MATH 107: Introduction to Partial Differential Equations

Sam K. Miller

Spring 2021

Contents

1 Introduction

1.1 What is a PDE?

What is a *Partial Differential Equation*, or a *PDE*? Recall that in your previous Ordinary Differential Equations (ODEs) course, you may have solved problems such as the following:

$$f''(x) - 3f'(x) + xf(x) = 10.$$

A partial differential equation is similarly a differential equation - the difference is that, similar to how *partial* derivatives involve differentiating a multivariable function, a **partial differential equation** involves solving a differential equation in which there are multiple independent variables.

Definition: A **partial differential equation (PDE)** is an equation which relates a multivariable function with its partial derivatives.

For example, for a function $u(x, t)$,

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} = 0$$

is a partial differential equation. A *solution* of this PDE would be some function which satisfies the equation.

PDEs describe many physical systems and phenomena, some of which we will explore in further detail in this course. There are a wide range of classes of PDEs, some of which have rich techniques for finding solutions, others of which are far beyond the scope of this course, and some which have no sufficiently "nice" solutions at all! In fact, a large majority of PDEs can only be solved numerically, if at all. PDEs can describe the spreading of heat in a conducting medium, they can describe the diffusion of a pollutant in a liquid, or can describe the waves propagating from a disturbance, such as shocks caused from earthquakes, the vibrations of a drum after being struck, or a stone being dropped in a still pond.

The mathematics of PDEs includes the formulation of techniques to find solutions, as well as the development of theoretical tools and results that address the properties of solutions, such as **uniqueness** and **existence**. The scope of this course will be primarily *linear PDEs* (a definition will soon come), which exhibit friendlier behavior. By the end of this course, my hope is to present to you the tools necessary to find the solutions of many types of linear PDEs as well as to describe those solutions. Along the way, I also hope for you to gain some appreciation of the wide range of applications of PDEs, and for you to be able to discuss in some depth one of those applications.

1.2 Linear vs. Nonlinear PDEs

For the rest of this course, we will denote partial derivatives via subscripts, i.e.for a *dependent variable* $u(x, y)$ with *independent variables* x, y,

$$\frac{\partial u}{\partial x} = u_x.$$

Example: Some examples of PDEs, all of which occur in physical theory, and some of which we will study further in this course, are:

1. $u_t + (\alpha + \beta u)u_x = \gamma u_{xxx}$ (Korteweg de Vries equation)
2. $u_t = \kappa \nabla^2 u (= \kappa(u_{xx} + u_{yy} + \cdots))$ (Heat diffusion equation)
3. $u_{tt} = c^2 \nabla^2 u$ (Wave equation)
4. $\nabla^2 u = 0$ (Laplace's equation)
5. $\nabla^2 u = f(x, y, z)$ (Poisson's equation)
6. $u_x^2 + u_y^2 + u_z^2 = 1/c^2$ (Eikonal equation)

Note that for all of these PDEs, there are a number of dependent variables $(x, y, t, \cdots)$ but only one dependent variable u. We will primarily focus on PDEs with only one dependent variable.

Definition: The **order** of a PDE is the highest derivative that appears.

Remark: One may note that by moving all the terms of a PDE to one side, every order 1 PDE with 2 independent variables x, t and 1 dependent variable $u(x, t)$ may be expressed in the form

$$F(x, t, u(x, t), u_x(x, t), x_t(x, t)) = F(x, t, u, u_x, u_t) = 0,$$

for some function F. In this sense, this is the most general PDE of two independent variables of first order. Similarly, the most general second order PDE in two independent variables is

$$F(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}) = 0.$$

Remark: One may also note that if we isolate all the terms of a PDE containing the dependent variable u on one side, and all the terms *not* containing u on the other, we may express any PDE (in this case, with 2 independent variables x, t of first order) as

$$F(x, t, u, u_x, u_t) = g(x, t).$$

This representation of the PDE will become important soon.

Some of the aforementioned PDEs satisfy an additional friendly property, called *linearity*, which makes finding solutions for those PDEs considerably less difficult. Recall from Linear Algebra that a linear map between vector spaces is a function that satisfies some nice additive and multiplicative properties.

Definition: (the brief version) A **vector space** over a field (such as $\mathbb{R}$ or $\mathbb{C}$) is a set of vectors for which addition of vectors and scalar multiplication are defined, and satisfy associativity, commutativity, and distributivity.

Example: Denote by $C^n(\mathbb{R})$ (respectively, $C^n(\mathbb{C})$) the set of all continuous functions $\mathbb{R} \to \mathbb{R}$ (respectively, $\mathbb{C} \to \mathbb{C}$) whose 1st through nth derivatives are also continuous. Denote by $C^\infty(\mathbb{R})$ and $C^\infty(\mathbb{C})$ the set of **smooth functions** on $\mathbb{R}$ or $\mathbb{C}$, that is, the set of all of whose derivatives are continuous. These are vector spaces over $\mathbb{R}$ or $\mathbb{C}$.

Definition: A **linear transformation** over a vector space V over field F is a function $T : V \to V$ that satisfies the following properties:

1. $T(v + w) = T(v) + T(w)$ for all $v, w \in V$.
2. $T(cv) = cT(v)$ for all $v \in V, c \in F$.

Example: Over $C^n(\mathbb{R})$ with $n \geq 1$ or $C^\infty(\mathbb{R})$, the operator $\frac{d}{dx}$, $f(x) \mapsto f'(x)$ is a linear transformation. Similarly, partial derivatives are linear transformations over their analogous vector space.

Remark: In a sense, we can think of each PDE as a transformation (maybe linear) over some vector space (think to yourself what that vector space would be). Recall from above that every PDE concerning some dependent variable u may be expressible in the form

$$F(u, \dots) = g(\dots),$$

where g is some function of the independent variables of u, but not u itself or any of its partial derivatives. We may consider a map sending a function to another function:

$$\mathcal{L} : u \mapsto F(u, \dots).$$

This map can be thought of as a transformation over a vector space! In fact, this map is always explicitly expressible - see the below example to understand how.

Definition: If the map $\mathcal{L}$ described above is a linear operator, then the PDE is called a **linear equation**. Furthermore, if $g = 0$, then the PDE is called a **homogeneous linear equation**. Otherwise, the PDE is called an **inhomogeneous linear equation**.

Example: For the PDE $u_x + yu_y = x$ above,

$$\mathcal{L} = \frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.$$

The PDE is a *inhomogeneous linear first order* PDE.

Exercise: For each PDE listed above, classify its order and determine whether it is linear or nonlinear. If it is linear, then determine whether it is homogeneous or inhomogeneous.

A quick guide to determining linearity: in a linear PDE, the dependent variable (u and its derivatives) appears linearly! That is, every summand contains at most one copy of u or a derivative multiplied to it. Additionally, in a homogeneous PDE, every nonzero term involves at least one dependent variable. Alternatively, the PDE has solution $u = 0$.

Why do we care about homogeneous vs. inhomogeneous PDEs? Well, if we have a solution u_1 to homogeneous PDE defined by $\mathcal{L}$,

$$\mathcal{L}u = 0,$$

and solution u_2 to an inhomogeneous PDE defined by $\mathcal{L}$,

$$\mathcal{L}u = g(x),$$

then $u_1 + u_2$ is again a solution to the inhomogeneous PDE! This is known as the **principle of superposition**.

Enough classification, let's get to solving some PDEs, shall we?

1.3 Our First PDEs + ICs & BCs

We begin with a first order PDE.

Exercise: Find solutions to the first order, linear, homogeneous PDE

$$u_t + u_x = 0,$$

where $u = u(x, t)$, by thinking of non-constant functions that make this equation true.

Solution: Playing around with this equation, we see that $x - t$ is a solution. If we are a bit savvy, we may notice that for any function f, $f(x-t)$ is a solution as well! This means there is a large class of functions which satisfy this PDE. We will see that indeed, all solutions of this PDE are of that form. Now, let's move on to a second order PDE.

Exercise: Let $u = u(x, y)$.

1. Find a solution to the PDE $u_{yy} = 0$ in the domain $-\infty < x < \infty$, $y > 0$.

2. Is this *really* a PDE??

3. Can you find *all* solutions to this PDE and justify that you've found all of them?

Solution: To answer part 3, if we have $u_{yy} = 0$, apply $\int dy$ to both sides to obtain

$$u_y = f(x).$$

Then do it again to obtain

$$u(x, y) = yf(x) + g(x)$$

to obtain the general solution to the PDE!

Again, we find existence of a PDE but not uniqueness - we've found a class of solutions, and two arbitrary functions in the general solution! In general, this will happen unless we specify boundary conditions (BCs) and initial conditions (ICs). The solutions to problems we study may be heavily reliant on ICs or BCs.

Definition: An **initial condition** specifies the physical state of a system at a particular time t_0, for example, for a function $u(x, t)$, $u(x, t_0) = g(x)$. A **boundary condition** specifies the physical state of a system at a specific boundary, for example for a function $u(x, y)$ defined on $y \geq 0$, $u(x, 0) = f(x)$. **Dirichlet boundary conditions** specify a condition for u, while **Neumann boundary conditions** specify a condition for a derivative of u.

Definition: An IC/BC is **homogeneous** if it is of the form $u(0, \dots) = 0$, and **inhomogeneous** otherwise.

Definition: An **initial value problem** (IVP) consists of a PDE with ICs or BCs.

Example: The unique solution to the IVP consisting of the previous PDE with boundary conditions $u(x, 0) = g(x)$ and $u_y(x, 0) = f(x)$ is $u = yf(x) + g(x)$.

Remark: Again, one may note that if u, v are solutions of a homogeneous PDE with homogeneous BCs, then $au + bv$ is also a solution for all scalars a, b, again demonstrating the **principle of superposition**. However, with inhomogeneous ICs or BCs, the principle of superposition *may no longer be satisfied.*

In general we typically impose as many boundary conditions or initial conditions as the order of the PDE to yield a unique solution. The moral of the story here is that general PDEs can have a large class of solutions without BCs or ICs, but with the right amount of initial data, we can find a uniquely determined solution. On the other hand, if there is too much

initial data, we can have an overdetermined solution, and will have to make a choice about how to proceed.

2 First Order PDEs & Characteristics

2.1 Introduction

Let's look at some slightly more difficult PDEs which we cannot solve at first glance. Let $u = u(x,t)$, $c \in \mathbb{R}$ a nonzero constant, and solve:

$$u_t + cu_x = 0, \qquad u(x,0) = f(x).$$

We present two ways of solving this problem and thinking about its general solution:

1. **Algebraic:** We introduce a (wisely chosen) change of coordinates:

   $$v(r,s) = u(x,t), \quad \text{where } r = t, s = x - ct,$$

   and perform a change of variables using the chain rule. (How we obtained these choices of r, s will be explained soon)

   $$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial t} = \frac{\partial v}{\partial r} \cdot \frac{\partial r}{\partial t} + \frac{\partial v}{\partial s} \cdot \frac{\partial s}{\partial t} = \frac{\partial v}{\partial r} \cdot 1 + \frac{\partial v}{\partial s} \cdot (-c)$$

   $$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial x} = \frac{\partial v}{\partial r} \cdot \frac{\partial r}{\partial x} + \frac{\partial v}{\partial s} \cdot \frac{\partial s}{\partial x} = \frac{\partial v}{\partial r} \cdot 0 + \frac{\partial v}{\partial s} \cdot 1$$

   Therefore, the original PDE is equivalently stating:

   $$0 = u_t + cu_x = (v_r - cv_s) + cv_s = v_r$$

   Thus we have turned the PDE into an ODE, $\frac{dv}{dr} = 0$, with adjusted initial condition $v(0,s) = f(s)$. This solves to

   $$v(r,s) = f(s) = f(x - ct) = u(x,t).$$

2. **Geometric:** The quantity $u_t + cu_x$ is the directional derivative of u in the direction of the vector $V = (c, 1)$, and the PDE specifies that said directional derivative is always 0. Therefore, $u(x,t)$ is constant in the direction of V, in other words, along the line moving in the direction in V. We may then note that any line parallel to V must have that the value $(x - ct)$ is constant (this may be computed by noting that $(1, -c)$ is orthogonal to $(c, 1)$). Indeed, we verify, for a line with parameterization $l(x,t) = (a,b) + t(c,1)$:

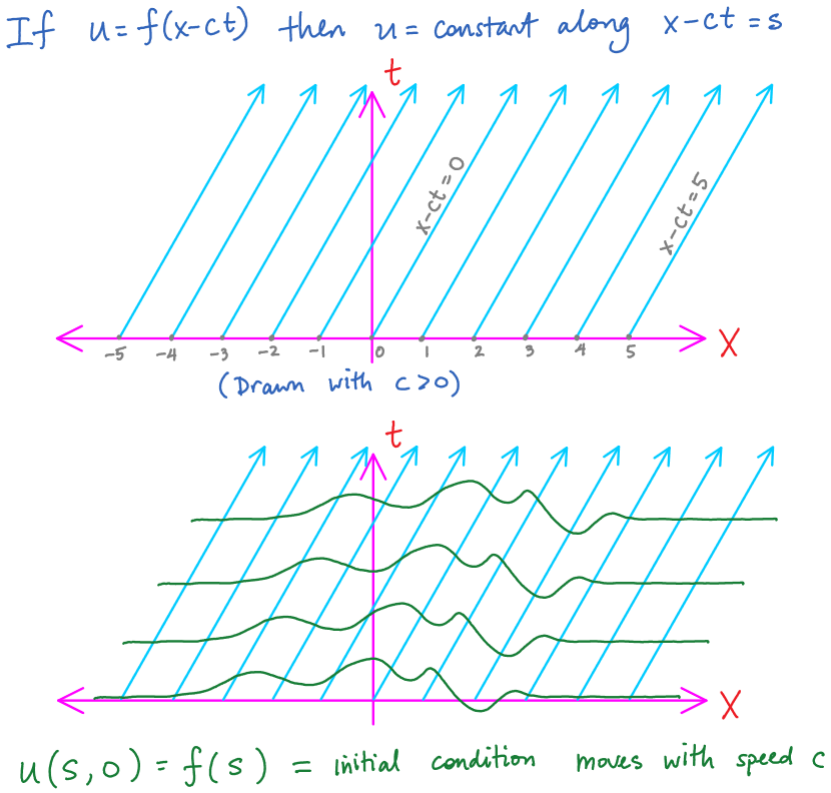
   $$(x - ct) = a + tc - c(b + t) = a - cb.$$

   So we see the value $(x - ct)$ is independent of t, as desired. These are called the **characteristic lines** or **curves** of the PDE, and we will take advantage of these to solve first order PDEs.

Since the solution is constant on each line, $u(x,t)$ only need depend on $s = x - ct$. Taking the initial condition into account, the solution must be

$$u(x,t) = f(x - ct),$$

for any smooth function f.



Definition: A **characteristic curve** is a curve along which boundary data or initial data propagates into the domain of the PDE. Solving PDEs using characteristic curves is known as using the **method of characteristics.**

Remark: By identifying a characteristic curve and restricting the problem to it, we will find that solving the PDE effectively becomes nothing more than solving a system of ODEs. Then, extending the characteristic curves over the entire space obtains a full solution (hopefully). While PDEs may not have constant behavior along a characteristic curve, they will generally have "nice" behavior.

2.2 The Method of Characteristics

Here we present two strategies for solving more general problems using the method of characteristics, both of which are valid, but one of which may be easier depending on the context. Ultimately it is up to personal preference which you choose to use, but it is best to understand

both. Suppose we wish to solve the following PDE:

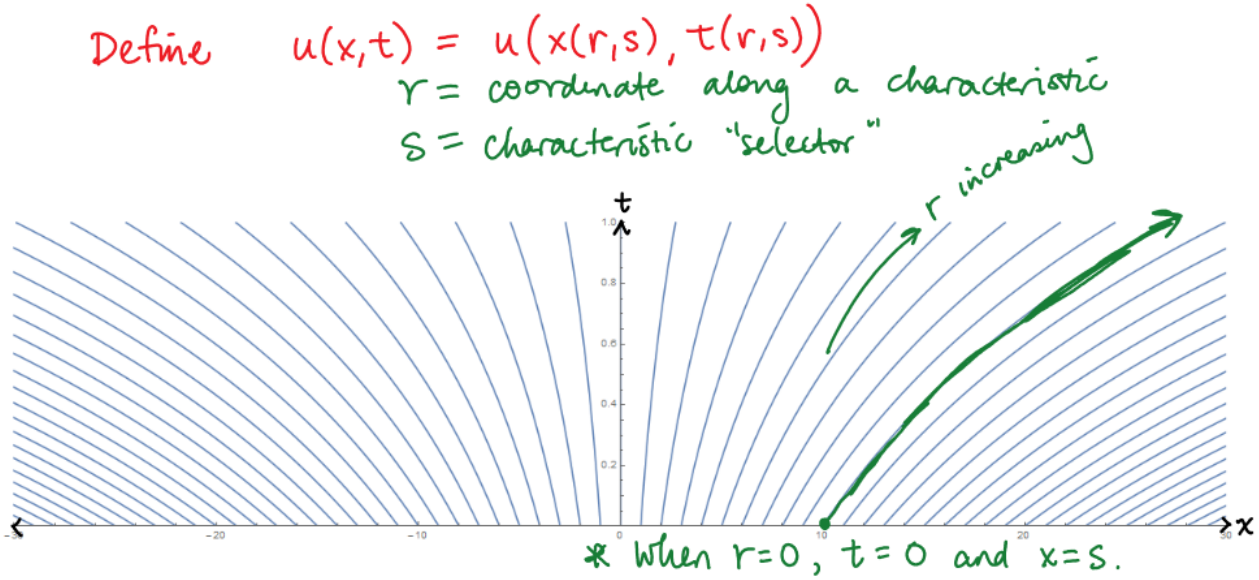
$$\frac{\partial u}{\partial t} + c \cdot \frac{\partial u}{\partial x} = g, u(x,0) = f(x) \quad \text{with } c, r \text{ functions of } x, t, u.$$

1. **Geometric:** Suppose $(x(t), t)$ is a smooth curve paremeterized by t, then observe that by the chain rule,
$$\frac{d}{dt}u(x(t),t) = u_t + \frac{dx}{dt}u_x.$$
Let's set $\frac{dx}{dt} = c$, then considering u only as a function of t, we have a system of ODEs:
$$\frac{dx}{dt} = c, \quad \frac{du}{dt} = g.$$
To set initial conditions, let us suppose $x(0) = s$, then $u(0) = f(s)$. Here, x is the characteristic curve passing through $x = s, t = 0$. In this sense, s propagates a characteristic curve over the entire space, by moving along the space where the initial condition is determined. Solve for x and u, these will be expressed in terms of s, then use the relation between x and s to express u in terms of x, eliminating s from the final expression of u.

   This is the technique taught in Shearer & Levy (as well as the one on Wikipedia, I think)

2. **Algebraic:** Define $u(x,t) = u(x(r,s), t(r,s))$. r will be the parameterization of the characteristic, and s will be the characteristic selector. s again is the Lagrangian variable (the "characteristic selector"), and r will describe the coordinate along a characteristic. This characterization gives us initial conditions as follows: s moves along the curve determined by the initial condition, so when $r = 0$, $t = 0$ and $x = s$. We have via the chain rule that:
$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial t}\frac{\partial t}{\partial r} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial r}$$
We wish to make this match the original PDE, which yields the following system of ODEs:
$$\frac{\partial t}{\partial r} = 1, \quad t|_{r=0} = 0$$
$$\frac{\partial x}{\partial r} = c, \quad x|_{r=0} = s$$
$$\frac{\partial u}{\partial r} = g, \quad u|_{r=0} = f(s)$$
As before, we solve this system and express u in terms of x and t.

Remark: Algebraically, characteristics can be thought of as the curves along which the PDE reduces to an ODE. Geometrically, characteristics are curves along which the boundary and/or initial data propagate into the solution domain.



Let's put these strategies to work!

Exercise: : Solve the PDE:

$$\frac{\partial u}{\partial t} + x\frac{\partial u}{\partial x} = 1, \quad u(x,0) = \sin(x), \quad -\infty < x < \infty, t > 0.$$

Solutions using the two methods:

1. Consider a characteristic curve of u, $(x(t), t)$, in other words, it satisfies

   $$u'(x(t), t) = 1 = u_t + xu_x.$$

   We set up the system of ODEs as derived previously:

   $$\frac{dx}{dt} = x, \quad x(0) = s, \qquad \frac{du}{dt} = 1, \quad u(0) = \sin(s).$$

   Solving these ODEs yields:

   $$x(t) = se^t, \quad u(t) = \sin(s) + t.$$

   Then, substituting $x(t)$ in, and eliminating s allows us to conclude that

   $$u(x,t) = \sin(xe^{-t}) + t$$

2. Suppose u is a solution to the PDE. Write $u(x,t) = u(x(r,s), t(r,s))$, where r is the coordinate along a characteristic curve and s is the Lagrangian variable. Then we have:

   $$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial t}\frac{\partial t}{\partial r} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} = u_t + xu_x = 1.$$

This gives us the following system of ODEs:

$$\frac{dt}{dr} = 1, \quad t(0) = 0, \qquad \frac{dx}{dr} = x, \quad x(0) = s, \qquad \frac{du}{dr} = 1, \quad u(0) = \sin(s).$$

Solving these yields the following respective solutions:

$$t = r, \qquad x = se^r, \qquad u = \sin(s) + r.$$

Putting these together yields:

$$u(x,t) = \sin(xe^{-t}) + t,$$

and check to make sure this works.

Let's look at another example that seems similar, but illustrates the fact that global closed form solutions may not necessarily exist.

Example: Let $u = u(x,y)$. Solve the first order PDE

$$u_y + xu_x = -\alpha u, \quad x, y > 0, \alpha > 0, \quad u(x, 1/x) = f(x)$$

As before, we let $u(x,y) = u(x(r,s), y(r,s))$, with s the characteristic selector. Then we have:

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = -\alpha u$$

As before, this gives us three ODEs to solve:

$$\frac{\partial x}{\partial r} = x, \quad \frac{\partial y}{\partial r} = 1, \quad , \frac{\partial u}{\partial r} = -\alpha u.$$

We determine the initial conditions from BC we are given - note that the BC runs along the curve $(x, 1/x)$. This translates to:

$$x|_{r=0} = s, \quad , y|_{r=0} = \frac{1}{s}, \quad u|_{r=0} = f(s).$$

This yields solutions to the ODEs:

$$x = se^r, \quad y = r + \frac{1}{s}, \quad u = f(s)e^{-\alpha r},$$

and putting this all together yields:

$$x = se^{y-\frac{1}{s}}, \quad u = f(s)e^{-\alpha(y-\frac{1}{s})}.$$

Verify the solution!

Remark: There are some things we must note here. Most importantly, there is no global closed-form solution - the solution is *implicitly defined*, that is given x and y, we may solve for s numerically. If you've taken a course in Analysis, you may recall the *implicit function theorem*, which states that in small neighborhoods, closed form solutions may exist. Next, we note that the value of u is not constant along the characteristic curves, since this is a inhomogeneous PDE. Last, we note that we may extend a solution to $y < 0$; whether solutions are extendable to a larger space is an important question!

Exercise: Solve $u_t + tu_x = x, u(x,0) = f(x)$ on $-\infty < x < \infty, t > 0$, and describe the characteristics of the PDE.

Solution: As usual, we write $u(x,t) = u(x(r,s), t(r,s))$ and expand out the u_r term to obtain system of PDEs:

$$\begin{aligned} \frac{\partial t}{\partial r} = 1, \quad t|_{r=0} = 0 \quad &\implies t = r \\ \frac{\partial x}{\partial r} = t, \quad x|_{r=0} = s \quad &\implies x = \frac{1}{2}r^2 + s \\ \frac{\partial u}{\partial r} = x, \quad u|_{r=0} = f(s) \quad &\implies u = \frac{1}{6}r^3 + sr + f(s) \end{aligned}$$

We may derive $s = x - \frac{1}{2}t^2$, and recalling the role of s as a characteristic selector, we may note that $s = x - \frac{1}{2}t^2$ describes the characteristic curves - parabolas. Putting this all together, we find:

$$u(x,t) = \frac{1}{6}t^3 + \left(x - \frac{1}{2}t^2\right)t + f\left(x - \frac{1}{2}t^2\right).$$

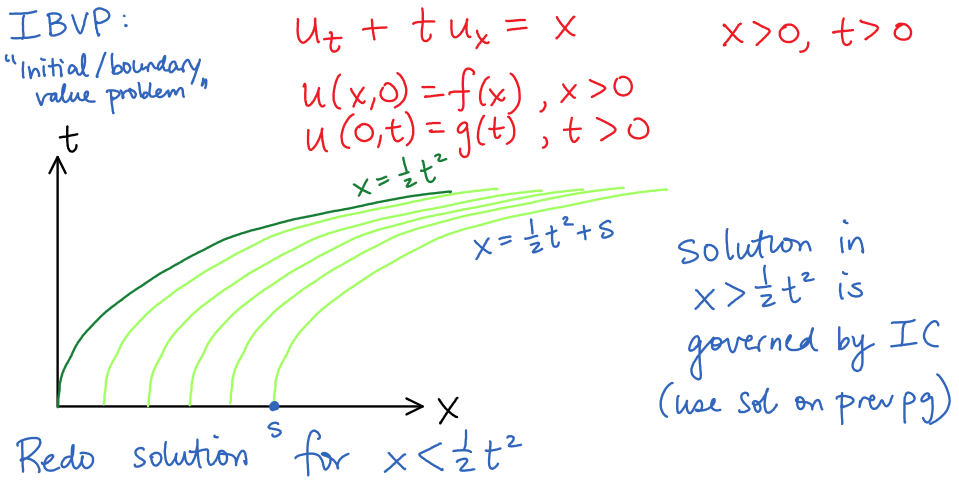
Let's change this problem up slightly by making the ICs/BCs give a bit more information.

Example: Solve $u_t + tu_x = x$ subject to ICs/BCs: $u(x,0) = f(x), x > 0$ and $u(0,t) = g(t), t > 0$ on the region $x > 0, t > 0$.

We may note that *the solution of a PDE along a characteristic curve is governed entirely by the curve and the ICs/BCs for which that curve intersects.* Recalling that in the previous version of the problem, the characteristic curves were given by

$$x = \frac{1}{2}t^2 + s,$$

we note that when $x > \frac{1}{2}t^2$, the characteristic curve intersects the same BC as the previous problem, so along these curves, the solution is the same as before. Therefore, we only must perform computations when $x < \frac{1}{2}t^2$.



Performing this computation yields the following system of ODEs - note that the ICs are different (determined by the new IC), but the ODEs themselves are unchanged!

$$\frac{\partial t}{\partial r} = 1, \quad t|_{r=0} = s \quad \Longrightarrow \ t = r + s$$
$$\frac{\partial x}{\partial r} = t, \quad x|_{r=0} \quad \Longrightarrow \ x = \frac{1}{2}r^2 + rs$$
$$\frac{\partial u}{\partial r} = x, \quad u|_{r=0} = g(s) \ \Longrightarrow \ u = \frac{1}{6}r^3 + \frac{1}{2}r^2 s + g(s)$$

There's a bit more work to be done here. $s = t - r$ implies

$$x = \frac{1}{2}r^2 + r(t-r) = -\frac{1}{2}r^2 + rt.$$

Solving the quadratic yields

$$r = t \pm \sqrt{t^2 - 2x},$$

recalling $t^2 - 2x > 0$, we obtain

$$s = t - r = \mp\sqrt{t^2 - 2x}.$$

However, $s > 0$, since $t > 0$ and s is the IC of t, so we choose

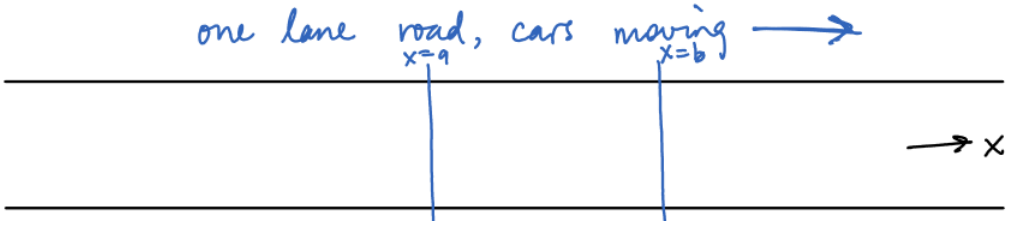
$$s = \sqrt{t^2 - 2x}, r = t - \sqrt{t^2 - 2x},$$

and hence we obtain that if $x < \frac{1}{2}t^2$,

$$u = \frac{1}{6}(t - \sqrt{t^2 - 2x})^3 + \frac{1}{2}(t - \sqrt{t^2 - 2x})^2\sqrt{t^2 - 2x} + g(\sqrt{t^2 - 2x}).$$

2.3 Application: Traffic Flow Modeling

Most PDEs are derived from physical principles like Newton's 2nd Law, or conservation principles. We provide an example, where we derive a PDE which describes traffic flow. Let us consider a one lane road.



Let $\rho(x,t)$ measure the density of cars at a point x at time t, a unit of *cars per mile* (it is always important to ask in a model what the units of any variable are!). If we pick a control volume, $a < x < b$, then the number of cars in this control volume is given by:

$$\int_a^b \rho(x,t)\,dx = M(t).$$

Since cars are neither created nor destroyed (at least in this model), the only way $M(t)$ changes is by cars entering the control volume at $x = a$ or leaving at $x = b$. This principle allows us to consider the rate of cars passing through a fixed location at a given time.

Definition: The rate at which cars pass by a fixed location x is defined as the **flux** of cars, denoted by $q(x,t)$. The units of q are *cars per hour*.

Using the conservation law here, we may relate ρ and q by:

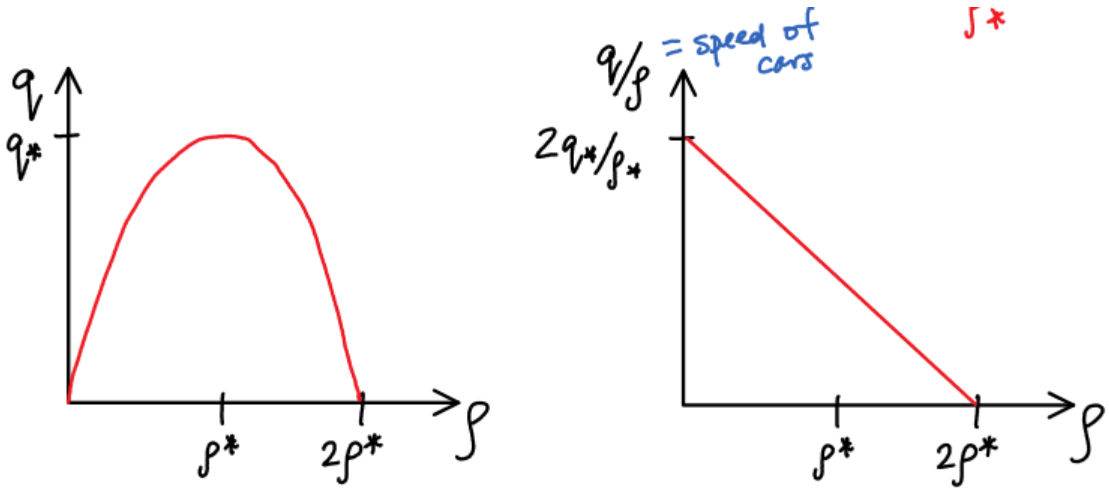
$$\frac{d}{dt}\int_a^b \rho(x,t)\,dx = q|_{x=a} - q|_{x=b}$$

In other words, the instantaneous change of the number of cars in the control volume is measured by the rate of cars entering the control volume at $x = a$, minus the rate of cars leaving the control volume at $x = b$. Question: what assumptions are we making here?

We now need a way of modeling the behavior of drivers. If we imagine there is a constant density of cars ρ (cars/mile) and flux q (cars/hour), then their speed is in fact, q/ρ (miles/hour)!

One traffic model is as follows: let q_* be the maximum flux of the roadway (it is a reasonable assumption that such a maximum flux exists), and let ρ_* be the optimal car density which obtains q_*. Then, we have:

$$q = q_* \frac{\rho(2\rho_* - \rho)}{\rho_*^2}.$$



We may compute that $2\rho_*$ is the density of a complete standstill of traffic, and $2q_*/\rho_*$ is the maximum speed on this road. As a logic check, this makes sense: we should expect an inverse relationship between speed of cars and car density.

As an (unrealistic) example, if an average car is 9 ft and in a complete standstill there is 1 ft between cars, then $2\rho_* = 1$ car/10 ft $= 528$ cars/mile. This gives us $\rho_* = 264$ cars/mile. Furthermore, if we assume the max speed of $2q_*/\rho_* = 30$ mph, then $q_* = 3960$ cars/hour.

How do we obtain a PDE from this? Recall, we have a relation in **integral form**:

$$\frac{d}{dt}\int_a^b \rho(x,t)\,dx = q|_{x=a} - q|_{x=b}.$$

Note that the right hand side can be expressed as

$$q|_{x=a} - q|_{x=b} = \int_b^a \frac{\partial q}{\partial x}\,dx,$$

and on the left hand side, we may reorder things:

$$\int_a^b \frac{\partial \rho}{\partial t}\,dx = \frac{d}{dt}\int_a^b \rho(x,t)\,dx.$$

Putting these together yields

$$\int_a^b \frac{\partial \rho}{\partial t}\,dx = \int_b^a \frac{\partial q}{\partial x}\,dx \iff \int_a^b \frac{\partial q}{\partial x} + \frac{\partial \rho}{\partial t}\,dx = 0.$$

However, the integral is zero *regardless* of a and b, so the only way this can be true is if

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = \rho_t + q_x = 0.$$

This is the **divergence form** of the conservation law.

Furthermore, assuming the flux is related to ρ by the simple model, i.e. $q = q_* \frac{\rho(2\rho_* - \rho)}{\rho_*^2}$, then we may substitute in the relation to obtain:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}\left[q_* \frac{\rho(2\rho_* - \rho)}{\rho_*^2}\right] = 0 \implies \rho_t + \frac{2q_*}{\rho_*^2}(\rho_* - \rho)\rho_x = 0.$$

This is a first order nonlinear PDE. Recall that ρ_* and q_* are constants which may vary by the road.

Exercise: Given the integral form of a conservation law:

$$\frac{d}{dt}\int_{x_1}^{x_2} u(x,t)\,dx + \frac{1}{2}[u(x_2,t)^2 - u(x_1,t)^2] = 0,$$

derive the corresponding divergence form of the conservation law.

Solution: This conservation law has flux function $q = \frac{1}{2}u^2$, so the PDE version is given by:

$$\frac{\partial u}{\partial t} + \frac{\partial q}{\partial x} = \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = u_t + uu_x = 0.$$

This is the **Inviscid Burgers' Equation**. It describes nonlinear wave propagation (and as we saw in the derivation, is a special case of the traffic flow modeling relation).

2.4 Shocks

So far, the PDEs we have encountered have behaved nicely, in that the characteristic curves of the PDEs have been nonintersecting. Is this guaranteed to always be the case? (Ponder...)

As it turns out, no, and the intersection of characteristics results in the breakdown of smooth solutions. We illustrate with an example following from the previously derived equation.

Exercise: Determine the characteristics of the PDE given by:

$$\rho_t + q_x = \rho_t + \frac{2q_*}{\rho_*^2}(\rho_* - \rho)\rho_x = 0, \quad \rho(x,0) = \begin{cases} 2\rho^* & x > 0 \\ \rho^* & x < 0 \end{cases}$$

We note that this this PDE could be a model of traffic on a road that is flowing normally at $x < 0$ and at a standstill at $x > 0$. As usual, write $\rho = \rho(x(r,s), t(r,s))$, so we have

$$\frac{\partial \rho}{\partial r} = \frac{\partial \rho}{\partial t}\frac{\partial t}{\partial r} + \frac{\partial \rho}{\partial x}\frac{\partial x}{\partial r}.$$

This yields the system of ODEs:

$$\frac{\partial t}{\partial r} = 1, \quad t|_{r=0} = 0 \quad \implies t = r$$

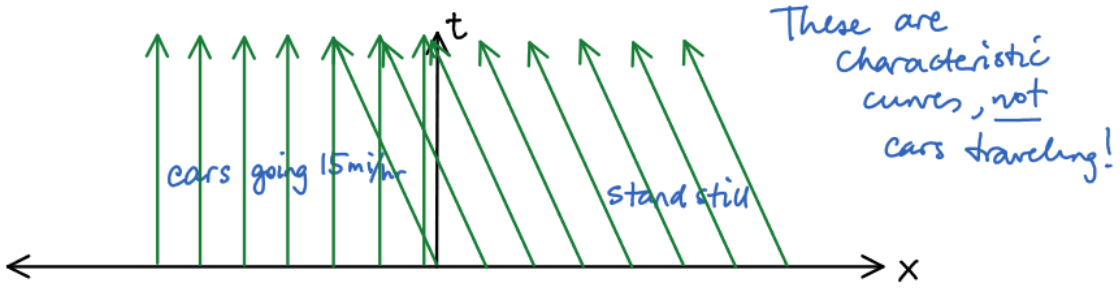
$$\frac{\partial \rho}{\partial r} = 0, \quad \rho|_{r=0} = \rho(s,0) = \begin{cases} 2\rho_* & s > 0 \\ \rho_* & s < 0 \end{cases} \implies \rho = \begin{cases} 2\rho_* & s > 0 \\ \rho_* & s < 0 \end{cases}$$

$$\frac{\partial x}{\partial r} = \frac{2q_*}{\rho_*^2}(\rho_* - \rho) = \begin{cases} \frac{2q_*}{\rho_*^2}(\rho_* - 2\rho_*) = -\frac{2q_*}{\rho_*} & s > 0 \\ \frac{2q_*}{\rho_*^2}(\rho_* - \rho_*) = 0 & s < 0 \end{cases}$$

So putting this all together, we yield:

$$x = \begin{cases} -\frac{2q_*}{\rho_*} r + s & s > 0 \\ s & s < 0 \end{cases}.$$

Thus the characteristic lines look something like this:



What happens when characteristic curves intersect? A **shock wave** forms, which causes the solution to be discontinuous. There are a few conceivable ways to rid ourselves of this issue. One may be to avoid intersections of characteristics, restricting only to a domain where the characteristics are consistent, perhaps in the example above, to only $x > 0$. A second option may be to allow our solution to the PDE to have discontinuities. A third solution could be to admit that the solution may only exist around $t = 0$, but then past some larger t breaks down.

Depending on the context of the problem, all three of these are plausibly acceptable responses. We will first focus here on computing the shock wave trajectory, in other words, describing where the discontinuity occurs. This shock trajectory divides the domain so the characteristics do not conflict within each domain.

Theorem: *Rankine-Hugoniot Jump Condition*: The shock speed, σ', is the ratio of the jump in flux, q, to the jump in the quantity being preserved, ρ. In other words,

$$\sigma' = \frac{[q]_-^+}{[\rho]_-^+} = \frac{q^+ - q^-}{\rho^+ - \rho^-}.$$

Proof. Let $\sigma(t)$ be the x coordinate of the shock. Then it still must be true that

$$\frac{d}{dt}\int_a^b \rho(x,t)\,dx = q(a,t) - q(b,t), \qquad \text{where } \rho, q = \begin{cases} \rho^+, q^+ & x > \sigma(t) \\ \rho^-, q^+ & x < \sigma(t) \end{cases}.$$

Then we have:

$$\frac{d}{dt}\int_a^b \rho(x,t)\,dx = \frac{d}{dt}\left(\int_a^{\sigma(t)} \rho^-(x,t)\,dx + \int_{\sigma(t)}^b \rho^+(x,t)\,dx\right)$$
$$= \rho^-(\sigma(t),t)\cdot\sigma'(t) + \int_a^{\sigma(t)} \frac{\partial \rho^-}{\partial t}dx + \int_{\sigma(t)}^b \frac{\partial \rho^+}{\partial t}dx - \rho^+(\sigma(t),t)\cdot\sigma'(t)$$

This follows from application of the chain rule. Now, let $a \to \sigma(t)$ from the left, and $b \to \sigma(t)$ from the right. Then, the left-hand side has limit

$$\text{LHS} \to (\rho^-(\sigma,t) - \rho^+(\sigma,t))\sigma'(t),$$

and the right-hand side has limit

$$\text{RHS} \to q^-(\sigma,t) - q^+(\sigma,t).$$

Thus, we have

$$\sigma'(t) = \frac{q^+(\sigma(t),t) - q^-(\sigma(t),t)}{\rho^+(\sigma(t),t) - \rho^-(\sigma(t),t)}$$

□

In the case of our example, we computed that $\rho^- = \rho_*$ and $\rho^+ = 2\rho_*$. Recall as well that in the traffic congestion model, that q and ρ had relation given by:

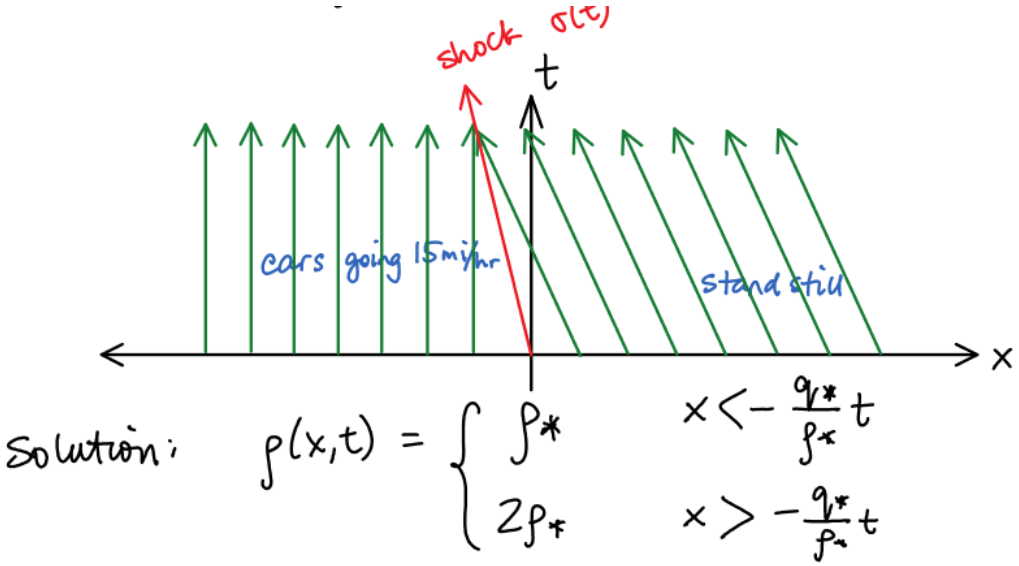
$$q = q_*\frac{\rho(2\rho_* - \rho)}{\rho_*^2}.$$

Therefore, $q^- = q_*$ and $q^+ = 0$. Thus, the Rankine-Hugoniot jump condition gives us that

$$\sigma'(t) = \frac{q^+ - q^-}{\rho^+ - \rho^-} = \frac{0 - q_*}{2\rho_* - \rho_*} = \frac{-q_*}{\rho_*}.$$

Since $\rho^\pm$ and $q^\pm$ are constant for all t, and the shock originates at $(x,t) = (0,0)$, we conclude

$$\sigma(t) = -\frac{q_*}{\rho_*}t.$$



Remark: How can we make sense of a function that is discontinuous or having discontinuous derivatives satisfying a differential equation? Especially when non-unique solutions may possibly arise? We introduce the notion of *test functions* and **weak solutions**. We introduce the notion for the integral conservation law posed above, and if we have time, may discuss weak solutions more generally later.

Definition: For the *integral conservation law*

$$\frac{\partial}{\partial t}\int_a^b \rho\, dx = q|_{x=a} - q|_{x=b},$$

with associated *divergence form*

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

we say that $\rho(x,t)$ is **weak solution** to the divergence form PDE if:

$$\int_0^\infty \int_{-\infty}^\infty [\rho\psi_t + q\psi_x]\, dxdt = 0,$$

for all **test functions** ψ, that is, a C^∞ function with *compact support*, i.e. the closure of the set

$$\{(x,t) : \psi(x,t) \neq 0\}$$

is closed and bounded, i.e. compact, in the xt plane.

Example: One such test function example is:

$$\psi(x,t) = \begin{cases} \exp\left(\frac{1}{x^2+t^2-1}\right) & \text{if } \sqrt{x^2+t^2} < 1 \\ 0 & \text{else} \end{cases}$$

Why does this definition make sense? For a complete discussion, look at chapter 9.2 in Levy & Shearer, but for now, an overview:
Suppose ψ is a test function and S is the support of ψ, that is, the set $S = \{(x,t) : \psi(x,t) \neq 0\}$. Since S has compact closure, we know we may find a rectangle $R = \{(x,t) : x_1 \leq x \leq x_2, t_1 \leq t \leq t_2\}$ such that $S \subset R$, i.e. R encloses S. Then, let us consider the integral

$$\iint_R \left[\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x}\right] \psi\, dxdt$$

We have:

$$\begin{aligned}\iint_R \left[\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x}\right] \psi\, dxdt &= \int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial \rho}{\partial t}\psi\, dxdt + \int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial q}{\partial x}\psi\, dxdt \\ &= \int_{x_1}^{x_2}\int_{t_1}^{t_2} \frac{\partial \rho}{\partial t}\psi\, dtdx + \int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial q}{\partial x}\psi\, dxdt\end{aligned}$$

Integrating by parts, we find:

$$= \int_{x_1}^{x_2} \left(\rho\psi|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{\partial \psi}{\partial t} \rho \, dt \right) dx + \int_{t_1}^{t_2} \left(q\psi|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{\partial \psi}{\partial x} q \, dx \right) dt$$

However, since $S \subset R$, $\psi(t, x_1) = \psi(t, x_2) = 0$, and similarly, $\psi(x, t_1) = \psi(x, t_2) = 0$!

$$= - \iint_R \left(\rho \frac{\partial \psi}{\partial t} + q \frac{\partial \psi}{\partial x} \right) dx \, dt$$

So if that latter integral vanishes for *all* possible test functions ψ, we can feel confident enough that the original PDE was also satisfied, that is, $\rho_t + q_x = 0$.

Let's work through another PDE with shocks.

Exercise: Recall that the conservation law

$$\frac{d}{dt} \int_{x_1}^{x_2} u(x,t) \, dx + \frac{1}{2} \left(u(x_2,t)^2 - u(x_1,t)^2 \right) = 0$$

has associated divergence form of the *Inviscid Burgers' Equation*

$$u_t + u u_x = 0.$$

Solve the PDE and determine the shock wave trajectories given domain $-\infty < x < \infty, t > 0$ and initial condition

$$u(x,0) = \begin{cases} 1 & x < -4 \\ 0 & |x| < 4 \\ -1 & x > 4 \end{cases}$$

Solution: Recall that the integral conservation law has flux function $q = \frac{1}{2}u^2$, so any shocks have speed

$$\sigma' = \frac{(\frac{1}{2}u^2)^+ - (\frac{1}{2}u^2)^-}{u^+ - u^-}.$$

We set up the problem as usual by letting $u(x,t) = u(x(r,s), t(r,s))$. Then,

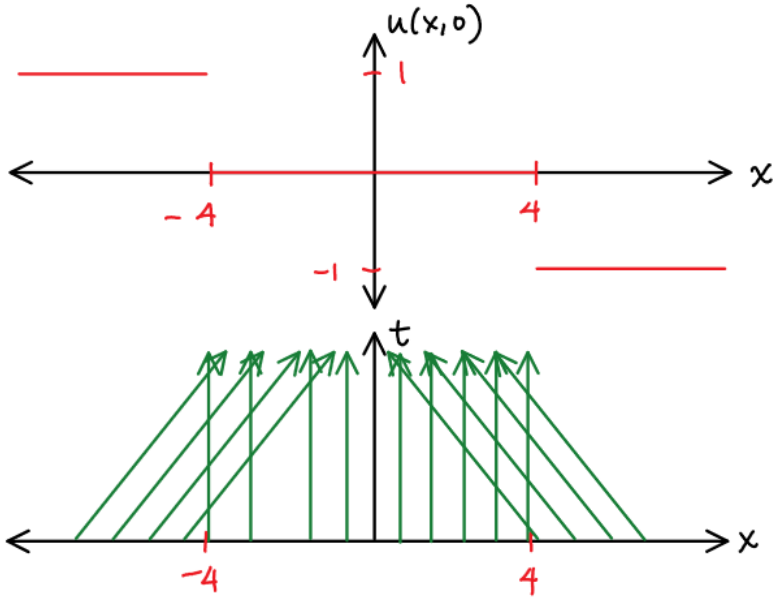
$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial r} \longleftrightarrow 1 u_t + u u_x = 0.$$

This yields system of ODEs:

$$\frac{\partial u}{\partial r} = 0, \quad u|_{r=0} = \begin{cases} 1 & s < -4 \\ 0 & |s| < 4 \\ -1 & s > 4 \end{cases} \implies u = \begin{cases} 1 & s < -4 \\ 0 & |s| < 4 \\ -1 & s > 4 \end{cases}$$

$$\frac{\partial t}{\partial r} = 1, \quad t|_{r=0} = 0 \quad \implies t = r$$

$$\frac{\partial x}{\partial r} = u, \quad x|_{r=0} = s \quad \Longrightarrow \; x = \begin{cases} r+s & s < -4 \\ s & |s| < 4 \\ -r+s & s > 4 \end{cases}$$



We see that two shocks form immediately (at $t = 0$), beginning at $x = \pm 4$. Let's handle each shock individually, however they should play out similarly by symmetry. First, the shock originating at $(-4, 0)$:

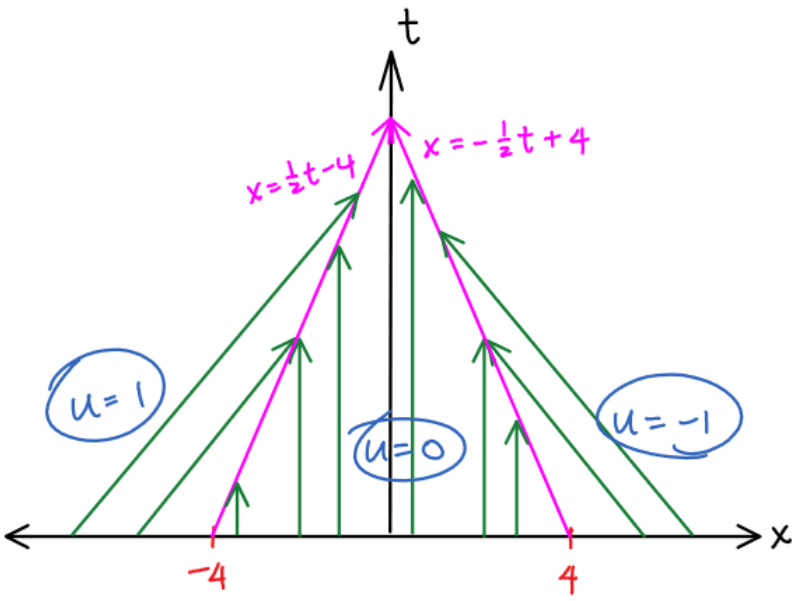
$$\sigma' = \frac{\frac{1}{2}0^2 - \frac{1}{2}1^2}{0 - 1} = \frac{1}{2},$$

and $\sigma(0) = 4$, so $\sigma(t) = \frac{1}{2}t - 4$. Remember that u^+ and u^- are constant in this case, so the shock speed is constant.

Similarly, from $(4, 0)$, we have

$$\sigma' = \frac{\frac{1}{2}(-1)^2 - \frac{1}{2}0^2}{-1 - 0} = -\frac{1}{2},$$

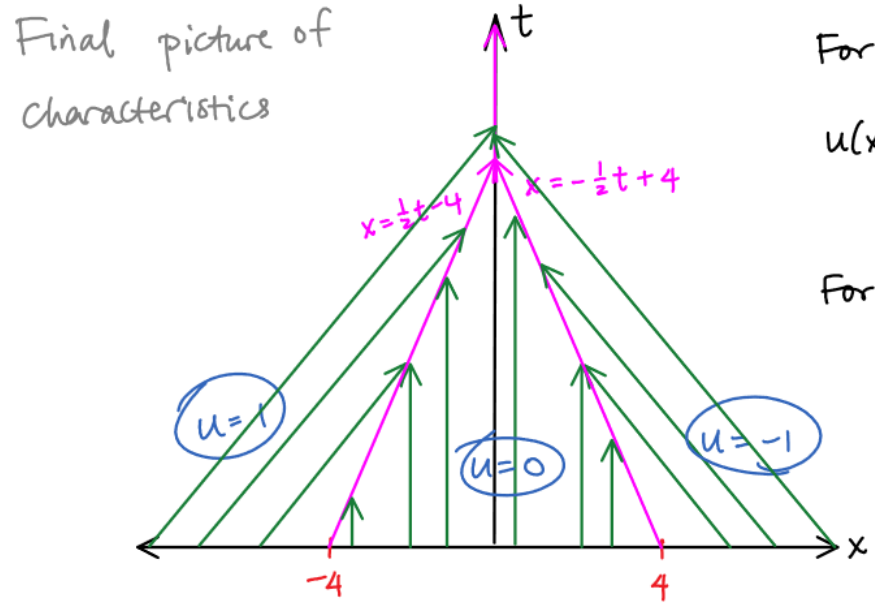
and $\sigma(0) = 4$, so $\sigma(t) = -\frac{1}{2}t + 4$. Visually:



But now at $t = 8$, the two shocks meet at $x = 0$, which will form a new shock! Here, we have $u^+ = -1$, $u^- = 1$. Then,

$$\sigma' = \frac{\frac{1}{2}(-1)^2 - \frac{1}{2}1^2}{-1-1} = 0.$$

So the third shock remains constant as t increases, dividing the region in two along the $x = 0$ line. Thus we conclude:



For $0 < t \leq 8$

$$u(x,t) = \begin{cases} 1 & x < \frac{1}{2}t - 4 \\ 0 & \text{else} \\ -1 & x > -\frac{1}{2}t + 4 \end{cases}$$

For $t \geq 8$

$$u(x,t) = \begin{cases} 1 & x < 0 \\ -1 & x > 0 \end{cases}$$

3 The Heat Equation, Separation of Variables, & Introducing Fourier Series

Note that these next two chapters will follow chapters 6.1 and 6.2, and chapter 7 in Levy & Shearer (though not necessarily in order). Chapter 5 also provides a useful extension to the topic.

We consider our first 2nd order PDE, which models the propagation of heat through a conducting materials, in a few different scenarios. Among the problems we will consider, we will cover the fundamental solution, which considers heat spread in infinite domains, as well as solutions on finite domains, which will motivate the strategy of *separation of variables*.

3.1 Derivation of the Heat Equation

Let us sketch a derivation of the heat equation, in 3 dimensions in this case, but the principle may be extended arbitrarily. Suppose we have a block of heat-conducting material, call it V. This will be the domain - let $u(x, y, z, t) = u(\vec{x}, t)$ be the temperature within any point of v. Moreover, let $E(t)$ be the total *heat energy* inside V (note that it only relies on t). We are given the relation:

$$E(t) = \iiint_V u(\vec{x}, t) \cdot \rho \cdot C \, dV,$$

where ρ is the *density* of the material (mass/volume) and C is the *specific heat* of the material (calories/degree $\cdot$ mass).

Fourier's Law of Heat Conduction states that heat energy flux is proportional to the outward normal gradient. In symbolic notation, this is:

$$\frac{dE}{dt} = \iint_{\partial V} K_0 \frac{\partial u}{\partial n} \, d\vec{S} = \iint_{\partial V} K_0 \vec{\nabla} u \cdot \hat{n} \, d\vec{S},$$

where K_0 is a proportional constant which gives us thermal conductivity (cal/length $\cdot$ time $\cdot$ temp), and $\frac{\partial u}{\partial n} = \vec{\nabla} u \cdot \hat{n}$, that is, the partial derivative is equivalently the gradient of u dotted with the outward unit normal. Then, we have:

$$\frac{dE}{dt} = \frac{d}{dt} \iiint_V u \cdot \rho C \, dV = \iint_{\partial V} K_0 \vec{\nabla} u \cdot \hat{n} \, d\vec{S}$$

By Gauss's Theorem, otherwise known as the Divergence Theorem (recall from MV Calc), we have that:

$$\iint_{\partial V} K_0 \vec{\nabla} u \cdot \hat{n} \, d\vec{S} = \iiint_V \operatorname{div}(K_0 \vec{\nabla} u) dV.$$

Therefore, we now have the equality:

$$\iiint_V \frac{\partial}{\partial t}(u \rho C) \, dV = \iiint_V \operatorname{div}(K_0 \vec{\nabla} u) dV.$$

Equivalently,

$$0 = \iiint_V \left(\frac{\partial}{\partial t}(u\rho C) - \operatorname{div}(K_0 \vec{\nabla} u) \right) dV.$$

But note that this integral is zero regardless of what volume V we choose, so the integrand itself must be zero. Thus, in 3 dimensions, the heat equation is given by:

$$\frac{\partial}{\partial t}(u\rho C) = \operatorname{div}(K_0 \vec{\nabla} u).$$

Moreover, if we assume ρ, C, K_0 to be constant throughout the material, then we have:

$$\rho C \frac{\partial u}{\partial t} = K_0 \cdot \operatorname{div}(\vec{\nabla} u) = K_0 \nabla^2 u,$$

where $\nabla^2 = \text{div grad}$. Let $\kappa = K_0/(\rho C)$, the "thermal diffusivity" (length^2/time), then we finally obtain

$$\frac{\partial u}{\partial t} = \kappa \nabla^2 u = \kappa(u_{xx} + u_{yy} + u_{zz}).$$

Let's now derive a law for boundary conditions. Newton's Law of Cooling through a surface states that the heat energy flux on the boundary condition surface B is given by the relation:

$$\frac{dE}{dt} = \iint_B H \cdot \Delta t \, dS,$$

where H is a heat transfer coefficient (higher values mean greater heat transfer) and Δt is the difference in temperature. However, Fourier's Law of Heat Conduction also relates heat energy flux, so we have:

$$\iint_B K_0 \frac{\partial u}{\partial n} \, dS = \iint_B H \cdot \Delta t \, dS.$$

Differentiating twice yields:

$$K_0 \frac{\partial u}{\partial n}|_B = H \cdot \Delta t.$$

This is a **Robin**, or **mixed**, boundary condition.

We have two special cases of the boundary condition which we will focus on first:

1. $H \to \infty$, a perfectly efficient heat transfer. This implies that Δt will always be 0, so $u|_B$ will always be the exterior temperature. This is a Dirichlet boundary condition!

2. $H = 0$, a perfectly insulated surface. Then, $\frac{\partial u}{\partial n}|_B = 0$. This is a Neumann boundary condition!

3.2 Separation of Variables

Exercise: Imagine a thin rod (1 dimensional) of heat-conducting material of length L, which is insulated everywhere except for at the two ends, which have perfectly efficient heat transfer. At $t = 0$, the rod is dumped into a large vat of ice water. First, set up the corresponding PDE and its initial and boundary conditions.
Solution: We use $u(x,t)$ to describe the temperature of the rod at length x at time t, where $t > 0$ and $0 < x < L$. From our previous work, we know that the temperature satisfies the heat equation, so our PDE will be:

$$u_t = \kappa u_{xx},$$

for some thermal diffusivity κ. Since the rod is insulated everywhere but the very edges, and assumed to be *very* thin, the only boundary we need consider is the two edges. Since the edges on the other hand have perfectly efficient heat transfer, we know that at all times, they will have temperature equal to the exterior temperature, which is (roughly) 0 degrees Celsius. Thus, we have boundary conditions:

$$u(0,t) = u(L,t) = 0.$$

Finally, we know that at time $t = 0$, there is some initial temperature distribution along the rod, which we will represent by a function $f(x)$. Thus, the initial condition is:

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

Let's now solve the PDE, using the **Separation of Variables** method.

Definition: The **Separation of Variables Method** is where we make the **product solution ansatz** (assumption) to solve a PDE, that is, we assume *the solution is a product of functions of each independent variable.*

In the case of this problem, our ansatz is

$$u(x,t) = X(t)T(t),$$

in other words, that the solution is a product of an x-related function and a t-related function. Plugging in $u(x,t) = X(x)T(t)$ into the PDE yields:

$$\frac{\partial}{\partial t}(XT) = \kappa \frac{\partial^2}{\partial x^2}(XT)$$

$$XT' = \kappa T X''$$

Dividing both sides by κXT yields:

$$\frac{1}{\kappa}\frac{T'}{T} = \frac{X''}{X}.$$

However, note that the left hand side of this equation is a function only of t, and the right side is a function only of x, and yet equality holds for all x and t in the domain. Therefore, this equality is *constant*, i.e.

$$\lambda = \frac{1}{\kappa}\frac{T'}{T} = \frac{X''}{X},$$

where λ is defined to be the **separation constant**. This gives us 2 ODEs:

$$T' = \kappa\lambda T, \qquad X'' = \lambda X.$$

Exercise: What are the boundary conditions for these ODEs?

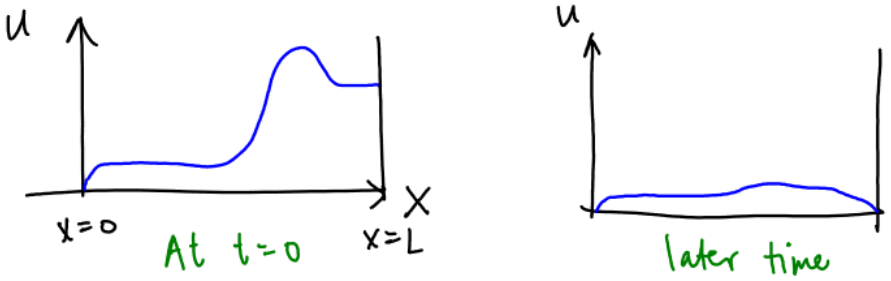
Solution: The boundary conditions $u(0,t) = u(L,t) = 0$, after substituting the product ansatz in, become

$$X(0)T(t) = X(L)T(t) = 0 \implies X(0) = X(L) = 0.$$

What about the initial condition, $u(x,0) = f(x)$? Suppose we try substituting the product ansatz in again, yielding

$$X(x)T(0) = f(x).$$

This isn't very useful for now - we will worry about this later (not very satisfying). However, what this tells us is that the temperature is just a scaled version of $f(x)$! That is,



Let us first solve the ODE given by X,

$$\begin{cases} X''(x) = \lambda X(x) \\ X(0) = 0 \\ X(L) = 0 \end{cases}$$

by considering the sign of λ. First however, note that if we consider $\frac{d^2}{dx^2}$ as a linear operator, this problem boils down to finding the *eigenfunctions* (X) and *eigenvalues* (λ) of the linear operator which satisfy the BCs. We call this an **Eigenvalue Problem**, for this reason. (Perhaps these cases will make good in-class exercises)

- Case 1: $\lambda = 0$. Then, we have:

$$X''(x) = 0 \implies X(x) = ax + b.$$

$X(0) = 0$ implies $B = 0$, and $X(L) = 0$ implies $AL = 0$, so $A = 0$. Thus $X(x) = 0$ in this case. This is BORING (not a mathematical term) - this solution gives us no new or useful information.

- Case 2: $\lambda > 0$. Then,

$$X''(x) - \lambda X(x) = 0$$

which has general solution

$$X(x) = c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x}.$$

The first BC gives us

$$X(0) = 0 = c_1 e^0 + c_2 e^0 = c_1 + c_2 \implies c_1 = -c_2,$$

then the second BC gives us

$$X(L) = 0 = -c_2 e^{\sqrt{\lambda}L} + c_2 e^{-\sqrt{\lambda}L} = -2c_2 \sinh(\sqrt{\lambda}L).$$

However since $\lambda \neq 0$ and $L \neq 0$, $\sinh(\sqrt{\lambda}L) \neq 0$. Therefore $c_2 = 0$, another trivial solution. BOOOOOOOOOO.

- Case 3: $\lambda < 0$. Let us redefine $\lambda = -\alpha, \alpha > 0$. Then,

$$X''(x) = \lambda X(x) = -\alpha X(x) \implies X''(x) + \alpha X = 0.$$

This has general solution:

$$X(x) = c_1 \cos(\sqrt{\alpha}x) + c_2 \sin(\sqrt{\alpha}x).$$

BC 1 gives us:

$$X(0) = 0 = c_1 \cos(0) + c_2 \sin(0) = c_1 \implies c_1 = 0.$$

BC2 gives us:

$$X(L) = 0 = c_2 \sin(\sqrt{\alpha}L) = 0.$$

In order to not obtain yet another trivial solution, then we need $\sin(\sqrt{\alpha}L) = 0$. Fortunately, since sin is a periodic function, this is possible - recall that $\sin(x)$ has zeroes at $n \cdot \pi$ for all $n \in \mathbb{Z}$. We thus choose $\sqrt{\alpha}L = n \cdot \pi$ for $n = 1, 2, 3, \ldots (n \in \mathbb{N})$ ($n \neq 0$ since we ruled out $\lambda = 0$), implying:

$$\alpha = \left(\frac{n\pi}{L}\right)^2 = -\lambda, \quad (n \in \mathbb{N}^+).$$

Therefore, we obtain in this case, solutions to X of the form

$$X(x) = c_2 \sin\left(\frac{n\pi x}{L}\right) \quad (n \in \mathbb{N}^+)$$

Now we solve for T, under the values of λ we have chosen:

$$T' = \lambda\kappa T = -\left(\frac{n\pi}{L}\right)^2 \kappa T.$$

This is pretty easily solved to be:

$$T(t) = C \cdot \exp\left(-\left(\frac{n\pi}{L}\right)^2 \kappa t\right)$$

3.3 Introducing a Fourier Series

First, let's recap. We have the "product solution"

$$u(x,t) = X(x)T(t) = C_0 \sin\left(\frac{n\pi x}{L}\right) \exp\left(-\left(\frac{n\pi}{L}\right)^2 \kappa t\right),$$

which satisfies the PDE and the two BCs for all $n \in \mathbb{N}^+$. Moreover, since the BCs and the PDE are homogeneous, we may apply the principle of superposition to conclude that arbitrary sums of terms of that form are also solutions to the PDE. However, this solution does *not* satisfy the IC... yet.

To satisfy the initial condition, we form a superposition of all possible product solution to form the "richest" solution possible. Remember the principle of superposition - since the PDEs and BCs are *linear and homogeneous*, arbitrary (converging) $\mathbb{R}$-linear combinations of solutions will again satisfy the PDE and BCs! So, we let:

$$u(x,t) = \sum_{n=1}^{\infty} D_n \sin\left(\frac{n\pi x}{L}\right) \exp\left(-\left(\frac{n\pi}{L}\right)^2 \kappa t\right),$$

with undetermined coefficients D_n. How can we find D_n in such a way that the IC, $u(x,0) = f(x)$, is satisfied? That is,

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

The task here is to represent $f(x)$ as a trigonometric series, in other words, a **Fourier Series**. To do this, we exploit a convenient property that sine functions have, *orthogonality* (recall from homework?).

$$\int_0^L \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx = \begin{cases} 0 & m \neq n \\ L/2 & m = n \end{cases} \qquad (m, n \in \mathbb{N}^+)$$

How does this help? We multiply both sides of the $u(x,0)$ equality by $\sin(m\pi x/L)$ (where $m \in \mathbb{N}$) and integrate by x from 0 to L.

$$\int_0^L f(x)\sin\left(\frac{m\pi x}{L}\right)\,dx = \int_0^L \sum_{n=1}^{\infty} D_n \sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right)\,dx$$

We may interchange the integral and the infinite sum when the sum converges uniformly, a notion one will learn about in a Real Analysis course. We will not prove it here, but this sum does indeed converge uniformly.

$$= \sum_{n=1}^{\infty} \int_0^L D_n \sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right)\,dx$$

All the terms of the sum are 0 except when $m = n$ by the orthogonality relation

$$= D_m \frac{L}{2}$$

Therefore, we obtain that for any $m \in \mathbb{N}$:

$$D_m = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{m\pi x}{L}\right)\,dx$$

Thus we finally conclude that the solution to the rod heat equation is:

$$u(x,t) = X(x)T(t) = D_n \sin\left(\frac{n\pi x}{L}\right)\exp\left(-\left(\frac{n\pi}{L}\right)^2 \kappa t\right),$$

$$\text{where} \quad D_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{m\pi x}{L}\right)\,dx$$

We cannot simplify this more without knowing $f(x)$, but the bottom line is that we do know the general form of the solution, and can find the coefficients corresponding to each term by computing an integral.

Observe here that we expressed $f(x)$ as an $\mathbb{R}$-linear sum of trigonometric functions. This is an example of a Fourier series, as noted before.
Definition: A Fourier series is a way of representing a periodic function or function defined on a finite length boundary as a (possibly infinite) sum of sine and cosine functions.

Example: Let $f(x) = x \cdot (L - x)$. We compute D_n.

$$\begin{aligned} D_n &= \frac{2}{L}\int_0^L x\cdot(L-x)\sin\left(\frac{n\pi x}{L}\right)\,dx \\ &= \frac{2}{L}\left(\int_0^L Lx\sin\left(\frac{n\pi x}{L}\right)\,dx - \int_0^L x^2\sin\left(\frac{n\pi x}{L}\right)\,dx\right) \end{aligned}$$

This is integrable using integration by parts, which is omitted.

$$= \frac{-2L^3}{n^3\pi^3}(2\cos(n\pi) + n\pi\sin(n\pi) - 2)$$

The sine term is 0 for all $n \in \mathbb{N}$, and the cosine term alternates between -1 and 1.

$$= \frac{4L^2}{n^3\pi^3}(1 - (-1)^n)$$
$$= \begin{cases} 0 & \text{even n} \\ (8L^2)/(n^3\pi^3) & \text{odd n} \end{cases}$$

In general, we know by the exponential term in the solution that the temperature will indeed decrease exponentially. In the case of this solution, the coefficient $D_n = 0$ for n even arises as a result of the symmetry of the initial temperature distribution. Each term in the series decays exponentially, and we may observe from the formula that as n increases, the corresponding term decreases exponentially faster in time. Moreover, the terms D_n get smaller with increasing n, which means that we may truncate the series and get a decent approximation of the solution.

3.4 Recap of the Separation of Variables Method

1. Plug in a product solution ansatz into the PDE and as many of the boundary conditions and initial conditions as you can. For example, $u(x, y, t) = X(x)Y(y)T(t)$.

2. Isolate the variables from each other (this may take a few passes) and set them to constants λ. These are the separation constants. If there are n independent variables, you'll get n ODEs and $n - 1$ separation constants.

3. Solve the ODE problems, noting that $n - 1$ of them will be eigenvalue problems.

4. Form the most general linear combination of the product solutions as you can. This step relies on the linearity of the differential equation so that the principle of superposition applies.

5. Use orthogonality of eigenfunctions to find the corresponding Fourier series/unknown constants.

When does separation of variables break down?

- When the DE is not linear, the principle of superposition will not apply.
- We require the PDE and boundary conditions to be homogeneous.

For example,

$$\begin{cases} u_t = \kappa u_{xx} & \leftarrow \text{homogeneous} \\ u(x,0) = f(x) & \leftarrow \text{not homogeneous, but still ok} \\ u(0,t) = 100 & \leftarrow \text{not homogeneous, not ok} \\ u(L,t) = 50 & \leftarrow \text{not homogeneous, not ok} \end{cases}$$

What happens? If we try the product solution ansatz, we obtain the system of ODEs as before, but when we try to find boundary conditions of the ODEs, we find:

$$X(0)T(t) = 100 \implies X(0) = 100/T(t),$$

which doesn't lead to a useful BC. The end result is that we need homogeneous BCs and PDEs for the separation of variables method to work. We will figure out how to deal with inhomogeneous equations soon!

4 Self-Adjoint Operators & More (Complex) Fourier Series

In the previous section, we introduced a homogeneous PDE which required the use of a Fourier series to obtain a complete solution, as well as separation of variables. In this section, we will develop more theory for Fourier series and consider strategies for more difficult PDEs of a similar nature. We first study eigenvalue problems more rigorously, and specifically, study the orthogonality of eigenvalues or eigenfunctions in a broader context.

4.1 Self-Adjoint Linear Operators

Definition: An **inner product**, or **Hermitian inner product** on a vector space V over $\mathbb{C}$ is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$$

which satisfies the following conditions: for $a \in \mathbb{C}$, $x, y, z \in V$

1. *Linearity* in the first argument:

$$\langle ax, y \rangle = a\langle x, y \rangle \qquad \text{and} \qquad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle,$$

2. *Antilinearity* in the second argument:

$$\langle x, ay \rangle = \overline{a}\langle x, y \rangle \qquad \text{and} \qquad \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$$

3. *Conjugate Symmetry*:

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

4. *Positive definiteness:*

$$0 \leq \langle x, x \rangle \in \mathbb{R},$$

with equality holding if and only if $x = 0$.

Remark: Occasionally, the definition of a Hermitian inner product will be changed to have Antilinearity in the first argument and Linearity in the second argument. In particular, this is common in physics. If this is the case, simply redefine the inner product by switching the first and second variables - nothing else changes.

A vector space equipped with an inner product is called an **inner product space**. Two vectors in an inner product space are **orthogonal** if their inner product is 0.

Remark: If instead V is a vector space over $\mathbb{R}$, the same definition works, except the inner product is a map:

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

instead. The same conditions hold, however because $\overline{x} = x$ for all $x \in \mathbb{R}$, the antilinearity condition simply becomes a linearity condition, and similarly, the conjugate symmetry condition becomes a regular symmetry condition.

Example: The **dot product** on $\mathbb{R}^n$ is an example of an inner product over $\mathbb{R}$. Exercise: verify this! Moreover, two vectors are orthogonal (in the usual sense, that they form a right angle) if and only if their dot product is zero.

However, the dot product on $\mathbb{C}^n$ is **not** an inner product space over $\mathbb{C}$. Instead, if $v_1, v_2 \in \mathbb{C}^n$ with $v_1 = (c_1, \ldots, c_n)$ and $v_2 = (d_1, \ldots, d_n)$, then

$$\langle v_1, v_2 \rangle = c_1\overline{d_1} + \cdots + c_n\overline{d_n}.$$

This is known as the **Hermitian Form**. In the case that you are using the alternative definition of an inner product, take the conjugate of the first vector's coordinates rather than the second vector's.

Remark: If V is an inner product space (in fact, V does not need to be a vector space), then we can define a **norm** (a "size" function) on V by the following formula:

$$||\vec{x}|| = \sqrt{\langle x, x \rangle},$$

and a **metric** (a distance function) on V by the following formula:

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

As an example, one can check that in $\mathbb{R}^2$, the metric and norm induced by the dot product are the usual Euclidean metric and norm (where distance may be computed using the

Pythagorean theorem).

We shall motivate the importance of inner product spaces, and two self-adjoint eigenvalue problems which arise, with two parallel examples.

Example: (1) Our first example comes from Linear Algebra. Consider the classic eigenvalue problem over the vector space $\mathbb{C}^n$,

$$A\vec{x} = \lambda\vec{x}.$$

- λ is an eigenvalue
- $\vec{x}$ is an eigenvector (must be nontrivial)
- A is a square matrix, which can be considered a linear operator.

We define an inner product over $\mathbb{C}^n$:

$$\langle \vec{u}, \vec{v} \rangle = \vec{u}^H \vec{v}, \qquad \text{where } A^H = (\overline{A})^T.$$

As an example, consider

$$A = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}.$$

We may compute that the matrix has two eigenvalues,

$$\lambda_1 = 1, \lambda_2 = -1,$$

with corresponding eigenvectors,

$$\vec{x_1} = \begin{bmatrix} i \\ 1 \end{bmatrix}, \vec{x_2} = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

We observe two things, first, that the eigenvalues of this complex matrix are real, and that $\vec{x_1}$ and $\vec{x_2}$ are orthogonal, that is,

$$\langle \vec{x_1}, \vec{x_2} \rangle = \begin{bmatrix} -i & 1 \end{bmatrix} \begin{bmatrix} -i \\ 1 \end{bmatrix} = i^2 + 1 = 0.$$

Example: (2) Our second example comes from Differential Calculus. Consider a standard ODE problem on the space $C^\infty([0, L], \mathbb{C})$,

$$\mathscr{D}y(x) = \lambda \cdot y(x)$$

- λ is an eigenvalue
- y is an eigenfunction (must be nontrivial)

- $\mathscr{D}$ is a differential operator, which is a linear operator. (ex. $\mathscr{D} = d^2/dx^2$)

We define an inner product over $C^\infty(\mathbb{C})$:

$$\langle u(x), v(x) \rangle = \int_0^L u\overline{v}\, dx$$

As an example, consider the eigenvalue problem

$$\mathscr{D} = \frac{d^2}{dx^2}, \qquad \text{so } y''(x) = \lambda y(x), \quad \begin{cases} y(0) = 0 \\ y(L) = 0 \end{cases} .$$

We find that this eigenvalue problem has eigenvalues

$$\lambda_n = \frac{\pi n^2}{L}, n \in \mathbb{N},$$

with corresponding eigenfunctions

$$y_n(x) = \sin\left(\frac{n\pi x}{L}\right).$$

Again, we observe that the eigenvalues of this linear operator over a complex space are real, and that all the eigenfunctions are pairwise orthogonal, that is,

$$\int_0^L \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) dx = 0 \quad \text{if } n \neq m.$$

What is it about these linear operators that causes these two conditions to arise? The answer arises from how the linear operators behave with respect to the inner products.

Definition: The **adjoint** of a linear operator $\mathscr{L} : V \to V$ is the operator $\mathscr{L}^* : V \to V$ such that

$$\langle \mathscr{L}u, v \rangle = \langle u, \mathscr{L}v \rangle$$

for all $u, v \in V$.

Definition: If $\mathscr{L} = \mathscr{L}^*$, then $\mathscr{L}$ is **self-adjoint**.

In an advanced course in linear algebra, one learns of the *spectral theorems*, powerful theorems which determine when linear transformations can be diagonalized, and how their eigenvalues appear. Though we will not cover these theorems in full, we state some examples of results of the spectral theorems, with proof, which apply directly to our study of self-adjoint eigenvalue problems.

Theorem: Self-adjoint operators have real eigenvalues.

Proof. Suppose $\mathscr{L}$ is a self-adjoint operator on V, and that $\mathscr{L}y = \lambda y$, that is, $\mathscr{L}$ has eigenvalue $\lambda \in \mathbb{C}$ with nonzero eigenvector $y \in V$. Then,

$$\langle \mathscr{L}y, y\rangle = \langle \lambda y, y\rangle = \lambda\langle y, y\rangle.$$

On the other hand, we have

$$\langle \mathscr{L}y, y\rangle = \langle y, \mathscr{L}y\rangle = \langle y, \lambda y\rangle = \overline{\lambda}\langle y, y\rangle.$$

Since $y \neq 0$, we have that $\overline{\lambda} = \lambda$. This statement is true if and only if $\lambda \in \mathbb{R}$. □

Theorem: Self-adjoint operators have orthogonal eigenvectors (for eigenvectors corresponding to different eigenvalues).

Proof. Suppose $\mathscr{L} = \mathscr{L}^*$, and $\mathscr{L}y_1 = \lambda_1 y_1$, $\mathscr{L}y_2 = \lambda_2 y_2$. Then,

$$\langle \mathscr{L}y_1, y_2\rangle = \langle \lambda_1 y_1, y_2\rangle = \lambda_1\langle y_1, y_2\rangle.$$

On the other hand, we find:

$$\langle \mathscr{L}y_1, y_2\rangle = \langle y_1, \mathscr{L}y_2\rangle = \langle y_1, \lambda y_2\rangle = \overline{\lambda_2}\langle y_1, y_2\rangle = \lambda_2\langle y_1, y_2\rangle.$$

The last equality comes from applying the previous theorem, as $\lambda_1, \lambda_2 \in \mathbb{R}$. Therefore, we see

$$(\lambda_1 - \lambda_2)\langle y_1, y_2\rangle = 0.$$

If $\lambda_1 \neq \lambda_2$, then $\langle y_1, y_2\rangle = 0$, as desired. □

The big picture here is to understand that the cosine identity we derived previously is not a coincidence - there is more going on underneath the hood here. More functions (such as sines) will also have a similar property, which we may be able to exploit for deriving solutions to other PDEs in the future!

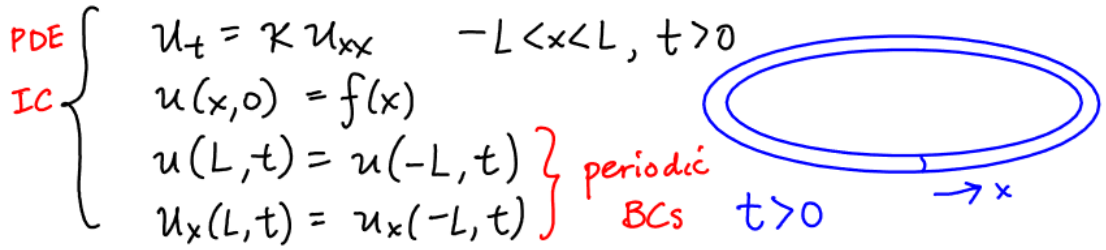
4.2 Another Heat PDE Problem

We next turn to another heat equation problem which will again require separation of variables. This time, we will

Example: Consider a thin rod of heat-conducting material whose temperature can be modeled by the 1-D heat equation. Let the length of the rod be $2L$, oriented $-L < x < L$, and moreover suppose the two ends of the rod are joined together to form a ring. We assume the ends of the rod satisfy the same conductive properties as the rest of the ring.

Exercise: Set up the IVP corresponding to this physical system.

Solution: As before, we let $u(x,t)$ represent the temperature inside the rod at point x at time t, where $t > 0$ and $-L < x < L$. We again give an initial heat distribution on the rod by $u(x,0) = f(x)$. This time however, because the ends of the ring are joined, our boundary temperatures must be equal, that is, $u(L,t) = u(-L,t)$. Moreover, since we want our solution to be smooth, the same must be true for the derivative, that $u_x(L,t) = u_x(-L,t)$. These are **periodic boundary conditions**.



We begin solving the PDE by using the product solution ansatz, i.e. separation of variables. Let $u(x,t) = X(x)T(t)$, then plugging into the PDE we obtain (after a bit of rearranging)

$$\frac{T'}{\kappa T} = \frac{X''}{X} = -\lambda \quad \Longrightarrow \quad \begin{cases} T' = -\kappa\lambda T \\ X'' + \lambda X = 0 \end{cases}$$

We can't use the inhomogeneous IC yet, but the boundary conditions give:

$$u(L,t) = u(-L,t) \implies X(L)T(t) = X(-L)T(T) \implies X(L) = X(-L)$$

$$u_x(L,t) = u_x(-L,t) \implies X'(L)T(t) = X'(-L)T(t) \implies X'(L) = X'(-L)$$

We end up with another eigenvalue problem for X, given by:

$$X'' + \lambda X = 0, \qquad \begin{cases} X(L) = X(-L) \\ X'(L) = X'(-L) \end{cases}$$

We again split into 3 cases, depending on the sign of λ.

- Case 1: $\lambda < 0$. Let $\alpha = -\lambda, \alpha > 0$. Then solving the 2nd order ODE yields:

  $$X'' - \alpha X = 0 \implies X = Ae^{\sqrt{\alpha}x} + Be^{-\sqrt{\alpha}x}, \quad (A, B \in \mathbb{R})$$

  Applying the Dirichlet boundary condition yields:

  $$X(L) = X(-L) \implies Ae^{\sqrt{\alpha}L} + Be^{-\sqrt{\alpha}L} = Ae^{-\sqrt{\alpha}L} + Be^{\sqrt{\alpha}L} \tag{1}$$

  Note $X' = A\sqrt{\alpha}\exp(\sqrt{\alpha}x) - B\sqrt{\alpha}\exp(\sqrt{\alpha}x)$. Applying the Neumann boundary condition yields:

  $$\begin{aligned} X'(L) = X'(-L) &\implies A\sqrt{\alpha}e^{\sqrt{\alpha}L} - B\sqrt{\alpha}e^{-\sqrt{\alpha}L} = A\sqrt{\alpha}e^{-\sqrt{\alpha}L} - B\sqrt{\alpha}e^{\sqrt{\alpha}L} \\ &\implies Ae^{\sqrt{\alpha}L} - Be^{-\sqrt{\alpha}L} = Ae^{-\sqrt{\alpha}L} - Be^{\sqrt{\alpha}L} \end{aligned} \tag{2}$$

Adding equations (1) and (2) gives us:

$$2Ae^{\sqrt{\alpha}L} = 2Ae^{-\sqrt{\alpha}L} \implies A = 0,$$

the implication is true since $e^x = e^{-x}$ if and only if $x = 0$, but $\sqrt{\alpha}L > 0$. Similarly, subtracting (1) and (2) gives us:

$$2Be^{-\sqrt{\alpha}L} = 2Be^{\sqrt{\alpha}L} \implies B = 0.$$

Thus if $\lambda < 0$, we only obtain the trivial solution, so we may otherwise ignore this case.

- Case 2: $\lambda = 0$. The IVP then becomes:

$$X'' = 0 \implies X = Ax + B, \quad (A, B \in \mathbb{R}).$$

Plugging in the Dirichlet BC yields

$$AL + B = -AL + B \implies A = 0.$$

However, the Neumann BC leaves no restriction on the value of B, so we conclude that $\lambda = 0$ gives $X = B$ for some constant B as a possible solution.

- Case 3: $\lambda > 0$. In this case, this 2nd order ODE has general solution:

$$X'' + \lambda X = 0 \implies X = A\cos(\sqrt{\lambda}x) + B\sin(\sqrt{\lambda}x).$$

Note that $X' = -A\sqrt{\lambda}\sin(\sqrt{\lambda}x) + B\sqrt{\lambda}\cos(\sqrt{\lambda}x)$. We next determine the restrictions of A, B, λ by applying the boundary conditions. For the Dirichlet condition, we find:

$$X(L) = X(-L) \implies A\cos(\sqrt{\lambda}L) + B\sin(\sqrt{\lambda}L) = A\cos(-\sqrt{\lambda}L) + B\sin(-\sqrt{\lambda}L)$$

Recall that cosine is an *even function*, i.e. $f(-x) = f(x)$, and sine is an *odd function*, i.e. $f(-x) = -f(x)$.

$$\begin{aligned} &\implies A\cos(\sqrt{\lambda}L) + B\sin(\sqrt{\lambda}L) = A\cos(\sqrt{\lambda}L) - B\sin(-\sqrt{\lambda}L) \\ &\implies 2B\sin(\sqrt{\lambda}L) = 0 \end{aligned}$$

For the Neumann condition, we find:

$$\begin{aligned} -A\sqrt{\lambda}\sin(\sqrt{\lambda}L) + B\sqrt{\lambda}\cos(\sqrt{\lambda}L) &= -A\sqrt{\lambda}\sin(-\sqrt{\lambda}L) + B\sqrt{\lambda}\cos(-\sqrt{\lambda}L) \\ &= A\sqrt{\lambda}\sin(\sqrt{\lambda}L) + B\sqrt{\lambda}\cos(\sqrt{\lambda}L) \\ &\implies 2A\sqrt{\lambda}\sin(\sqrt{\lambda}L) = 0 \end{aligned}$$

We want nontrivial solutions, so we require $\sin(\sqrt{\lambda}L) = 0$, or in other words, for $\sqrt{\lambda}L = n \cdot \pi$ for some $n = 1, 2, 3 \ldots$. Equivalently,

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad (n \in \mathbb{N})$$

So we conclude that when $\lambda > 0$, the most general solution for X is:

$$X = A\cos\left(\frac{n\pi x}{L}\right) + B\sin\left(\frac{n\pi x}{L}\right) \qquad (n = 1, 2, 3 \ldots)$$

We may use the principle of superposition to obtain the most general form of X, (since we may note that the BCs behave homogeneously, that is, the sum of two solutions still satisfy both BCs). Combining the $\lambda = 0$ and $\lambda > 0$ solutions together gives the full expression:

$$X_n = A_n \cos\left(\frac{n\pi x}{L}\right) + B_n \sin\left(\frac{n\pi x}{L}\right), \qquad \lambda_n = \left(\frac{n\pi}{L}\right)^2, \qquad (n = 0, 1, 2, 3 \dots)$$

Note that when $n = 0$, the cosine term is equivalently 1 and the sine term is equivalently 0, which covers the $\lambda = 0$ case.

Each λ_n now gives us a solution to the T differential equation - let T_n be the solution of:

$$T_n' = -\kappa \lambda_n T_n.$$

It is a quick computation that

$$T_n = C_n e^{-\kappa \lambda_n t} \qquad (C_n \in \mathbb{R})$$

Using the principle of superposition, we make the most general linear combination of product solutions as we can as follows. It remains to satisfy the inhomogeneous initial condition.

$$\begin{aligned} u(x,t) &= \sum_{n=0}^{\infty} C_n e^{-\kappa\left(\frac{n\pi}{L}\right)^2 t} \left(A_n \cos\left(\frac{n\pi x}{L}\right) + B_n \sin\left(\frac{n\pi x}{L}\right)\right) \\ &= \sum_{n=0}^{\infty} e^{-\kappa\left(\frac{n\pi}{L}\right)^2 t} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \end{aligned}$$

We have set $\alpha_n = A_n C_n$ and $\beta_n = B_n C_n$ for ease of computation. Now, we wish to solve for α_n and β_n for all $n \geq 0$ such that $u(x,t)$ satisfies the remaining initial condition, that is,

$$\begin{aligned} u(x,0) = f(x) &= \sum_{n=0}^{\infty} e^{-\kappa\left(\frac{n\pi}{L}\right)^2 0} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \\ &= \sum_{n=0}^{\infty} \alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right) \end{aligned}$$

In other words, we need to find the **Fourier Series** of $f(x)$. We do so, again, by orthogonality of sine and cosine functions.

Theorem: We have the following identities (one of which we've seen before) ($n, m \in \mathbb{N}$):

$$\int_{-L}^{L} \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) \, dx = \begin{cases} 0 & n \neq m \\ L & n = m \neq 0 \\ 2L & n = m = 0 \end{cases}$$

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

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

To find α_m, multiply both sides by an appropriate cosine term and integrate over the length:

$$\begin{aligned}
\int_{-L}^{L} f(x) \cos\left(\frac{m\pi x}{L}\right)\, dx &= \int_{-L}^{L} \sum_{n=0}^{\infty} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \cos\left(\frac{m\pi x}{L}\right)\, dx \\
&= \sum_{n=0}^{\infty} \int_{-L}^{L} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \cos\left(\frac{m\pi x}{L}\right)\, dx \\
&= \begin{cases} \alpha_m L & m > 0 \\ 2\alpha_m L & m = 0 \end{cases}
\end{aligned}$$

So therefore,

$$\alpha_m = \begin{cases} \dfrac{1}{L} \displaystyle\int_{-L}^{L} f(x) \cos\left(\frac{m\pi x}{L}\right)\, dx & m > 0 \\ \dfrac{1}{2L} \displaystyle\int_{-L}^{L} f(x) \cos\left(\frac{m\pi x}{L}\right)\, dx & m = 0 \end{cases}$$

We apply a similar procedure to find β_m, multiplying both sides by an appropriate sine term and integrating over the length.

$$\begin{aligned}
\int_{-L}^{L} f(x) \sin\left(\frac{m\pi x}{L}\right)\, dx &= \int_{-L}^{L} \sum_{n=0}^{\infty} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \sin\left(\frac{m\pi x}{L}\right)\, dx \\
&= \sum_{n=0}^{\infty} \int_{-L}^{L} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \sin\left(\frac{m\pi x}{L}\right)\, dx \\
&= \beta_m L
\end{aligned}$$

$$\beta_m = \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{m\pi x}{L}\right)\, dx$$

Thus we have a complete characterization of our solution of $u(x,t)$, which satisfies all the boundary and initial conditions given! (though it should be noted that we are making some very large assumptions - what are those?) Note also we may rewrite it slightly:

$$u(x,t) = \alpha_0 + \sum_{n=1}^{\infty} \left(\alpha_n \cos\left(\frac{n\pi x}{L}\right) + \beta_n \sin\left(\frac{n\pi x}{L}\right)\right) \exp\left(-\kappa \left(\frac{n\pi}{L}\right)^2 t\right)$$

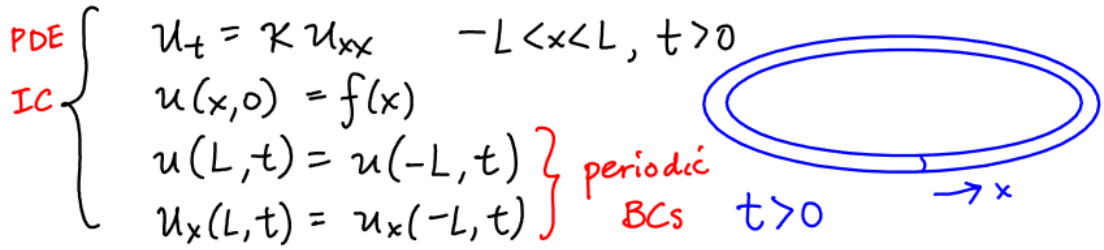
This constant α_0 may be shown to be both the average temperature of the system initially, and the final, constant temperature of the system as $t \to \infty$. As before, the solution cannot be expressed more simply (without some terms of integration appearing). However once $f(x)$ is known, the solution may be explicitly calculated.

4.3 A Complex Solution Method

There is an alternative approach to solving this PDE, using complex exponentials instead of cosines and sines. Recall Euler's formula:

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

Let us restate the problem:



As before, we use separation of variables to obtain the system of ODEs:

$$T' = -\kappa\lambda T \qquad X'' + \lambda X = 0 \qquad \begin{cases} X(L) = X(-L) \\ X'(L) = X'(-L) \end{cases}$$

We proceed by first solving for X as before. Recall that we had three cases, $\lambda < 0, \lambda = 0, \lambda > 0$. As before, the first two cases yield respectively the trivial solution and a constant solution. Let's work through the third case again, $\lambda > 0$, this time using complex functions.

For the ODE $X'' + \lambda X = 0$, we may find that there is a general complex solution of:

$$X = C\exp(i\sqrt{\lambda}x) + D\exp(-i\sqrt{\lambda}x).$$

We have that the initial conditions give relations:

$$C\exp(i\sqrt{\lambda}L) + D\exp(-i\sqrt{\lambda}L) = C\exp(-i\sqrt{\lambda}L) + D\exp(i\sqrt{\lambda}L)$$

$$Ci\sqrt{\lambda}\exp(i\sqrt{\lambda}L) - iD\sqrt{\lambda}\exp(-i\sqrt{\lambda}L) = Ci\sqrt{\lambda}\exp(-i\sqrt{\lambda}L) - iD\sqrt{\lambda}\exp(i\sqrt{\lambda}L)$$

Adding the equations yields:

$$2C\exp(i\sqrt{\lambda}L) = 2C\exp(-i\sqrt{\lambda}L)$$

and subtracting yields:

$$2D\exp(-i\sqrt{\lambda}L) = 2D\exp(i\sqrt{\lambda}L)$$

To avoid C, D being zero and obtaining trivial solutions, we require:

$$\begin{aligned}
&\implies \exp(i\sqrt{\lambda}L) = \exp(-i\sqrt{\lambda}L) \\
&\implies \exp(2i\sqrt{\lambda}L) = 1 \\
&\implies \cos(2\sqrt{\lambda}L) + i\sin(2\sqrt{\lambda}L) = 1 \\
&\implies 2\sqrt{\lambda}L = 2n\pi, \qquad n = 1, 2, 3\ldots \\
&\implies \lambda = \left(\frac{n\pi}{L}\right)^2
\end{aligned}$$

So we obtain for X:

$$X = C\exp\left(\frac{in\pi x}{L}\right) + D\exp\left(-\frac{in\pi x}{L}\right), \qquad n = 1, 2, 3\ldots$$

Finding T as before and combining the $\lambda = 0$ case obtains a general solution of $u(x,t)$:

$$u(x,t) = \sum_{n=-\infty}^{\infty} Q_n \exp\left(-\kappa\left(\frac{n\pi}{L}\right)^2 t\right)\exp\left(\frac{in\pi x}{L}\right)$$

We still need to satisfy the inhomogeneous initial condition,

$$u(x,0) = f(x) = \sum_{-\infty}^{\infty} Q_n \exp\left(\frac{in\pi x}{L}\right).$$

This is a **Complex Fourier series**. To find Q_n we need an inner product which produces orthogonality with respect to the exponential functions. Fortunately such an inner product exists:

$$\langle u, v\rangle = \int_{-L}^{L} u\cdot\overline{v}\,dx$$

This inner product indeed demonstrates orthogonality for exponentials:

$$\left\langle \exp\left(\frac{in\pi x}{L}\right), \exp\left(\frac{im\pi x}{L}\right)\right\rangle = \begin{cases} 0 & m \neq n \\ 2L & m = n \end{cases}$$

Now we take the inner product of the equality with respect to $\exp\left(i\,m\,\pi\,x/L\right)$:

$$\begin{aligned}
\left\langle f, \exp\left(\frac{im\pi x}{L}\right)\right\rangle &= \left\langle \sum_{-\infty}^{\infty} Q_n \exp\left(\frac{in\pi x}{L}\right), \exp\left(\frac{im\pi x}{L}\right)\right\rangle \\
&= \sum_{-\infty}^{\infty} Q_n \left\langle \exp\left(\frac{in\pi x}{L}\right), \exp\left(\frac{im\pi x}{L}\right)\right\rangle \\
&= 2L\cdot Q_m
\end{aligned}$$

So therefore, we obtain:

$$Q_m = \frac{1}{2L} \int_{-L}^{L} f(x) \exp\left(\frac{-im\pi x}{L}\right) \, dx$$

It follows that our final solution for u is:

$$u(x,t) = \sum_{n=-\infty}^{\infty} \left(\frac{1}{2L} \int_{-L}^{L} f(x) \exp\left(\frac{-in\pi x}{L}\right) \, dx \right) \exp\left(-\kappa \left(\frac{n\pi}{L}\right)^2 t\right) \exp\left(\frac{in\pi x}{L}\right)$$

5 Understanding Solutions to the Heat/Diffusion Equation

5.1 Convergence of Fourier Series

In general, if $f(x)$ is a function which we want to represent using a Fourier series on the interval $[0, 2\pi]$, then we have that its **Fourier series** is:

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}, \qquad \text{where} \qquad a_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{inx}\, dx$$

However, we have been assuming that this process of approximating a function by an infinite series of trigonometric functions is mathematically rigorous. There is no de facto reason why this approximation can be performed. Fortunately, under most conditions, we are indeed able to do this.

Q1: Under what conditions can we use the Fourier transformation? The only condition which we need in order to derive the Fourier coefficients a_n is for $f(x)$ to be integrable on the interval $0 \leq x \leq 2\pi$. In other words, this formula always makes sense, so long as $f(x)$ is integrable on that region.

If you have taken a course in Analysis, you may have studied the notions of *pointwise convergence* and *uniform convergence*, two notions of convergence for infinite sequences or series of functions. If not, do not fret. We will not cover precise definitions but give the general ideas of each notion.

- A sequence of functions $f_n(x)$ *converges pointwise* to a function $f(x)$ if for every a in the domain of all f_i and f, the following limit is satisfies:

  $$\lim_{n \to \infty} f_n(a) = f(a).$$

  Visually, you can picture a 2D graph of a curve, where over time, every point on that curve eventually converges to the limit curve.

- Uniform convergence of functions is a strictly stronger notion. For a sequence of functions $f_n(x)$ to *converge uniformly* to $f(x)$, not only must the sequence converge pointwise, but all of the points on the functions must "converge at roughly the same speed." A more technical definition will not be provided here, but can certainly be found in any Real Analysis textbook, such as Rudin.

These two modes of convergence have an important distinction. One reasonable question to ask is if a converging sequence of continuous functions (such as a Fourier series) converges

to a continuous function. As it turns out, pointwise convergence is not enough for this to be true! As an example, consider the following sequence:

$$\{x^n\}_{n\in\mathbb{N}}, x \in [0,1] \to \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases}$$

This is a pointwise convergence of continuous functions converging to a discontinuous function! On the other hand, uniform convergence guarantees convergence.

Theorem: If a sequence of continuous functions $\{f_n(x)\}_{n\in\mathbb{N}}$ converges uniformly to a function $f(x)$, then $f(x)$ is continuous.

There exist similar statements for differentiability and integrability, but require a bit more precision. For the sake of brevity, they will not be mentioned here. To learn more about convergence and uniform convergence of functions, I would recommend that you take a course in real analysis, *especially* if you want to continue studying mathematics or applied mathematics beyond your undergraduate education.

Q2: Under what conditions does the Fourier series converge pointwise? One necessary condition is for $f(x)$ to have bounded variation. In fact, even function which are discontinuous can be used. The Fourier series will converge to the average of the left- and right- hand limits on either side of the discontinuity.

Q3: Under what conditions does the Fourier series converge uniformly? One necessary condition is for $f(x)$ to satisfy a Holder condition:

$$|f(x) - f(y)| \leq C|x - y|^\alpha,$$

for some positive $C, \alpha > 0$. This is a more restrictive condition than just being continuous, but not as restrictive as being continuously differentiable!

Q4: How quickly does the Fourier series for a function converge? Surprisingly, the answer depends on "how differentiable" the function is. If the function $f(x)$ is p times differentiable, i.e. $f \in C^p(\mathbb{R})$, then

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

That is, there exists some constants $C > 0$ and $N > 0$ large for which, for $n > N$, $a_n \leq C/n^p$. For more information, see:

- *An Introduction to Harmonic Analysis* by Yitzhak Katznelson
- *Fourier Analyis: An Introduction* by Elias Stern, Rami Shakarchi

5.2 More Intuition on the Behavior of the Heat Equation

Let's take a look at the solution of the Heat equation in one dimension to gain some intuition on how it behaves, and to confirm it with our physical intuition. Let's set up the problem as before, with an explicit initial temperature distribution:

$$\begin{cases} u_t = \kappa u_{xx}, & 0 < x < L, t > 0 \\ u(0,t) = 0 \\ u(L,t) = 0 \\ u(x,0) = 10\sin(\pi x/L) + 5\sin(3\pi x/L) \end{cases}$$

From earlier, we found that the general solution for an arbitrary initial condition $f(x)$ is:

$$u(x,t) = D_n \sin\left(\frac{n\pi x}{L}\right) \exp\left(-\left(\frac{n\pi}{L}\right)^2 \kappa t\right).$$

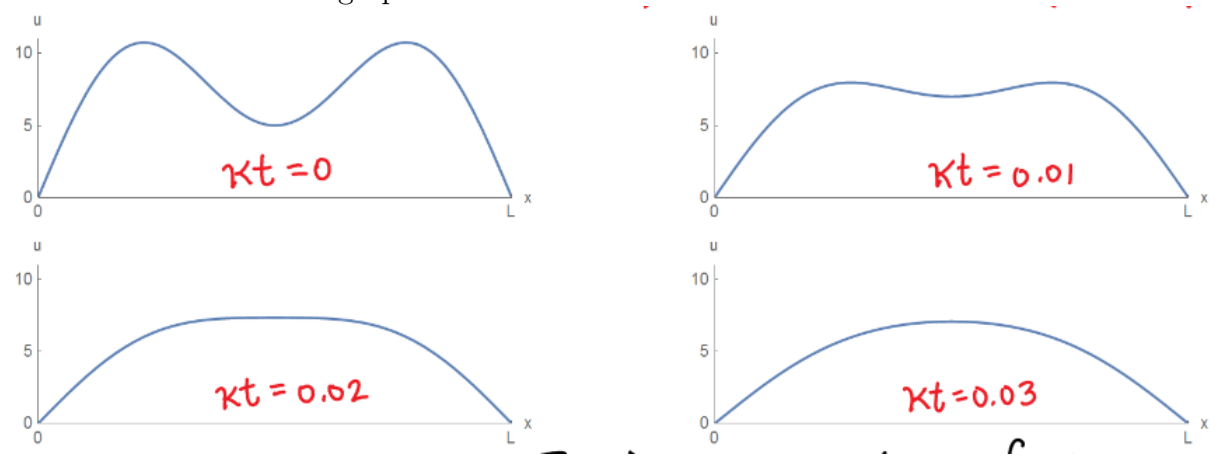
In this case, we may compute each D_n term explicitly by:

$$D_n = \frac{2}{L}\int_0^L \left(10\sin\left(\frac{\pi x}{L}\right) + 5\sin\left(\frac{3\pi x}{L}\right)\right)\sin\left(\frac{n\pi x}{L}\right)\,dx$$

By orthogonality, we see that $D_n = 0$ for any $n \neq 1, 3$, $D_1 = 10$ and $D_3 = 5$. So explicitly,

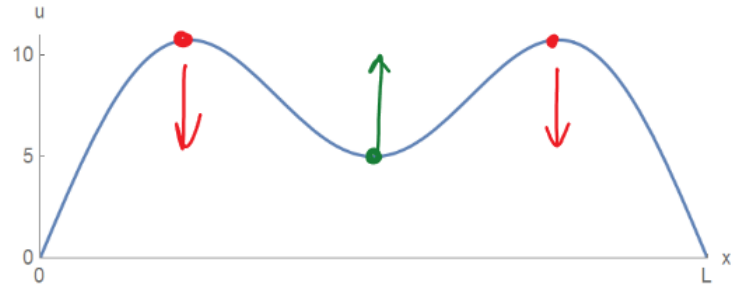
$$u(x,t) = 10\sin\left(\frac{x\pi}{L}\right)\exp\left(-\frac{\pi^2\kappa t}{L^2}\right) + 5\sin\left(\frac{3x\pi}{L}\right)\exp\left(-\frac{9\pi^2\kappa t}{L^2}\right)$$

Let's take a look at some graphs for fixed values of κt.



As we can see, over time the distribution heat becomes more uniform, and there are multiple ways of reasoning why this is. One reason is that, if we analyze the solution, we see that higher frequency Fourier modes (the terms corresponding to large n) decay faster, due to the corresponding large n in the exponent term. Essentially, as time passes, the waves of greater frequency become less significant.

Another way we can think about the relation $u_t = \kappa u_{xx}$ is from a graphical standpoint, relating concavity of u with respect to x to how u changes as t increases. $u_{xx} = \frac{\partial^2 u}{\partial x^2}$ is the concavity of u as a function of x, and $u_t = \frac{\partial u}{\partial t}$ is the rate of change of u with respect to t. κ is a diffusion constant which is greater than zero.



concave down $\frac{\partial^2 u}{\partial x^2} < 0$ so u decreases

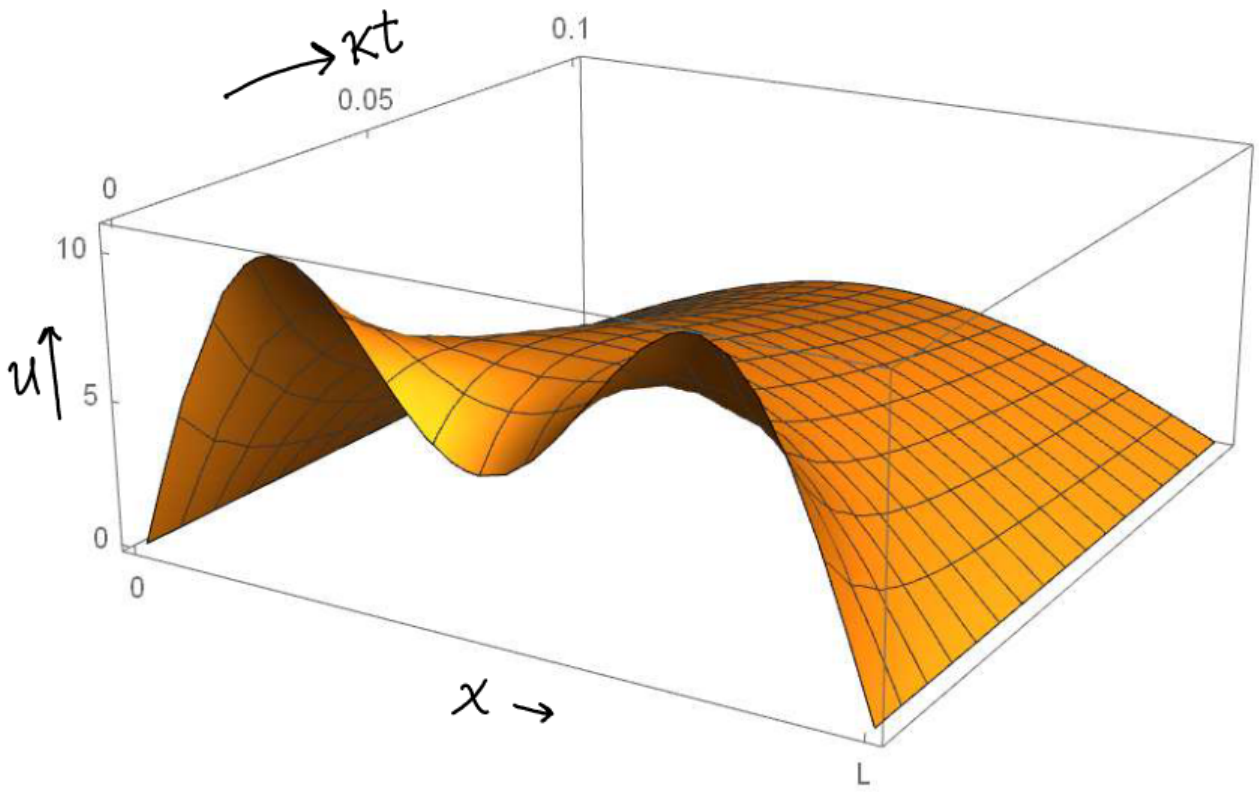
Concave up $\frac{\partial^2 u}{\partial x^2} > 0$ so u increases

For values of x near the middle, u is concave up, implying that $u_{xx} > 0$. The PDE relation implies that u will increase at that point x. On the other hand, for values of x around the first and third quarters of the domain, u is concave down, so the PDE implies that u should decrease there. Therefore as time passes, we see that as before, the solution will smooth out over time!

We may ask what the role of κ is in the solution. Observing $u(x,t)$ again, we see that κ only arises with t, that is, κt always appears together. So κ scales time. What if $\kappa < 0$? In this case, we get "negative diffusion," which would result in an unstable PDE in the sense that $u \to \infty$ as $t \to \infty$, which would break physical sense.

5.3 The Maximum Principle for the Heat Equation

Notice how the maximum of $u(x,t)$ (used from the previous subsection) occurs at $t = 0$, which is along the boundary of the solution domain, $\{0 \leq x \leq L, t \geq 0\}$. This is not a coincidence!



Theorem: (Weak Maximum Principle) Let $u(\vec{x}, t)$ be a continuous solution of the heat equation $u_t = \kappa \nabla^2 u$ with $\kappa > 0$ in a bounded region of $\vec{x}$ and $0 \leq t \leq T$. Then, the maximum and minimum values of u are attained along the boundary of $\vec{x}$ or at $t = 0$.

We will not provide a complete proof, but a full proof can be found in Levy & Shearer, chapter 5.4. However, we will provide an incomplete sketch in the case of 1 spatial dimension.

Proof. (Incomplete sketch in 1 spatial dimension) Suppose for contradiction that $u(x, t)$ attains a maximum at an interior point (x_0, t_0), so $u_t(x_0, t_0) = 0$ and $u_x(x_0, t_0) = 0$. We must have that $u_{xx}(x_0, t_0) < 0$, that is, the function is concave down. However, this is a contradiction, as then, the PDE cannot be satisfied unless $\kappa = 0$. Thus the function cannot have a true maximum at an interior point. □

Why is this not actually a complete proof? Counter-intuitively, a function can indeed have a true maximum at an interior point of its domain and still satisfy $u_{xx}(x_0, t_0) = 0$. For a simple example, consider the function $y = -x^4$. The actual proof uses the idea stated above though, it simply uses an argument to perturb, and thus eliminate, the degenerate interior maximum cases where $u_{xx} = 0$.

It is also worth noting that we must bound the region for t, as this makes the domain region a *compact* topological space. For those of you who have taken Real Analysis, one fact about

compact spaces is that any continuous function on a compact space must attain a maximum and minimum. Otherwise, there is no justification for a global maximum or minimum existing! We need the assumption that such a maximum exists.

The name "Weak Maximum Principle" suggests that there is a corresponding "strong" principle which should exist. There is in fact, an even stronger statement we can make about maxima of solutions to the heat equation.

Theorem: (Strong Maximum Principle) Let $u(\vec{x}, t)$ be a continuous solution of the heat equation $u_t = \kappa \nabla^2 u$ in a bounded region of $\vec{x}$ and $0 \leq t \leq T$. If the maximum of $u(\vec{x}, t)$ is attained at an interior point in the solution domain, then u must be a constant function.

The general proof of this theorem is quite involved. However, let's provide another more physical justification for why this should be so, again in 1 spatial dimension. Let $u(x, t)$ be a solution to the homogeneous Dirichlet problem:

$$\begin{aligned} u_t &= \kappa u_{xx} && \text{for } 0 < x < L, t > 0 \\ u(0,t) = u(L,t) &= 0 && \text{for } t > 0 \\ u(x,0) &= f(x) && \text{for } 0 < x < L \end{aligned}$$

We define an "energy" function $W(t)$ as follows:

$$W(t) = \frac{1}{2} \int_0^L u(x,t)^2 \, dx$$

Note if u has units of temperature, then W is proportional to the total heat energy. Let's compute the change in "energy" over time.

$$\begin{aligned} W'(t) &= \frac{d}{dt} \left(\frac{1}{2} \int_0^L u(x,t)^2 \, dx \right) \\ &= \int_0^L \frac{1}{2} \cdot \frac{\partial}{\partial t} u(x,t)^2 \, dx \\ &= \int_0^L u \cdot u_t \, dx \\ &= \kappa \int_0^L u \cdot u_{xx} \, dx \end{aligned}$$

Integrating by parts, we find:

$$= \kappa (u \cdot u_x)_{x=0}^{x=L} - \kappa \int_0^L u_x^2 \, dx$$

However, the boundary conditions imply that $u \cdot u_x$ is 0 at $x = 0$ and $x = L$. Therefore, we conclude:

$$W'(t) = -\kappa \int_0^L u_x^2 \, dx \leq 0$$

This implies that the total energy of the system must be strictly decreasing unless $u_x = 0$ for all values of x. This only happens if the temperature of the system is constant. The only constant solution which satisfies the boundary conditions is the zero solution! In any other case, the total energy is decreasing, and therefore we should expect that any maxima which occurs happens at the initial time of the PDE (with massive handwaving). Moreover, in any other case, since the energy is strictly decreasing, this could provide an argument for why solutions of this problem tend to 0 as $t \to \infty$ (again, massive handwaving).

5.4 Uniqueness of Solutions to the (original) Heat Equation

The previous argument also demonstrated that the only solution which is possible if the initial condition is given by $f(x) = 0$ is the zero solution. This important result also allows us to prove uniqueness of solutions to the given heat PDE problem.

Theorem: Let $u(x,t)$ be a solution to the homogeneous Dirichlet problem

$$\begin{aligned} u_t &= \kappa u_{xx} && \text{for } 0 < x < L, t > 0 \\ u(0,t) = u(L,t) &= 0 && \text{for } t > 0 \\ u(x,0) &= f(x) && \text{for } 0 < x < L \end{aligned}$$

It is unique.

Proof. Suppose that $v(x,t)$ and $w(x,t)$ are two solutions to the homogeneous Dirichlet problem above. We wish to show they are necessarily equivalent. We know the following:

$$\begin{aligned} v_t &= \kappa v_{xx} \\ v(0,t) = v(L,t) &= 0 \\ v(x,0) &= f(x) \end{aligned} \qquad \begin{aligned} w_t &= \kappa w_{xx} \\ w(0,t) = w(L,t) &= 0 \\ w(x,0) &= f(x) \end{aligned}$$

Because we have a homogeneous PDE with homogeneous boundary conditions, we can superimpose these solutions to find a solution to a PDE problem with the same boundary conditions! We find:

$$\begin{gathered} (v-w)_t = \kappa (v-w)_{xx} \\ v(0,t) - w(0,t) = 0 \\ v(L,t) - w(L,t) = 0 \\ v(x,0) - w(x,0) = f(x) - f(x) = 0 \end{gathered}$$

So $z(x,t) = v(x,t) - w(x,t)$ satisfies the same PDE problem, but with initial condition $f(x) = 0$. We already know this PDE problem can only have $z(x,t) = 0$ as the solution, so therefore $v(x,t) = w(x,t)$, as desired. $\square$

This proof technique works for any such heat PDE problem which satisfies homogeneous boundary conditions, with the proof following similarly as above, that is, applying the strong maximum principle.

5.5 Orthogonal Bases and Least-Squares Minimization

We finish up this section by asking how "good" of an approximation Fourier series are for general functions. Consider a finite set of nonzero objects $x_1, x_2, \ldots, x_n$ in a real inner product space which are orthogonal to each other, and let f be another object in the vector space. We would like to express f as a linear combination of the x_i's as closely as possible. In other words, we find to find scalar coefficients d_i such that the error, defined by

$$e = f - d_1 x_2 + d_2 x_2 + \cdots + d_n x_n$$

is as small as possible, in other words, that

$$||e|| = \sqrt{\langle e, e \rangle}$$

is minimized. First, we let

$$w = d_1 x_1 + \cdots + d_n x_n.$$

First, let us compute the norm of w.

$$\begin{aligned} ||w||^2 = \langle w, w \rangle &= \left\langle \sum_{i=1}^{n} d_i x_i, \sum_{j=1}^{n} d_j x_j \right\rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} d_i d_j \langle x_i, x_j \rangle \\ &= \sum_{i=1}^{n} d_i^2 ||x_i||^2 \end{aligned}$$

The last equality comes from the fact that all x_i are orthogonal. Now, let us compute the norm of $e = f - w$.

$$\begin{aligned} ||e||^2 = ||f - w||^2 &= \langle f - w, f - w \rangle \\ &= \langle f, f \rangle + \langle w, w \rangle - 2\langle f, w \rangle = ||f||^2 + ||w||^2 - 2\langle f, w \rangle \\ &= ||f||^2 + \sum_{i=1}^{n} d_i^2 ||x_i||^2 - 2 \sum_{i=1}^{n} d_i \langle f, x_i \rangle \end{aligned}$$

So in order to minimize $||e||^2$, we differentiate with respect to d_i and find when that equation is zero.

$$\frac{\partial(||e||^2)}{\partial d_i} = 2d_i||x_i||^2 - 2\langle f, x_i\rangle = 0$$

Therefore, we determine the d_i which minimizes error to be:

$$d_i = \frac{\langle f, x_i\rangle}{||x_i||^2} = \frac{\langle f, x_i\rangle}{\langle x_i, x_i\rangle}$$

Does this seem familiar? Recall from the first heat PDE problem we studied in class that we determined:

$$\sum_{n=1}^{\infty} D_n \sin\left(\frac{n\pi x}{L}\right) = f(x), \qquad D_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{n\pi x}{L}\right)\,dx$$

On the other hand, from what we have just derived, we should see that to minimize the error between $f = f(x)$ and $w = \sum_{n=1}^{N} D_n \sin\left(\frac{n\pi x}{L}\right)$ (for some finite N), then we should have

$$D_n = \frac{\langle f, x_i\rangle}{\langle x_i, x_i\rangle} = \frac{\int_0^L f(x)\sin\left(\frac{n\pi x}{L}\right)\,dx}{\int_0^L \sin^2\left(\frac{n\pi x}{L}\right)\,dx} = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{n\pi x}{L}\right)\,dx$$

In the case of Fourier series, we have an infinite number of orthogonal functions $x_i = \sin\left(\frac{i\pi x}{L}\right)$. As we increase the size of the set of functions, the error should thus minimize. It turns out that this infinite set of orthogonal functions can be used to represent *any* L^2 function $f(x)$ on the interval $0 \le x \le L$ with $f(0) = f(x) = 0$.

Definition: A L^2 function is one that is square integrable, that is, the integral:

$$\int_0^L |f(x)|^2\,dx < \infty$$

exists and is finite. (Note that we make no assumptions about the continuity of f.)

In other words, if we let

$$e_N = f(x) - \sum_{i=1}^{N} d_n x_n$$

be the error of the truncated sum using the first N functions, with d_n as above, then we have

$$\lim_{n\to\infty} ||e_n|| = 0$$

for any L^2 function $f(x)$ on $0 \le x \le L$ with $f(0) = f(L) = 0$. This is what is referred to as the **completeness** of the Fourier basis. More formally,

Definition: An orthogonal (usually orthonormal) basis $\{f_i\}_{i=1}^{\infty}$ is **complete** with respect to the inner product space V if for every $v \in V$, there exists a sequence $\{a_i\}_{i=1}^{\infty}$ such that the following limit is satisfied:

$$\lim_{n\to\infty} ||f - \sum_{i=1}^{n} a_i f_i|| = 0$$

Note that the completeness result is for the Fourier sine basis functions. The same is true for the usual Fourier basis:

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

or the complex version

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \exp(inx)$$

for L^2 functions on the interval $-\pi \leq x \leq \pi$.

We finish this section off with some identities which arise from the considerations we have made. Recall that we had the relation

$$||e||^2 = ||f||^2 + \sum_{i=1}^{n} d_i^2 ||x_i||^2 - 2\sum_{i=1}^{n} d_i \langle f, x_i \rangle.$$

If we make the error-minimizing substitution

$$d_i = \frac{\langle f, x_i \rangle}{||x_i||} \iff \langle f, x_i \rangle = d_i ||x_i||,$$

we obtain:

$$||e||^2 = ||f||^2 - \sum_{i=1}^{n} d_i^2 ||x_i||^2.$$

However, since $||e||^2 \geq 0$, we obtain the following inequality:

Bessel's Inequality:

$$||f||^2 \geq \sum_{i=1}^{n} d_i^2 ||x_i||^2$$

In the case when the orthogonal family of functions is complete, then as $n \to \infty$, then $||e|| \to \infty$ (by definition), so we obtain the following general identity:

Parseval's Identity:

$$||f||^2 = \sum_{i=1}^{\infty} d_i^2 ||x_i||^2$$

As an extra fun fact, this can be used to compete the well-known sum of inverse squares.

Theorem: (Basel Problem)

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

Proof. Consider the function $f(x) = x$ over the interval $0 \leq x \leq 1$ with $x_n(x) = \sin(n\pi x)$. We compute:

$$||f(x)||^2 = \langle f, f \rangle = \int_0^1 x^2 \, dx = \frac{1}{3}$$

$$||x_n||^2 = \langle x_n, x_n \rangle = \int_0^1 \sin(n\pi x)^2 dx = \frac{1}{2}$$

$$d_i = \frac{\langle f, x_n \rangle}{||x_n||} = 2 \int_0^1 x \sin(n\pi x) \, dx = \frac{2(-1)^{n+1}}{\pi n}$$

Then, Parseval's Identity becomes:

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

This reduces to

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

□

6 The Wave Equation, and A Very Common Eigenvalue Problem

6.1 A Very Common Eigenvalue Problem

It seems we've been running into eigenvalue problems over and over again when we use separation of variables. Let's recap the types of eigenvalue problems we've seen, and create a library of solutions which we can refer to for future problems.

Exercise: For the eigenvalue problem $y''(x) + \lambda y(x) = 0$ on the finite domain $0 \leq x \leq L$, pick one of the listed initial condition pairs and find all eigenvalue/eigenfunction pairs:

- $y(0) = 0$ Dirichlet, $y(L) = 0$ Dirichlet
- $y'(0) = 0$ Neumann, $y'(L) = 0$ Neumann
- $y(0) = 0$ Dirichlet, $y'(L) = 0$ Neumann
- $y'(0) = 0$ Neumann, $y'(L) = 0$ Dirichlet

In addition, find the eigenvalue/eigenfunctions of the eigenvalue problem $y''(x) + \lambda y(x) = 0$ on the finite domain $-L \leq x \leq L$ with periodic initial conditions $y(-L) = y(L)$ and $y'(-L) = y'(L)$.

Solution:

- $$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \qquad y_n = \sin\left(\frac{n\pi x}{L}\right), \qquad n = 1, 2, 3\ldots$$
- $$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \qquad y_n = \cos\left(\frac{n\pi x}{L}\right), \qquad n = 0, 1, 2\ldots$$
- $$\lambda_n = \left(\frac{(n+1/2)\pi}{L}\right)^2, \qquad y_n = \sin\left(\frac{(n+1/2)\pi x}{L}\right), \qquad n = 0, 1, 2\ldots$$
- $$\lambda_n = \left(\frac{(n+1/2)\pi}{L}\right)^2, \qquad y_n = \cos\left(\frac{(n+1/2)\pi x}{L}\right), \qquad n = 0, 1, 2\ldots$$
- $$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \qquad y_n = \exp\left(\frac{in\pi x}{L}\right), \qquad n = 0, \pm 1, \pm 2\ldots$$

  Cosines and sines can also be used instead of complex exponentials.

See the Canvas site for a nice table with these identities. You may use them at will.

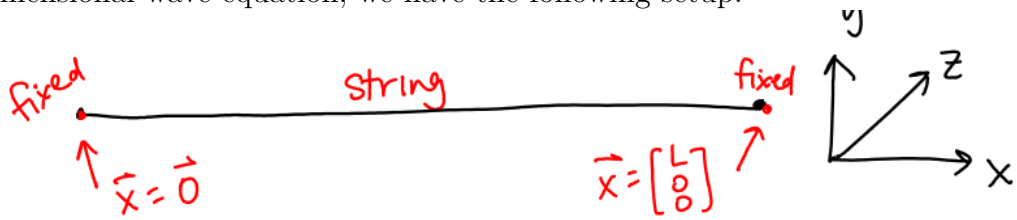
6.2 The Wave Equation

The **linear wave equation**, which describes the propagation of sound waves, electromagnetic waves, or waves in elastic solids, in what is assumed to be a perfectly elastic material, is given by the following PDE:

$$u_{tt} = c^2 \nabla^2 u$$

For the sake of time, I will not be deriving the wave equation. For a derivation in the one spatial dimension case, see Shearer & Levy chapter 4. Here, $c^2 = k/\rho$ is a constant given by the material, where ρ is the density of the material and k is a constitutive parameter called the *elastic modulus* (in general, it may not be constant). It depends on the elastic properties of the material, and may be measured in experiments. c has dimensions of speed, that is, length/time.

In general, a PDE describing wave motion in a material will not be linear, depending on the properties of the material. To obtain a linear PDE, we make the assumption that the tension is proportional to the stretched length of each infinitesimal section of the string. For the 1 dimensional wave equation, we have the following setup:



Then when the string vibrates, we model by u the difference from its resting mode.

This 1-D version of the wave equation, $u_{tt} = c^2 u_{xx}$ models the vibrations of a tautly held string with the following assumptions:

- Gravity is ignored, as it is minimal compared to the restoring force of the string.
- The string is made from a material that can be described by a linear restorative force.
- The string is displaced in such a way that horizontal displacements of each particle of the string can be neglected.
- The string is displaced in such a way that motion happens in one plane.

Exercise: Pick one PDE problem below. Work out the solution using separation of variables. Use your graphing software of choice (Desmos, Mathematica...) to visualize the solution,

using $f(x) = \max(0, 1 - 4(x-1)^2)$ with $L = 4$ and $c = 1$. What do you notice about the behavior of the solution?

DD:

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

DN:

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

NN:

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

ND:

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

We notice that in each case, the initial condition splits in half, one half going left and one going right with speed c. When the wave hits the boundary, it reflects back. If the boundary condition is Dirichlet, it will reflect back flipped upside down. If the boundary condition is Neumann, it will reflect back the way it was!

Suppose we have two Dirichlet boundary conditions, but no initial conditions, so our job is only to find all possible product solutions of the problem. This could be a crude model for a guitar string. Each product solution describes a different *mode of vibration* of the guitar string. Product solutions will be of the form

$$X(x)T(t) = \sin\left(\frac{n\pi x}{L}\right)\left(c_1\cos\left(\frac{cn\pi t}{L}\right) + c_2\sin\left(\frac{cn\pi t}{L}\right)\right),$$

where each of these modes is a *standing wave*. However, while the complete solution is a superposition of all these standing waves, it itself is not necessarily a standing wave! We will revisit this idea later when we derive d'Alembert's solution to the wave equation.

7 Inhomogeneous Boundary Conditions

7.1 An Inhomogeneous Problem

Inhomogeneous differential equations or boundary conditions often pose a slight difficulty when using separation of variables. Here are two examples:

1. The heat PDE problem:

$$\begin{aligned} u_t &= \kappa u_{xx}, \qquad 0 \le x \le L, \quad t \ge 0 \\ u(x,0) &= f(x) \\ u_x(L,t) &= 0 \\ u_x(0,t) &= \alpha(u(0,t) - T_0) \end{aligned}$$

   The last boundary condition is Robin and inhomogeneous. If we try to perform separation of variables, so $u = X(x)T(t)$, the second boundary condition becomes:

$$X'(0)T(t) = \alpha(X(0)T(t) - T_0 \implies X'(0) = \alpha\left(X(0) - \frac{T_0}{T(t)}\right)$$

   We can't eliminate $T(t)$ here, so there's nothing we can do using only tools we know about.

2. The Poisson equation, or the inhomogeneous Laplace equation:

$$\nabla^2\phi(x,y) = f(x,y)$$

   Trying $\phi = X(x)Y(y)$ gives:

$$X''Y + Y''X = f \implies \frac{X''}{X} + \frac{Y''}{Y} = \frac{f}{XY}$$

   We can't separate here either, boooooo.

There are two methods for dealing with problems involving inhomogeneous PDEs and/or boundary conditions:

- Change variables to eliminate the inhomogeneity.
- The Eigenfunction expansion method

Example: (1) Let's consider the heat PDE from before:

$$\begin{aligned} u_t &= \kappa u_{xx}, \qquad 0 \le x \le L, \quad t \ge 0 \\ u(x,0) &= f(x) \\ u_x(L,t) &= 0 \\ u_x(0,t) &= \alpha(u(0,t) - T_0) \end{aligned}$$

We need to deal with the T_0 term. Let's try making a change of variables to make our life easier.

Exercise: What change of variables could we make which makes sense here? Play around with this PDE and see if you can come up with something! We would ideally like some other variable v related to u, so that if u is a solution to the PDE, then v satisfies the first three constraints, but the second boundary condition becomes homogeneous.

Solution: If we want to eliminate the T_0 term, let's try adding T_0 to a new variable. Suppose $u(x,t)$ is a solution, then set:

$$u(x,t) = T_0 + v(x,t).$$

Let's see what happens:

$$\textbf{PDE: } u_t = \kappa u_{xx} \implies \frac{\partial}{\partial t}(T_0 + v) = \kappa \frac{\partial^2}{\partial x^2}(T_0 + v(x,t)) \implies v_t = \kappa v_{xx}$$

$$\textbf{IC: } u(x,0) = T_0 + v(x,0) = f(x) \implies v(x,0) = f(x) - T_0$$

$$\textbf{BC1: } u_x(L,t) = \frac{\partial}{\partial x}(T_0 + v(L,t)) = 0 \implies v_x(L,t) = 0$$

$$\textbf{BC2: } u_x(0,t) = \alpha(u(0,t) - T_0) \implies \frac{\partial}{\partial x}(T_0 + v(0,t)) = \alpha(v(0,t) + T_0 - T_0) \implies v_x(0,t) = \alpha v(0,t)$$

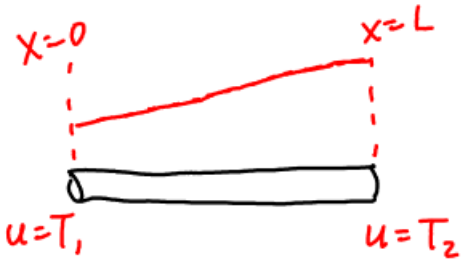
So to sum up, we have a new system:

$$\begin{aligned} v_t &= \kappa v_{xx}, \qquad 0 \le x \le L, \quad t \ge 0 \\ v(x,0) &= f(x) - T_0 \\ v_x(L,t) &= 0 \\ v_x(0,t) &= \alpha v(0,t) \end{aligned}$$

The inhomogeneous boundary condition is now a homogeneous one! Now we just need to solve for v using usual separation of variables, then substitute back to get u. This particular change of variables is something akin to changing Kelvin to Celsius, except that we're defining the T_0 temperature to be zero in some new temperature scale.

Example: (2) Consider a new heat PDE with Dirichlet boundary conditions as follows:

$$\begin{aligned} u_t &= \kappa u_{xx}, \qquad 0 \le x \le L, \quad t \ge 0 \\ u(x,0) &= f(x) \\ u(0,t) &= T_1 \\ u(L,t) &= T_2 \end{aligned}$$



Here, we use the substitution:

$$u(x,t) = T_1 + \frac{x}{L}(T_2 - T_1) + v(x,t)$$

Again, assuming $u(x,t)$ is a solution to the PDE, let's see how the IVP translates to $v(x,t)$.

$$\textbf{PDE: } u_t = v_t \text{ and } u_{xx} = v_{xx} \implies v_t = \kappa v_{xx}$$

$$\textbf{IC: } u(x,0) = f(x) = T_1 + \frac{x}{L}(T_2 - T_1) + v(x,0) \implies v(x,0) = f(x) - T_1 - \frac{x}{L}(T_2 - T_1)$$

$$\textbf{BC1: } u(0,t) = T_1 = T_1 + \frac{0}{L}(T_2 - T_1) + v(0,t) \implies v(0,t) = 0$$

$$\textbf{BC2: } u(L,t) = T_2 = T_1 + \frac{L}{L}(T_2 - T_1) + v(L,t) \implies v(L,t) = 0$$

We have found a new IVP with homogeneous boundary conditions! Solve for v, and we obtain u as well.

Definition: Define $u_s(x) = \lim_{t\to\infty} u(x,t)$ to be the **steady-state solution** of a PDE. If such a thing exists, then it would satisfy the same PDE that $u(x,t)$ does, except all t-derivatives vanish.

Often, the steady-state solution, if one exists, is a good candidate for a substitution that "homogenizes" the PDE, or BCs. From the example above, we see that $u_t = \kappa u_{xx}$ has a steady state solution $u_s(x)$ which satisfies

$$(u_s)_t = 0 = \kappa (u_s)_{xx} \implies u_s(x) = ax + b.$$

Since $u_s(0) = T_1$ and $u_s(L) = T_2$, we can easily solve to find

$$u_s(x) = T_1 + \frac{x}{L}(T_2 - T_1)$$

7.2 Eigenfunction Expansions

However, a change of variables may not always work for turning an inhomogeneous problem into a homogeneous one. In particular, if the *PDE* is where the inhomogeneity resides, rather than the initial data, we instead will have to turn to the method of **eigenfunction expansions.** We introduce and illustrate the method with an example.

Example: Solve $\nabla^2 u(x,y) = x - L/2$, a Poisson equation, in the domain $0 \le x \le L$, $y \ge 0$, subject to homogeneous Neumann boundary conditions at $x = 0$ and $x =$), a homogeneous Dirichlet boundary condition at $y = 0$, and a boundedness condition $\lim_{y\to\infty} u(x,y) < \infty$.

Solution: We first ignore the inhomogeneity in the PDE, so we pretend we are solving **Laplace's Equation** for a different function $v(x, y)$

$$\nabla v(x, y) = v_{xx} + v_{yy} = 0,$$

using separation of variables. As usual, let $v(x, y) = X(x)Y(y)$ and plug in to get

$$\frac{X''}{X} + \frac{Y''}{Y} = 0$$

This is easily separated so we set $X''/X = -\lambda$ and obtain the system of ODEs:

$$\begin{cases} X'' + \lambda X = 0 \\ Y'' - \lambda Y = 0 \end{cases}$$

The boundary conditions become:

$$\begin{cases} v_x(0, y) = 0 & \implies X'(0) = 0 \\ X'(L) = 0 & \implies v_x(L, y) = 0 \\ v(x, 0) = 0 & \implies Y(0) = 0 \end{cases}$$

The X ODE is 2nd order and has two initial conditions, $X'(0) = X'(L) = 0$, so it is our eigenfunction problem. By a process which by now, we are well-acquainted with, we determine that the ODE gives us corresponding eigenfunction and eigenvalue pairs

$$X_n(x) = \cos\left(\frac{n\pi x}{L}\right), \qquad \lambda_n = \left(\frac{n\pi x}{L}\right)^2 \quad (n = 0, 1, 2, \dots)$$

(side remark) If we were to continue solving Laplace's equation, then we would next solve the other ODE, $Y'' - \left(\frac{n\pi}{L}\right)^2 Y = 0$. If we were to do this, by a similar process we would find that

$$Y_n(y) = C_1 \exp\left(\frac{n\pi y}{L}\right) + C_2 \exp\left(-\frac{n\pi y}{L}\right) \quad (n > 0).$$

Since $Y \not\to \infty$ as $y \to \infty$, we must have that $C_1 = 0$ and the Dirichlet boundary condition $Y(0) = 0$ then gives that $C_2 = 0$, and hence $Y(y) = 0$ (for $n \neq 0$), a trivial solution.

HOWEVER, this is all moot, because we are solving $\nabla^2 u(x, y) = x - L/2$, not $\nabla^2 v(x, y) = 0$. This is where the method of eigenfunction expansions differs. The big idea is as follows:

Method of Eigenfunction Expansions: *Use the eigenfunctions of the homogeneous problem as an inspiration for a different solution.* Explicitly in this case, where

$$v(x, y) = \sum_{n=0}^{\infty} \cos\left(\frac{n\pi x}{L}\right) \cdot A_n \cdot Y_n(y),$$

(where A_n is a constant coefficient) we make the ansatz that $u(x, y)$ has a similar form,

$$u(x, y) = \sum_{n=0}^{\infty} \cos\left(\frac{n\pi x}{L}\right) \cdot A_n(y)$$

(where A_n here is a function of y.) This is a plausible guess for the solution, since it is not hard to show that $u_x(0, y) = u_x(L, y) = 0$ (because of the eigenfunctions). So $u(x, y)$ must solve

- $u_{xx} + u_{yy} = x - L/2$ (1)
- $u(x, 0) = 0$ (2)
- $u(x, y) \not\to \infty$ as $y \to \infty$ (3)

If we plug the ansatz into (1), we get:

$$u_{xx} = -\sum_{n=0}^{\infty} A_n(y) \cos\left(\frac{n\pi x}{L}\right) \cdot \left(\frac{n\pi}{L}\right)^2$$
$$u_{yy} = \sum_{n=0}^{\infty} A_n''(y) \cos\left(\frac{n\pi x}{L}\right)$$

Putting this together, we find:

$$u_{xx} + u_{yy} = \sum_{n=0}^{\infty} \cos\left(\frac{n\pi x}{L}\right) \cdot \left(A_n''(y) - \left(\frac{n\pi}{L}\right)^2 A_n(y)\right) = x - \frac{L}{2}$$

We can use orthogonality of cosines here! Let's try multiplying by another eigenfunction and integrating from 0 to L, to isolate one term in the sum. For the case when $m > 0$,

$$\int_0^L \left(x - \frac{L}{2}\right) \cos\left(\frac{m\pi x}{L}\right) dx = \int_0^L \left(\sum_{n=0}^{\infty} \cos\left(\frac{n\pi x}{L}\right) \cdot \left(A_n''(y) - \left(\frac{n\pi}{L}\right)^2 A_n(y)\right)\right) \cos\left(\frac{m\pi x}{L}\right) dx$$
$$= \frac{L}{2}\left(A_m''(y) - \left(\frac{m\pi}{L}\right)^2 A_m(y)\right)$$

On the other hand, the integral on the left hand side can be explicitly computed using separation of variables.

$$\int_0^L \left(x - \frac{L}{2}\right) \cos\left(\frac{m\pi x}{L}\right) dx = \frac{L^2((-1)^m - 1)}{\pi^2 m^2}$$
$$= \begin{cases} \frac{-2L^2}{\pi^2 m^2} & m \text{ odd} \\ 0 & m \text{ even} \end{cases}$$

For the $m = 0$ case, we get:

$$0 = \int_0^L \left(x - \frac{L}{2}\right) dx = \int_0^L \sum_{n=0}^{\infty} \cos\left(\frac{n\pi x}{L}\right) \cdot \left(A_n''(y) - \left(\frac{n\pi}{L}\right)^2 A_n(y)\right) dx$$
$$= L(A_0''(y) - 0 \cdot A_0(y))$$

So to sum up, we obtain an infinite set of linear, constant coefficient ODEs to solve:

$$A_n''(y) - \left(\frac{n\pi}{L}\right)^2 A_n(y) = \begin{cases} -\frac{4L}{\pi^2 n^2} & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$$

We will find the initial conditions from condition (2), $u(x, 0) = 0$.

$$u(x, 0) = \sum_{n=0}^{\infty} A_n(0) \cos\left(\frac{n\pi x}{L}\right) = 0 \implies A_n(0) = 0$$

Lastly, condition (3), that $u(x, y) \not\to \infty$ as $y \to \infty$ implies that $A_n(y) \not\to \infty$ as well. So in total we have three cases of ODEs to solve:

1. $n = 0$ case:

$$A_0''(y) = 0 \implies A_0 = a_0 y + b_0$$
$$A_0(0) = 0 \implies b_0 = 0$$
$$|A_0(\infty)| < \infty \implies a_0 = 0$$

   So $A_0(y) = 0$.

2. $n > 0$ is even case:

$$A_n''(y) - \left(\frac{n\pi}{L}\right)^2 A_n(y) = 0$$

   This ODE has general solution:

$$A_n = a_n \exp\left(-\frac{n\pi y}{L}\right) + b_n \exp\left(\frac{n\pi y}{L}\right)$$

   The boundedness condition implies $b_n = 0$, and plugging in the initial condition $A_n(0) = 0$ implies $a_n = 0$ as well. Thus, $A_n(y) = 0$ if n is even.

3. n is odd case:

$$A_n''(y) - \left(\frac{n\pi}{L}\right)^2 A_n(y) = -\frac{4L}{\pi^2 n^2}$$

   This ODE has general solution similar to the even case, only differing by a constant term:

$$A_n = \frac{4L^3}{n^4 \pi^4} + a_n \exp\left(-\frac{n\pi y}{L}\right) + b_n \exp\left(\frac{n\pi y}{L}\right)$$

As before, the boundedness condition implies $b_n = 0$. Plugging in the initial condition $A_n(0) = 0$ implies

$$a_n = -\frac{4L^3}{n^4\pi^4}$$

Hence, we conclude for n odd,

$$A_n(y) = \frac{4L^3}{n^4\pi^4}\left(1 - \exp\left(-\frac{n\pi y}{L}\right)\right)$$

So to conclude, our final answer is:

$$u(x,t) = \sum_{n=1,\text{ odd}}^{\infty} \frac{4L^3}{n^4\pi^4}\left(1 - \exp\left(-\frac{n\pi y}{L}\right)\right)\cos\left(\frac{n\pi x}{L}\right)$$

7.3 Choosing Strategies

It is not always clear which strategy one should use when attempting to solve an inhomogeneous PDE problem. Let's work through examples of problems where we have to make some adjustments to the usual separation of variables method we use.

Exercise: For each of the PDE problems posed, make sense of what the equations represent, and outline a solution strategy with as much detail as possible while doing as few calculations as possible. The solution should either present a way of homogenizing the problem, or describe a way we can use eigenfunctions to find a solution.

$$\begin{cases} u_t = \kappa\nabla^2 u(x,t) \\ u(0,t) = f(t) \\ u_x(L,t) = 0 \\ u(x,0) = 0 \end{cases} \qquad \begin{cases} w_t = \kappa\nabla^2 w(x,y,t) \\ w(0,y,t) = w(L,y,t) = T \\ w_x(x,0,t) = w_x(x,H,t) = 0 \\ w(x,y,0) = h(x,y) \end{cases}$$

$$\begin{cases} v_t = \kappa\nabla^2 v(x,t) + s(x,t) \\ v(0,t) = 0 \\ v_x(L,t) = 0 \\ v(x,0) = g(x) \end{cases} \qquad \begin{cases} z_t = \kappa\nabla^2 z(x,y,t) \\ z(0,y,t) = z(L,y,t) = T(t) \\ z_x(x,0,t) = z_x(x,H,t) = 0 \\ z(x,y,0) = 0 \end{cases}$$

Solution:

- This is a deceptive scenario, initially it may appear that all is well. However, for $u(x,t)$, we see that the inhomogeneity lies in an x boundary condition rather than a t initial condition. As eigenfunctions may not necessarily arise when working with the t ODE which arises, we should perform a change of variables which homogenizes the boundary conditions, and this will as a result make the initial condition inhomogeneous (which we know how to deal with). We perform the substitution $u(x,t) = f(t) + v(x,t)$. Working

out the corresponding IVP for $v(x,t)$, we see that it is actually the same formatted (up to names) IVP as $v(x,t)$ in the next case, which we now detail how to work with.

- For $v(x,t)$, the new issue lies in the PDE itself, rather than the ICs and BCs. Since the inhomogeneity lies in the PDE, we will have to use the method of eigenfunction expansions. The corresponding eigenvalue problem will have eigenfunctions of the form

$$\sin\left(\frac{(n+1/2)\pi x}{L}\right),$$

so we make the ansatz that the problem has solution of the form

$$v(x,t) = \sum_{n=0}^{\infty} \sin\left(\frac{(n+1/2)\pi x}{L}\right) A_n(t)$$

and work through the problem in the same way as before, using orthogonality to obtain an infinite set of ODEs.

- For $w(x,t)$, we see that the problem is nearly solvable, except for the Dirichlet boundary condition on x, which is raised by a constant. This is not so hard to resolve. We make a substitution $w = T + \tilde{w}$, which resolves the inhomogeneous boundary condition and does not affect the PDE. The only change to the other boundary and initial conditions is that the inhomogeneous initial condition is affected by a factor of T, but this is fine.

- The $z(x,y,t)$ case is similar to the u case, where because the inhomogeneity is in a boundary condition, we may not be able to find the proper eigenfunctions without first making a substitution. We make the substitution $z = T(t) + \tilde{z}$ - this homogenizes the inhomogeneous boundary condition, dehomogenizes the initial condition (which is fine), but also dehomogenizes the PDE. However, from there, we can use eigenfunction expansions to solve this PDE. (but it'll take some work!)

Some general observations which can be made:

- When separating variables, only homogeneous boundary conditions yield corresponding boundary conditions for the associated ODEs. We require homogeneous boundary conditions to have valid eigenfunction problems.

- If we encounter inhomogeneous boundary conditions, sometimes we can use a change of variables which can homogenize them. As much as possible, we want to try not to mess up other parts fo the problem with the change of variables.

- Eigenfunction expansion method involves coming up with a representation of the solution using eigenfunctions of the homogeneous version of the problem.

7.4 D'Alembert's Solution

Recall back from homework one, where I had you verify that the so-called *d'Alembert's Solution*,

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

satisfies the one-dimensional wave equation $u_{tt} = c^2 u_{xx}$, with initial conditions $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$. Let's derive this solution!

Example: Solve the PDE problem:

$$\begin{aligned} u_{tt} &= c^2 u_{xx} \qquad \text{on} \qquad -\infty < x < \infty, t > 0 \\ u(x,0) &= f(x) \\ u_t(x,0) &= g(x) \end{aligned}$$

We define a change of variables, $\xi = x - ct$ and $\eta = x + ct$. Let $U(\xi,\eta) = u(x,t)$. Then, the chain rule gives us:

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \xi}\frac{\partial \xi}{\partial t} + \frac{\partial}{\partial \eta}\frac{\partial \eta}{\partial t} = -c\frac{\partial}{\partial \xi} + c\frac{\partial}{\partial \eta}$$

This implies:

$$u_t = -cU_\xi + cU_\eta,$$

Similarly, we have for x,

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial}{\partial \eta}\frac{\partial \eta}{\partial x} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}$$

This implies:

$$u_x = U_\xi + U_\eta.$$

Let us now compute u_{tt} and u_{xx}. We do so by noting $u_{tt} = (u_t)_t = \frac{\partial}{\partial t}(u_t)$. Then, we have:

$$\begin{aligned} (u_t)_t &= \left[-c\frac{\partial}{\partial \xi} + c\frac{\partial}{\partial \eta}\right](-cU_\xi + cU_\eta) \\ &= c^2 U_{\xi\xi} - 2c^2 U_{\xi\eta} + c^2 U_{\eta\eta} \end{aligned}$$

$$\begin{aligned} c^2(u_x)_x &= c^2\left[\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right](U_\xi + U_\eta) \\ &= c^2(U_{\xi\xi} + 2U_{\xi\eta} + U_{\eta\eta}) \end{aligned}$$

However, we also have that $u_{tt} = c^2 u_{xx}$ therefore,

$$c^2 U_{\xi\xi} - 2c^2 U_{\xi\eta} + c^2 U_{\eta\eta} = c^2(U_{\xi\xi} + 2U_{\xi\eta} + U_{\eta\eta}) \implies U_{\xi\eta} = 0.$$

Therefore, we have that $U(\xi,\eta)$ is of the form:

$$U(\xi,\eta) = A(\xi) + B(\eta) \iff u(x,t) = A(x-ct) + B(x+ct).$$

Furthermore, the initial conditions give us:

$$u(x,0) = f(x) = A(x) + B(x)$$
$$u_t(x,0) = g(x) = -cA'(x) + cB'(x)$$

Equivalently,

$$-\frac{1}{c}g(x) = A'(x) - B'(x).$$

Let's integrate each side by x from 0 to z (for any arbitrary real z) - we find:

$$-\frac{1}{c}\int_0^z g(\tau)d\tau = A(z)-A(0)-B(z)+B(0) \iff -\frac{1}{c}\int_0^z g(\tau)d\tau+A(0)-B(0) = A(z)-B(z).$$

Also,

$$f(z) = A(z) + B(z),$$

so we have a system of equations which we can solve for $A(z), B(z)$. We find by adding the equations that:

$$2A(z) = f(z) - \frac{1}{c}\int_0^z g(\tau)d\tau + A(0) - B(0)$$

and similarly, by subtracting, we find

$$2B(z) = f(z) + \frac{1}{c}\int_0^z g(\tau)d\tau - A(0) + B(0)$$

Putting everything together, we have:

$$\begin{aligned} u(x,t) &= A(\xi) + B(\eta) = A(x-ct) + B(x+ct) \\ &= \frac{1}{2}(f(x-ct) + f(x+ct)) - \frac{1}{2c}\int_0^{x-ct} g(\tau)d\tau + \frac{1}{2c}\int_0^{x+ct} g(\tau)d\tau \\ &= \frac{1}{2}(f(x-ct) + f(x+ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\tau)d\tau \end{aligned}$$

8 Introduction to Integral Transforms

Integral transforms (like the Fourier and Laplace transforms) have multitudes of uses, such as signal processing and the study of control systems. In this class, we will use integral transforms to solve differential equations. First, here are a few transforms:

- **Fourier**:
$$F(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx = \mathscr{F}(f(x))$$

- **Inverse Fourier**:
$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(k)e^{ikx}\,dk = \mathscr{F}^{-1}(F(k))$$

- **Laplace**:
$$F(s) = \int_{0}^{\infty} f(t)e^{-st}\,dt = \mathscr{L}(f(t))$$

- **Inverse Laplace**:
$$f(t) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} F(s)e^{st}\,dx = \mathscr{L}^{-1}(F(s)),$$
where the integration is performed on the real line $\mathrm{Re}(z) = a$, all singularities of $F(s)$ occur to the left of the line, and $F(s)$ is bounded on the line.

- **Hankel**:
$$F(\lambda) = \int_{0}^{\infty} f(r)J_n(\lambda r)\,dr,$$
where J_n is a Bessel function of the first kind.

- **General form**:
$$F(\xi) = \int_{a}^{b} f(x)K(\xi, x)\,dx,$$
where K is the **kernel** of the transform.

8.1 Introduction

To motivate the Fourier transform, let's compare it to the Fourier series:

Fourier Series

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

The *wave number* n is discrete, and x is continuous.

Fourier Transform

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(k)e^{ikx}\,dk$$
$$F(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx$$

The *wave number* k is continuous, and x is continuous.

For notation and physical intuition, if we have $f(x)$ a spatial function with x the spatial variable in a physical domain, then in the Fourier domain, $F(k) = \mathscr{F}(F(x))$ (sometimes written $\hat{f}(k)$) is a function of k, the **wave number**. On the other hand, if $f(t)$ is a function of t time in the physical domain, then in the Fourier domain, $F(\omega) = \mathscr{F}(f(x))$ has ω as a variable of frequency. Essentially, spatial or time functions are decomposed by the Fourier transform into functions of spatial or temporal frequency.

Usually the Laplace transformation is applied to functions of time $f(t)$, and in the Laplace domain, we generally write $L(s) = \mathscr{L}(f(t))$. Generally, we will use lower-case letters to denote functions in physical variables, and upper-case letters, or letters with a carat on top, for functions that have been transformed. It should be clear which transformation has been applied.

Definition: A function $f(x)$ on $\mathbb{R}$ is **absolutely integrable** if $\int_{-\infty}^{\infty} |f(x)|\,dx < \infty$. $f(x)$ is **piecewise continuous** if it is continuous except at a finite number of places, and each discontinuity is finite.

Theorem: (Fourier Integral Theorem)
Suppose $f(x)$ is piecewise continuous, absolutely integrable, and has the property that $F(k) = \mathscr{F}(f(x)) \le \frac{M}{|k|}$ for some M and sufficiently large k. Then,

$$\begin{aligned}\mathscr{F}^{-1}(\mathscr{F}(f(x))) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx\,e^{ikx}\,dk \\ &= \frac{f(x^+) + f(x^-)}{2}\end{aligned}$$

where $f(x^{\pm}) = \lim_{a\to x^{\pm}} f(a)$. In particular, if $f(x)$ is continuous, then $f(x^+) = f(x^-) = f(x)$, so indeed, the Fourier Transform and Inverse Fourier Transform are inverses.

Given these transforms, we should be wary of applying them to any function, as the infinite integrals may not converge! Under what conditions will the transformations be well-behaved?

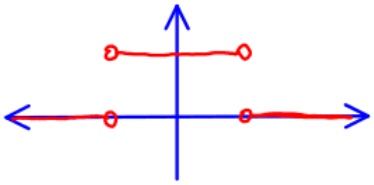
Question: The Fourier transform is an integral over all of $\mathbb{R}$. How can we be sure the integral will converge?
Answer: If $f(x)$ is Riemann integrable on every interval $[a, b]$ and absolutely integrable, then the Fourier transform will be well-defined. However, Fourier transforms for functions like $e^{ik_0 x}$ that aren't absolutely integrable can still be defined in a meaningful way.
If a function $f(x)$ is absolutely integrable, then $F(k) = \mathscr{F}(f(x)) \to 0$ as $k \to \pm\infty$. Since $\mathscr{F}$ and $\mathscr{F}^{-1}$ are essentially the same operation, the analogous statement applies to $\mathscr{F}^{-1}$ as well.

Question: How about the Laplace transform?
Answer: Laplace transform can be applied to functions that are piecewise continuous and grow no faster than an exponential, that is, $|f(t)| \leq Me^{at}$ for some $M, a \in \mathbb{R}$.

Example: Let's compute some Fourier transforms!

1. Boxcar function:

$$I_a(x) = \begin{cases} 1 & |x| < a \\ 0 & |x| > a \end{cases}$$



$$I_a(x) = \begin{cases} 1 & |x| < a \\ 0 & |x| > a \end{cases}$$

$$\begin{aligned}
\mathscr{F}(I_a(x)) &= \int_{-\infty}^{\infty} I_a(x) e^{-ikx}\, dx \\
&= \int_{-a}^{a} 1 \cdot e^{-ikx}\, dx \\
&= \left(\frac{e^{-ikx}}{-ik} \right)_{x=-a}^{x=a} \\
&= \frac{e^{-ika} - e^{ika}}{-ik} \\
&= \frac{(\cos(ka) - i\sin(ka)) - (\cos(ka) + i\sin(ka))}{-ik} \\
&= \frac{-2i\sin(ka)}{-ik} = \frac{2\sin(ka)}{k}
\end{aligned}$$

On the other hand, it may be computed that

$$\begin{aligned}\mathscr{F}^{-1}\left(\frac{2\sin(ka)}{k}\right) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2\sin(ka)}{k}e^{ikx}\,dk \\ &= \begin{cases}1 & |x|<a \\ 1/2 & |x|=a \\ 0 & |x|>a\end{cases} \\ &= \frac{I_a(x^+)+I_a(x^-)}{2}\end{aligned}$$

On the other hand, what is the inverse Fourier transform of $I_a(k)$? It should probably look similar:

$$\begin{aligned}\mathscr{F}^{-1}(I_a(k)) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}I_a(k)e^{ikx}\,dk \\ &= \frac{1}{2\pi}\int_{-a}^{a}e^{ikx}\,dk \\ &= \left(\frac{e^{ikx}}{2\pi ix}\right)_{k=-a}^{k=a} \\ &= \frac{\sin(ax)}{\pi x}\end{aligned}$$

So we have a correspondence as follows:

$$\begin{array}{rcl}\text{Physical domain} & \Longleftrightarrow & \text{Fourier domain} \\ I_a(x) & \Longleftrightarrow & \dfrac{2\sin(ka)}{k} \\ \dfrac{\sin(ax)}{\pi x} & \Longleftrightarrow & I_a(k)\end{array}$$

2. Heaviside Function/Step Function: Let $f(t)=e^{-at}H(t)$, where $H(t)$ is the **Heaviside Function**

$$H(t)=\begin{cases}0 & t<0 \\ 1 & t>0\end{cases},$$

and $Re(a)>0$. Calculate $\hat{f}(\omega)=\mathscr{F}(f(t))$ and $F(s)=\mathscr{L}(f(t))$.

$$\begin{aligned}\hat{f}(\omega) &= \int_{-\infty}^{\infty} f(t)e^{i\omega t}\,dt \\ &= \int_{-\infty}^{\infty} e^{-at}H(t)e^{i\omega t}\,dt \\ &= \int_{0}^{\infty} e^{-(a+i\omega)t}\,dt \\ &= \left(-\frac{e^{(a+i\omega)t}}{a+i\omega}\right)\end{aligned}$$

Since we have made the assumption that $Re(a) > 0$, we have that the limit $\lim_{t\to\infty} e^{-(a+i\omega)t} = 0$ (where ω is real).

$$= \frac{1}{a+i\omega}$$

Next, we compute $F(s)$.

$$\begin{aligned}F(s) &= \int_{0}^{\infty} f(t)e^{-st}\,dt \\ &= \int_{0}^{\infty} e^{-at}e^{-st}\,dt \\ &= \int_{0}^{\infty} e^{-(a+s)t}\,dt \\ &= \left(-\frac{e^{-(a+s)t}}{a+s}\right)_{t=0}^{t\to\infty}\end{aligned}$$

Again, since $Re(a) > 0$, $\lim_{t\to\infty} e^{-(a+s)t} = 0$ in the complex domain $Re(s) > 0$.

$$= \frac{1}{a+s}$$

So in the complex domain $Re(s) > 0$, we find that $F(s)$ converges to the function $\frac{1}{a+s}$.

8.2 Properties of Fourier & Laplace

1. These integral transform operators are linear, in other words,

   $$\mathscr{F}(af(x) + bg(x)) = a\mathscr{F}(f(x)) + b\mathscr{F}(g(x)),$$

   and same for the Laplace transform, and their inverses.

2. How does translation affect the transforms? Given some function $f(x)$ for which we assume $f(x) = 0$ for $x < 0$, let's compute $\mathscr{L}(f(t-T))$.

   $$\mathscr{L}(f(t-T)) = \int_{0}^{\infty} f(t-T)e^{-st}dt.$$

Let $\tau = t - T$. Then $d\tau = dt$. Perform a substitution for t, dt.

$$\begin{aligned} &= \int_{\tau=-T}^{\infty} f(\tau) e^{-s(\tau+T)} d\tau \\ &= \int_{0}^{\infty} f(\tau) e^{-s\tau} e^{-sT} d\tau \\ &= e^{-sT} \mathscr{L}(f(t)) = e^{-sT} F(s) \end{aligned}$$

So translation of the function in the time domain is equivalent to multiplication by an exponential in the Laplace domain. A similar effect occurs for the Fourier transform, one may compute by a similar computation that:

$$\mathscr{F}(f(x - x_0)) = e^{-ikx_0} \mathscr{F}(f(x)) = e^{-ikx_0} F(k)$$

In fact, one does not need the support of the function f to be positive for the Fourier transform.

3. Derivatives transform to algebra! This makes these integral transforms very useful for solving differential equations. First in the case of the Laplace transform, assuming a certain growth condition, we find:

$$\begin{aligned} \mathscr{L}(f'(t)) &= \int_0^{\infty} f'(t) e^{-st}\, dt \\ &= (f(t)e^{-st})_{t=0}^{t=\infty} + s \int_0^{\infty} f(t) e^{-st}\, dt \end{aligned}$$

We assume $\lim_{t\to\infty} f(t)e^{-st} = 0$, in other words, that the growth of $f(t)$ isn't too fast.

$$\begin{aligned} &= 0 - f(0) + s\mathscr{L}(f(t)) \\ &= sF(s) - f(0) \end{aligned}$$

Exercise: Suppose $\mathscr{F}(f(x)) = F(k)$ and $f(x) \to 0$ as $x \to \pm\infty$. Compute $\mathscr{F}(f'(x))$.

Solution: We compute, using integration by parts:

$$\begin{aligned} \mathscr{F}(f'(x)) &= \int_{-\infty}^{\infty} f'(x) e^{-ikx}\, dx \\ &= (e^{-ikx} f(x))_{x=-\infty}^{x=\infty} - \int_{-\infty}^{\infty} f(x) e^{-ikx}(-ik)\, dx \\ &= 0 + ik \int_{-\infty}^{\infty} f(x) e^{-ikx}\, dx = ikF(k) \end{aligned}$$

One may note that we must assume $f(x) \to 0$ as $x \to \pm\infty$, because we need $f(x)$ to be absolutely integrable to compute its Fourier transform, just as we did for the Fourier series.

8.3 The Dirac Delta Function

Definition: Let

$$\delta_\epsilon(x) = \begin{cases} 0 & |x| > \epsilon \\ \frac{1}{2\epsilon} & |x| < \epsilon \end{cases}$$

Then, the **Dirac delta function** $\delta(x)$ is defined as

$$\delta(x) = \lim_{\epsilon \to 0^+} \delta_\epsilon(x)$$

Delta functions model point charges and force "impulses" - concentration of force at an infinitesimally short amount of time. Observe that the following property holds for all $\epsilon > 0$:

$$\int_{-\infty}^{\infty} \delta_\epsilon(x)\, dx = 1.$$

We constrain the Dirac delta function to hold the same property,

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1.$$

Caution: As an aside, the approach we are taking to defining the Dirac delta function is not quite rigorous, as there is clearly no real function defined on the real numbers which satisfies the above property. To define the Dirac delta function, we may do so in two equivalent ways, defining it as a **measure**, a function which inputs a subset of $\mathbb{R}$ and outputs a real number, or as a **distribution**, essentially a generalized function which is defined in relation to how it affects other "test functions" when integrated against them (this is the same notion as when we talked about "weak solutions" to a PDE). For a technical rundown (I do recommend this, especially if you're interested in studying analysis or PDEs further), the Wikipedia page https://en.wikipedia.org/wiki/Dirac_delta_function has a very accessible rundown.

Recall the Heaviside function:

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

Notice:

$$\int_{-\infty}^{x} \delta(\xi) d\xi = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$

This means that in a sense, $H'(x) = \delta(x)$. Again, this is not rigorously true, but if we define $H(x)$ and $\delta(x)$ as distributions, this will have a more precise meaning.

Exercise: What is $\int_{-\infty}^{\infty} 5\delta(\xi - x_0)\, d\xi$, a shifted delta function?

Solution: Intuitively, we see that the shifting does not change the value of the integral, and the scaling can be pulled out. So it is 5.

Exercise: What is $\int_{-\infty}^{\infty} f(\xi)\delta(\xi - x_0)\, d\xi$?

Solution:

$$\begin{aligned}\int_{-\infty}^{\infty} f(\xi)\delta(\xi - x_0)\, d\xi &= \int_{-\infty}^{\infty} f(x_0)\delta(\xi - x_0)\, d\xi \\ &= f(x_0)\int_{-\infty}^{\infty} \delta(\xi - x_0) d\xi \\ &= f(x_0)\end{aligned}$$

So integration is super easy whenever there's a delta function involved - just evaluate at x_0! In fact, the technical definitions of $\delta(x)$ hinge on this property - $\delta(x)$ is defined as the function which satisfies this property for any test function $f(x)$ using Lebesgue integration (to learn more about Lebesgue integration, take a Measure Theory course such as Math205).

This property makes evaluating $\mathscr{F}(\delta(x - x_0))$ very easy. We find:

$$\mathscr{F}(\delta(x - x_0)) = \int_{-\infty}^{\infty} \delta(x - x_0)e^{-ikx}\, dx = e^{-ikx_0}.$$

Similarly, one has

$$\mathscr{F}^{-1}(e^{-ikx_0}) = \delta(x - x_0).$$

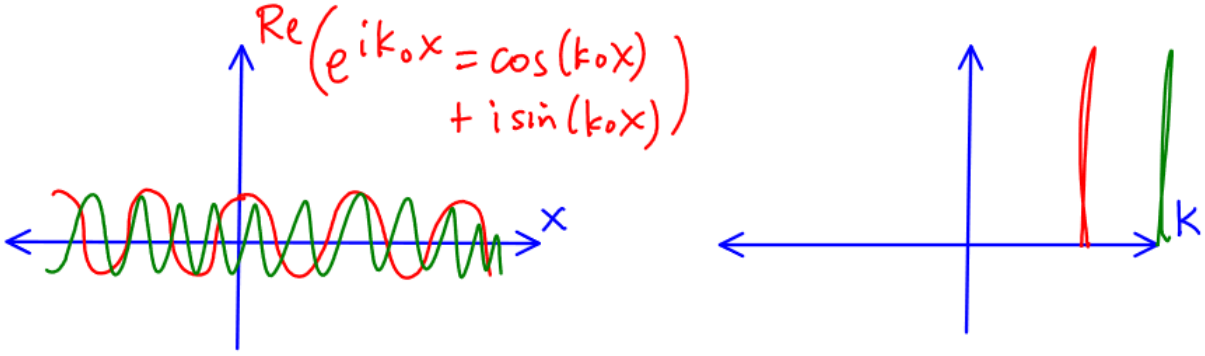
Likewise,

$$\mathscr{F}^{-1}(\delta(k - k_0)) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta(k - k_0)e^{ikx}\, dk = \frac{1}{2\pi}e^{ik_0x},$$

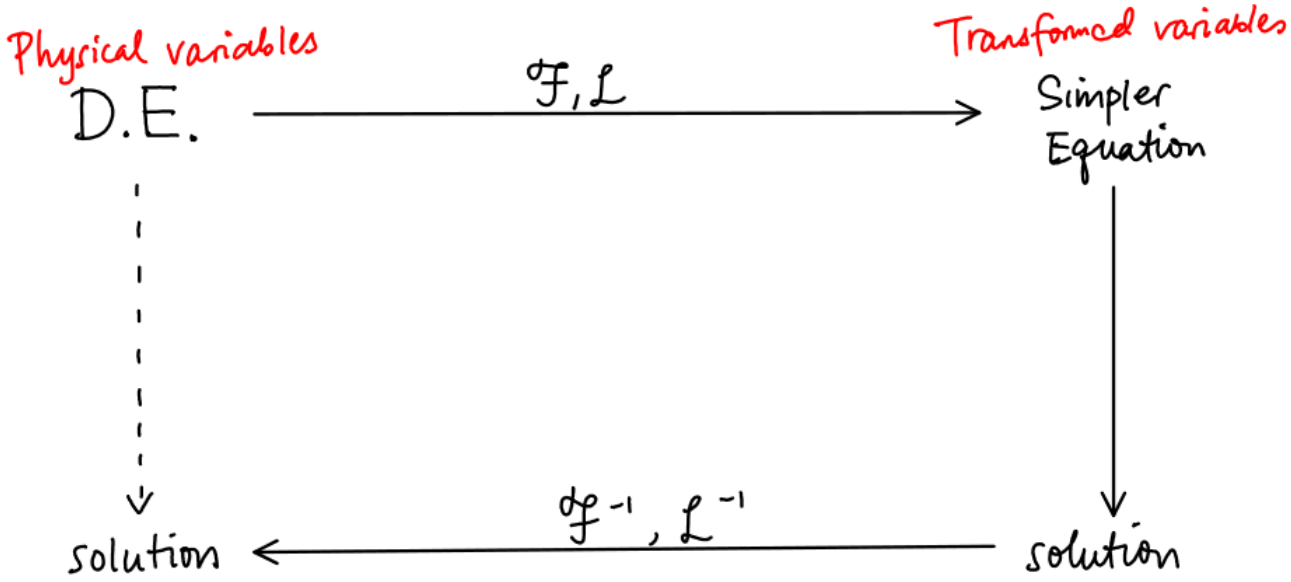
so by inversion,

$$\mathscr{F}(e^{ik_0x}) = 2\pi\delta(k - k_0).$$

This is a demonstration of how the Fourier transform tells us information about the frequency and amplitudes of trig functions that make up a function. Recall $Re(e^{ikx_0}) = \cos(k_0x)$.



8.4 Solving Differential Equations using integral transforms



Example: Solving an ODE using Fourier.
Solve the following:

$$\begin{cases} g''(x) - g(x) = \delta(x) & -\infty < x < \infty \\ g(x) \to 0 \quad \text{as } x \to \pm\infty \end{cases}$$

First, let $G(k) = \mathscr{F}(g(x))$, and apply $\mathscr{F}$ to the DE.

$$\begin{aligned} \mathscr{F}(g''(x) - g(x)) &= \mathscr{F}(\delta(x)) = 1 \\ \mathscr{F}(g''(x)) - \mathscr{F}(g(x)) &= 1 \\ (ik)^2 G(k) - G(k) &= 1 \end{aligned}$$

Then, solve for $G(k)$.

$$((ik)^2 - 1)G(k) = 1 \implies G(k) = -\frac{1}{k^2+1}$$

Then, transform back to the original variables. Observe on the guide that $e^{-a|x|}$ in physical variables corresponds to $\frac{2a}{a^2+k^2}$ in Fourier variables. Let's use this, with $a = 1$. We find:

$$g(x) = \mathscr{F}^{-1}(G(k)) = \mathscr{F}^{-1}\left(-\frac{1}{1+k^2}\right) = -\frac{1}{2}\mathscr{F}^{-1}\left(\frac{2}{1+k^2}\right) = -\frac{1}{2}e^{-|x|}$$

And we are done! No calculus necessary. The solution $g(x)$ is known as a **Green's function**, the response of a differential equation to a delta function forcing. We will refer to it as $g(x)$ generally in this chapter.

Side note: the solution $g(x)$ is not differentiable at $x = 0$ (and neither is the ODE, for that matter). What does it mean for $g(x)$ to "solve the ODE?"
One way to think about how $g(x)$ solves the ODE is to rewrite the ODE so that it doesn't involve $\delta(x)$. Given the original ODE $g''(x) - g(x) = \delta(x)$, let's integrate over a small interval:

$$\int_{-\epsilon}^{\epsilon} g''(x)\,dx - \int_{-\epsilon}^{\epsilon} g(x)\,dx = \int_{-\epsilon}^{\epsilon} \delta(x)\,dx$$

Working this out, we find:

$$(g'(x))_{-\epsilon}^{\epsilon} - \int_{-\epsilon}^{\epsilon} g(x)\,dx = 1.$$

Now, we take the limit as $\epsilon \to 0$, and assume $g(x)$ is continuous at $x = 0$, so we may have as a result that

$$\int_{-\epsilon}^{\epsilon} g(x)\,dx \to 0 \quad \text{as} \quad \epsilon \to 0.$$

This gives us

$$g'(0^+) - g'(0^-) = 1,$$

where $g'(0^{\pm})$ is the limit as x approaches 0 from the left or right, as usual. This is enough to give us an equivalent formulation of the ODE, in two pieces.

$$\text{For } x < 0: \begin{cases} g_L''(x) - g_L(x) = 0 \\ g_L(-\infty) = 0 \end{cases} \qquad \text{For } x > 0: \begin{cases} g_R''(x) - g_R(x) = 0 \\ g_R(\infty) = 0 \end{cases}$$

With the additional stipulation that $g_R'(0) - g_L'(0) = 1$ - this is what encodes the $\delta(x)$ condition at $x = 0$. Then, we define our final solution to be:

$$g(x) = \begin{cases} g_L(x) & x < 0 \\ g_R(x) & x > 0 \end{cases}$$

One can check that with this setup, one obtains the same solution as before.

Example: Solve the more general ODE:

$$\begin{cases} y''(x) - y(x) = f(x) \\ y, f \to 0 \text{ as } x \to \pm\infty \end{cases}$$

As before, we apply $\mathscr{F}$ to the problem. Let $Y(k) = \mathscr{F}(y(x))$ and $F(k) = \mathscr{F}(f(x))$. Then,

$$\mathscr{F}(y'' - y) = \mathscr{F}(f) \implies (ik)^2 Y - Y = F.$$

Solving this algebraic equation yields:

$$Y(k) = -\frac{F(k)}{1 + k^2}.$$

We now invert with $\mathscr{F}^{-1}$:

$$y(x) = \mathscr{F}^{-1}(Y) = \mathscr{F}^{-1}\left(-\frac{F(k)}{1+k^2}\right).$$

But wait, is $\mathscr{F}$ or $\mathscr{F}^{-1}$ linear over multiplication? That is, can we say:

$$\mathscr{F}^{-1}\left(-\frac{F(k)}{1+k^2}\right) = \mathscr{F}^{-1}(F(k)) \cdot \mathscr{F}^{-1}\left(-\frac{1}{1+k^2}\right)?$$

NO. This is generally false.

Definition: Define the **convolution** of two functions $a(x) * b(x)$ to be:

$$a(x) * b(x) = \int_{-\infty}^{\infty} a(\xi) b(x-\xi)\, d\xi.$$

Theorem: Multiplication of functions in the Fourier domain translates to convolution in the physical domain. Explicitly:

$$\mathscr{F}^{-1}(A(k) \cdot B(k)) = \mathscr{F}^{-1}(A(k)) * \mathscr{F}^{-1}(B(k))$$

So to solve the problem, we conclude by using convolution that

$$y(x) = \mathscr{F}^{-1}(F(k)) * \mathscr{F}^{-1}\left(-\frac{1}{1+k^2}\right) = \int_{-\infty}^{\infty} f(\xi) \cdot \left(-\frac{1}{2}e^{-|x-\xi|}\right)\, d\xi$$

In short, the result to a more general forcing for this linear DE involves the convolution of the Green's function $g(x)$ of the system and the forcing. This is generally true for Green's functions.

Exercise: What is the solution to

$$\begin{cases} y_1''(x) - y_1(x) = \delta(x-3) \\ y_1(x) \to 0 \quad \text{as} \quad x \to \pm\infty \end{cases}$$

Using the previous convolution answer, we see that it is:

$$y_1(x) = \delta(x-3) * g(x) = \int_{-\infty}^{\infty} \delta(\xi - 3) g(x-\xi)\, d\xi = g(x-3) = -\frac{1}{2}e^{-|x-3|}$$

Exercise: What is the solution to

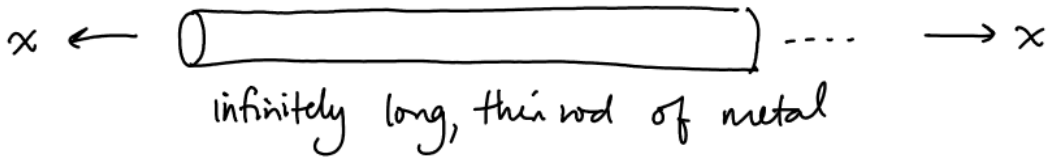
$$\begin{cases} y_2''(x) - y_2(x) = 2\delta(x) \\ y_2(x) \to 0 \quad \text{as} \quad x \to \pm\infty \end{cases}$$

We compute:

$$y_2(x) = 2\delta(x) * g(x) = \int_{-\infty}^{\infty} 2\delta(\xi) g(x-\xi) d\xi = 2g(x)$$

8.5 Free-Space Green's function for the 1D heat equation

Consider the following heat PDE problem concerning an infinitely long and thin rod of metal. (see Chapter 5 of Levy & Shearer for another derivation not using Fourier!)



$$\begin{cases} u_t = \kappa u_{xx} & -\infty < x < \infty, t > 0 \\ u(x,t) \to 0 & \text{as } x \to \pm\infty \\ u(x,0) = f(x) \end{cases}$$

We will solve this PDE using the Fourier transform, in the process demonstrating how to use integral transforms to solve PDEs - not just ODEs.

1. First, we apply the Fourier transform in the variable x. Define:

$$U(k,t) = \mathscr{F}_x[u(x,t)] = \int_{-\infty}^{\infty} u(x,t)e^{-ikx}\,dx$$

   Apply Fourier to the PDE as wel, with respect to x.

$$\mathscr{F}_x[u_t] = \mathscr{F}_x(\kappa u_{xx})$$

   We can pull the t partial derivative on the left-hand side out since the Fourier transform is with respect to x, and on the right, we can simplify using the derivative rule.

$$\frac{\partial}{\partial t}(U) = \kappa(ik)^2 U$$

   The initial condition of this ODE comes from the initial condition of the PDE:

$$u(x,0) = f(x) \implies U(k,0) = \mathscr{F}[f(x)] := F(k)$$

   We now have a well-posed ODE - using the Fourier transform eliminated the partial derivative with respect to x in the PDE!

2. Solve the resulting ODE problem. We have a linear first order ODE:

$$\frac{\partial U}{\partial t} = -\kappa k^2 U, \qquad U(k,0) = F(k)$$

   This is easily solved to be:

$$U(k,t) = F(k)e^{-\kappa k^2 t}$$

3. Invert: we have

$$u(x,t) = \mathscr{F}^{-1}[U] = \mathscr{F}^{-1}\left[F(k)e^{-\kappa k^2 t}\right] = \mathscr{F}^{-1}[F(k)] * \mathscr{F}^{-1}[e^{-\kappa k^2 t}]$$

We have (from the handout) the identity:

$$\mathscr{F}\left[\frac{1}{\sqrt{4\pi a}}e^{-x^2/4a}\right] = e^{-ak^2}$$

Identify $a = \kappa t$, then we conclude:

$$u(x,t) = f(x) * \left(\frac{1}{\sqrt{4\pi\kappa t}}e^{-\frac{x^2}{4\kappa t}}\right)$$

The function

$$g(x,t) = \frac{1}{\sqrt{4\pi\kappa t}}e^{-\frac{x^2}{4\kappa t}}$$

is known as **The Fundamental Solution**, the Green's Function of the 1-D Heat PDE problem. Here, it satisfies the property that $g(x,0) = \delta(x)$, and like before, a solution to the PDE problem with forcing $f(x)$ will be of the form $u(x,t) = f(x) * g(x,t)$.

To sum up, the Green's function for this problem is the solution to the problem with a delta function as its initial condition. The solution to the problem with a general initial condition is the convolution of that Green's function with the initial temperature function.

However, we can solve this problem in a different way - let's do it using Laplace instead!

1. Apply Laplace transform to the problem, with respect to t. Define:

$$U(x,s) = \mathscr{L}[u(x,t)] = \int_0^\infty u(x,t)e^{-st}\,dt.$$

Applying $\mathscr{L}$ to the PDE gives:

$$\mathscr{L}[u_t] = \mathscr{L}[\kappa u_{xx}] \implies sU - u(x,0) = \kappa\frac{\partial^2 U}{\partial x^2}$$

The BCs of the PDE give us $U \to 0$ as $x \to \pm\infty$.

2. Solve the simpler equation:

$$\frac{\partial^2 U}{\partial x^2} - \frac{s}{\kappa}U = -\frac{1}{\kappa}f(x), \qquad , \qquad U(x,s) \to 0 \text{ as } x \to \pm\infty$$

Using Fourier, one may solve this ODE (omitted):

$$U(x,s) = \frac{f(x)}{\kappa} * \left(\frac{1}{2\sqrt{s/\kappa}}e^{-|x|\sqrt{s/\kappa}}\right)$$

Notice the similarity to the previous problem we worked through:

$$y''(x) - y(x) = f(x) \implies y(x) = f(x) * \left(-\frac{1}{2}e^{-|x|}\right)$$

3. Perform the Inverse Laplace Transform:

$$\begin{aligned} u(x,t) &= \mathscr{L}^{-1}[U(x,s)] \\ &= \mathscr{L}^{-1}\left[\int_{-\infty}^{\infty} \frac{f(\xi)}{\kappa} \cdot \frac{1}{2\sqrt{s/\kappa}} e^{-|x-\xi|\sqrt{s/\kappa}} d\xi\right] \end{aligned}$$

Recall that $\mathscr{L}^{-1}$ is an integral in s, so we may reorder terms like so:

$$\begin{aligned} &= \int_{-\infty}^{\infty} \frac{f(\xi)}{\kappa} \mathscr{L}^{-1}\left[\frac{1}{2\sqrt{s/\kappa}} e^{-|x-\xi|\sqrt{s/\kappa}}\right] d\xi \\ &= \frac{f(x)}{\kappa} * \mathscr{L}^{-1}\left[\frac{1}{2\sqrt{s/\kappa}} e^{-|x|\sqrt{s/\kappa}}\right] \end{aligned}$$

Notice in the handout, we have the identity:

$$\mathscr{L}\left[\frac{1}{\sqrt{\pi t}} e^{-a^2/4t}\right] = \frac{1}{\sqrt{s}} e^{-a\sqrt{s}}.$$

Identify $a = |x|/\sqrt{\kappa}$, and this allows us to compute the Laplace term. We find:

$$\mathscr{L}^{-1}\left[\frac{1}{2\sqrt{s/\kappa}} e^{-|x|\sqrt{s/\kappa}}\right] = \frac{\sqrt{\kappa}}{2} \mathscr{L}^{-1}\left[\frac{1}{\sqrt{s}} e^{-|x|\sqrt{s/\kappa}}\right] = \frac{\sqrt{\kappa}}{2} \frac{1}{\sqrt{\pi t}} e^{-x^2/4\kappa t}$$

Hence, we obtain final solution:

$$u(x,t) = \frac{f(x)}{\kappa} * \frac{\sqrt{\kappa}}{2} \frac{1}{\sqrt{\pi t}} e^{-x^2/4\kappa t} = f(x) * \frac{1}{\sqrt{4\pi\kappa t}} e^{-x^2/4\kappa t}$$

This is the same result as before!

8.6 Half-space Green's Diffusion Equation with Neumann BC

Problems with nonzero initial conditions can often be rewritten with delta function forcing. Consider the following:

$$\begin{cases} u_t = \kappa u_{xx} + f(x)\delta(t) & -\infty < x < \infty, t > 0 \\ u(x, 0^-) = 0 \end{cases}$$

Apply $\int_{-\epsilon}^{\epsilon} - dt$ to the PDE:

$$\int_{-\epsilon}^{\epsilon} u_t \, dt = \kappa \int_{-\epsilon}^{\epsilon} u_{xx} \, dt + \int_{-\epsilon}^{\epsilon} f(x)\delta(t) \, dt$$

This becomes:

$$u(x,\epsilon) - u(x,-\epsilon) = \kappa \frac{\partial^2}{\partial x^2} \int_{-\epsilon}^{\epsilon} u \, dt + f(x),$$

however by continuity of u (an assumption), the remaining integral will vanish as $\epsilon \to 0^+$. Hence we obtain after taking the limit $\epsilon \to 0^+$:

$$u(x, 0^+) - u(x, 0^-) = f(x)$$

but by assumption, $u(x, 0^-) = 0$. We end with two equivalent PDE problems: (this alone isn't enough to demonstrate equivalence, but it can be shown)

$$\begin{cases} u_t = \kappa u_{xx} + f(x)\delta(t) & -\infty < x < \infty, t > 0 \\ u(x, 0^-) = 0 \end{cases} \iff \begin{cases} u_t = \kappa u_{xx} & -\infty < x < \infty, t > 0 \\ u(x, 0^+) = f(x) \end{cases}$$

This idea leads to Duhamel's Principle, a general method for obtaining solutions to inhomogeneous linear PDEs such as these, or even nonlinear PDEs. See Wikipedia for more details.

Now, consider the Neumann, half-space (referring to the domain being restricted to $x > 0$) PDE problem:

$$\begin{cases} u_t = \kappa u_{xx} + h(x, t) & x > 0, t > 0 \\ u(x, 0) = f(x) \\ u_x(0, t) = 0 \end{cases}$$

Here $h(x, t)$ is the source, or forcing term. To solve this, we decompose this problem into 2 separate problems:

$$\begin{cases} v_t = \kappa v_{xx} + h(x, t) \\ v(x, 0) = 0 \\ v_x(0, t) = 0 \end{cases} + \begin{cases} w_t = \kappa w_{xx} \\ w(x, 0) = f(x) \\ w_x(0, t) = 0 \end{cases}$$

Both these problems in fact have the same Green's function (we will see how to use this later), which satisfies the PDE problem, with a point source forcing at (x_0, t_0):

$$\begin{cases} g_t = \kappa g_{xx} + \delta(x - x_0)\delta(t - t_0) & x > 0, t > 0 \\ g(x, 0) = 0 & x_0 > 0 \\ g_x(0, t) = 0 & t_0 > 0 \end{cases}$$

We first solve for $g(x, t)$, then use g to find v and w. To solve for g, here we will use the Fourier cosine transform, given by:

$$\mathscr{F}_c[f(x)] = \int_0^\infty f(x) \cdot cos(kx)\, dx$$

$$\mathscr{F}_c^{-1}[F(k)] = \frac{2}{\pi} \int_0^\infty F(k) \cdot \cos(kx)\, dk$$

1. Define $G(k,t) = \mathscr{F}_c[g(x,t)]$ (so the transformation is taken with respect to x). Then applying $\mathscr{F}_c$ to the PDE and ICs, we obtain:

$$\begin{aligned}
\mathscr{F}_c[g_t] &= \mathscr{F}_c[\kappa g_{xx} + \delta(x-x_0)\delta(t-t_0)] \\
G_t(k,t) &= \kappa \mathscr{F}_c[g_{xx}(x,t)] + \mathscr{F}_c[\delta(x-x_0)\delta(t-t_0)] \\
G_t(k,t) &= \kappa \mathscr{F}_c[g_{xx}(x,t)] + \cos(kx_0)\delta(t-t_0)
\end{aligned}$$

To compute $\mathscr{F}_c[g_{xx}]$, we'll have to use integration by parts, twice. Here, we assume $g, g_x \to 0$ as $x \to \infty$. Note that we use the Neumann BC here as well.

$$\begin{aligned}
\mathscr{F}_c[g_{xx}] &= \int_0^\infty g_{xx}(x,t)\cos(kx)\,dx \\
&= g_x(x,t)\cos(kx)|_0^\infty + \kappa \int_0^\infty g_x(x,t)\sin(kx)\,dx \\
&= \kappa \int_0^\infty g_x(x,t)\sin(kx)\,dx \\
&= kg(x,t)\sin(kx)|_0^\infty - k^2 \int_0^\infty g(x,t)\cos(kx)\,dx \\
&= -k^2 \mathscr{F}_c[g(x,t)] = -k^2 G(k,t)
\end{aligned}$$

So our reduced equation is:

$$G_t = -\kappa k^2 G + \cos(kx_0)\delta(t-t_0)$$

It has initial condition $G(k,0) = 0$.

2. Solve the reduced equation: here we use the integrating factor $e^{\kappa k^2 t}$. We obtain:

$$\begin{aligned}
G_t e^{\kappa k^2 t} + \kappa k^2 G e^{\kappa k^2 t} &= \cos(kx_0)\delta(t-t_0)e^{\kappa k^2 t} \\
\frac{d}{dt}\left(G e^{\kappa k^2 t}\right) &= \cos(kx_0)\delta(t-t_0)e^{\kappa k^2 t}
\end{aligned}$$

Integrate both sides from 0 to t, and use $G(k,0) = 0$

$$\begin{aligned}
G(k,t)e^{\kappa k^2 t} - 0 &= \int_0^t \cos(kx_0)\delta(\tau - t_0)e^{\kappa k^2 \tau}\,d\tau \\
&= \begin{cases} \cos(kx_0)e^{\kappa k^2 (t-t_0)} & t > t_0 \\ 0 & t < t_0 \end{cases}
\end{aligned}$$

Thus, we obtain:

$$G(k,t) = \begin{cases} \cos(kx_0)e^{\kappa k^2 (t-t_0)} & t > t_0 \\ 0 & t < t_0 \end{cases}$$

3. Apply $\mathscr{F}_c^{-1}$ to obtain $g(x,t)$. Unfortunately, we have to do this manually.

$$g(x,t) = \mathscr{F}_c^{-1}[G(k,t)] = \begin{cases} \mathscr{F}_c^{-1}\left(\cos(kx_0)e^{\kappa k^2(t-t_0)}\right) & t > t_0 \\ 0 & t < t_0 \end{cases}$$

Explicitly, we have:

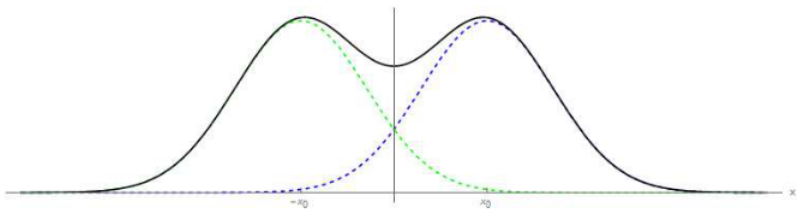
$$\begin{aligned} \mathscr{F}_c^{-1}\left(\cos(kx_0)e^{\kappa k^2(t-t_0)}\right) &= \frac{2}{\pi}\int_0^\infty \cos(kx_0)e^{\kappa k^2(t-t_0)}\cos(kx)\,dk \\ &= \frac{1}{\pi}\int_{-\infty}^\infty \cos(kx_0)e^{\kappa k^2(t-t_0)}\cos(kx)\,dk \\ &= \frac{1}{2\pi}\int_{-\infty}^\infty \left(\cos(k(x-x_0)) + \cos(k(x+x_0))\right)e^{-\kappa k^2(t-t_0)}\,dk \end{aligned}$$

For notational purposes, we will split this into two separate integrals and evaluate them simultaneously. By Euler's formula, we may note that:

$$\begin{aligned} \frac{1}{2\pi}\int_{-\infty}^\infty \cos(k(x\pm x_0))e^{-\kappa k^2(t-t_0)}\,dk &= \mathrm{Re}\left(\frac{1}{2\pi}\int_{-\infty}^\infty e^{ik(x\pm x_0)}e^{-\kappa k^2(t-t_0)}\,dk\right) \\ &= \mathrm{Re}\left(\mathscr{F}^{-1}\left(e^{\pm ikx_0}e^{-\kappa k^2(t-t_0)}\right)\right) \\ &= \mathrm{Re}\left(\frac{1}{\sqrt{4\pi\kappa(t-t_0)}}\exp\left(\frac{-(x\pm x_0)^2}{4\kappa(t-t_0)}\right)\right) \\ &= \frac{1}{\sqrt{4\pi\kappa(t-t_0)}}\exp\left(\frac{-(x\pm x_0)^2}{4\kappa(t-t_0)}\right) \end{aligned}$$

Hence, we conclude:

$$g(x,t) = \begin{cases} \dfrac{1}{\sqrt{4\pi\kappa(t-t_0)}}\left(\exp\left(\dfrac{-(x+x_0)^2}{4\kappa(t-t_0)}\right) + \exp\left(\dfrac{-(x-x_0)^2}{4\kappa(t-t_0)}\right)\right) & t > t_0 \\ 0 & t < t_0 \end{cases}$$



As an aside, we see that the Green's function for the diffusion equation on the half-space $x > 0$ with Neumann BCs is the sum of two copies of the usual free-space Green's

function. This is not a coincidence, and the solution could have been obtained by the method of images, which will not be covered here. Here, the $(x - x_0)$ function (blue) is responsible for the forcing $\delta(x - x_0)\delta(t - t_0)$, and the other function is responsible for the sum of slopes being 0 at $x = 0$, and makes the function symmetric about the y axis.

In general, the method of images takes advantage of the symmetry of the free-space Green's function and the symmetry of the domain to construct new Green's functions for that domain.

4. Determine $v(x,t)$ and $w(x,t)$. Recall,

$$\begin{cases} v_t = \kappa v_{xx} + h(x,t) \\ v(x,0) = 0 \\ v_x(0,t) = 0 \end{cases}$$

Here, we may rewrite $h(x,t)$:

$$\int_0^\infty \int_0^\infty h(x_0,t_0)\delta(x - x_0)\delta(t - t_0)\, dx_0\, dt_0$$

Then in fact, using Green's functions theory, it can be determined that

$$v(x,t) = \int_0^\infty \int_0^\infty h(x_0,t_0)g(x,t;x_0,t_0)\, dx_0\, dt_0$$

This can be thought of as a more general convolution as well. This can be simplified further since $g = 0$ when $t_0 > t$, to

$$v(x,t) = \int_0^\infty \int_0^t h(x_0,t_0)g(x,t;x_0,t_0)\, dx_0\, dt_0.$$

Next, recall for $w(x,t)$,

$$\begin{cases} w_t = \kappa w_{xx} \\ w(x,0) = f(x) \\ w_x(0,t) = 0 \end{cases}$$

However, recall that we can express this identically as:

$$\begin{cases} w_t = \kappa w_{xx} + f(x)\delta(t) \\ w(x,0^+) = 0 \\ w_x(0,t) = 0 \end{cases} .$$

It follows as it did with $v(x,t)$ that we may express

$$w(x,t) = \int_0^\infty \int_0^t f(x_0)\delta(t_0)g(x,t;x_0,t_0)\, dx_0\, dt_0 = \int_0^\infty f(x_0)g(x,t;x_0,0)$$

Our final solution is $u(x,t) = v(x,t) + w(x,t)$.

9 Vibrations of a circular membrane

Why is it that when we hit a timpani in the center, we get a dull sound? Let's model the system. Let $u(r,\theta,t)$ model the displacement of the membrane, and c be the wave speed. We obtain:

$$PDE: u_{tt} = c^2\nabla^2 u, \quad r < a, t > 0, -\pi < \theta < \pi$$

$$BCs: \begin{cases} u(a,\theta,t) = 0 \\ |u(r,\theta,t)| < \infty \\ u(r,\pi,t) = u(r,-\pi,t) \\ u_\theta(r,\pi,t) = u_\theta(r,-\pi,t) \end{cases}$$

$$ICs: \begin{cases} u(r,\theta,0) = f(r,\theta) \\ u_t(r,\theta,0) = g(r,\theta) \end{cases}$$

In polar coordinates, the PDE is:

$$u_{tt} = c^2\left(\frac{1}{r}\frac{\partial}{\partial r}\left(ru_r\right) + \frac{1}{r^2}u_{\theta\theta}\right)$$

Begin by assuming a product solution: let $u(r,\theta,t) = R(r)H(\theta)T(t)$. We obtain:

$$RHT'' = c^2\left(\frac{1}{r}\frac{\partial}{\partial r}\left(rR'\right)HT + \frac{1}{r^2}H''RT\right)$$

$$\implies -\lambda = \frac{T^2}{c^2T} = \frac{1}{rR}\frac{\partial}{\partial r}(rR') + \frac{1}{r^2}\frac{H''}{H}, \qquad \frac{H''}{H} = -n^2$$

So in total, we obtain 3 ODEs with 2 separation constants (chosen carefully):

$$\begin{cases} H'' + n^2H = 0 \\ T'' + c^2\lambda T = 0 \\ r^2R'' + rR' + (\lambda r^2 - n^2)R = 0 \end{cases}$$

As it turns out, the eigenvalue problems here are H and R, as these give us allowable values for λ or n (the issue with the T problem is with the way the ICs/BCs play out). Solving these generally yields:

$$\begin{cases} H(\theta) = A\cos(n\theta) + B\sin(n\theta) & n = 0,1,2,\ldots \\ T(t) = C\cos(c\sqrt{\lambda}t) + D\sin(c\sqrt{\lambda}t) \end{cases}$$

$$R(r) = \begin{cases} C_1r^n + C_2r^{-n} & \lambda = 0, n = 0 \\ C_1 + C_2\ln r & \lambda = 0, n > 0 \\ C_1J_n(\sqrt{\lambda}r) + C_2Y_n(\sqrt{\lambda}r) & \lambda > 0 \\ C_1I_n(\sqrt{-\lambda}r) + C_2K_n(\sqrt{-\lambda}r) & \lambda < 0 \end{cases}$$

In this case, J_n, U_n, I_n, K_n are all well-known functions which are solutions to that general ODE. They are known as **Bessel Functions**, solutions to **Bessel's Equation** (the r ODE). The BC that $|R(0)| < \infty$ implies C_2 is 0 in all of the above cases, and the BC that $R(a) = 0$ implies that only $\lambda > 0$ is allowed. Hence,

$$R(r) = C_1 J_n(\sqrt{\lambda} r), \qquad \text{with } J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(n\tau - x \sin\tau)\, d\tau$$

To avoid the trivial solution, we need $J_n(\sqrt{\lambda} a) = 0$. If we define z_{mn} to be the mth positive zero of $J_n(z)$, we have $\sqrt{\lambda} a = z_{mn}$ for $m = 1, 2, 3, \ldots$, in other words,

$$\lambda = \left(\frac{z_{mn}}{a}\right)^2$$

The functions $J_n(x)$ are known as **Bessel Functions of the first kind**, and is the only set of Bessel functions here for which $|R(0)| < \infty$.

So the final product solution is of the form

$$RHT = \alpha J_n\left(\frac{z_{mn}}{a} r\right) \left(A\cos(n\theta) + B\sin(n\theta)\right) \left(C\cos\left(\frac{c z_{mn}}{a} t\right) + D\sin\left(\frac{c z_{mn}}{a} t\right)\right)$$

with $n = 0, 1, 2, 3, \ldots$, and $m = 1, 2, 3, \ldots$. Each of these product solutions is a vibrational mode of the circular membrane. Each vibrational mode vibrates with a different frequency. The general motion of the membrane is a superposition of all these product solutions over m and n. The period of each vibration is $T = \frac{2\pi a}{c z_{mn}}$, and the frequency is the inverse of the period.

Our superposition solution can be written as:

$$\begin{aligned} u(r,\theta,t) = &\sum_{m=1}^{\infty} \sum_{n=0}^{\infty} J_n\left(\frac{z_{mn}}{a} r\right) \cos\left(\frac{c z_{mn}}{a} t\right) \left(A_{mn}\cos(n\theta) + B_{mn}\sin(n\theta)\right) \\ &+ \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} J_n\left(\frac{z_{mn}}{a} r\right) \sin\left(\frac{c z_{mn}}{a} t\right) \left(C_{mn}\cos(n\theta) + D_{mn}\sin(n\theta)\right) \end{aligned}$$

When $t = 0$,

$$u(r,\theta,0) = \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} J_n\left(\frac{z_{mn}}{a} r\right) \left(A_{mn}\cos(n\theta) + B_{mn}\sin(n\theta)\right) = f(r,\theta)$$

What can we do when we're stuck with two coefficients? Is there an equivalent orthogonality result for J_n? Stay tuned... but first a discussion.

If one hits the timpani in the center, the (1,0), (2,0), (3,0), etc. modes are preferentially excited, but no modes (m, n) with $n > 0$ are excited because these modes all have node lines

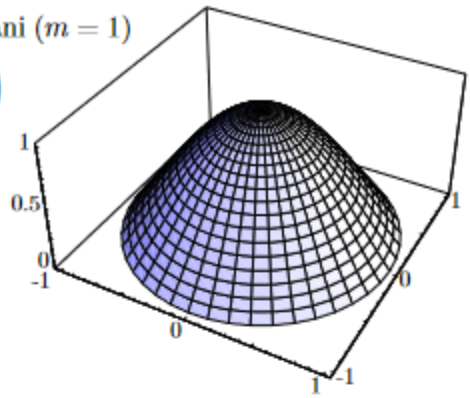
that cross at the center of the timpani - in other words, they are have zeroes in the center. Because the energy is directed into a small set of modes, the sound is somewhat hollow as a result. Furthermore, the ratio of the frequency of the (2,0) mode to the (1,0) mode is about 2.29542, so if the (1,0) mode is tuned to A3 at 220 Hz, the (2,0) mode sounds at about 504.992 Hz, which is between B4 and C5, a very discordant musical interval.

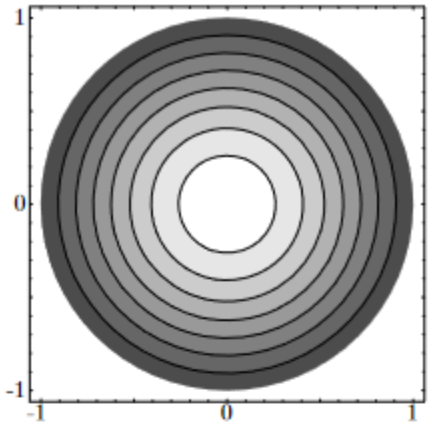
A qualified percussionist will tell you that the best place to hit the timpani is along a circle with about 2/3 or 3/4 the radius of the entire membrane (about 3-4 inches in front of the edge of the timpani). In 1877, Lord Rayleigh observed that when the timpani is struck in this spot, the (1,0) mode is not excited - in this case, the principal vibration is the (1,1) mode.

Vibrational modes of a timpani ($m = 1$)

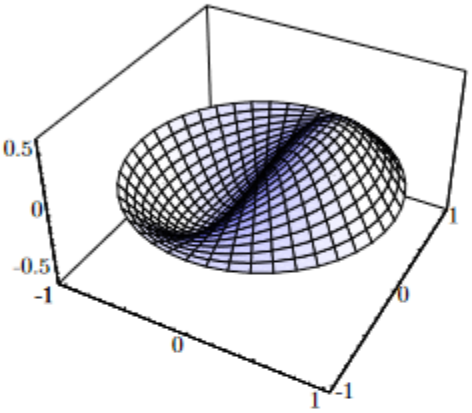
$$J_n\left(\frac{z_{mn}r}{a}\right)\cos(n\theta)\cos\left(\frac{cz_{mn}t}{a}\right)$$

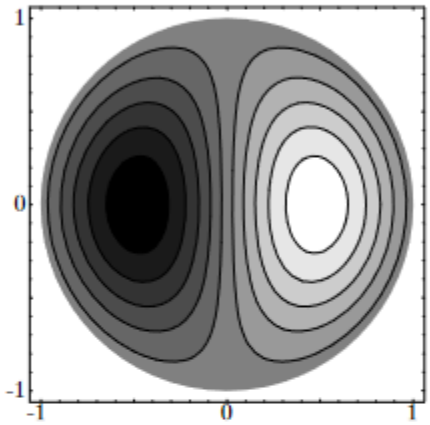
$(m, n) = (1, 0)$



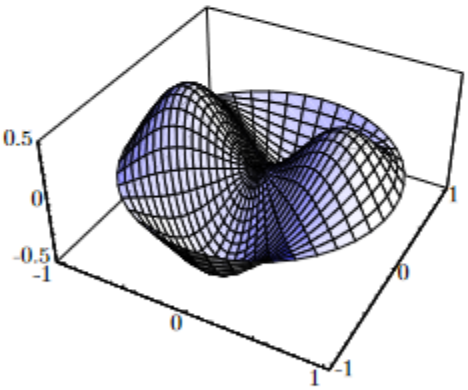


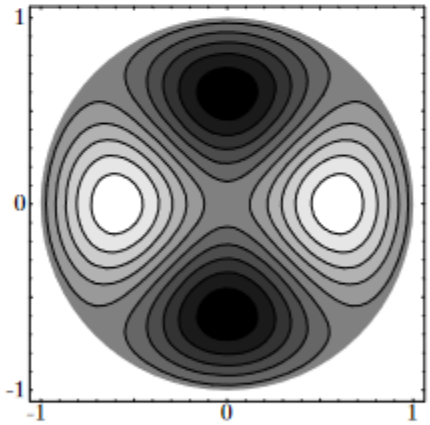
$(m, n) = (1, 1)$



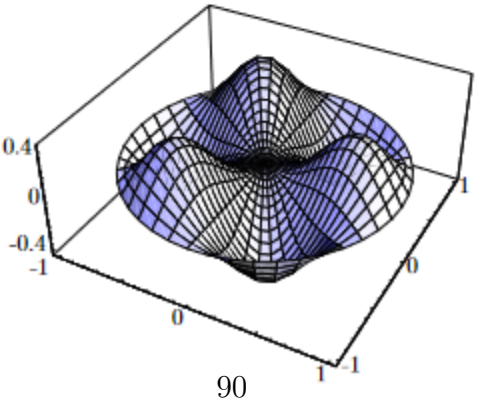


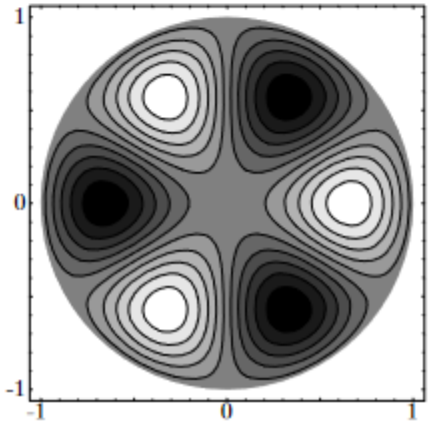
$(m, n) = (1, 2)$





$(m, n) = (1, 3)$

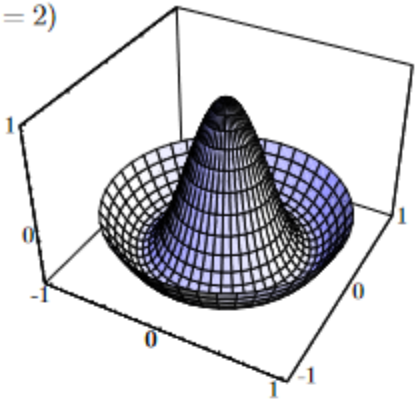


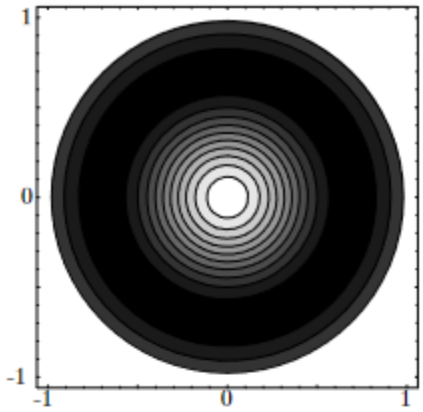


Vibrational modes of a timpani ($m = 2$)

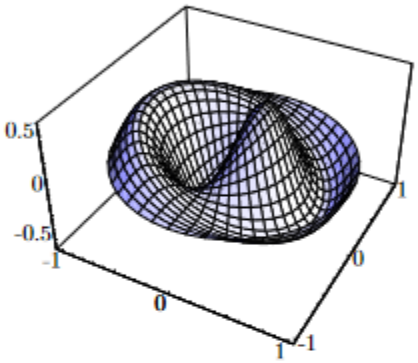
$$J_n\left(\frac{z_{mn}r}{a}\right)\cos(n\theta)\cos\left(\frac{cz_{mn}t}{a}\right)$$

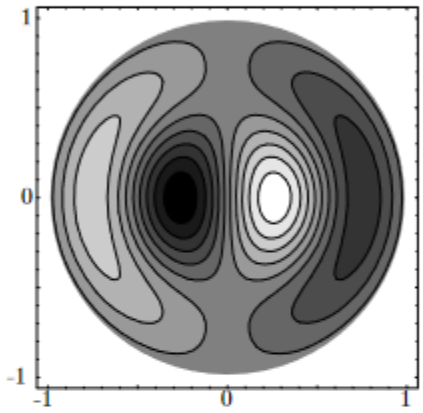
$(m, n) = (2, 0)$



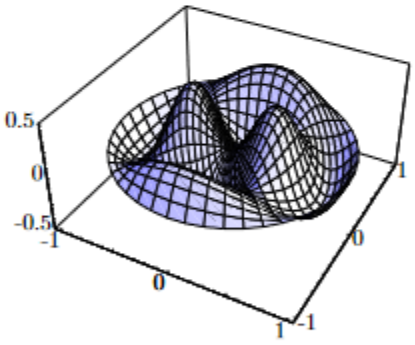


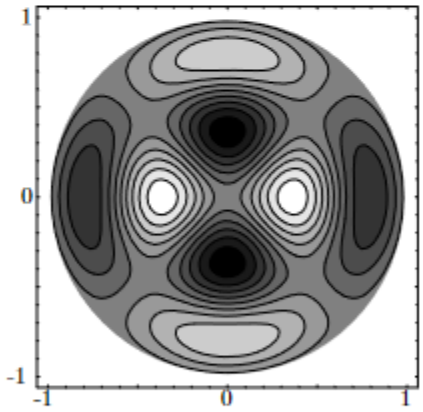
$(m, n) = (2, 1)$



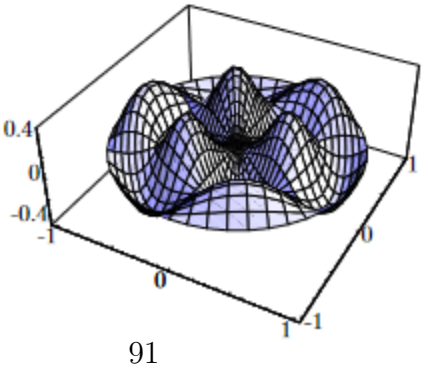


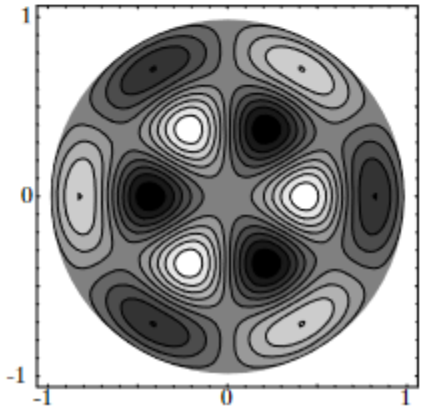
$(m, n) = (2, 2)$





$(m, n) = (2, 3)$

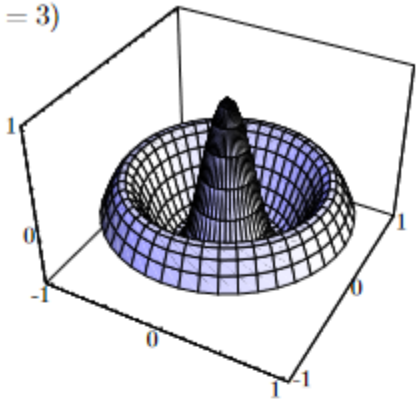


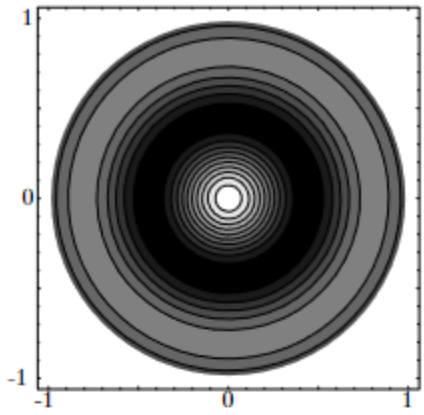


Vibrational modes of a timpani ($m = 3$)

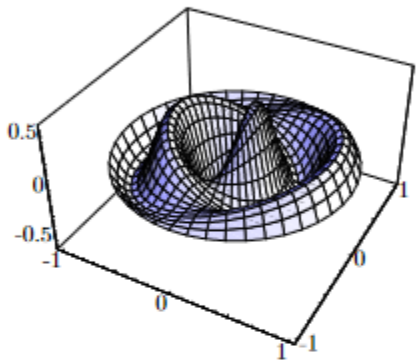
$$J_n\left(\frac{z_{mn}r}{a}\right)\cos(n\theta)\cos\left(\frac{cz_{mn}t}{a}\right)$$

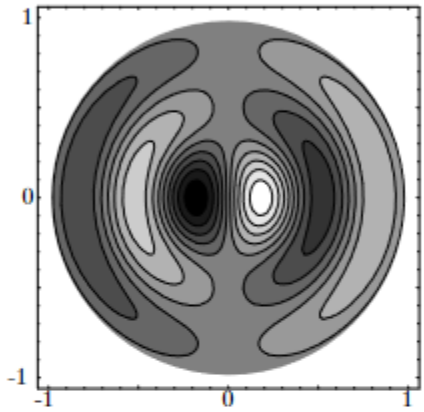
$(m, n) = (3, 0)$



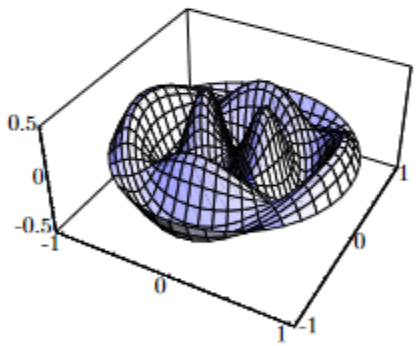


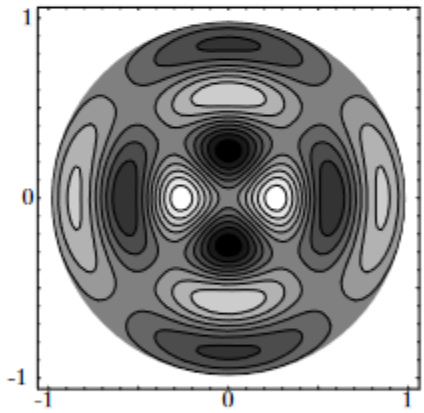
$(m, n) = (3, 1)$



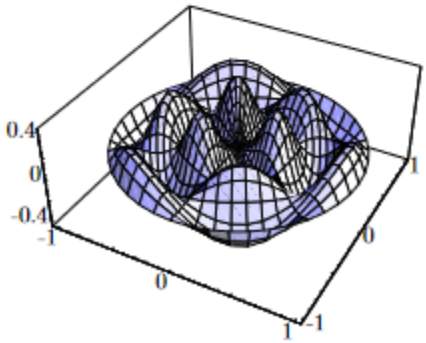


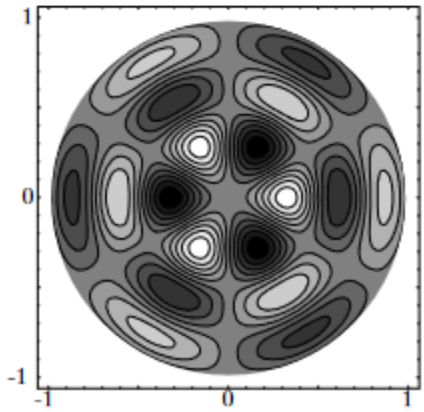
$(m, n) = (3, 2)$





$(m, n) = (3, 3)$





Anyways, our work with the vibrating circular membrane shows the need for a more general theory on eigenvalue problems.

Definition: Sturm-Liouville Eigenvalue Problems are eigenvalue problems of the form:

$$\frac{d}{dx}\left(p(x)\cdot\frac{dy}{dx}\right)+q(x)y(x)+\lambda\sigma(x)y(x)=0, \qquad a\leq x\leq b$$

The most important thing here is that when the differential equation is written in Sturm-Liousville form, the weight function $\sigma(x)$ can be identified. We say the SLEP is *regular* if:

- the interval $[a,b]$ is finite
- p, p', q, σ are continuous on $[a,b]$
- p, σ are positive on $[a,b]$
- the boundary conditions are homogeneous Dirichlet, Robin, or Neumann at each end of the interval - in particular, they may be identified with the following:

$$\alpha y(a)+\beta y'(a)=0 \qquad \alpha^2+\beta^2\neq 0$$

$$\gamma y(b)+\delta y'(b)=0 \qquad \gamma^2+\delta^2\neq 0$$

A SLEP is *singular* if:

- either a or b is infinite
- p or σ are zero or infinite at $x=a$ or $x=b/$

Example:

1. The common eigenvalue problem $X''+\lambda X=0$ on $0\leq x\leq L$ with $X(0)=X(L)=0$ is a RSLEP. We see this may be rewritten as:

   $$\frac{d}{dx}\left(1\cdot\frac{dX}{dx}\right)+0\cdot X+\lambda\cdot 1\cdot X=0,$$

   so $p(x)=1$, $q(x)=0$, $\sigma(x)=1$.

   Recall we had a countable set of eigenvalues $\lambda=\left(\frac{n\pi}{L}\right)^2$ and the corresponding eigenfunctions, sines, formed a countable basis of L^2.

2. The eigenvalue problem which arose from the vibrating membrane problem,

$$r^2R'' + rR' + (\lambda r^2 - n^2)R = 0, \qquad 0 \le r \le a,$$

with $|R(0)| < \infty$ and $R(a) = 0$ is a singular SLEP. We may rewrite it as:

$$\frac{d}{dr}(rR') - \frac{n^2}{r}R + \lambda rR = 0,$$

so $p(r) = r, q(r) = -\frac{n^2}{r}, \sigma(r) = r$.

What do SLEPs tell us? Here's the kicker...

Theorem: All Regular SLEPs have the following properties:

1. All eigenvalues are real.
2. There is a smallest eigenvalue and no largest eigenvalue. They form a countable set, and therefore can be listed in increasing order.
3. There is a unique eigenfunction ϕ_n up to a multiplicative factor for every eigenvalue λ_n.
4. The eigenfunctions form an orthonormal family of functions relative to the inner product

$$\langle u, v\rangle = \int_a^b u\overline{v}\sigma\, dx.$$

In toher words, $\langle \sigma_m, \sigma_n\rangle = 0$ if $m \neq n$.

5. The mth eigenfunction ϕ_m has exactly $m - 1$ zeros in the open interval (a, b).
6. The eigenfunctions form a "complete" set, meaning that any piecewise smooth function $f(x)$ can be approximated by an infinite series of the eigenfunctions. Specifically, there exist unique a_n such that

$$||f - \sum_{n=1}^{N} a_n\phi_n|| \to 0 \quad \text{as } N \to \infty$$

If $f(x)$ is continuous, the series converges uniformly. At a discontinuity, the series converges to the average of the left-hand and right-hand limits.

7. Any eigenvalue λ is related to its eigenfunction ϕ by

$$\lambda = \frac{-p\phi\phi'|_a^b + \int_a^b \left(p(\phi')^2 - q\phi^2\right)\, dx}{\int_a^b \phi^2\sigma\, dx}$$

This is called the **Rayleigh quotient**. In fact, each eigenvalue/function pair solves this minimization problem:

$$\lambda_n = \min \frac{-puu'|_a^b + \int_a^b \left(p(u')^2 - qu^2\right)\, dx}{\int_a^b u^2\sigma\, dx}$$

over all $u(x)$ satisfying the BCs and that are orthogonal to $\phi_1, \ldots, \phi_{n-1}$.

Note that Singular and periodic SLEPs often share some of these same features. Going back to our problem, we expect that

$$\int_0^a J_n\left(\frac{z_{mn}r}{a}\right) J_n\left(\frac{z_{kn}r}{a}\right) r\, dr = 0 \quad \text{if } m \neq k.$$

And in fact, one may prove that

$$\int_0^a J_n\left(\frac{z_{mn}r}{a}\right)^2 r\, dr = \frac{a^2}{2} J_{n+1}(z_{mn})^2$$

In fact, we can obtain the constants $(A, B, C, D)_{mn}$ using both orthogonality of sines/cosines and orthogonality of Bessel functions. For example, with A_{mn}: recall we have

$$u(r,\theta,0) = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} J_n\left(\frac{z_{mn}}{a}r\right)\left(A_{mn}\cos(n\theta) + B_{mn}\sin(n\theta)\right) = f(r,\theta) \qquad (*)$$

To get A_{mn},

$$\int_0^a \int_{-\pi}^{\pi} (*) J_n\left(\frac{z_{m'n}r}{a}\right)\cos(n'\theta)\, d\theta r\, dr$$

The orthogonality of sines and cosines eliminate the B_{mn} and A_{kn} for $k \neq m'$, leaving only A_{mn}! The rest of the details are left as an exercise to the reader.