2,\ldots \tag{6.29}$$

$$\psi_n(x) = c_1\sin(n\pi x/L) + c_2\cos(n\pi x/L)\,, \quad n = 1,2,\ldots \tag{6.30}$$

and c_1, c_2 are arbitrary constants.

($\lambda = 0$) In this case since the root has algebraic multiplicity two it follows that

$$\psi(x) = c_1 x + c_2$$

Next, apply the BCs: $\psi(-L) = -c_1 L + c_2 = c_1 L + c_2 = \psi(L) \iff c_1 = 0$. Moreover, $\psi'(-L) = c_2 = \psi'(L)$ which is true since c_2 is constant. Thus, $\lambda = 0$ is an eigenvalue with eigenfunction c_2.

($\lambda < 0$) In this case since the roots are real and distinct it follows that

$$\psi(x) = c_1 \cosh(\sqrt{|\lambda|}x) + c_2 \sinh(\sqrt{|\lambda|}x)\,,$$

and

$$\psi'(x) = c_1\sqrt{|\lambda|}\sinh(\sqrt{|\lambda|}x) + c_2\sqrt{|\lambda|}\cosh(\sqrt{|\lambda|}x)\,.$$

Next, apply the BCs:

$$\psi(-L) = \psi(L) \iff 2c_2 \sinh(\sqrt{|\lambda|}L) = 0 \iff c_2 = 0\,,$$

and

$$\psi'(-L) = \psi'(L) \iff 2c_1 \sinh(\sqrt{|\lambda|}L) = 0 \iff c_1 = 0\,.$$

Thus, we only get the trivial solution. Thus there are no eigenvalues for $\lambda < 0$.

We have arrived at the following solution after applying the superposition principle:

$$u(x,t) = \Big(a_0 + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L) + \sum_{n=1}^{\infty} b_n \sin(n\pi x/L)\Big)e^{-\kappa(n\pi/L)^2 t}\,. \tag{6.31}$$

6.5 Laplace's equation in a circular disk

Consider a thin circular disk of radius a with constant thermal properties and no sources. Assume that the temperature is prescribed at the boundary and is independent of time (see Fig. 9). Then we expect that there will be a steady-state solution that satisfies Laplace's equation:

$$\Delta u = 0 \tag{6.32}$$

The geometry of the problem suggests that we use polar coordinates, i.e., $u = u(r,\theta)$. Thus, we have the following BVP:

$$\begin{cases} \Delta u = \frac{1}{r}\frac{\partial}{\partial r}\big(r\frac{\partial u}{\partial r}\big) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0 \quad \text{[PDE]} \\ u(a,\theta) = f(\theta) \quad \text{[BC]} \end{cases} \tag{6.33}$$

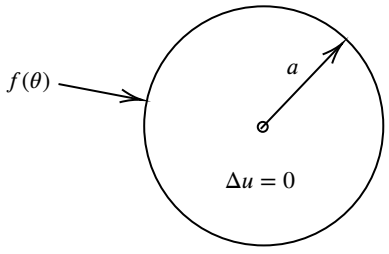

Figure 9: Laplace's equation in a thin circular disk.

Note that in polar coordinates we have $0 \leq r \leq a$ and $-\pi \leq \theta \leq \pi$. Mathematically, we need conditions at the endpoints of the coordinate system. We already have that $u(a, \theta) = f(\theta)$. Since polar coordinates are not well-defined for $r = 0$ we demand boundedness at the origin, i.e., $|u(0, \theta)| < \infty$. Moreover, we demand periodicity in θ. Thus

$$u(r, -\pi) = u(r, \pi) \tag{6.34a}$$
$$u_\theta(r, -\pi) = u_\theta(r, \pi) \tag{6.34b}$$

which implies $u(r, \theta + 2\pi) = u(r, \theta)$.

Next, let's solve (6.33) with the periodic BCs (6.34) and bounded at the origin. To this end we make the ansatz: $u(r, \theta) = g(r)\psi(\theta)$. Then it follows that

$$\begin{aligned} &\frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{dg}{dr}\Big)\psi(\theta) + \frac{g(r)}{r^2}\psi''(\theta) = 0 \\ \iff\ &r\frac{\partial}{\partial r}\Big(r\frac{dg}{dr}\Big)\psi(\theta) + g(r)\psi''(\theta) = 0 \\ \iff\ &\frac{r}{g(r)}\frac{\partial}{\partial r}\Big(r\frac{dg}{dr}\Big) = -\frac{\psi''(\theta)}{\psi(\theta)} = \lambda \end{aligned}$$

Thus, we get two ODE problems. Let' begin with

$$\begin{aligned} &\psi'' + \lambda\psi = 0 \\ &\psi(-\pi) = \psi(\pi) \\ &\psi'(-\pi) = \psi(\pi) \end{aligned}$$

Luckily, this problem has already been solved. We know that:

$$\begin{aligned} \lambda &= \Big(\frac{n\pi}{\pi}\Big)^2 = n^2, \quad n = 0, 1, \ldots \quad \text{(eigenvalues)} \\ \psi_n(x) &= \{\cos(n\theta), \sin(n\theta)\}_{n=0}^{\infty}, \quad n = 0, 1, \ldots \quad \text{(eigenfunctions)} \end{aligned}$$

Next, we consider the r-dependent problem:

$$\begin{aligned} &\frac{r}{g(r)}\frac{\partial}{\partial r}\Big(r\frac{dg}{dr}\Big) = \lambda \\ \iff\ &r(g' + rg'') = n^2 g \\ \iff\ &r^2 g'' + rg' - n^2 g = 0. \end{aligned}$$

Importantly, ODEs such as

$$r^2 g'' + rg' - n^2 g = 0 \tag{6.36}$$

are called Cauchy-equidimensional equations. They are one of the very few ODEs with variable coefficients that can be solved in closed-form.

Consider the ansatz:

$$g(r) = r^p \tag{6.37}$$

so that it follows

$$g'(r) = pr^{p-1}, \qquad g''(r) = p(p-1)r^{p-2}.$$

Substitution into (6.36) then gives

$$\begin{aligned} r^2 p(p-1) r^{p-2} + r p r^{p-1} - n^2 r^p = 0 \\ \iff p(p-1) r^p + p r^p - n^2 r^p = 0 \\ \iff [p(p-1) + p - n^2] r^p = 0 \end{aligned}$$

Thus, $g(r) = r^p$ is a solution if and only if

$$p(p-1) + p - n^2 = 0. \tag{6.38}$$

Equation (6.38) is referred to as the "indicial equation". Thus, for $n > 0$ we have

$$g(r) = c_1 r^n + c_2 r^{-n} \tag{6.39}$$

and to remain bounded as $r \to 0^+$ we must have $c_2 = 0$. Moreover, for $n = 0$ it follows that

$$\begin{aligned} r \frac{\partial}{\partial r}(r g'(r)) = 0 &\iff \frac{\partial}{\partial r}(r g'(r)) = 0 \\ &\iff g'(r) = \frac{c}{r} \\ &\iff g(r) = c \ln(r) + d \end{aligned}$$

Again, by the boundedness condition we must have $c = 0$. Thus, in this case we only have a constant solution. Putting everything together and by the superposition principle it follows that

$$u(r, \theta) = \sum_{n=0}^{\infty} c_n r^n \cos(n\theta) + \sum_{n=1}^{\infty} b_n r^n \sin(n\theta). \tag{6.40}$$

Moreover,

$$u(a, \theta) = \sum_{n=0}^{\infty} c_n a^n \cos(n\theta) + \sum_{n=1}^{\infty} b_n a^n \sin(n\theta) = f(\theta), \tag{6.41}$$

where

$$c_o = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) d\theta \tag{6.42a}$$

$$c_n = \frac{1}{\pi a^n} \int_{-\pi}^{\pi} f(\theta) \cos(n\theta) \tag{6.42b}$$

$$b_n = \frac{1}{\pi a^n} \int_{-\pi}^{\pi} f(\theta) \sin(n\theta) \tag{6.42c}$$

6.6 Qualitative properties of Laplace's equation

Mean Value Theorem: Consider a smooth bounded region $\mathcal{R}$ (see Fig. 10). Then the value of u at any point in $\mathcal{R}$ is the average value of u along any circle of radius r contained in the region $\mathcal{R}$. In fact, this can be seen easily with the above solution formula:

$$u(0, \theta) = c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) d\theta \tag{6.43}$$

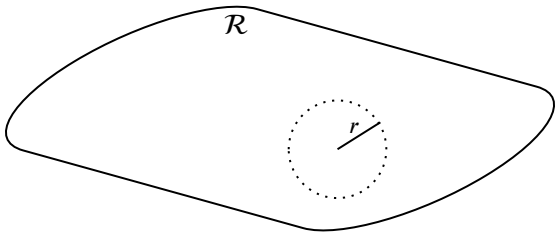

Figure 10: The region $\mathcal{R}$.

Maximum principle: Let $\mathcal{R}$ be a connected and bounded open set. Let $u(\mathbf{x})$ be a harmonic function, i.e., $\Delta u = 0$ that is continuous on $\mathcal{R} \cup \partial\mathcal{R}$. Then the maximum and the minimum values of u are attained on $\partial\mathcal{R}$ and nowhere inside (unless $u \equiv$ constant).

Well-posedness and uniqueness: The maximum principle is a very important tool for further analysis of PDEs, especially in establishing qualitative properties. We say that a problem is **well-posed** if there exists a unique solution that depends continuously on the data.

Consider Laplace's equation $\Delta u = 0$ with $u = f(\mathbf{x})$ on the boundary. Suppose that we vary the boundary data a small amount, i.e., $\Delta v = 0$ with $v = g(\mathbf{x})$ on the boundary and $f(\mathbf{x}) \approxeq g(\mathbf{x})$. Consider the difference $w = u - v$ so that $\Delta w = \Delta(u - v) = \Delta u - \Delta v = 0$. The maximum (and minimum) principle for Laplace's equation imply that the maximum and minimum occur on the boundary. Thus at any point inside,

$$\min(f(\mathbf{x}) - g(\mathbf{x})) \leq w \leq \max(f(\mathbf{x}) - g(\mathbf{x}))$$

and since $f(\mathbf{x}) \approxeq g(\mathbf{x})$ it follows that w is small, and thus the solution v is nearly the same as the solution u.

Solvability condition: If on the boundary the flow $-K_o \nabla \cdot \mathbf{n}$ $(= \frac{\partial u}{\partial n})$ is specified instead of the temperature, Laplace's equation may have no solutions. To show this we integrate over the entire region (say 2D):

$$\begin{aligned} 0 = \int\int_{\mathcal{R}} \Delta u \, dx dy &= \int\int_{\mathcal{R}} \nabla \cdot (\nabla u) \, dx dy \\ &= \oint_{\partial\mathcal{R}} \nabla u \cdot \mathbf{n} \, ds \end{aligned}$$

Since $\nabla u \cdot \mathbf{n}$ is proportional to the heat flow through $\partial\mathcal{R}$, it follows that $\oint_{\partial\mathcal{R}} \nabla u \cdot \mathbf{n} \, ds = 0$ implies that the net heat flow through the boundary must be zero in order for a solution to exist. This makes sense as if heat is flowing across the boundary we do not expect there to exist a steady-state solution.

Poisson integral formula: For Laplace's equation inside a circular disk ($r \leq a$) the solution can be written as

$$u(r, \theta) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\bar{\theta}) \left[\frac{1}{2} + \sum_{n=1}^{\infty} \left(\frac{r}{a} \right)^n \cos n(\theta - \bar{\theta}) \right] d\bar{\theta}.$$

Using $\cos z = \mathrm{Re}(e^{iz})$, sum the resulting geometric series to obtain Poisson's integral formula:

$$u(r, \theta) = \frac{a^2 - r^2}{2\pi} \int_{-\pi}^{\pi} \frac{f(\bar{\theta})}{a^2 - 2ar\cos(\theta - \bar{\theta}) + r^2} d\bar{\theta}.$$

We focus on the summation:

$$\left[\frac{1}{2}+\sum_{n=1}^{\infty}\left(\frac{r}{a}\right)^n\cos n(\theta-\bar{\theta})\right].$$

First note that since $e^{i\theta}=\cos\theta+i\sin\theta$ it is clear that $\mathrm{Re}(e^{i\theta})=\cos\theta$. Thus, $\cos n(\theta-\bar{\theta})=\mathrm{Re}(e^{in(\theta-\bar{\theta})})$. Moreover, since r/a is real it follows that

$$\begin{aligned}\sum_{n=1}^{\infty}\left(\frac{r}{a}\right)^n\mathrm{Re}(e^{in(\theta-\bar{\theta})}) &= \mathrm{Re}\left(\sum_{n=1}^{\infty}\left(\frac{r}{a}\right)^n e^{in(\theta-\bar{\theta})}\right)\\ &= \mathrm{Re}\left(\sum_{n=1}^{\infty}\left(\left[\left(\frac{r}{a}\right)e^{i(\theta-\bar{\theta})}\right]^n\right).\end{aligned}$$

Then since $0<r<a$ and $|e^{i(\theta-\bar{\theta})}|\leq 1$ it follows that the infinite series is summable. Recall that the sum of a geometric series is given by

$$\sum_{n=0}^{\infty}s^n=\frac{1}{1-s}\implies\sum_{n=1}^{\infty}s^n=\frac{1}{1-s}-1=\frac{s}{1-s}.$$

Thus,

$$\begin{aligned}\mathrm{Re}\left(\sum_{n=1}^{\infty}\left(\left[\left(\frac{r}{a}\right)e^{i(\theta-\bar{\theta})}\right]^n\right) &= \mathrm{Re}\left(\frac{\left(\frac{r}{a}\right)e^{i(\theta-\bar{\theta})}}{1-\left(\frac{r}{a}\right)e^{i(\theta-\bar{\theta})}}\right)\\ &= \mathrm{Re}\left(\frac{re^{i(\theta-\bar{\theta})}}{a-re^{i(\theta-\bar{\theta})}}\cdot\frac{a-re^{-i(\theta-\bar{\theta})}}{a-re^{-i(\theta-\bar{\theta})}}\right)\\ &= \mathrm{Re}\left(\frac{are^{i(\theta-\bar{\theta})}-r^2}{a^2-2ar\cos(\theta-\bar{\theta})+r^2}\right)\\ &= \frac{ar\cos(\theta-\bar{\theta})-r^2}{a^2-2ar\cos(\theta-\bar{\theta})+r^2}\end{aligned}$$

Finally,

$$\begin{aligned}\frac{1}{2}+\frac{ar\cos(\theta-\bar{\theta})-r^2}{a^2-2ar\cos(\theta-\bar{\theta})+r^2} &= \frac{a^2-2ar\cos(\theta-\bar{\theta})+r^2}{2(a^2-2ar\cos(\theta-\bar{\theta})+r^2)}+\frac{ar\cos(\theta-\bar{\theta})-r^2}{a^2-2ar\cos(\theta-\bar{\theta})+r^2}\cdot\left(\frac{2}{2}\right)\\ &= \frac{a^2-r^2}{2(a^2-2ar\cos(\theta-\bar{\theta})+r^2)}\end{aligned}$$

Thus, we get the **Poisson integral formula** for Laplace's equation in a disk:

$$u(r,\theta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\bar{\theta})P_{r,\theta}(\bar{\theta})\,d\bar{\theta} \tag{6.44}$$

where

$$P_{r,\theta}(\bar{\theta})=\frac{a^2-r^2}{a^2-2ar\cos(\theta-\bar{\theta})+r^2} \tag{6.45}$$

is the **Poisson kernel**.

6.7 Summary of separation of variables

Let us summarize the method of separation of variables:

1 Make sure that you have a linear and homogeneous PDE with linear and homogeneous BCs.

2 Temporarily ignore the initial data.

3 Separate variables (determine the differential equations implied by the assumption of product solutions) and introduce a separation constant.

4 Determine the separation constants as eigenvalues of a boundary value problem.

5 Solve the other differential equations. Record all product solutions of the PDE obtained by this method.

6 Apply the superposition principle (form a linear combination of all product solutions).

7 Attempt to satisfy the initial condition.

8 Determine coefficients using orthogonality of the eigenfunctions.

These steps should be understood, not memorized. It is important to note that:

1 The principle of superposition applies to solutions of the PDE (do not add up solutions of various different differential equations).

2 Do not apply the initial conditions until *after* the principle of superposition.

7 Wave equation

7.1 Derivation from Newton's second law of motion

In this section we consider the mathematical model of vibrations of perfectly elastic strings and membranes. Assume that we have a horizontally stretched string with the ends tied down in some way (e.g. a musical instrument).

We track the motion of each particle that comprises the string. Let α be the x-coordinate of a particle when the string is in the horizontal position (see Fig. 11). The string moves in time.

- Assume the slope of the string is small which implies v (horizontal displacement) can be neglected.
- As an approximation the motion is entirely vertical, i.e., $x = \alpha$.
- The vertical displacement thus depends on (x, t), i.e., $u = u(x, t) = y$.

We Newton's second law of motion, i.e., $\mathbf{F} = m\mathbf{a}$. Consider the interval $[x, x + \Delta x]$ and assume the mass density is $\rho_o(x)$. Thus, in this segment total mass is $m = \rho_o(x)\Delta x$. To find the displacement we must find the forces acting on the string.

Assume perfectly flexible (no resistance to bending). Thus, the force exerted on the endpoints of the segment of the string is in the direction tangent to the string. This tangential force is known as **tension** and $T(x, t)$ denotes its magnitude.

$$\text{slope of the string} \quad = \frac{dy}{dx} = \tan\theta = \frac{\partial u}{\partial x}.$$

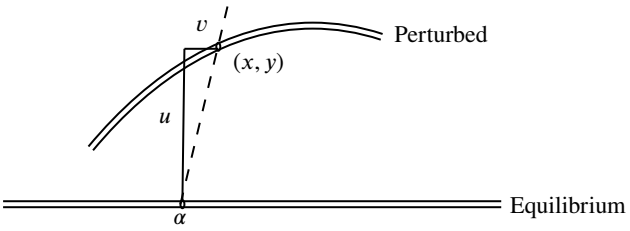

Figure 11: Vertical and horizontal displacements of a particle on a highly stretched string.

Since the horizontal component of the force is assumed small it can be neglected. Thus,

$$\begin{aligned}
\rho_o(x)(\Delta x)u_{tt} &= T(x+\Delta x,t)\sin\theta(x+\Delta x,t) - T(x,t)\sin\theta(x,t) + \rho_o(x)(\Delta x)Q(x,t) \\
\rho_o(x)u_{tt} &= \lim_{\Delta x\to 0}\frac{T(x+\Delta x,t)\sin\theta(x+\Delta x,t) - T(x,t)\sin\theta(x,t)}{\Delta x} + \rho_o(x)Q(x,t) \\
\rho_o(x)u_{tt} &= \frac{\partial}{\partial x}[T(x,t)\sin\theta(x,t)] + \rho_o(x)Q(x,t)
\end{aligned}$$

Then for small angles θ (as we have assumed) it follows that $u_x = \tan\theta = \frac{\sin\theta}{\cos\theta} \approx \sin\theta$. Thus,

$$\rho_o(x)u_t t = \frac{\partial}{\partial x}\big(T(x,t)\frac{\partial u}{\partial x}\big) + \rho_o(x)Q(x,t).$$

Since real strings are nearly perfectly elastic we can assume further that $T(x,t) \approx T_o$. Also, the body force due to gravity $Q(x,t) = -g$ can be neglected. Thus, we arrive at the waves equation:

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}, \tag{7.1}$$

where $c^2 = \frac{T_o}{\rho_o(x)}$.

7.2 Boundary conditions

As in the heat equation we apply one boundary condition at each endpoint. The major cases are the following:

- Prescribed displacement - $u(0,t) = \alpha(t)$ and $u(L,t) = \beta(t)$.
- Free end (vertical component of tensile force must vanish) - $T_0\frac{\partial u}{\partial x}(0,t) = 0$.
- Vibrations (Hooke's law) - $T_0\frac{\partial u}{\partial x}(0,t) = k(u(0,t) - u_E(t))$.

7.3 Wave equation in higher dimensions

7.4 d'Alembert's formula

Consider the wave equation of $-\infty < x < \infty$. Real physical situations are usually on finite intervals, but far away from the boundary these effects are negligible. Mathematically, the absence of BCs is a big

simplification. The most fundamental properties of PDEs can be found most easily without the complication of BCs. Thus, we consider

$$u_{tt} = c^2 u_{xx}, \quad -\infty < x < \infty \tag{7.2}$$

which can be rewritten as

$$(\partial_t - c\partial_x)(\partial_t + c\partial_x)u = 0. \tag{7.3}$$

This motivates a change to "characteristic coordinates"

$$\xi = x + ct, \qquad \eta = x - ct. \tag{7.4}$$

Then taking derivatives gives

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta} \\
\frac{\partial u}{\partial t} &= c\frac{\partial u}{\partial \xi} - c\frac{\partial u}{\partial \eta} \\
\frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial \xi^2} + 2\frac{\partial^2 u}{\partial \eta \partial \xi} + \frac{\partial^2 u}{\partial \eta^2} \\
\frac{\partial^2 u}{\partial t^2} &= c^2\frac{\partial^2 u}{\partial \xi^2} - 2c^2\frac{\partial^2 u}{\partial \eta \partial \xi} + c^2\frac{\partial^2 u}{\partial \eta^2}
\end{aligned}$$

Thus, plugging into the wave equation gives

$$-4c^2 u_{\eta\xi} = 0 \iff u_{\eta\xi} = 0.$$

which can be easily solved by integration. Consider the following calculation:

$$\begin{aligned}
u_{\eta\xi} = 0 &\iff \frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right) = 0 \\
&\iff \int \frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right) d\eta = \int 0 d\eta \\
&\iff \frac{\partial u}{\partial \xi} = f(\xi) \\
&\iff \int \frac{\partial u}{\partial \xi} d\xi = \int f(\xi) d\xi \\
&\iff u(\xi, \eta) = F(\xi) + G(\eta)
\end{aligned}$$

where $F' = f$. Thus, we get the "general" solution of the wave equation on an infinite domain

$$u(x,t) = F(x + ct) + G(x - ct). \tag{7.5}$$

Hence, the general solution is a sum of two arbitrary twice continuously differentiable functions:

(i) A wave of arbitrary shape traveling to the right with speed c.

(ii) A wave of arbitrary shape traveling to the left with speed c.

Next, we consider the Cauchy problem (IVP) for the wave equation:

$$\begin{cases}
u_{tt} = c^2 u_{xx}, \quad -\infty < x < \infty \\
u(x,0) = f(x) \\
u_t(x,t) = h(x)
\end{cases}$$

Thus, using the general solution we get

$$u(x,t) = F(x+ct) + G(x-ct)$$
$$u_t(x,t) = cF'(x+ct) - cG'(x-ct)$$

and using the initial data we get a system of two equations in two unknowns, namely,

$$F(x) + G(x) = f(x) \implies F'(x) + G'(x) = f'(x)$$
$$F'(x) - G'(x) = \frac{1}{c}h(x).$$

Thus, adding and subtracting these equations gives

$$2F'(s) = f'(s) + \frac{1}{c}h(s)$$
$$2G'(s) = f'(s) - \frac{1}{c}h(s).$$

Hence,

$$F' = \frac{1}{2}\big(f' + \frac{h}{c}\big) \implies F = \frac{1}{2}\big(f + \int \frac{h}{c}ds\big) + A$$
$$G' = \frac{1}{2}\big(f' - \frac{h}{2c}\big) \implies G = \frac{1}{2}\big(f - \int \frac{h}{2c}ds\big) + B$$

Note that from the first IC it follows that $A + B = 0$. Hence,

$$F(x+ct) = \frac{f(x+ct)}{2} + \frac{1}{2c}\int_0^{x+ct} h(s)ds + A$$
$$G(x-ct) = \frac{f(x-ct)}{2} - \frac{1}{2c}\int_0^{x-ct} h(s)ds + B$$

Thus, we have arrived at d'Alembert's formula:

$$u(x,t) = \frac{f(x+ct) + f(x-ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} h(s)ds. \tag{7.6}$$

It says:

(1) Effect of initial position = pair of waves traveling in either direction at speed c and half amplitude.

(2) Effect of initial velocity = wave spreading out at speed $\leq c$ in both directions.

So part of the wave may lag behind (if there is an initial velocity), but no part goes faster than speed c. This is the *principle of causality*.

An initial condition (position, velocity, or both) at the point $(x_o, 0)$ can affect the solution for $t > 0$ only in the shaded sector, called the *domain of influence* (see Fig. 12). If f, h vanish for $|x| > R$, then $u(x,t) = 0$ for $|x| > R + ct$. Thus, the domain of influence of an interval $|x| \leq R$ is a sector $|x| \leq R + ct$.

Alternatively, fix a point (x,t) for $t > 0$. How does $u(x,t)$ depend on the data? It depends on values of f only at two points $x \pm ct$, and it depends only on the values of h within the interval $[x-ct, x+ct]$. Thus, $(x-ct, x+ct)$ is the interval of dependence of the point (x,t) at $t = 0$.

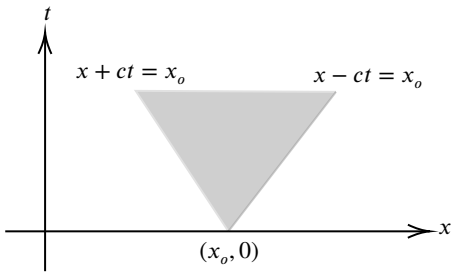

Figure 12: The domain of influence.

7.5 Energy

We will use conservation of energy to derive a conserved quantity for the wave equation (a quantity that remains constant under the flow of the wave equation). Consider an infinite string with constants ρ and T. Hence $\rho u_{tt} = T u_{xx}$ for $-\infty < x < \infty$. The kinetic energy is then

$$KE = \tfrac{1}{2} m v^2 = \tfrac{1}{2} \rho \int_{\mathbb{R}} u_t^2 dx.$$

Assume that ϕ and ψ vanish for $|x| \leq R$. This implies $u(x,t) = 0$ and $u_t(x,t) = 0$ for $|x| > R + ct$.

Thus, taking the derivative and using the wave equation gives

$$\begin{aligned}
\frac{dKE}{dt} &= \rho \int_{\mathbb{R}} u_t u_{tt} \\
&= \int_{\mathbb{R}} u_t \rho u_{tt} dx \\
&= \int_{\mathbb{R}} u_t T u_{xx} dx \\
&= T u_t u_x \Big|_{-\infty}^{\infty} - T \int_{\mathbb{R}} u_{tx} u_x dx \\
&= -\frac{T}{2} \int_{\mathbb{R}} \frac{\partial}{\partial t} (u_x)^2 dx \\
&= -\frac{T}{2} \frac{d}{dt} \int_{\mathbb{R}} u_x^2 dx \\
&= -\frac{dPE}{dt}
\end{aligned}$$

Hence the "energy" is given by

$$E(t) = \int_{\mathbb{R}} \Big(\frac{\rho}{2} u_t^2 + \frac{T}{2} u_x^2 \Big) dx. \tag{7.7}$$

7.6 Normal modes of vibration

Consider the problem of a vibrating string of length L with fixed ends:

$$\begin{cases} u_{tt} = c^2 u_{xx} \\ u(x,0) = f(x) \\ u_t(x,0) = h(x) \\ u(0,t) = 0, \quad u(L,t) = 0 \end{cases} \tag{7.8}$$

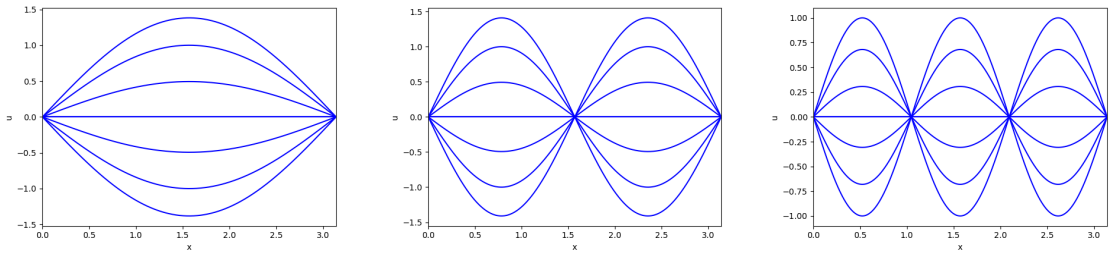

Figure 13: Normal modes of vibration for a string. (Left) $n = 1$, (Center) $n = 2$, and (Right) $n = 3$.

Applying the method of separation of variables the solution is represented as an infinite superposition of sines and cosines, namely,

$$u(x,t) = \sum_{n=1}^{\infty} \Big(A_n \cos(n\pi ct/L) + B_n \sin(n\pi ct/L) \Big) \sin(n\pi x/L), \tag{7.9}$$

where

$$A_n = \frac{2}{L} \int_0^L f(x) \sin(n\pi x/L) dx, \tag{7.10a}$$

$$B_n = \frac{2}{n\pi c} \int_0^L h(x) \sin(n\pi x/L) dx. \tag{7.10b}$$

We can interpret these results in the context of stringed musical instruments with fixed ends. The vertical displacement is composed of a linear combination of simple product solutions

$$\sin(n\pi x/L) \Big(A_n \cos(n\pi ct/L) + B_n \sin(n\pi ct/L) \Big).$$

These are called the **normal modes** of vibration. The intensity of the sound produced depends on the amplitude $(A_n^2 + B_n^2)^{1/2}$. The time dependence is simple harmonic with **circular frequency** (the number of oscillations in 2π units of time) equaling $n\pi c/L$, where $c = \sqrt{T_o/\rho_o}$. The sound produced consists of the superposition of these infinite number of natural frequencies ($n = 1, 2, \dots$). The normal mode $n = 1$ is called the first harmonic or fundamental. The larger the natural frequency, the higher the pitch of the sound produced.

Note also that to understand the motion at fixed time t each mode ($n = 1, 2, \dots$) looks like a simple oscillation in x. The amplitude varies periodically in time. These are called **standing waves**. In all cases there is no displacement at both ends due to the boundary conditions. For the second harmonic ($n = 2$), the displacement is also zero for all time at in the middle $x = L/2$. We thus call $x = L/2$ a **node** for the second harmonic. Similarly there are two nodes for the third harmonic (see Fig. 13)

8 Regular Sturm-Liouville eigenvalue problems

Notice that in separation of variables we faced with analyzing an eigenvalue problem for a linear differential operator. To show that this is the rule and not the exception consider heat flow in a non-uniform rod and external forcing $Q(x,t) = \alpha u(x,t)$:

$$c\rho u_t = \big(K_o u_x \big)_x + \alpha u. \tag{8.1}$$

Consider the separation anasatz: $u(x,t) = g(t)\psi(x)$. Then taking derivatives gives

$$c\rho\psi g' = \Big(K_o\psi_x g\Big)_x + \alpha g\psi \iff \frac{g'}{g} = \frac{1}{c\rho\psi}\Big(K_o\psi_x\Big)_x + \frac{\alpha}{c\rho} = -\lambda$$

Thus, in the spatial variable x we have the following eigenvalue problem:

$$-\frac{d}{dx}\big(K_o(x)\frac{d\psi}{dx}\big) - \alpha\psi = \lambda c\rho\psi. \tag{8.2}$$

Regular Sturm-Liouville eigenvalue problem:

$$L\psi = -\frac{d}{dx}\Big(p(x)\frac{d\psi}{dx}\Big) - q(x)\psi = \lambda\sigma(x)\psi, \quad a < x < b \tag{8.3a}$$

$$\beta_1\psi(a) + \beta_2\psi'(a) = 0 \tag{8.3b}$$

$$\beta_3\psi(b) + \beta_4\psi(b) = 0 \tag{8.3c}$$

where $\beta_i \in \mathbb{R}$. Moreover, p, q, and σ are real-valued and continuous. Also, $p > 0$, and $\sigma > 0$ for $a \le x \le b$. Note that this exlcudes periodic and singular bounary conditions.

A regular Sturm-Liouville eigenvalue problem has many usefule properties which are discussed in the following theorem:

Theorem 8.1. For any regular Sturm-Liouville problem

1. All eigenvalues are real.

2. There exists an infinite sequence of eigenvalues

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n < \lambda_{n+1} < \cdots$$

3. Corresponding to each eigenvalue λ_n, there is an eigenfunction, denoted $\psi_n(x)$ (unique up to a multiplicative constant). Also, $\psi_n(x)$ has exactly $n-1$ zeros for $a < x < b$.

4. The eigenfunctions for a "complete set", meaning any piecewise smooth $f(x)$ can be represented by a generalized Fourier series of eigenfunctions:

$$f(x) \sim \sum_{n=1}^{\infty} a_n\psi_n(x).$$

   Furthermore, this series converges to $[f(x^+) + f(x^-)]/2$ at points where there is a jump discontinuity.

5. Eigenfunctions corresponding to different eigenvalues are orthogonal relative to the weight function $\sigma(x)$. Thus, $\langle\psi_n, \psi_m\rangle_{L^2_\sigma(a,b)} = \int_a^b \psi_n(x)\psi_m(x)dx = 0$.

6. Any eigenvalue can be related to its eigenfunctions by the Rayleigh quotient

$$\lambda = \frac{-p\psi\psi'|_a^b + \int_a^b [p(\psi')^2 - q\psi^2]dx}{\int_a^b \psi^2\sigma dx}. \tag{8.4}$$

The following identities are useful in proving some of the above results.

Lagrange's identity: Using operator notation $Lu = (p(x)u')' + qu$ it follows that

$$uLv - vLu = u(pv')' + quv - v(pu')' - quv = u(pv')' - v(pu')'.$$

Also, note that

$$(u(pv'))' = u(pv')' + (pv')'.$$

Therefore,

$$\begin{aligned} uLv - vLu &= u(pv')' - v(pu')' \\ &= (u(pv'))' - (pv')u' - (v(pu'))' + (pu')v' \\ &= (p(uv' - vu'))' \end{aligned}$$

Thus, Lagrange's indentity is given by

$$uLv - vLu = \frac{d}{dx}\left[p(x)\left(u\frac{dv}{dx} - v\frac{du}{dx}\right)\right]. \tag{8.5}$$

Green's formula: The integral form of Lagrange's identity is also known as Green's formula. Green's formula is given by

$$\int_a^b [uLv - vLu]dx = p(x)\left(u\frac{dv}{dx} - v\frac{du}{dx}\right)\Big|_a^b. \tag{8.6}$$

Note that if $p \equiv 1$ and $q \equiv 0$ so that $L = \frac{d^2}{dx^2}$ the formulae are:

$$uv'' - vu'' = \left(uv' - vu'\right)' \qquad \text{(Lagrange)} \tag{8.7a}$$

$$\int_a^b (uv'' - vu'')dx = (uv' - vu')|_a^b \qquad \text{(Green)} \tag{8.7b}$$

Next, we prove point 1. in Theorem 8.1:

Lemma 8.2. Let $L : L^2_\sigma(a,b) \to L^2_\sigma(a,b)$ be the linear operator

$$L = -\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) - q(x).$$

corresponding to a regular Sturm-Liouville problem. If λ is an eigenvalue, then $\lambda \in \mathbb{R}$.

Proof. Suppose $\lambda \in \mathbb{C}$ is an eigenvalue corresponding to the eigenfunction $\psi_n(x)$. Then $L\psi = \lambda\sigma\psi$ and $\overline{L\psi} = \overline{\lambda}\sigma\overline{\psi}$ (where overbar denotes complex conjugation). Since the coefficients of L are real it follows that $\overline{L\psi} = L\overline{\psi}$. Moreover, if ψ satisfies BCs with real coefficients so does $\overline{\psi}$. Thus, if (λ, ψ) is an eigenpair, then so to is $(\overline{\psi}, \overline{\lambda})$. Hence,

$$\begin{aligned} \int_a^b [\psi(L\overline{\psi} - \overline{\lambda}\sigma\overline{\psi}) - \overline{\psi}(L\psi - \lambda\sigma\psi)]dx &= \int_a^b [\psi L\overline{\psi} - \overline{\psi}L\psi]dx + (\lambda - \overline{\lambda})\int_a^b \psi\overline{\psi}\sigma dx \\ &= (\lambda - \overline{\lambda})\int_a^b |\psi|^2\sigma dx \\ &= (\lambda - \overline{\lambda})\|\psi\|_{L^2_\sigma(a,b)} \\ &= 0 \end{aligned}$$

Finally, by continuity $\|\psi\|_{L^2_\sigma(a,b)} \not\equiv 0$. Thus $\overline{\lambda} = \lambda$ which implies $\lambda \in \mathbb{R}$. □

8.1 Asymptotic behavior of eigenvalues

9 Separation of variables in higher dimensions

Next we consider the technique of separation of variables in higher spatial dimensions.

- Laplace's equation: $\Delta u = 0, \quad \Delta := \partial_{x_1}^2 + \cdots + \partial_{x_n}^2$
- Heat equation: $u_t = \kappa \Delta u$
- Heat equation with convection: $u_t = \kappa \Delta u + au$
- Wave equation $u_{tt} = c^2 \Delta u$

Consider a vibrating membrane (any shape):

$$\begin{cases} u_{tt} = c^2 \Delta u \\ u(\mathbf{x}, 0) = \alpha(\mathbf{x}) \\ u_t(\mathbf{x}, 0) = \beta(\mathbf{x}) \end{cases} \tag{9.1}$$

Then the separation ansatz gives $u(\mathbf{x}, t) = g(t)\psi(\mathbf{x})$. Plugging into the PDE gives

$$g''(t)\psi(\mathbf{x}) = c^2 g(t) \Delta\psi(\mathbf{x}) \iff \frac{g''}{c^2 g} = \frac{\Delta\psi}{\psi} = -\lambda.$$

Thus, we get the two problems:

(i) $g'' + \lambda c^2 g = 0$

(ii) $\Delta\psi + \lambda\psi = 0$

Next, consider heat conduction (any region)

$$\begin{cases} u_t = \kappa \Delta u \\ u(\mathbf{x}, 0) = \alpha(\mathbf{x}) \end{cases} \tag{9.2}$$

Then the separation ansatz gives $u(\mathbf{x}, t) = g(t)\psi(\mathbf{x})$. Plugging into the PDE gives

$$g'(t)\psi(\mathbf{x}) = \kappa g(t) \Delta\psi(\mathbf{x}) \iff \frac{g'}{\kappa g} = \frac{\Delta\psi}{\psi} = -\lambda.$$

Thus, we get the two problems:

(i) $g' + \lambda\kappa g = 0$

(ii) $\Delta\psi + \lambda\psi = 0$

In both cases we get an eigenvalue problem of the form $\Delta\psi = \lambda\psi$. This eigenvalue problem is central to the method of separation of variables. In particular the eigenvalue problem that we are interested in is give by

$$\Delta\psi + \lambda\psi = 0 \tag{9.3a}$$

$$a\psi + b\nabla\psi \cdot \mathbf{n} = 0 \tag{9.3b}$$

9.1 Vibrating rectangular membrane

Let's solve the following IBVP corresponding to a vibrating rectangular membrane:

$$\begin{cases} u_{tt} = c^2\Delta u, \quad 0 < x < L, \quad 0 < y < H \\ u(x,y,0) = \alpha(x,y) \\ u_t(x,y,0) = \beta(x,y) \\ u(0,y,t) = 0, \quad u(x,0,t) = 0 \\ u(L,y,t) = 0, \quad u(x,H,t) = 0 \end{cases} \tag{9.4}$$

We begin as usual with the separation ansatz: $u(x,y,t) = g(t)\psi(x,y)$. Plugging into the wave equation gives

$$g''(t)\psi(x,y) = c^2 g(t)\Delta\psi(x,y) \iff \frac{g''}{c^2 g} = \frac{\Delta\psi}{\psi} = -\lambda.$$

Thus, we have two problems:

(i) $g'' + \lambda c^2 g = 0$

(ii) $\Delta\psi + \lambda\psi = 0$

Remark 9.1. Notice that problem (ii) is a PDE and not an ODE like in the case with two independent variables. On the other hand, it is still an eigenvalue problem.

Since (ii) is a PDE with two independent variables we separate variables further. Thus, we make a second separation ansatz, namely, $\psi(x,y) = f(x)h(y)$. Then plugging into (ii) gives

$$f''(x)h(y) + f(x)h''(y) + \lambda f(x)h(y) = 0 \iff \lambda + \frac{h''}{h} = -\frac{f''}{f} = \mu.$$

Moreover, $g(t)\psi(0,y) = g(t)f(0)h(y) = 0$ and $g(t)\psi(L,y) = g(t)f(L)h(y) = 0$ implies that we have the following eigenvalues problem:

$$f'' + \mu f = 0, \quad f(0) = 0, \quad f(L) = 0. \tag{9.5}$$

Thus,

$$\mu_n = (n\pi/L)^2, \qquad f_n(x) = \sin(n\pi x/L), \quad n = 1, 2, \ldots. \tag{9.6}$$

Next, the BCs $g(t)\psi(x,0) = g(t)f(x)h(0) = 0$ and $g(t)\psi(x,H) = g(t)f(x)h(H) = 0$ implies that we have the following eigenvalue problems:

$$h'' + (\lambda - \mu)h = 0, \quad h(0) = 0, \quad h(H) = 0. \tag{9.7}$$

Importantly, there are infinitely many eigenvalue problems, one for each value of μ_n. Moreover, we know for each fixed μ_n the above eigenvalue problem has infinitely many eigenvalues $(m\pi/H)^2$. Thus, there is a doubly infinite sequence of λ eigenvalues:

$$\lambda_{nm} = (n\pi/L)^2 + (m\pi/H)^2, \quad \psi_{nm}(x,y) = \sin(n\pi x/L)\sin(m\pi y/H), \quad n, m = 1, 2, \ldots. \tag{9.8}$$

Next, the time-dependent part of the product solutions are $\sin(c\sqrt{\lambda_{nm}}t)$ and $\cos(c\sqrt{\lambda_{nm}}t)$, i.e., oscillations with natural frequencies $c\sqrt{\lambda_{nm}} = c\sqrt{(n\pi/L)^2 + (m\pi/H)^2}$. Thus, by superposition we have

$$u(x,y,t) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left(A_{nm}\cos(c\sqrt{\lambda_{nm}}t) + B_{nm}\sin(c\sqrt{\lambda_{nm}}t)\right)\sin(n\pi x/L)\sin(m\pi y/H). \tag{9.9}$$

The last step is to obtain formulae for the unknown constants. Here we use the initial conditions.

$$\begin{aligned} u(x,y,0) &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{nm}\sin(n\pi x/L)\sin(m\pi y/H) \\ &= \sum_{m=1}^{\infty}\Big(\sum_{n=1}^{\infty} A_{nm}\sin(n\pi x/L)\Big)\sin(m\pi y/H) \\ &= \alpha(x,y) \end{aligned}$$

Thus, we have a double Fourier series, i.e., a Fourier series expansion in two variables. It then follows that

$$\sum_{n=1}^{\infty} A_{nm}\sin(n\pi x/L) = \frac{2}{H}\int_0^H \alpha(x,y)\sin(m\pi y/H)dy.$$

Thus, applying the Fourier series formula for the constants one more time gives

$$A_{nm} = \frac{2}{L}\int_0^L \Big(\frac{2}{H}\int_0^H \alpha(x,y)\sin(m\pi y/H)dy\Big)\sin(n\pi x/L)dx. \tag{9.10}$$

To find B_{nm} we take the derivative termwise of the solution and evaluate at $t = 0$. Thus,

$$\begin{aligned} u_t(x,y,0) &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} B_{nm}c\sqrt{\lambda_{nm}}\sin(n\pi x/L)\sin(m\pi y/H) \\ &= \sum_{m=1}^{\infty}\Big(\sum_{n=1}^{\infty} B_{nm}c\sqrt{\lambda_{nm}}\sin(n\pi x/L)\Big)\sin(m\pi y/H) \\ &= \beta(x,y) \end{aligned}$$

It then follows that

$$\sum_{n=1}^{\infty} B_{nm}c\sqrt{\lambda_{nm}}\sin(n\pi x/L) = \frac{2}{H}\int_0^H \beta(x,y)\sin(m\pi y/H)dy.$$

Thus, applying the Fourier series formula for the constants one more time gives

$$B_{nm} = \frac{2}{Lc\sqrt{\lambda_{nm}}}\int_0^L \Big(\frac{2}{H}\int_0^H \beta(x,y)\sin(m\pi y/H)dy\Big)\sin(n\pi x/L)dx. \tag{9.11}$$

9.2 Vibrating circular membrane

Let's solve the following IBVP corresponding to a vibrating rectangular membrane:

$$\begin{cases} u_{tt} = c^2\Delta u, 0 < r < a \\ u(r,\theta,0) = \alpha(r,\theta) \\ u_t(r,\theta,0) = \beta(r,\theta) \\ u(a,\theta,t) = 0 \end{cases} \tag{9.12}$$

We begin as usual with the separation anasatz: $u(r,\theta,t) = g(t)\psi(r,\theta)$. Plugging into the wave equation gives

$$g''(t)\psi(x,y) = c^2\Delta\psi(x,y) \iff \frac{g''}{c^2 g} = \frac{\Delta\psi}{\psi} = -\lambda.$$

Thus, we have two problems:

(i) $g'' + \lambda c^2 g = 0$

(ii) $\Delta\psi + \lambda\psi = 0$

Since (ii) is a PDE with two independent variables we separate variables further. Thus, we make a second separation ansatz, namely, $\psi(r,\theta) = f(r)h(\theta)$. Also, recall the Laplacian in polar coordinates:

$$\Delta\psi = \frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial\psi}{\partial r}\Big) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2}.$$

Plugging the separation ansatz into the eigenvalue problem (ii) gives

$$\frac{h(\theta)}{r}\frac{\partial}{\partial r}\Big(r\frac{df(r)}{dr}\Big) + \frac{f(r)}{r^2}\frac{d^2h(\theta)}{d\theta^2} + \lambda f(r)h(\theta) = 0 \iff \frac{r}{f}(rf')' + \lambda r^2 = -\frac{h''}{h} = \mu.$$

Moreover, due to continuity we expect $u(r,-\pi,t) = u(r,\pi,t)$ and $u_t(r,-\pi,t) = u_t(r,\pi,t)$. Thus we have the following eigenvalue problem:

$$h'' + \mu h = 0, \quad h(-\pi) = h(\pi), \quad h'(-\pi) = h'(\pi) \tag{9.13}$$

which is a regular Sturm-Liouville eigenvalue problem with periodic boundary conditions. Thus, it follows that

$$\mu_m = m^2, \qquad h_m(\theta) = c_m\cos(m\theta) + d_m\sin(m\theta), \quad m = 0,1,\dots. \tag{9.14}$$

The other problem that results from the separation ansatz is given by

$$\begin{cases} (rf')' + (\lambda r - \frac{m^2}{r})f = 0 \\ f(a) = 0 \\ |f(0)| < \infty \end{cases} \tag{9.15}$$

which is a "singular" Sturm-Liouville problem due to the second boundary condition.

Remark 9.2. Let $Lf = -(rf')' + m^2 f/r$. Then we have $Lf = \lambda rf$ with $p(r) = \sigma(r) = r$, $q(r) = m^2/r$, and the above boundary conditions. This is not a regular Sturm-Liouville problem though due to the BCs, since $q(r) \to \infty$ as $r \to 0$, and since $p(r) \to 0$ and $q(r) \to 0$ as $r \to 0$. This is still an eigenvalue problem though, and some of the statements corresponding to a regular Sturm-Liouville problem still hold.

The eigenvalue problem (9.15) corresponds to an ODE with variable coefficients. Let's make the change of variable $z := \sqrt{\lambda}r$. Then it follows that

$$f'(r) = \frac{df}{dz}\frac{dz}{dr} = \sqrt{\lambda}f'(z)$$
$$f''(r) = \frac{d^2f}{dz^2}\Big(\frac{dz}{dr}\Big)^2 + \frac{df}{dz}\frac{d^2z}{dr^2} = \lambda f''(z)$$

Thus, the ODE in the eigenvalue problem (9.15) becomes

$$z^2f''(z) + zf'(z) + (z^2 - m^2)f = 0 \tag{9.16}$$

which is known as **Bessel's equation of order m**.

Consider a linear homogeneous second-order ODE with variable coefficients in standard form:

$$y'' + p(z)y' + q(z)y = 0 \tag{9.17}$$

It is very rare to find closed-form solutions to (9.17). It turns out though that if p and q are analytic at z_o, then the solution can be represented by a power series centered at z_o. We call points where p and q are analytic **ordinary points**. On the other hand if p and q are not analytic at z_o, but the limits

$$\lim_{z\to z_o}(z-z_o)p(z) \qquad \text{(pole of order 1)} \tag{9.18a}$$

$$\lim_{z\to z_o}(z-z_o)^2 q(z) \qquad \text{(pole of order 2)} \tag{9.18b}$$

exist, then the solution can be represented by a Frobenius series centered at z_o and we call z_o a **regular singular point**.

Thus, writing Bessel's equation (9.16) in standard form

$$f'' + \frac{1}{z}f' + \Big(1 - \frac{m^2}{z^2}\Big)f = 0$$

it follows that all $z \neq 0$ are ordinary points and $z = 0$ is a regular singular point since

$$\lim_{z\to 0} z\Big(\frac{1}{z}\Big) = 1,$$
$$\lim_{z\to 0} z^2\Big(\frac{z^2-m^2}{z^2}\Big) = -m^2.$$

Thus, solutions to Bessel's equation are well-behaved for all $z \neq 0$. At $z = 0$ we need to perform a closer analysis.

For $m > 0$ and $z \approx 0$ we expect that $(z^2 - m^2)f \approx -m^2 f$ since m is fixed. Thus $z^2 f'' + zf' - m^2 f \approx 0$ which is the Cauchy-Euler equidimensional equation. Thus, the indicial equation is $s(s-1)+s-m^2 = 0 \iff s^2-m^2 = 0$. Hence, $f \approx z^m$, or $f \approx z^{-m}$. On the other hand if $m = 0$, then $z^2 f''+zf' \approx 0 \iff z(zf')' \approx 0$. Thus, $f \approx 1$ or $f \approx \ln z$. The key is that in both cases there exist well-behaved solutions of Bessel's equation at the regular singular point $z = 0$. We have the following:

$$J_m(z) = \text{Bessel function of the first kind of order } m \tag{9.19a}$$

$$Y_m(z) = \text{Bessel function of the second kind of order } m \tag{9.19b}$$

Therefore,

$$f(z) = c_1 J_m(z) + c_2 Y_m(z) \tag{9.20}$$

where

$$J_m(z) \sim \begin{cases} 1 & m = 0 \\ \frac{z^m}{2^m m!} & m > 0 \end{cases} \tag{9.21a}$$

$$Y_m(z) \sim \begin{cases} \frac{2}{\pi}\ln z & m = 0 \\ -\frac{2^m(m-1)!z^{-m}}{\pi} & m > 0 \end{cases} \tag{9.21b}$$

as $z \to 0$.

Big O and little o notation: In asymptotic analysis big O and little o notation is extremely useful for finding rigorous error estimates to approximate solutions of differential equations.

Definition 9.3. Let f be a function of a real, or complex variable and let g be a comparison function. Then we say that $f(z) = O(g(z))$ as $z \to z_o$ if there exists $\delta > 0$ and $M > 0$ such that for all z satisfying $0 < |z - z_o| < \delta$ it follows that

$$|f(z)| \leq M|g(z)|. \tag{9.22}$$

The definition can be extended easily to limits $z \to \infty$.

Definition 9.4. Let f be a function of a real, or complex variable and let g be a comparison function. Then we say that $f(z) = o(g(z))$ as $z \to z_o$ if for any $\epsilon > 0$ there exists $\delta > 0$ such that for all z satisfying $0 < |z - z_o| < \delta$ it follows that

$$|f(z)| \leq \epsilon |g(z)|. \tag{9.23}$$

The definition can be extended easily to limits $z \to \infty$.

Definition 9.5. We say that f is asymptotic to g as $z \to z_o$ denoted $f \sim g$ if

$$f(z) = g(z)(1 + o(1)) \quad \text{as} \quad z \to z_o. \tag{9.24}$$

Note that if $g(z) \neq 0$ in a neighborhood of z_o, then (9.24) is equivalent to

$$\lim_{z \to z_o} \left| \frac{f(z)}{g(z)} - 1 \right| = 0. \tag{9.25}$$

10 Nonhomogeneous problems

10.1 Method of eigenfunction expansions–time dependent problems

10.2 Method of eigenfunction expansions–time independent problems

11 Fourier transform

Most problems we have analyzed previously were defined on finite regions, i.e., IBVPs. Now we will analyze problems that extend to infinity. Physical problems never are infinite, but by introducing a mathematical model with infinite extent, we are able to determine the behavior of problems in situations in which the influence of actual boundaries is expected to be negligible.

Let's consider the following Cauchy problem (initial-value problem):

$$\begin{cases} u_t = \kappa u_{xx}, \quad -\infty < x < \infty, \quad t > 0, \\ u(x, 0) = f(x) \end{cases} \tag{11.1}$$

Let's assume that $f(x) \to 0$ as $x \to \pm\infty$. Physically, for all time the temperature approaches zero as $x \to \pm\infty$: $u(\pm\infty, t) = 0$. Thus we have "homogeneous boundary conditions". Hence let's see what happens if we separate variables: $u(x, t) = g(t)\psi(x) \implies g'/\kappa g = \psi''/\psi = -\lambda$.

Eigenvalue problem: $\psi'' + \lambda\psi = 0$, $\psi(\pm\infty) = 0$.

($\lambda > 0$): $\psi(x) = c_1 \cos(\sqrt{\lambda}x) + c_2 \sin(\sqrt{\lambda}x)$– no eigenfunctions since the limit does not exist.

($\lambda = 0$): $\psi(x) = c_1 x + c_2$– no eigenfunctions.

($\lambda > 0$): $c_1 e^{\sqrt{|\lambda|}x} + c_2 e^{-\sqrt{|\lambda|}x}$– no eigenfunctions.

Thus, we get no information from this problem. The key is to consider bounded eigenfunctions, i.e., $|\psi(\pm x)| < \infty$. In this case we have the following:

Eigenvalue problem: $\psi'' + \lambda\psi = 0$, $|\psi(\pm\infty)| < \infty$.

($\lambda > 0$): $\psi(x) = c_1 \cos(\sqrt{\lambda}x) + c_2 \sin(\sqrt{\lambda}x)$– bounded for all $\lambda > 0$.

($\lambda = 0$): $\psi(x) = c_1 x + c_2$– bounded for $c_1 = 0$.

($\lambda > 0$): $c_1 e^{\sqrt{|\lambda|}x} + c_2 e^{-\sqrt{|\lambda|}x}$– no eigenfunctions.

Hence the set of eigenvlues is given by $\{\lambda \in \mathbb{R} : \lambda \geq 0\}$. We have the following:

$$\begin{cases} \psi'' + \lambda\psi = 0, \\ \psi(0) = \psi(L) = 0 \end{cases} \tag{11.2}$$

"Discrete spectrum" given by $\sigma = \{(n\pi/L)^2\}_{n=1}^{\infty}$. Also,

$$\begin{cases} \psi'' + \lambda\psi = 0, \\ |\psi(\pm\infty)| < \infty \end{cases} \tag{11.3}$$

"Continuous spectrum" given by $\sigma = \{\lambda \in \mathbb{R} : \lambda \geq 0\}$. Then by the superposition principle we have

$$u(x,t) = \int_0^{\infty} [A(\omega)\cos(\omega x) + B(\omega)\sin(\omega x)]e^{-\kappa\omega^2 t}\, d\omega, \quad \omega = \sqrt{\lambda}. \tag{11.4}$$

where now $A(\omega)$ and $B(\omega)$ are now arbitrary functions.

Note that the x-dependent eigenfunctions were determined from the basis $\{\cos(\sqrt{\lambda}x), \sin(\sqrt{\lambda}x)\}$. For some problems it is preferable to use the basis $\{e^{i\sqrt{\lambda}x}, e^{-i\sqrt{\lambda}x}\}$. If we introduce $\omega = \sqrt{\lambda}$ we have as the basis $\{e^{i\omega x}, e^{-i\omega x}\}$.

11.1 Complex Fourier series

In solving problems on infinite domains $-\infty < x < \infty$, it is usual to analyze $e^{-i\omega x}$ ($\omega \in \mathbb{R}$) rather than $\{\cos(\sqrt{\lambda}x), \sin(\sqrt{\lambda}x)\}$ ($\lambda \geq 0$).

Recall the Fourier series:

$$f(x) \sim \frac{a_o}{2} + \sum_{n=1}^{\infty}(a_n \cos(n\pi x/L) + b_n \sin(n\pi x/L)), \tag{11.5}$$

where

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(n\pi x/L)\, dx, \tag{11.6a}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(n\pi x/L)\, dx. \tag{11.6b}$$

From Euler's formula $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ it follows that

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

Thus,

$$f(x) \sim \frac{a_o}{2} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n - ib_n)e^{in\pi x/L} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n + ib_n)e^{-in\pi x/L}. \tag{11.7}$$

Let $n \mapsto -n$ in the first sum. This gives

$$f(x) \sim \frac{a_o}{2} + \frac{1}{2}\sum_{n=-1}^{-\infty}(a_{-n} - ib_{-n})e^{-in\pi x/L} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n + ib_n)e^{-in\pi x/L}. \tag{11.8}$$

Then noting that $a_{-n} = a_n$ and $b_{-n} = -b_n$ it follows that

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{-in\pi x/L}, \tag{11.9}$$

where

$$c_n = \frac{a_n + ib_n}{2} = \frac{1}{2L}\int_{-L}^{L} f(x)e^{in\pi x/L}. \tag{11.10}$$

This is the complex form of a Fourier series.

Using the complex form of a Fourier series we can re-express our solution (11.4) as

$$u(x,t) = \int_{-\infty}^{\infty} c(\omega)e^{-i\omega x}e^{-\kappa\omega^2 t}. \tag{11.11}$$

The issue is that at this point we do not know how to obtain $c(\omega)$. This is how Fourier transforms come into the picture.

11.2 Fourier transform pair

In this section we derive the Fourier transform pair. For simplicity let's assume that f is continuous and absolutely integrable. To this end consider the Fourier indentity:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{-in\pi x/L} = \sum_{n=-\infty}^{\infty}\left[\frac{1}{2L}\int_{-L}^{L} f(\bar{x})e^{in\pi\bar{x}/L}\,d\bar{x}\right]e^{-in\pi x/L}. \tag{11.12}$$

We want to consider (11.12) as $L \to \infty$. In other words, functions defined for $-\infty < x < \infty$ may be thought of in some sense as periodic functions with an infinite period.

For periodic functions, $-L < x < L$, the allowable wave numbers ω (number of waves in 2π distance) are the infinite set of discrete values $\omega = n\pi/L = 2\pi\frac{n}{2L}$. Thus, the wave lengths are $2L/n$, integral partitions of the original region of length $2L$. The distance between successive values of the wave number is $\Delta\omega = (n+1)\pi/L - n\pi/L = \pi/L$. Thus, from (11.12)

$$f(x) = \sum_{n=-\infty}^{\infty}\frac{\Delta\omega}{2\pi}\int_{-L}^{L} f(\bar{x})e^{i\omega\bar{x}}\,d\bar{x}e^{-i\omega x}. \tag{11.13}$$

Note that as $L \to \infty$ it follows that $\Delta\omega \to 0$ and the wave numbers ω are more closely spaced. Thus, the eigenvalues approach a continuum; all possible wave numbers are allowable. Equation (11.13) represents a sum of rectangles (from $-\infty$ to ∞) of base $\Delta\omega$ and height $\int_{-L}^{L} f(\bar{x})e^{i\omega\bar{x}}\,d\bar{x}e^{-i\omega x}/2\pi$. As $L \to \infty$ these heights are approximately equal $\int_{-\infty}^{\infty} f(\bar{x})e^{i\omega\bar{x}}\,d\bar{x}e^{-i\omega x}$. Thus, we expext as $L \to \infty$ that the areas of the rectangles approach the Riemann sum. Since $\Delta\omega \to 0$ as $L \to \infty$ (11.13) becomes the Fourier integral identity.

Fourier integral identity:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\bar{x})e^{i\omega\bar{x}}\,d\bar{x}\right)e^{-i\omega x}\,d\omega. \tag{11.14}$$

Hence, we have arrived at the Fourier transform pair.

Fourier transform pair:

$$\hat{f}(\omega) = \mathcal{F}[f](\omega) := \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{i\omega x}\,dx, \tag{11.15a}$$

$$f(x) = \mathcal{F}^{-1}[f](x) := \int_{-\infty}^{\infty} \hat{f}(\omega)e^{-i\omega x}\,d\omega. \tag{11.15b}$$

Importantly, from the Fourier integral identity we have the the Fourier transform is a bijection (one-to-one and onto) on the space of square-integrable functions, i.e., $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ with inverse $\mathcal{F}^{-1}$. Also, $\hat{f}(\omega)$ represents the amplitude of the wave with wave number ω. Finally, the Fourier transform and its inverse are linear operators, i.e.,

$$\mathcal{F}[cf + g](\omega) = c\mathcal{F}[f](\omega) + \mathcal{F}[g](\omega) = c\hat{f}(\omega) + \hat{g}(\omega), \tag{11.16a}$$

$$\mathcal{F}^{-1}[c\hat{f} + \hat{g}](x) = c\mathcal{F}^{-1}[\hat{f}](x) + \mathcal{F}^{-1}[\hat{g}](x) = cf(x) + g(x). \tag{11.16b}$$

11.3 Inverse Fourier transform–Gaussian

Consider the Gaussian in Fourier space

$$\hat{g}(\omega) = e^{-\alpha\omega^2}, \qquad \alpha > 0. \tag{11.17}$$

Thus, by the Fourier integral identity it follows that

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} \hat{g}(\omega)e^{-i\omega x}\,d\omega \\ &= \int_{-\infty}^{\infty} e^{-\alpha\omega^2}e^{-i\omega x}\,d\omega \\ &= \int_{-\infty}^{\infty} e^{-\alpha\omega^2 - i\omega x}\,d\omega. \end{aligned}$$

We complete the square: $-\alpha\omega^2 - i\omega x = -\alpha(\omega^2 + ix\omega/\alpha) = -\alpha(\omega^2 + ix\omega/\alpha + (ix/2\alpha)^2 - (ix/2\alpha)^2) = -\alpha((\omega + ix/\alpha)^2 + x^2/4\alpha^2)$. Thus,

$$g(x) = e^{x^2/4\alpha}\int_{-\infty}^{\infty} e^{-\alpha(\omega+ix/\alpha)^2}\,d\omega.$$

Next, let $s = \omega + ix/\alpha$ and $ds = d\omega$. Thus,

$$\begin{aligned} g(x) &= e^{x^2/4\alpha}\int_{-\infty+ix/\alpha}^{\infty+ix/\alpha} e^{-\alpha s^2}\,d\omega \\ &= e^{x^2/4\alpha}\int_{①} e^{-\alpha s^2}\,ds. \end{aligned}$$

where ① is displayed in Fig. **??**. Then Cauchy's theorem in complex analysis says that if $f(z)$ in an analytic function and γ a simple closed contour we have

$$\int_{\gamma} f(z)\,dz = 0 \tag{11.18}$$

Thus,

$$\int_{①+②+③+④} e^{-\alpha s^2}\, ds = \left(\int_{①} + \int_{②} + \int_{③} + \int_{④} \right) e^{-\alpha s^2}\, ds = 0$$

Moreover, it can be shown that

$$\left(\int_{②} + \int_{④} \right) e^{-\alpha s^2}\, ds = 0.$$

Thus, letting the length of the horizontal edges tend to infinity it follows that

$$\int_{①} e^{-\alpha s^2}\, ds = -\int_{③} e^{-\alpha s^2}\, ds = \int_{-\infty}^{\infty} e^{-\alpha s^2}\, ds.$$

Hence,

$$g(x) = e^{-x^2/4\alpha} \int_{-\infty}^{\infty} e^{-\alpha s^2}\, ds.$$

Let $z = \sqrt{\alpha} s \implies dz/\sqrt{\alpha} = ds$. Thus,

$$g(x) = \frac{e^{-x^2/4\alpha}}{\sqrt{\alpha}} \int_{-\infty}^{\infty} e^{-z^2}\, dz.$$

Consider the following:

$$I = \int_{-\infty}^{\infty} e^{-z^2}\, dz.$$

This implies

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2}\, dy \right) \\ &= \int_{-\infty}^{\infty} e^{-(x^2+y^2)}\, dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r\, dr d\theta \\ &= \pi \int_{-\infty}^{0} e^{u}\, du \\ &= \pi \end{aligned}$$

Hence, it follows that $I = \sqrt{\pi}$. Thus, we have arrived at the following important result:

$$\mathcal{F}^{-1}[e^{-\alpha\omega^2}] = \sqrt{\frac{\pi}{\alpha}} e^{-x^2/4\alpha}. \tag{11.19}$$

Moreover, by the bijectivity of the Fourier transform it follows that

$$\mathcal{F}[e^{-\beta x^2}] = \frac{1}{\sqrt{4\pi\beta}} e^{-\omega^2/4\beta}. \tag{11.20}$$

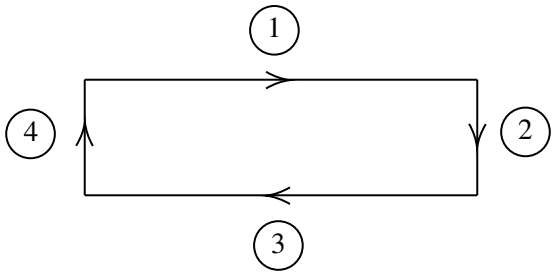

Figure 14: Integral contour in the complex plane

11.4 Fourier transform–derivatives

The key property of Fourier transforms (besides the bijectivity) is that it maps differentiation in the physical variables to multiplication in the Fourier (or spectral) variables. This is easy to confirm using integration by parts.

$$\begin{aligned}\mathcal{F}[f'](\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f'(x)e^{i\omega x}\,dx \\ &= \frac{e^{i\omega x}}{2\pi}f(x)\Big|_{-\infty}^{\infty} - i\omega\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{i\omega x}\,dx \\ &= -i\omega\mathcal{F}[f](\omega) \\ &= -i\omega\hat{f}(\omega).\end{aligned}$$

By induction it follows that

$$\mathcal{F}[f^{(n)}](\omega) = (-i\omega)^n\hat{f}(\omega). \tag{11.21}$$

11.5 Heat equation–infinite line

Let's see how to obtain the solution to the heat equation on the infinite line by using Fourier transforms. Consider the following Cauchy problem:

$$\begin{cases} u_t = \kappa u_{xx}, \quad -\infty < x < \infty, \quad t > 0, \\ u(x,0) = f(x) \end{cases} \tag{11.22}$$

Step one: Take the Fourier transform of the PDE in the spatial variable and treat time as a parameter. Thus,

$$\begin{aligned}\mathcal{F}[u_t] &= \mathcal{F}[\kappa u_{xx}] \\ \hat{u}_t &= \kappa(-i\omega)^2\hat{u} \\ \hat{u}_t &= -\kappa\omega^2\hat{u}.\end{aligned}$$

Thus, we have an ODE in the spectral variable:

$$\hat{u}_t = -\kappa\omega^2\hat{u}. \tag{11.23}$$

Solve the ODE: Since this is a first-order ODE we can solve using integrating factors:

$$\hat{u}(\omega,t) = c(\omega)e^{-\kappa\omega^2 t}. \tag{11.24}$$

Note that our integration constant is dependent on the spectral variable ω.

Invert the Fourier transform: Next, we use the inverse Fourier transform to obtain a representation of the solution

$$u(x,t) = \mathcal{F}^{-1}[\hat{u}(\omega,t)] = \int_{-\infty}^{\infty} c(\omega) e^{-\kappa\omega^2 t} e^{-i\omega x}\, d\omega. \tag{11.25}$$

Remark 11.1. Earlier in the Fourier transform section we obtained this formula using separation of variables, but we did not know how to obtain $c(\omega)$. We can now obtain $c(\omega)$ using the Fourier transform.

Use the initial data to obtain $c(\omega)$: Finally, consider the initial condition which gives

$$u(x,0) = \int_{-\infty}^{\infty} c(\omega) e^{-i\omega x} = \mathcal{F}^{-1}[c(\omega)](x) = f(x).$$

Hence,

$$c(\omega) = \mathcal{F}[f(x)](\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{i\omega x}\, dx. \tag{11.26}$$

Thus, $c(\omega)$ is simply the Fourier transform of the initial data. Putting everything together gives

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\bar{x}) e^{i\omega\bar{x}}\, d\bar{x} \right) e^{-i\omega x} e^{-\kappa\omega^2 t}\, d\omega \\ &= \int_{-\infty}^{\infty} f(\bar{x}) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\kappa\omega^2 t} e^{-i\omega(x-\bar{x})}\, d\omega \right) d\bar{x} \\ &= \int_{-\infty}^{\infty} f(\bar{x}) \frac{1}{\sqrt{4\pi\kappa t}} e^{-(x-\bar{x})^2/4\kappa t}\, d\bar{x}. \end{aligned}$$

Thus,

$$u(x,t) = \int_{-\infty}^{\infty} f(\bar{x}) G(x,t;\bar{x},0)\, d\bar{x}, \tag{11.27}$$

where

$$G(x,t;\bar{x},0) = \frac{e^{-(x-\bar{x})^2/4\kappa t}}{\sqrt{4\pi\kappa t}} \tag{11.28}$$

is the influence, or Green's function.

11.6 Dirac delta

Consider a force $f(x)$ on a physical system. In order to isolate the effect of each point we decompose $f(x)$ into a linear combination of unit pulses of width Δx. Hence,

$$f(x) \approx \sum_j f(x_j)(\text{unit pulse starting at } x_j). \tag{11.29}$$

Equation (11.29) resembles the definition of an integral with Δx missing. Hence, we write

$$f(x) = \lim_{\Delta x \to 0} \sum_j f(x_j) \frac{\text{unit pulse}}{\Delta x} \Delta x. \tag{11.30}$$

This gives a rectangular pulse of width Δx and height $1/\Delta x$ (unit area). In the limit $\Delta x \to 0$ it approaches an infinitely concentrated pulse (not an actual function) denoted $\delta(x - x_j)$ which is zero everywhere except at $x = x_j$ where it is infinite, still with unit area. Thus,

$$\delta(x - x_j) = \begin{cases} 0, & x \neq x_j \\ \infty, & x = x_j \end{cases} \tag{11.31}$$

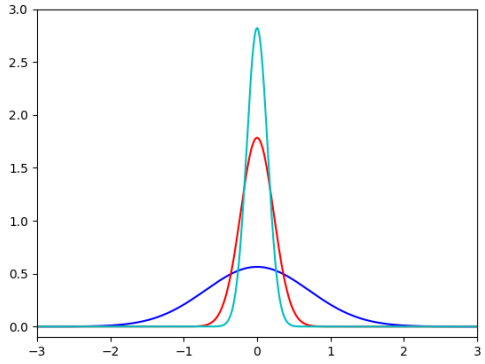

Figure 15: Sequence of concentrated pulse $\lim_{t\to 0^+} \frac{e^{-x^2}/t}{t\sqrt{\pi}}$.

and

$$\int_{-\infty}^{\infty} \delta(x - x_j)\, dx_j = 1, \tag{11.32}$$

and finally $\delta(x - x_j) = \delta(x_j - x)$. We call $\delta(x - x_j)$ the Dirac delta. Physically we can think of $\delta(x - x_j)$ as a concentrated source or impulsive force at $x = x_j$. Mathematically, the Dirac delta is a measure, or generalized function. For simplicity we define $\delta(x - x_j)$ as an operator with the property that for any continuous function $f(x)$:

$$f(x) = \int_{-\infty}^{\infty} f(x_j)\delta(x - x_j)\, dx_j. \tag{11.33}$$

In fact, the Dirac delta may be motivated by the limiting function of any sequence of concentrated pulses. In this derivation we took rectangular concentrated pulses. (For another example-relevant to the heat equation-see Fig. 15.) Importantly, the Dirac delta is so concentrated that it "sifts" out the value at $x = x_j$.

We note that the mathematical operation of integration adds smoothness, while the mathematical operation of differentiation removes smoothness. To see this more clearly consider the Heaviside step function:

$$H(x - x_j) := \begin{cases} 0, & x < x_j \\ 1, & x > x_j \end{cases} \tag{11.34}$$

which has a finite jump discontinuity. Interestingly,

$$\frac{d}{dx} H(x - x_j) = \delta(x - x_j), \tag{11.35}$$

that is the derivative of the Heaviside step function is the Dirac delta.

It turns out that the Dirac delta is related to Green's functions. What is a Green's function? Mathematically, it is the kernel of an integral operator that represents the inverse of a differential operator; physically, it is the response of a system when a unit point source is applied to the system. For example, it is the electric field induced by a single point charge.

Let's consider a linear ODE written in operator notation:

$$(Lu)(x) = f(x). \tag{11.36}$$

Formally, inverting the operator then gives the solution, i.e.,

$$u(x) = (L^{-1} f)(x). \tag{11.37}$$

Naturally, the inverse of a differential operator should be an intergral operator. It turns out that

$$u(x) = (L^{-1}f)(x) = \int_a^b f(x_o)G(x, x_o)\mathrm{d}x_o, \tag{11.38}$$

where $G(x, x_o)$ is the kernel of the integral operator, or the Green's function. Suppose that $f(x)$ is concentrated at some point $x = x_s$. Hence, $f(x) = \delta(x - x_s)$. Then the response at x, i.e., $u(x)$, satisfies

$$u(x) = \int_a^b \delta(x_o - x_s)G(x, x_o)\, dx_o = G(x, x_s). \tag{11.39}$$

Thus, the Green's function $G(x, x_s)$ physically is the response at x due to a concentrated source at x_s. Thus it must be that

$$(LG)(x, x_s) = \delta(x - x_s). \tag{11.40}$$

Thus, the Green's function is the solution to the ODE when the external forcing term is the Dirac delta. In fact a simple computation shows

$$\begin{aligned}(Lu)(x) &= L\Big(\int_a^b f(x_o)G(x, x_o)\, dx_o\Big)\\ &= \int_a^b f(x_o)(LG)(x, x_o)\, dx_o\\ &= \int_a^b f(x_o)\delta(x - x_o)\, dx_o\\ &= f(x).\end{aligned}$$

11.7 Laplace's equation in the upper half plane

In this section we consider the solution to the following boundary value problem:

$$\begin{cases} u_{xx} + u_{yy} = 0 \\ u(x, 0) = f(x) \end{cases} \tag{11.41}$$

where $f(x) \to 0$ as $x \to \pm\infty$. The assumption $f(x) \to 0$ as $x \to \pm\infty$ implies that there are three other "boundary conditions", namely,

- $\lim_{x\to\infty} u(x, y) = 0$,
- $\lim_{x\to-\infty} u(x, y) = 0$,
- $\lim_{y\to\infty} u(x, y) = 0$.

Since there are homogeneous BCs as $x \to \pm\infty$ we try the Fourier transform:

$$\begin{aligned}\mathcal{F}[u_{xx} + u_{yy}] &= \mathcal{F}[u_{xx}] + \mathcal{F}[u_{yy}]\\ &= (-i\omega)^2\hat{u}(\omega, y) + \hat{u}_{yy}(\omega, y)\end{aligned}$$

Thus, we get the ODE:

$$\hat{u}_{yy} - \omega^2\hat{u} = 0. \tag{11.42}$$

Remark 11.2. Note that we took the Fourier transform in x variable since the data is given along the x-axis, i.e., we know $u(x,0) = f(x)$.

Thus, solving the ODE gives us

$$\hat{u}(\omega, y) = a(\omega)e^{\omega y} + b(\omega)e^{-\omega y}. \tag{11.43}$$

We note the following:

1. $\hat{u}(\omega, 0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,0)e^{i\omega x}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{i\omega x}\,dx$
2. $\lim_{y\to\infty}\hat{u}(\omega, y) = \lim_{y\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,y)e^{i\omega x}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{y\to\infty} u(x,y)e^{i\omega x}\,dx = 0.$

From point two above $\hat{u}(\omega, y) \to 0$ as $y \to \infty$ it follows that

$$\hat{u}(\omega, y) = \begin{cases} a(\omega)e^{\omega y}, & \omega < 0 \\ b(\omega)e^{-\omega y}, & \omega > 0 \end{cases} \tag{11.44}$$

Thus,

$$\hat{u}(\omega, y) = c(\omega)e^{-|\omega| y}, \quad \omega \in \mathbb{R}. \tag{11.45}$$

It then follows that

$$u(x, y) = \mathcal{F}^{-1}[c(\omega)e^{-|\omega| y}]. \tag{11.46}$$

In order to apply the convolution theorem, i.e., the inverse Fourier transform of the product of two Fourier transfroms is $1/2\pi$ times the convolution of two functions. Indeed,

$$e^{-|\omega| y} = \mathcal{F}[g(x,y)] \iff g(x,y) = \mathcal{F}^{-1}[e^{-|\omega| y}],$$

where $g(x,y)$ is a function to be determined. We compute the following integral:

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

Hence,

$$u(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\bar{x})g(x-\bar{x}, y)\,d\bar{x}, \tag{11.47}$$

where

$$g(x-\bar{x}, y) = \frac{2y}{(x-\bar{x})^2+y^2} \tag{11.48}$$

Thus, the solution to Laplace's equation in the upper half plane is given by

$$u(x,y) = \frac{y}{\pi}\int_{-\infty}^{\infty}\frac{f(\bar{x})}{(x-\bar{x})^2+y^2}\,d\bar{x}. \tag{11.49}$$

Remark 11.3. Note that although we initially assumed $f(x) \to 0$ as $x \to \pm\infty$ all we need in order for the solution to be defined is that the integral in (11.49) be convergent. This is a common practice in PDEs: Begin by making certain strong assumptions so that a solution formula can be derived. Then try to weaken these assumptions.

11.8 Multidimensional Fourier transform–Bessel potential

Since we derived the Fourier transform from the method of separation of variables it is natural that the Fourier transform would exist in higher dimensions. Since the process of deriving the multi-dimensional Fourier transform is completely analagous to the one dimensional case here we just provide the definition and present an interesting example.

Multi-dimensional Fourier transform pair: Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$. Then

$$\hat{f}(\boldsymbol{\omega}) = \mathcal{F}[f](\boldsymbol{\omega}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(\mathbf{x}) e^{i\boldsymbol{\omega}\cdot\mathbf{x}}\, d^n\mathbf{x}, \tag{11.50a}$$

$$f(\mathbf{x}) = \mathcal{F}^{-1}[f](\mathbf{x}) := \int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\omega}) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\, d^n\boldsymbol{\omega}, \tag{11.50b}$$

where $\boldsymbol{\omega}\cdot\mathbf{x} = \omega_1 x_1 + \dots \omega_n x_n$ is the dot product. Thus, this is just an iterated integral as illustrated in the following important Fourier transform identity. Recall the Laplacian

$$\Delta f = (\partial_{x_1}^2 + \cdots + \partial_{x_2}^2) f = f_{x_1x_1} + \cdots f_{x_nx_n}.$$

We have the following:

$$\begin{aligned}
\mathcal{F}[\Delta f](\boldsymbol{\omega}) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} (\Delta f)(\mathbf{x}) e^{i\boldsymbol{\omega}\cdot\mathbf{x}}\, d^n\mathbf{x} \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} (f_{x_1x_1} + \cdots + f_{x_nx_n}) e^{i(\omega_1x_1+\cdots+\omega_nx_n)}\, d^n\mathbf{x} \\
&= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (f_{x_1x_1} + \cdots + f_{x_nx_n}) e^{i(\omega_1x_1+\cdots\omega_nx_n)}\, dx_1 \cdots dx_n \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f_{x_1x_1} e^{i(\omega_1x_1+\cdots+\omega_nx_n)}\, d^n\mathbf{x} + \cdots + \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f_{x_nx_n} e^{i(\omega_1x_1+\cdots+\omega_nx_n)}\, d^n\mathbf{x} \\
&= [(-i\omega_1)^2 + \cdots (-i\omega_n)^2]\hat{u}(\boldsymbol{\omega}) \\
&= -|\boldsymbol{\omega}|^2 \hat{u}(\boldsymbol{\omega})
\end{aligned}$$

where $|\boldsymbol{\omega}|^2 = \omega_1^2 + \cdots \omega_2^2$.

As an application of Fourier transforms in multiple dimensions we consider the following problem: Solve the following PDE:

$$-\Delta u + u = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n. \tag{11.51}$$

Begin by taking Fourier transforms:

$$\mathcal{F}[-\Delta u + u] = \mathcal{F}[f] \iff -\mathcal{F}[\Delta u] + \hat{u} = \hat{f}.$$

Note that unlike in the previous examples here we are taking the Fourier trasform of all indenpendent variables. Moreover,

$$-\mathcal{F}[\Delta u] = |\boldsymbol{\omega}|^2.$$

Thus, we get the equation

$$|\boldsymbol{\omega}|^2\hat{u} + \hat{u} = \hat{f} \iff \hat{u} = \frac{\hat{f}}{1 + |\boldsymbol{\omega}|^2}.$$

Thus,

$$u(\mathbf{x}) = \mathcal{F}^{-1}\left[\frac{\hat{f}}{1 + |\boldsymbol{\omega}|^2}\right]. \tag{11.52}$$

As usualy we apply the convolution theorem by noting that

$$\frac{1}{1 + |\boldsymbol{\omega}|^2} = \mathcal{F}[B(\mathbf{x})] \iff B(\mathbf{x}) = \mathcal{F}^{-1}\left[\frac{1}{1 + |\boldsymbol{\omega}|^2}\right],$$

where $B(\mathbf{x})$ is a function to be determined. Begin by noting that $\frac{1}{a} = \int_0^\infty e^{-a\tau}\, d\tau$ for each $a > 0$. Thus,

$$\begin{aligned} B(\mathbf{x}) &= \int_{\mathbb{R}^n} \frac{e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}}{1 + |\boldsymbol{\omega}|^2}\, d^n\boldsymbol{\omega} \\ &= \int_{\mathbb{R}^n} \left(\int_0^\infty e^{-(1+|\boldsymbol{\omega}|^2)\tau}\, d\tau \right) e^{-i\boldsymbol{\omega}\cdot\mathbf{x}}\, d^n\boldsymbol{\omega} \\ &= \int_0^\infty e^{-\tau} \left(\int_{\mathbb{R}^n} e^{-\tau|\boldsymbol{\omega}|^2 - i\boldsymbol{\omega}\cdot\mathbf{x}}\, d^n\boldsymbol{\omega} \right) d\tau \\ &= \int_0^\infty e^{-\tau} \left(\frac{\pi}{\tau}\right)^{n/2} e^{-|\mathbf{x}|^2/4\tau}\, d\tau \end{aligned}$$

Thus,

$$B(\mathbf{x}) = \pi^{n/2} \int_0^\infty \frac{e^{-\tau - |\mathbf{x}|^2/4\tau}}{\tau^{n/2}}\, d\tau, \tag{11.53}$$

which is known as the **Bessel potential**. Hence, by the convolution theorem:

$$u(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(\bar{\mathbf{x}}) B(\mathbf{x} - \bar{\mathbf{x}})\, d^n\bar{\mathbf{x}}. \tag{11.54}$$

11.9 Sine and cosine transforms