

An Introduction to Partial differential equations

Per Kristen Jakobsen

Department of Mathematics and Statistics, the Arctic University
of Norway, 9019 Tromsø, Norway

August 11, 2021

Contents

1	Introduction	3
2	First notions	4
3	PDEs as mathematical models of physical systems	8
3.1	A traffic model	8
3.2	Diffusion	12
3.3	Heat conduction	14
3.4	A random walk	15
3.5	The vibrating string	16
4	Initial and boundary conditions	19
4.1	Well posed Problems	21
5	Numerical methods	22
5.1	The finite difference method for the heat equation	23
5.2	The Crank-Nicholson scheme	29
5.2.1	Finite difference method for the heat equation: Neumann boundary conditions	30
5.3	Finite difference method for the wave equation	31
5.4	Finite difference method for the Laplace equation	35
5.4.1	Iterations in \mathbb{R}^n	38
5.4.2	Gauss-Jacobi	41
5.4.3	Gauss-Seidel	46
6	First order PDEs	48
6.1	Higher order equations and first order systems	49
6.2	First order linear scalar PDEs	51
6.3	First order Quasilinear scalar PDEs	62

7	Classification of equations: Characteristics	72
7.1	Canonical forms for equations of hyperbolic type	75
7.2	Canonical forms for equations of parabolic type	78
7.3	Canonical forms for equations of elliptic type	79
7.4	Characteristic curves	82
7.5	Characteristic curves again: A different point of view	85
7.6	Stability theory, energy conservation and dispersion	94
8	Adjoint differential operators	101
9	Initial and boundary value problems in bounded domains	105
9.1	Separation of variables	108
9.2	Self-adjoint and positive operators	111
10	The Sturm-Liouville problem and Fourier series	117
10.1	Properties of eigenvalues and eigenfunctions	121
10.1.1	Trigonometric eigenfunctions	122
10.1.2	Bessel eigenfunctions	127
10.1.3	Legendre polynomial eigenfunctions	130
10.2	Series solutions to initial boundary value problems	131
10.2.1	Vibrations of a string with fixed endpoints	133
10.2.2	Heat conduction in a finite rod	137
10.2.3	The Laplace equation in a rectangle	139
10.3	Inhomogenous equations - Duhamel's principle	141
10.3.1	The inhomogeneous wave equation	142
10.3.2	The inhomogeneous heat equation	144
10.4	The finite Fourier transform	146
10.4.1	Hyperbolic equations	151
10.4.2	The Poisson equation in a disk	152
10.5	Nonlinear stability theory	160
10.5.1	The infinite case	163
10.5.2	The finite case	165
11	Integral transforms	170
11.1	The 1D Fourier transform	174
11.1.1	A boundary value problem for an ODE	179
11.1.2	The Cauchy problem for the heat equation	187
11.1.3	The Cauchy problem for the wave equation	190
11.1.4	The Laplace equation in a half-plane	193
11.2	The Fourier sine- and cosine transform	196
11.2.1	The heat equation on a semi-infinite interval	200
11.3	The Laplace transform	203
11.3.1	The wave equation on a half interval	206
11.3.2	The heat equation on a finite interval	208
11.4	Group velocity 1	210
11.5	Method of stationary phase	215

11.6 Group velocity 2	217
12 Projects	221
12.1 Project1	221
12.2 Project2	223
12.3 Project3	225

1 Introduction

The field of partial differential equations (PDEs) is vast in size and diversity. The basic reason for this is that essentially all fundamental laws of physics are formulated in terms of PDEs. In addition, approximations to these fundamental laws, that form a patchwork of mathematical models covering the range from the smallest to the largest observable space-time scales, are also formulated in terms of PDEs. The diverse applications of PDEs in science and technology testify to the flexibility and expressiveness of the language of PDEs, but it also makes it a hard topic to teach right. Exactly because of the diversity of applications, there are just so many different points of view when it comes to PDEs. These lecture notes view the subject through the lens of applied mathematics. From this point of view, the physical context for basic equations like the heat equation, the wave equation and the Laplace equation are introduced early on, and the focus of the lecture notes are on methods, rather than precise mathematical definitions and proofs. With respect to methods, both analytical and numerical approaches are discussed.

These lecture notes has been successfully used as the text for a master class in partial differential equations for several years. The students attending this class are assumed to have previously attended a standard beginners class in ordinary differential equations and a standard beginners class in numerical methods. It is also assumed that they are familiar with programming at the level of a beginners class in informatics at the university level.

While writing these lecture notes we have been influenced by the writings of some of the many authors that previously have written textbooks on partial differential equations [3, 6, 5, 8, 7, 2, 4]. However, for the students that these lecture notes are aimed at, books like [3],[6],[5],[8] and [7] are much too advanced either mathematically [3],[6],[8],[8] or technically [5],[7]. The books [2],[4] would be accessible to the students we have in mind, but in [2] there is too much focus on precise mathematical statements and proofs and less on methods, in particular numerical methods. The book [4] is a better match than [2] for the students these lecture notes has been written for, but it is a little superficial and does not reach far enough. With respect to reach and choice of topics, the book “Partial Differential Equations of Applied Mathematics”, written by Erich Zauderer[1] would be a perfect match for the students we have mind. However, the students taking the class covered by these notes does not have any previous exposure to PDEs, and the book [1] is for the most part too difficult for them to follow. The book is also very voluminous and contain far to much material for

a class covering one semester with five hours of lecture each week. The lecture notes for the most part follow the structure of [1], but simplify the language and makes a selection of topics that can be covered in the time available. The lecture notes deviate from [1] at several points, in particular in the section covering physical modeling examples, integral transforms and in the treatment of numerical methods. After mastering these lecture notes, the students should be able to use [1] as a reference for more advanced methods that are not covered by the notes. In order for students to pass the master class these notes has been written for, the students must complete the three projects included in the last section of the lecture notes and pas an oral exam where they must defend their work and also answer questions based on the content of these lecture notes.

2 First notions

A partial differential equation(PDE), is an equation involving one or more functions of two or more variables, and their partial derivatives, up to some finite order. Here are some examples

$$u_x + u_y = 0, \quad (1)$$

$$u_x + yu_y = 0, \quad (2)$$

$$(u_x)^2 + (u_y)^2 = 1, \quad (3)$$

where

$$u_x \equiv \partial_x u \equiv \frac{\partial u}{\partial x}.$$

The highest derivative that occur in the equation is the *order* of the equation. All equations (1,2,3) are of order one.

A PDE is *scalar* if it involves only one unknown function. In this course we will mostly work with scalar PDEs of two independent variables. This is hard enough as we will see!

Most important theoretical and numerical issues can be illustrated using such equations. It is also a fact that many physical systems can be modeled by equations of this type. We will also for the most part restrict our attention on equations of order two or less. This is the type of equations that most frequently occur in modeling situations. This fact is linked to the use of second order Taylor expansions and to the use of Newton's equations in one form or another.

The most general first order scalar PDE, in two independent variables, is of the form.

$$F(x, y, u(x, y), u_x(x, y), u_y(x, y)) = 0, \quad (4)$$

$$(\text{or short } F(x, y, u, u_x, u_y) = 0)$$

where F is some function of five variables.

Let $D \in \mathbb{R}^2$ be some region in the xy-plane. Then $u(x, y)$, defined on D , is a *solution* of equation (4) if u_x, u_y exists and are continuous on D and if u satisfy

the equation for all $(x, y) \in D$. PDEs typically have many different solutions. The different solutions can be defined on different, even disjoint domains.

We will later relax this notion of solutions somewhat. In more advanced textbooks what we call a solution here is called a *classical* solution.

The most general second order scalar PDE of two independent variables is of the form

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0. \quad (5)$$

A solution in $D \in \mathbb{R}^2$ is now a function where all second order partial derivation are continuous and where the function satisfy equation (5) for all $(x, y) \in D$.

Here are some examples

$$u_{xx} + u_{yy} = 0, \quad (6)$$

$$u_{tt} - u_{xx} = 0, \quad (7)$$

$$u_t + uu_x - u_{xx} = 0. \quad (8)$$

Equations (1),(2),(6) and (7) are *linear* equations whereas equations (3) and (8) are not linear. Equations that are not linear are said to be *nonlinear*. The importance of the linearity/nonlinearity distinction in the theory of PDEs can hardly be overstated. We have already seen the importance of this notion in the theory of ODEs.

The precise meaning of linearity is best stated using the language of linear algebra. Let us consider equation (1).

Define an *operator* \mathcal{L} by

$$\mathcal{L} = \partial_x + \partial_y.$$

Equation (1) can then be written as

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

The operator \mathcal{L} has the following important properties:

1) For all differentiable functions u, v

$$\begin{aligned} \mathcal{L}(u + v) &= \partial_x(u + v) + \partial_y(u + v) \\ &= \partial_x u + \partial_x v + \partial_y u + \partial_y v \\ &= \partial_x u + \partial_y u + \partial_x v + \partial_y v \\ &= \mathcal{L}u + \mathcal{L}v, \end{aligned}$$

2) For all differentiable functions u and $c \in \mathbb{R}$ (or \mathbb{C})

$$\begin{aligned} \mathcal{L}(cu) &= \partial_x(cu) + \partial_y(cu) \\ &= c\partial_x u + c\partial_y u \\ &= c(\partial_x u + \partial_y u) \\ &= c\mathcal{L}(u). \end{aligned}$$

Properties 1 and 2 show us that \mathcal{L} is a *linear* differential operator.

Definition 1. A scalar PDE is linear if it can be written in the form

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

for some linear differential operator \mathcal{L} and given function g . The equation is homogeneous if $g = 0$ and inhomogeneous if $g \neq 0$.

Let us consider equation (3) from page 1. Define a differential operator \mathcal{L} by

$$\mathcal{L}(u) = (\partial_x u)^2 + (\partial_y u)^2.$$

Then the equation can be written as

$$\mathcal{L}(u) = 1.$$

Note that

$$\begin{aligned}\mathcal{L}(cu) &= (\partial_x(cu))^2 + (\partial_y(cu))^2 \\ &= c^2((\partial_x u)^2 + (\partial_y u)^2) \\ &= c^2 \mathcal{L}(u).\end{aligned}$$

\mathcal{L} fails property 2 and is thus not a linear operator (It also fails property 1). Thus the equation

$$(\partial_x u)^2 + (\partial_y u)^2 = 1,$$

is nonlinear. It is not hard to see that, basically, a PDE is linear if it can be written as a first order polynomial in u and its derivatives. All other equations are nonlinear.

The most general first order linear scalar PDE of two independent variables is of the form

$$a(x, y)u_x + b(x, y)u_y + c(x, y)u = d(x, y).$$

The most important property of homogeneous linear equations is that they satisfy the *superposition principle*.

Let u, v be solutions of a linear homogeneous scalar PDE. Thus

$$\begin{aligned}\mathcal{L}u &= 0, \\ \mathcal{L}v &= 0,\end{aligned}$$

where \mathcal{L} is the linear differential operator defining the equation. Let $w = au + bv$ where $a, b \in \mathbb{R}$ (or \mathbb{C}). Then

$$\begin{aligned}\mathcal{L}(w) &= \mathcal{L}(au + bv) \\ &\stackrel{1}{=} \mathcal{L}(au) + \mathcal{L}(bv) \\ &\stackrel{2}{=} a\mathcal{L}(u) + b\mathcal{L}(v) \\ &= 0.\end{aligned}$$

Thus w is also a solution for all choices of a and b . This is the superposition principle. We have already seen this principle at work in the theory of ODEs.

Generalizing the argument just given, we obviously have that if u_1, \dots, u_n are solutions then

$$w = \sum_{i=1}^n a_i u_i, \quad a_i \in \mathbb{R} \text{ (or } \mathbb{C}),$$

is also a solution for all choices of the constants $\{a_i\}$.

In the theory of ODEs we found that any solution of a homogeneous linear ODE of order n is of the form

$$y = c_1 y_1 + \dots + c_n y_n,$$

where y_1, \dots, y_n are n linearly independent solutions of the ODE and c_1, \dots, c_n are some constants. Thus the general solution has n free *constants*. For linear PDEs, the situation is not that simple.

Let us consider the PDE

$$u_{xx} = 0. \tag{9}$$

Integrating once with respect to x we get

$$u_x = f(y),$$

where $f(y)$ is an arbitrary function of y . Integrating one more time we get the general solution

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

where $g(y)$ is another arbitrary function of y . Thus the general solution of equation (9) depends on two arbitrary *functions*. The corresponding ODE

$$y'' = 0,$$

has a general solution $y = ax + b$ that depends on two arbitrary *constants*. The equation (9) is very special but the conclusion we can derive from it is very general. (But not universal!)

“The general solution of PDEs depends on one or more arbitrary functions.”

This fact makes it much harder to work with PDEs than ODEs.

Let us consider another example that will be important later in the course. The PDE that we want to solve is

$$u_{xy} = 0. \tag{10}$$

Integrating with respect to x we get

$$u_y = f(y), \tag{11}$$

where $f(y)$ is arbitrary. Integrating equation (11) with respect to y gives us the general solution to equation (10) in the form

$$u = F(x) + G(y),$$

where $G(y) = \int f(y) dy$ and $F(x)$ are arbitrary functions of y and x .

3 PDEs as mathematical models of physical systems

Methods for solving PDEs is the focus of this course. Deriving approximate description of physical systems in terms of PDEs is less of a focus and is in fact better done in specialized courses in physics, chemistry, finance etc.

It is however important to have some insight into the physical modeling context for some of the main types of equations we discuss in the course. This insight will make the methods we introduce to solve PDEs more natural and easy to understand.

3.1 A traffic model

We will introduce a model for cars moving on a road. In order for the complexity not to get out of hand, we start by making some assumptions about the constructions of the road.

- i) There is only one lane, so that all cars move in the same direction (no passing is allowed).
- ii) There are no intersections or off/on ramps. Thus cars can not enter or leave the road.

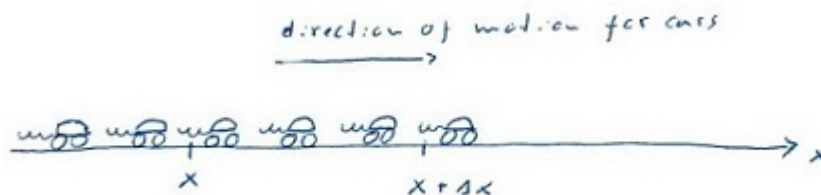


Figure 1: A simple traffic model for a one-lane road

Because of i) and ii), we can represent the road by a line that we coordinatize by a variable x . Let

$\Delta N(x, t) = \#$ of cars between x and $x + \Delta x$ at time t .

$f(x, t) = \#$ of cars passing a point x from left to right, per unit time, at time t .

From the definition of ΔN and f and assumption ii) we get the identity

$$\Delta N(x, t + \Delta t) = \Delta N(x, t) - f(x + \Delta x, t)\Delta t + f(x, t)\Delta t. \quad (12)$$

Equation 12 express the conservation of the number of cars on the road.

Both ΔN and f are integer valued and thus change by whole numbers. After all, they are counting variables! Define the *density* of cars, $n(x, t)$, by

$$\Delta N(x, t) = n(x, t)\Delta x.$$

We now make the fundamental continuum assumption: $n(x, t)$ and $f(x, t)$ are smooth functions. This is a reasonable approximation if the number of cars in each interval $[x, x + \Delta x]$ is very large and change slowly when x and t vary. $f(x, t)$ is also assumed to be slowly varying. (Imagine observing a crowded road from a great height, say from a helicopter). Rewrite equation (12) as

$$\frac{n(x, t + \Delta t) - n(x, t)}{\Delta t} = -\frac{f(x + \Delta x, t) - f(x, t)}{\Delta x}. \quad (13)$$

Taylor expanding to first order we get

$$\begin{aligned} \frac{1}{\Delta t} \{n(x, t) + \partial_t n(x, t) \Delta t - n(x, t)\} \\ \approx -\frac{1}{\Delta x} \{f(x, t) + \partial_x f(x, t) \Delta x - f(x, t)\}, \end{aligned}$$

\Downarrow

$$\partial_t n \approx -\partial_x f.$$

Letting Δt and Δx approach zero, we get a PDE

$$\partial_t n + \partial_x f = 0. \quad (14)$$

This type of PDE is called a *conservation law*. It express the conservation of cars in differential form. The notion of conservation laws is fundamental in the mathematical description of nature. Many of the most fundamental mathematical models of nature express conservation laws.

Note that equation (14) is a single first order PDE for two unknown functions, n and f . In order to have a determinate mathematical model we must have the same number of equations and unknowns.

Completing the model involves making further physical assumptions. We will assume that the *flux*, f , of cars at (x, t) depends on x and t only through the density of cars, $n(x, t)$. Thus

$$f(x, t) = \alpha(n(x, t)), \quad (15)$$

where α is some function whose precise specification requires further modeling assumptions.

As noted above, the quantity f in this model is called a flux. Such quantities are very common in application and typically measures the flow of some quantity through a point, surface or volume.

In equation (15), we assume that the flux of cars through a point only depends on the density of cars. This is certainly not a universal law, but it is a reasonable assumption (is it?). The chain rule now gives

$$\partial_x f = \partial_x \alpha(n) = \alpha'(n) \partial_x n.$$

Thus our model of the traffic flow is given by the PDE

$$\partial_t n + c(n) \partial_x n = 0. \quad (c(n) \equiv \alpha'(n)) \quad (16)$$

This equation describes transport of cars along a road and is an example of a *transport equation*. Transport equations occur often as mathematical models because conserved quantities are common in our description of the world.

The model is still not fully specified because the function $c(n)$ is arbitrary at this point. If we assume that $c(n)$ has a convergent power series expansion, we have

$$c(n) = c_0 + c_1 n + c_2 n^2 + \cdots. \quad (17)$$

If the density of cars is sufficiently low we can use the approximation

$$c(n) \approx c_0, \quad (18)$$

and our transport equation is a *linear* equation

$$\partial_t n + c_0 \partial_x n = 0. \quad (19)$$

For somewhat higher densities we might have to use the more precise approximation

$$c(n) \approx c_0 + c_1 n, \quad (20)$$

which gives us a nonlinear transport equation

$$\partial_t n + (c_0 + c_1 n) \partial_x n = 0.$$

For the linear transport equation (19), let us look for solutions of the form

$$n(x, t) = \varphi(x - c_0 t).$$

The chain rule gives

$$\begin{aligned} \partial_t n &= -c_0 \varphi', \\ \partial_x n &= \varphi', \\ \Downarrow \\ \partial_t n + c_0 \partial_x n &= -c_0 \varphi' + c_0 \varphi' = 0. \end{aligned}$$

Thus

$$n(x, t) = \varphi(x - c_0 t), \quad (21)$$

is a solution of equation (19) for any function φ . Let us assume that the density of cars is known at one time, say at $t = 0$. Thus

$$n(x, 0) = f(x), \quad (22)$$

where f is a given function. Using equation (22) and (21) we have

$$n(x, 0) = f(x),$$

$$\begin{aligned}
&\Downarrow \\
&\varphi(x - c_0 \cdot 0) = f(x), \\
&\Downarrow \\
&\varphi(x) = f(x).
\end{aligned}$$

Thus the function

$$n(x, t) = f(x - c_0 t), \quad (23)$$

is a solution of

$$\begin{aligned}
\partial_t n + c_0 \partial_x n &= 0, \\
n(x, 0) &= f(x).
\end{aligned} \quad (24)$$

(24) is an example of an initial value problem for a PDE. Initial value problems for PDEs are often, for historical reasons, called *Cauchy problems*.

Note that at this point we have not shown that (23) is the only solution to the Cauchy problem (24). Proving that there are no other solutions is obviously important because a deterministic model (and this is!) should give us exactly one solution, not several.

The solution (23) describe an interesting behavior that is common to *very* many physical systems. It describe a *wave*. From (23), it is evident that

$$n(x, t) = n(x - c_0 t, 0).$$

Thus we get the values of n at time t by translating the values of n at $t = 0$ along the x -axis.

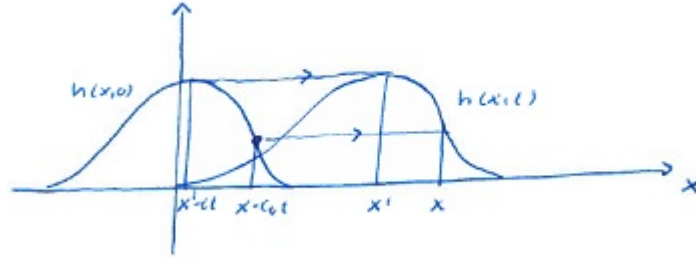


Figure 2: A traffic disturbance propagating like a wave.

We are used to observing waves when looking at the surface of the sea, but might not have expected to find it in the description of traffic on a one-lane road.

Traffic waves are in fact common and play no small part in the occurrence of traffic jams.

3.2 Diffusion

Consider a narrow pipe oriented along the x -axis containing water that is not moving. Assume there is a material substance, a dye for example, of mass density $u(x, t)$ dissolved in the water. Here we assume that the mass density is uniform across the pipe so that the density only depends on x which by choice is the coordinate along the length of the pipe.



Figure 3

The total mass of the dye between x_1 and x_2 is

$$M(t) = \int_{x_1}^{x_2} u(x, t) dx.$$

If the pipe is not leaking, the mass inside the section of the pipe between x_1 and x_2 can only change by dye entering or leaving through the points x_1 and x_2 . This is the law of *conservation of mass* which is the first great pillar of classical physics.

Let $f(x, t)$ be the amount of mass flowing through a point x at time t pr. unit time. This is the *mass flux*. By convention f positive means that mass is flowing to the right.

Mass conservation then gives the balance equation

$$\frac{dM}{dt} = f_1 - f_2, \quad f_i = f(x_i, t), \quad i = 1, 2, \dots \quad (25)$$

Fick's law states that we have the following relation between mass flux and mass density.

$$f(x, t) = -k u_x(x, t), \quad k > 0 \quad (26)$$

Thus mass flow from high density areas towards low density areas.

Fick's law is only a phenomenological approximate relation based on the empirical observations that mass flux is for the most part a function of (x, t) , thorough the density gradient, thus $f(x, t) = f(u_x(x, t))$. Assuming that the flux is a smooth function of the density gradient, we can represent the flux as a power series in the density gradient

$$f(u_x(x, t)) = f_0 + f_1 u_x(x, t) + f_2 (u_x(x, t))^2 + \dots \quad (27)$$

Since there is no mass flux if the density gradient is zero we must have $f_0 = 0$. Furthermore, since mass flows from high density domains to low density domains

we must have $f_1 < 0$. We can therefore write $f_1 = -k$, where $k > 0$. Clearly, if the density gradient is large enough, the third term in the power series 27 will be significant and must be included. Like for the traffic model, this will lead to a nonlinear equation instead of the linear equation that will be the result of truncating the power series at the second term. In general we will assume that the parameter k in Fick's law depends on x . This corresponds to the assumption that the relation between flux and density gradient depends on position. This is not uncommon. Using Fick's law we can write the mass balance equation as

$$\begin{aligned}
\frac{dM}{dt} &= ku_x \Big|_{x_1}^{x_2}, \\
&\Updownarrow \\
\int_{x_1}^{x_2} u_t dx &= \int_{x_1}^{x_2} (ku_x)_x dx, \\
&\Updownarrow \\
\int_{x_1}^{x_2} (u_t - (ku_x)_x) dx &= 0, \quad \forall x_1, x_2 \\
&\Updownarrow \\
u_t - (ku_x)_x &= 0.
\end{aligned} \tag{28}$$

This is the 1D *diffusion equation*. By using similar modeling in 2D and 3D, we get the 2D and 3D diffusion equations

$$u_t = \nabla \cdot (k \nabla u) \quad \begin{cases} \nabla = (\partial_x, \partial_y) & 2D \\ \nabla^2 = (\partial_x, \partial_y, \partial_z) & 3D \end{cases}.$$

If the mass flow coefficient, k , is constant in space, the diffusion equation takes the form

$$u_t = k \nabla^2 u. \tag{29}$$

These equations and their close cousins describe a huge range of physical, chemical, biological and economic phenomena.

If the mass flow is stationary (time independent) and the mass flow coefficient independent of space, we get the simplified equation

$$\nabla^2 u = 0, \quad u = u(\mathbf{x}). \tag{30}$$

This is the *Laplace equation* and occur in the description of many stationary phenomena.

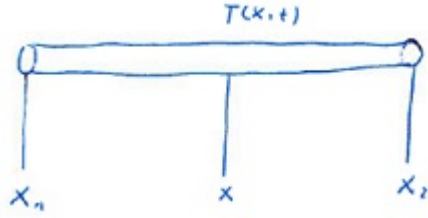


Figure 4

3.3 Heat conduction

Let us consider a narrow bar of some material, say a metal. We assume that the temperature of the bar, T , only depends on the coordinate along the bar which we call x .

Let the temperature of the bar at a point x at time t be $T(x, t)$. Let $e(x, t)$ be the energy density at x at time t . The total energy in the bar between x_1 and x_2 is

$$E(t) = \int_{x_1}^{x_2} e(x, t) dx.$$

The second great pillar of classical physics is the law of *conservation of energy*. Assuming no energy is gained or lost along the length of the bar, the conservation of energy gives us the balance equation

$$\frac{dE}{dt} = f_1 - f_2, \quad (31)$$

where $f_i = f_i(x_i, t)$ and $f(x, t)$ is the *energy flux* at x at time t . Here we use here the convention that positive energy flux means that energy flow to the right. In order to get a model for the bar, we must relate $e(x, t)$ and $f(x, t)$ to $T(x, t)$.

For many materials we have the following approximate identity

$$e(x, t) = C\rho T(x, t), \quad (32)$$

where ρ is the mass density of the bar and C is the heat capacity at constant volume. For (32) to hold the temperature must not vary too quickly. Energy is always observed to flow from high to low temperature regions. If the temperature gradient is not too large, *Fourier's law* hold for the energy flux

$$f(x, t) = -qT_x, \quad q > 0 \quad (33)$$

The energy balance equation then become

$$\frac{dE}{dt} = f_1 - f_2,$$

\Updownarrow

$$\begin{aligned}
\int_{x_1}^{x_2} C\rho T_t dx &= qT_x \Big|_{x_1}^{x_2} = \int_{x_1}^{x_2} (qT_x)_x dx, \\
&\Downarrow \\
\int_{x_1}^{x_2} (C\rho T_t - (qT_x)_x) dx &= 0, \quad \forall x_1, x_2 \\
&\Downarrow \\
C\rho T_t - (qT_x)_x &= 0.
\end{aligned}$$

If the heat conduction coefficient, q , does not depend on x , we get

$$T_t = DT_{xx}, \quad D = \frac{q}{C\rho}, \quad (34)$$

Again we get the diffusion equation. In this context the equation is called the *equation of heat conduction*. Energy flow in 2D and 3D similarly leads to 2D and 3D version of the heat conduction equations

$$T_t = D\nabla^2 T, \quad (35)$$

and for stationary heat flow, $T_t = 0$, we get the Laplace equation

$$\nabla^2 T = 0. \quad (36)$$

The wave equation, diffusion equation and Laplace equations and their close cousins will be the main focus in this course.

In the next example we will show that the diffusion equation also arise from a seemingly different context. This example is the beginning of a huge application that have ramifications in all areas of science, technology, economy, etc. I am talking about *stochastic processes*.

3.4 A random walk

Let x be the position of a particle restricted to moving along a line. We assume that the particle, if left alone, will not move. At random intervals we assume that the particle experience a collision that makes it shift its position by a small amount δ . The collision is equally likely to occur from the left or the right. Thus the shifts in position caused by a collision is equally likely to be $-\delta$ or $+\delta$. The collisions occur at random times, but let us assume that during some small time period, τ , a collision will occur for sure.

Let $P(x, t)$ be the probability of finding the particle at a point x at time t . We must then have

$$P(x, t + \tau) = \frac{1}{2}P(x - \delta, t) + \frac{1}{2}P(x + \delta, t). \quad (37)$$

This is just saying that *probability is conserved*. From this point of view (37) is a balance equation derived from a conservation law just like equations (11), (38), (25) and (31). Using Taylor expansions

$$\begin{aligned} P(x, t + \tau) &= P(x, t) + \partial_t P(x, t)\tau + O(\tau^2), \\ P(x \pm \delta, t) &= P(x, t) \pm \partial_x P(x, t)\delta \\ &\quad + \frac{1}{2}\partial_{xx}P(x, t)\delta^2 + O(\delta^3), \end{aligned}$$

we get from equation (37)

$$\begin{aligned} &P(x, t) + \partial_t P\tau + O(\tau^2) \\ &= \frac{1}{2}P(x, t) - \frac{1}{2}\partial_x P\delta + \frac{1}{4}\partial_{xx}P\delta^2 + \frac{1}{2}P(x, t) \\ &\quad + \frac{1}{2}\partial_x P\delta + \frac{1}{4}\partial_{xx}P\delta^2 + O(\delta^4). \end{aligned}$$

Thus by truncating the expansions we get

$$\partial_t P = D\partial_{xx}P, \quad D = \frac{\delta^2}{2\tau}.$$

This is again the diffusion equation. A more general and careful derivation of the random walk equation can be found in section 1.1 in [1].

3.5 The vibrating string

Let us consider a piece of string of length l that is fixed at $x = 0$ and $x = l$. We make the following modeling assumption

- i) The string can only move in a single plane, and in this plane it can only move transversally.

Using i) the state of the string can be described by a single function $u(x, t)$ measuring the height of the deformed string over the x -axis at time t . Negative values of u means that the string is below the x -axis.

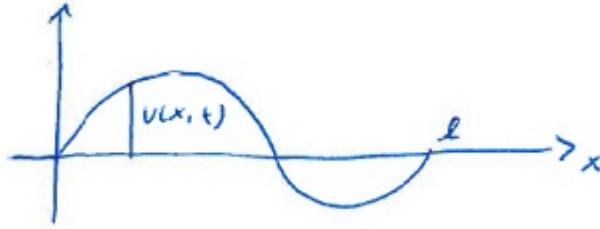


Figure 5

- ii) Any real string has a finite thickness, and bending of the string result in a tension (negative pressure) on the outer edge of the string and a compression (pressure) on the inner edge of the bend.



Figure 6

We model the string as infinitely thin and thus assume no compression force. The tension, \mathbf{T} , in a deformed string is always tangential to the string and assumed to be of constant length and time independent

$$T = |\mathbf{T}| = \text{const.}$$

- iii) The string has a mass density given by ρ which is independent of position and also time independent.

Note: ii) and iii) are not totally independent because a typical cause of nonuniform tension, T , is that the string has a variable material composition and this usually imply varying mass density.

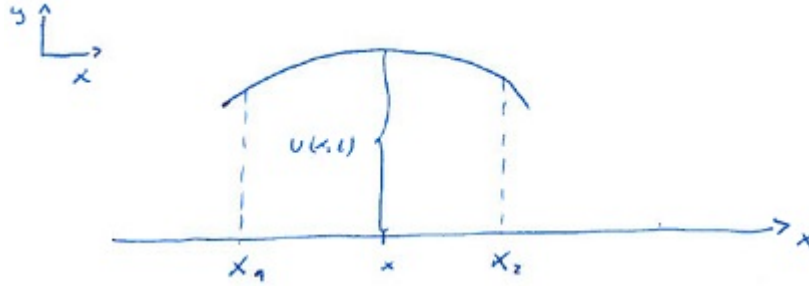


Figure 7: A piece of the string above the interval $[x_1, x_2]$.

Let us consider a piece of string above the interval $[x_1, x_2]$. The total vertical momentum (horizontal momentum is by assumption i) equal to zero), $P(t)$, in this piece of string at time t is

$$P(t) = \int_{x_1}^{x_2} dx \rho u_t. \quad (38)$$

Recall that momentum is conserved in classical (non relativistic) physics. This is the third great pillar of classical physics. Note that with the y -axis oriented like in figure 5, a piece of the string moving vertically up gives a positive contribution to the total momentum $P(t)$. Recall also that force is by definition change in momentum per unit time. Thus, if we assume there are no forces acting on the string inside the interval $[x_1, x_2]$, then momentum inside the interval can only change through the action of the tension force at the bounding points x_1 and x_2 . If we orient the coordinate axes as in figure 5, a positive vertical component, f_1 , of the tension force at x_1 , will lead to an increase the total momentum $p(t)$. A similar convention apply for f_2 at x_2 .

Conservation of momentum for the string is then expressed by the following balance equation

$$\frac{dP}{dt} = f_1 + f_2.$$



Figure 8

$$\begin{aligned} f_1 &= -T \cos \theta_1 = -T \cos\left(\frac{\pi}{2} - \theta\right) \\ &= -T \sin \theta \\ &= -T \frac{u_x}{\sqrt{1 + u_x^2}} \Big|_{x=x_1}. \end{aligned}$$

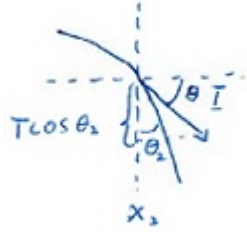


Figure 9

$$\begin{aligned} f_2 &= -T \cos \theta_2 = -T \cos\left(\frac{\pi}{2} + \theta\right) \\ &= T \sin \theta \\ &= T \frac{u_x}{\sqrt{1 + u_x^2}} \Big|_{x=x_2}. \end{aligned}$$

Our last modelling assumption is

- iv) The string perform small oscillations

$$|u_x| \ll 1.$$

Applying iv) gives us

$$\begin{aligned} f_1 &\approx -Tu_x \Big|_{x=x_1}, \\ f_2 &\approx Tu_x \Big|_{x=x_2}. \end{aligned} \tag{39}$$

Thus our fundamental momentum balance equation is well approximated by

$$\begin{aligned} \frac{dP}{dt} &= -Tu_x \Big|_{x_1} + Tu_x \Big|_{x_2} = Tu_x \Big|_{x_1}^{x_2}, \\ &\Downarrow \\ \int_{x_1}^{x_2} (\rho u_t)_t dx &= \int_{x_1}^{x_2} (Tu_x)_x dx, \\ &\Downarrow \\ \int_{x_1}^{x_2} (\rho u_{tt} - Tu_{xx}) dx &= 0, \quad \forall x_1, x_2 \\ &\Downarrow \\ \rho u_{tt} - Tu_{xx} &= 0, \\ &\Downarrow \\ u_{tt} - c^2 u_{xx} &= 0, \quad c = \sqrt{\frac{T}{\rho}} > 0. \end{aligned} \tag{40}$$

This is the *wave equation*. Similar modeling of small vibrations of membranes (2D) and solids (3D) give wave equations

$$u_{tt} - c^2 \nabla^2 u = 0 \quad \begin{cases} \nabla^2 = \partial_{xx} + \partial_{yy} & 2D \\ \nabla^2 = \partial_{xx} + \partial_{yy} + \partial_{zz} & 3D \end{cases}.$$

Wave equations describe small vibrations and wave phenomena in all sorts of physical systems, from stars to molecules.

4 Initial and boundary conditions

PDEs typically have a very large number of solutions (free functions, remember?). We usually select a unique solution by specifying extra conditions that the solutions need to satisfy. Picking the right conditions is part of the modeling and will be guided by the physical system that is under consideration.

There are typically two types of such conditions:

- i) initial conditions
- ii) boundary conditions

An initial condition consists of specifying the state of the system at some particular time t_0 . For the diffusion equation, we would specify

$$u(\mathbf{x}, t_0) = \varphi(\mathbf{x}),$$

for some given function φ . This would for example describe the distribution of dye at time t_0 . For the wave equation, we need *two* initial conditions since the equation is of second order in time (same as for ODEs)

$$\begin{aligned} u(\mathbf{x}, t_0) &= \varphi(\mathbf{x}), \\ \partial_t u(\mathbf{x}, t_0) &= \psi(\mathbf{x}). \end{aligned}$$

The modelling typically leads us to consider the PDE on some domain D (bounded or unbounded).

For the vibrating string it is natural to assume that the string has a finite length, l , so $D = [0, l]$. It is then intuitively obvious that in order to get a unique behavior we must specify what the conditions on the string are at the endpoints $x = 0$ and $x = l$. After all, a clamped string certainly behave different from a string that is loose at one or both ends! Here are a couple of natural conditions for the string:

- i) $u(0, t) = 0, u(l, t) = 0$,
- ii) $u(0, t) = g(t), u_x(l, t) = b(t)$.

For case i) the string is clamped at both ends and for case ii) the string is moved in a specified way at $x = 0$, and is acted upon by a force $Tb(t)$ at the right end, $x = l$.

The three types of boundary conditions that most often occur are

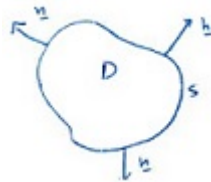
- D) u specified at the boundary.
- N) $\partial_{\mathbf{n}}u$ specified at the boundary.
- R) $\partial_{\mathbf{n}}u + au$ specified at the boundary.

The first is a *Dirichlet condition*, the second a *Neumann condition* and the third a *Robin condition*. Here \mathbf{n} is a normal to the curve (2D) or surface (3D) that defines the boundary of our domain and a is a function defined on the boundary. Let us consider the temperature profile of a piece of material covering a domain D . We want to model heat conduction in this material.

- i) Assume that the material is submerged in a large body of water kept at a temperature $T = 0$. The problem we must solve is then

$$\begin{aligned} \partial_t u &= D\nabla^2 u, & \text{in } D, \\ u(\mathbf{x}, 0) &= \varphi(\mathbf{x}), & \mathbf{x} \in D, \\ u(\mathbf{x}, t) &= 0, & \forall \mathbf{x} \in \partial D, \end{aligned}$$

where $\varphi(\mathbf{x})$ is the initial temperature profile.



\mathbf{n} is by convention the *outward pointing* normal.

Figure 10

- ii) If the material is perfectly isolated (no heat loss), we must solve the problem

$$\begin{aligned}\partial_t u &= D\nabla^2 u, & \text{in } D, \\ u(\mathbf{x}, 0) &= \varphi(\mathbf{x}), & \mathbf{x} \in D, \\ \partial_{\mathbf{n}} u(\mathbf{x}, t) &= 0, & \forall \mathbf{x} \in \partial D.\end{aligned}$$

The general morale here is that the choice of initial and boundary conditions are part of the modelling process just like the equation themselves.

4.1 Well posed Problems

Let us assume that the modelling of some physical system has produced a PDE, initial conditions and/or boundary conditions.

How can we know whether or not the model is any good? After all, the model is always a caricature of the real system. There are always many simplifying assumptions and approximations involved.

We could of course try to solve it and compare solutions to actual observation of the system. However, if this should even make sense to try, there are some basic requirements that *must* be satisfied.

- i) Existence: There is at least one solution that satisfy all the requirements of the model.
- ii) Uniqueness: There exists no more than one solution that satisfy all the requirements of the model.
- iii) Stability: The unique solution depends in a stable manner on the data of the problem. Thus, small changes in boundary conditions, initial conditions or parameters in equations, lead to small changes in the solutions.

Producing a formal mathematical proof that a given model is well posed can be a hard problem and belongs to the area of pure mathematics. In applied mathematics we derive the models using our best physical intuition and sound methods of approximation and assume as a “working hypothesis” that they are well posed. If they are not, the models will eventually break down and in this way let us know. The way they break down is often a hint we can use to modify them such that the modified models *are* well posed.

We thus use an “empirical” approach to the question of well-posedness in applied mathematics. We will return to this topic later in the class.

5 Numerical methods

In most realistic modelling, the equations and/or the geometry of the domains are too complicated for solution “by hand” to be possible. In the “old days” this would mean that the problems were unsolvable. Today, the ever increasing memory and power of computers offer a way to solve previously unsolvable problems. Both exact methods and approximate numerical methods are available, often combined in easy to use systems, like Matlab, Mathematica, Maple etc.

In this section we will discuss numerical methods. The use of numerical methods to solve PDEs does not make more theoretical methods obsolete. On the contrary, heavy use of mathematics is often required to reformulate the PDEs and write them in a form that is well suited for the architecture of the machine we will be running it on, which could be serial, shared memory, cluster etc.

Solving PDEs numerically has two sources of error.

- 1) Truncation error: This is an error that arises when the partial differential equation is approximated by some (multidimensional) difference equation. Continuous space and time can not be represented in a computer and as a consequence there will *always* be truncation errors. The challenge is to make them as small as possible and at the same time stay within the available computational resources.
- 2) Rounding errors: Any computer memory is finite and therefore only a finite set of real numbers can be represented. After any mathematical operation on the available numbers we must select the “closest possible” available number to represent the result. In this way we accumulate errors after every operation. In order for the accumulated errors not to swamp the result we are trying to compute, we must be careful about *how* the mathematical operations are performed and *how many* operations we perform.

1) and 2) apply to *all* possible numerical methods. Heavy use of mathematical methods are often necessary to build confidence that some numerical solution is an accurate representation of the true mathematical solution. The accuracy of numerical solutions are often tested by running them against known exact solutions to special initial and/or boundary conditions and to “stripped down” versions of the equations.

The simplest truncation method to describe is *the method of finite differences*. This is the only truncation method we will discuss in this class. We will introduce the method by applying it to the 1D heat equation, the 1D wave equation and the 2D Laplace equation

5.1 The finite difference method for the heat equation

Consider the following initial/boundary value problem for the heat equation

$$\begin{aligned} u_t &= u_{xx}, & 0 < x < 1, t > 0, \\ u(0, t) &= u(1, t) = 0, \\ u(x, 0) &= \varphi(x). \end{aligned}$$

We introduce a uniform spacetime grid (x_i, t_n) , where

$$\begin{aligned} x_i &= hi, & i = 0, \dots, N, \\ t_n &= kn, & n \in \mathbb{N}, \end{aligned}$$

where $x_0 = 0, x_N = 1$ and thus $h = \frac{1}{N}$. The solution will only be computed at the grid points.

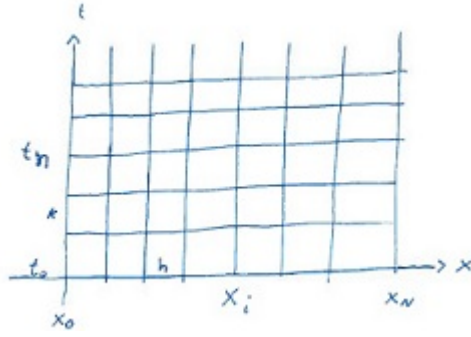


Figure 11

Evaluating the equation and the initial/boundary conditions at the grid points we find

$$\begin{aligned} (u_t)_i^n &= (u_{xx})_i^n, \\ u_0^n &= 0, \quad u_N^n = 0, \\ u_i^0 &= \varphi_i = \varphi(x_i), \end{aligned} \tag{41}$$

where $u_i^n = u(x_i, t_n)$.

We now must find expressions for u_t and u_{xx} , using *only* function values at the grid points. Our main tool for this is Taylors theorem.

$$\begin{aligned} u(x_i, t_n + k) &= u_i^n + \left(\frac{\partial u}{\partial t} \right)_i^n k + \frac{1}{2} \left(\frac{\partial^2 u}{\partial t^2} \right)_i^n k^2 + O(k^3), \\ u(x_i, t_n - k) &= u_i^n - \left(\frac{\partial u}{\partial t} \right)_i^n k + \frac{1}{2} \left(\frac{\partial^2 u}{\partial t^2} \right)_i^n k^2 + O(k^3). \end{aligned}$$

Using the Taylor expansions we easily get the following expressions

$$\begin{aligned}\frac{u(x_i, t_n + k) - u_i^n}{k} &= \left(\frac{\partial u}{\partial t} \right)_i^n + O(k), \\ \frac{u_i^n - u(x_i, t_n - k)}{k} &= \left(\frac{\partial u}{\partial t} \right)_i^n + O(k), \\ \frac{u(x_i, t_n + k) - u(x_i, t_n - k)}{2k} &= \left(\frac{\partial u}{\partial t} \right)_i^n + O(k^2).\end{aligned}$$

Since $t_n \pm k = t_{n \pm 1}$, we get the following three approximations to the first derivative with respect time

$$\left. \begin{aligned} \left(\frac{\partial u}{\partial t} \right)_i^n &\approx \frac{u_i^{n+1} - u_i^n}{k} && \text{forward difference} \\ \left(\frac{\partial u}{\partial t} \right)_i^n &\approx \frac{u_i^n - u_i^{n-1}}{k} && \text{backward difference} \\ \left(\frac{\partial u}{\partial t} \right)_i^n &\approx \frac{u_i^{n+1} - u_i^{n-1}}{2k} && \text{center difference} \end{aligned} \right\} \text{approx. has an error of order } k, \quad \text{order } k^2.$$

(42)

The second derivative in space is also approximated using Taylor theorem

$$\begin{aligned}u(x_i + h, t_n) &= u_i^n + \left(\frac{\partial u}{\partial x} \right)_i^n h + \frac{1}{2} \left(\frac{\partial^2 u}{\partial x^2} \right)_i^n h^2 + \frac{1}{6} \left(\frac{\partial^3 u}{\partial x^3} \right)_i^n h^3 + O(h^4), \\ u(x_i - h, t_n) &= u_i^n - \left(\frac{\partial u}{\partial x} \right)_i^n h + \frac{1}{2} \left(\frac{\partial^2 u}{\partial x^2} \right)_i^n h^2 - \frac{1}{6} \left(\frac{\partial^3 u}{\partial x^3} \right)_i^n h^3 + O(h^4).\end{aligned}$$

From these two expressions we get

$$\frac{u(x_i + h, t_n) - 2u_i^n + u(x_i - h, t_n)}{h^2} = \left(\frac{\partial^2 u}{\partial x^2} \right)_i^n + O(h^2).$$

Thus we get the following approximation to the second derivative in space

$$\left(\frac{\partial^2 u}{\partial x^2} \right)_i^n \approx \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \quad (\text{center difference}) \quad \text{order } h^2. \quad (43)$$

Using the forward difference in time and center difference in space in (41) gives

$$\frac{u_i^{n+1} - u_i^n}{k} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2}.$$

Define $s = \frac{k}{h^2}$. Then the problem we must solve is

$$\begin{aligned} u_i^{n+1} &= s(u_{i+1}^n + u_{i-1}^n) + (1 - 2s)u_i^n, \quad i = 1, \dots, N-1, \\ u_0^n &= u_N^n = 0, \\ u_i^0 &= \varphi_i. \end{aligned} \tag{44}$$

This is now a (partial) difference equation.

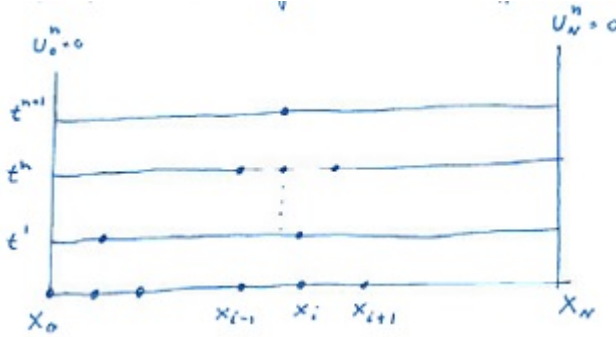


Figure 12

Let us use the initial condition

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

The exact solution is found by separation of variables which we will discuss later in this class.

$$u(x, t) = \frac{8}{\pi^2} \sum_{m=1}^{\infty} \frac{1}{m^2} \sin \frac{1}{2} m \pi e^{-m^2 \pi^2 t} \sin m \pi x.$$

In order to compute the numerical solution, we must choose values for the spacetime discretization parameters h and k . In the figures 13 and 14, we plot the initial condition, the exact solution and the numerical solution for $h = 0.05$, thus $N = 50$, and $k = h^2 s = 0.0004s$ for $s = 0.49$ and 0.51 . We use 100 iterations in n .

It is clear from the figure 14 that the numerical solution for $s = 0.51$ is very bad. This is an example of a *numerical instability*. Numerical instability is a constant danger when we compute numerical solutions to PDEs.

For a simple equation like the heat equation we can understand this instability, why it appears and how to protect against it.

Let us try to solve the difference equation in (44) by separating variables. We thus seek a solution of the form

$$u_j^n = X_j T_n,$$

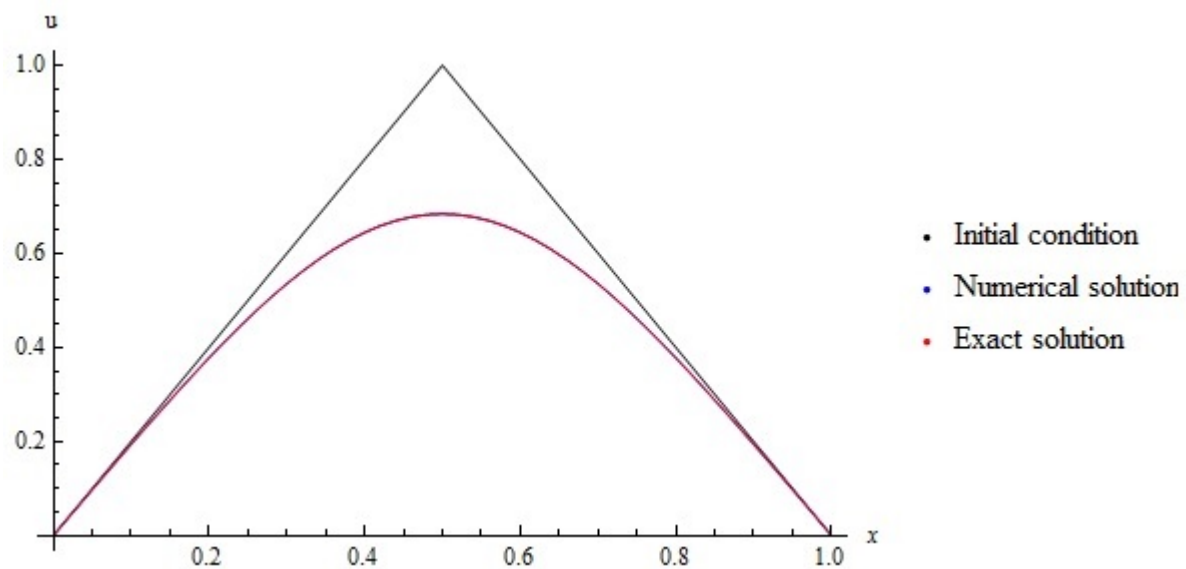


Figure 13: In this figure we have $s=0.49$

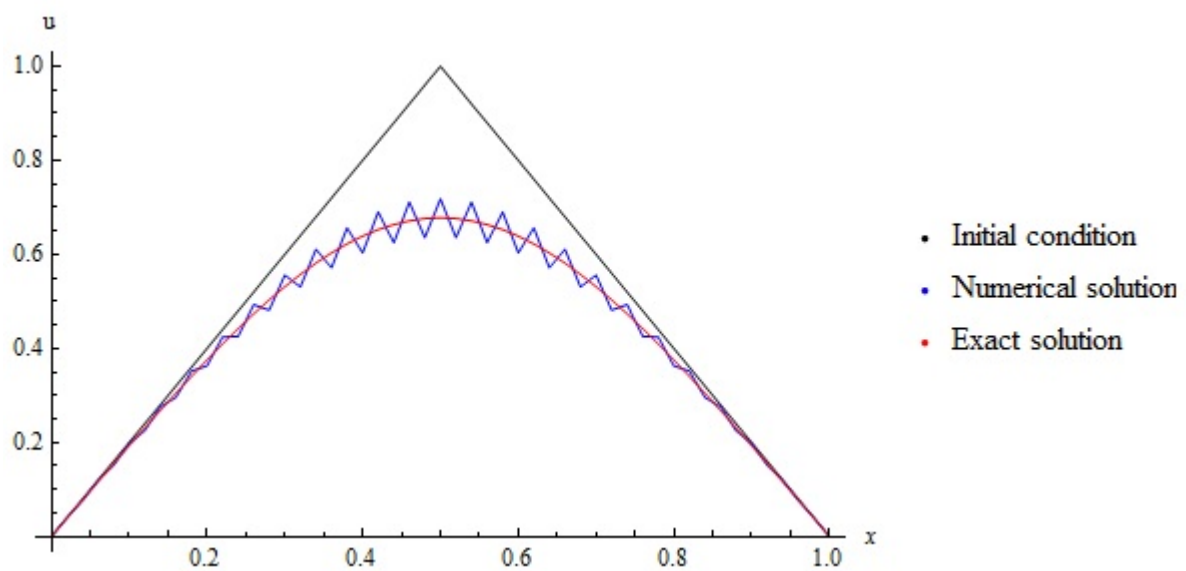


Figure 14: In this figure $s=0.51$

$$\begin{aligned}
& \Downarrow \\
X_j T_{n+1} &= s(X_{j+1} T_n + X_{j-1} T_n) + (1 - 2s) X_j T_n, \\
& \Downarrow \\
\frac{T_{n+1}}{T_n} &= (1 - 2s) + s \frac{X_{j+1} + X_{j-1}}{X_j}.
\end{aligned}$$

Both sides must be equal to the same constant ξ . This gives us two difference equations coupled through the separation constant ξ .

$$\begin{aligned}
1) \quad & T_{n+1} = \xi T_n. \\
2) \quad & s \frac{X_{j+1} + X_{j-1}}{X_j} + (1 - 2s) = \xi.
\end{aligned}$$

The solution of equation 1) is simply

$$T_n = \xi^n T_0.$$

The second equation is a boundary value problem

$$s(X_{j+1} + X_{j-1}) + (1 - 2s)X_j = \xi X_j, \quad (45)$$

$$X_0 = X_N = 0,$$

We look for solutions of the form $X_j = z^j$. This implies that

$$\begin{aligned}
X_{j-1} &= z^{-1} X_j, \\
X_{j+1} &= z X_j,
\end{aligned}$$

Thus (45) implies that

$$\begin{aligned}
\xi &= s(z + z^{-1}) + (1 - 2s), \\
& \Updownarrow \\
z^2 + \left(\frac{1}{s} - 2 - \frac{\xi}{s}\right)z + 1 &= 0.
\end{aligned} \quad (46)$$

This equation has two solutions z_1 and z_2 which we in the following assume to be different. (Show that the case when $z_1 = z_2$ does not lead to a solution of the boundary value problem (45).)

The general solution of (45) is thus

$$X_j = A z_1^j + B z_2^j.$$

The general solution must satisfy the boundary conditions. This implies that

$$\begin{aligned}
X_0 = 0 &\Leftrightarrow A + B = 0, \\
X_N = 0 &\Leftrightarrow A z_1^N + B z_2^N = 0.
\end{aligned}$$

Nontrivial solutions exists only if

$$\left(\frac{z_1}{z_2}\right)^N = 1. \quad (47)$$

This condition can not be satisfied if $z_1 \neq z_2$ are real. Since equation (46) has real coefficient, we must have

$$\begin{aligned} z_1 &= z = \rho e^{i\theta}, \\ z_2 &= \bar{z} = \rho e^{-i\theta}. \end{aligned}$$

But equation (46) $\Rightarrow z_1 z_2 = 1 \Leftrightarrow \rho^2 = 1 \Leftrightarrow \rho = 1$. Thus $z_1 = e^{i\theta}$, $z_2 = e^{-i\theta}$. The condition (47) then implies that

$$e^{2i\theta N} = 1, \quad \Leftrightarrow \quad \theta_r = \frac{\pi r}{N}, \quad r = 1, 2, \dots, N-1.$$

The corresponding solutions of the boundary value problem (45) are

$$X_j^{(r)} = A(e^{i\theta_r j} - e^{-i\theta_r j}) = 2iA \sin \theta_r j.$$

A basis of solutions can be choosen to be

$$X_j^{(r)} = \sin \frac{r\pi j}{N}, \quad r = 1, 2, \dots, N-1.$$

It is easy to verify that

$$c_1 X_j^{(1)} + \dots + c_{N-1} X_j^{(N-1)} = 0, \quad \forall j$$

\Downarrow

$$c_1 = c_2 = \dots = c_{N-1} = 0.$$

The vectors $\{X_j^{(r)}\}_{r=1}^{N-1}$ are thus linearly independent and indeed form a basis because \mathbb{R}^{N-1} has dimension $N-1$.

Each vector $X_j^{(r)}$ gives a separated solution.

$$u_{rj}^n = (\xi_r)^n \sin \frac{r\pi j}{N},$$

to the boundary value problem

$$\begin{aligned} u_i^{n+1} &= s(u_{i+1}^n + u_{i-1}^n) + (1-2s)u_i^n, \quad i = 1, \dots, N-1, \\ u_0^n &= u_N^n = 0. \end{aligned}$$

where we have defined

$$\xi_r = (1-2s) + s(e^{i\theta_r} + e^{-i\theta_r}) = 1 - 2s(1 - \cos \frac{r\pi}{N}).$$

These solutions give exponential growth if $|\xi_r| > 1$ for some r . Thus the requirement for *stability* is

$$|\xi_r| \leq 1, \quad \forall r.$$

If this holds, the numerical scheme is stable.

Now

$$\begin{aligned} -1 &\leq \cos \frac{r\pi}{N} \leq 1, \quad \forall r, \\ &\Downarrow \\ 0 &\leq 1 - \cos \frac{r\pi}{N} \leq 2, \quad \forall r. \end{aligned}$$

Therefore $-1 \leq \xi^r \leq 1$ if

$$\begin{aligned} 1 - 4s &\geq -1, \\ &\Downarrow \\ s &\leq \frac{1}{2}. \end{aligned}$$

This is the stability condition indicated by the numerical test illustrated in figures (13) and (14).

5.2 The Crank-Nicholson scheme

It is possible to avoid the stability condition by modifying the simple scheme discussed earlier. Define

$$(\delta^2 u)_j^n = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2}.$$

and let $0 \leq Q \leq 1$. For each such number, we define a “ Q -scheme”

$$\frac{u_j^{n+1} - u_j^n}{k} = (1 - Q)(\delta^2 u)_j^n + Q(\delta^2 u)_j^{n+1}. \quad (48)$$

When $Q = 0$, we get the explicit scheme we have discussed. For $Q > 0$ the scheme is implicit; At each step we must solve a linear system.

We analyze the stability of the scheme by inserting a separated solution of the form

$$u_j^n = (\xi_r)^n e^{i \frac{2r\pi j}{N}}. \quad (49)$$

Note: This separated solution will not satisfy Dirichlet boundary conditions but *periodic* boundary conditions, $u_{j+N}^n = u_j^n$. In this type of stability analysis we disregard the actual boundary conditions of the problem, using instead periodic boundary conditions. A numerical scheme is *Von Neumann Stable* if

$$|\xi_r| \leq 1, \quad \forall r$$

There are other notions of stability where the actual boundary conditions are taken into account, but we will not discuss them in this course. Usually Von Neumann Stability gives a good indicator of stability, also when the actual boundary conditions are used.

If we insert (49) into the Q -scheme we get

$$\xi_r = \frac{1 - 2(1 - Q)s(1 - \cos \frac{2r\pi}{N})}{1 + 2Qs(1 - \cos \frac{2r\pi}{N})}, \quad s = \frac{k}{h^2}.$$

Evidently $\xi_r \leq 1$. The condition $\xi_r \geq -1$ requires that

$$s(1 - 2Q)(1 - \cos \frac{2r\pi}{N}) \leq 1.$$

This is *always* true for $1 - 2Q \leq 0$. Thus if $\frac{1}{2} \leq Q \leq 1$ the numerical scheme is stable for *all* choices of grid parameters k and h . The price we pay is that the scheme is implicit.

For $\frac{1}{2} \leq Q \leq 1$ we say that the Q -scheme is *unconditionally stable*. For $Q = \frac{1}{2}$, the Q -scheme is called the Crank-Nicolson scheme.

Observe that when N is large, the angles

$$\theta_r = \frac{2\pi r}{N}, r = 1, 2, \dots, N,$$

form a very dense set of points in the angular interval $[0, 2\pi]$. Thus, in the argument above we might for simplicity take θ to be a continuous variable in the interval $[0, 2\pi]$. We will therefore in the following investigate Von Neuman stability by inserting

$$u_j^n = (\xi)^n (\eta)^j,$$

where $\eta = e^{i\theta}$, into the numerical scheme of interest. Verify that using this simplified approach leads to the same stability conditions for the two schemes (44) and (48).

5.2.1 Finite difference method for the heat equation: Neumann boundary conditions

Let us consider the heat equation as before but now with Neumann boundary conditions.

$$\begin{aligned} u_t &= u_{xx}, & 0 < x < 1, \\ u(x, 0) &= \varphi(x), \\ u_x(0, t) &= f(t), & u_x(1, t) = g(t). \end{aligned}$$

One possible discretization for the boundary conditions is to use forward difference at $x = 0$ and backward difference at $x = 1$.

$$\begin{aligned} \frac{u_1^n - u_0^n}{h} &= f^n, \\ \frac{u_N^n - u_{N-1}^n}{h} &= g^n. \end{aligned}$$

These approximations are only of order h . The space derivative in the equation was approximated to second order h^2 . We get second order also at the boundary



Figure 15

by using center difference. However we then need to extend our grid by “ghost points” x_{-1} and x_{N+1}

$$(u_x)_0 = \frac{u_1 - u_{-1}}{2h}, \quad (u_x)_N = \frac{u_{N+1} - u_{N-1}}{2h}.$$

Thus our numerical scheme is

$$\begin{aligned} u_i^{n+1} &= s(u_{i+1}^n + u_{i-1}^n) + (1 - 2s)u_i^n, \quad i = 0, \dots, N, \\ u_{-1}^n &= u_1^n - 2hf^n, \\ u_{N+1}^n &= u_{N-1}^n + 2hg^n. \end{aligned}$$

Introduction of ghost points is the standard way to treat derivative boundary conditions for 1D, 2D and 3D problems.

5.3 Finite difference method for the wave equation

We consider the following initial and boundary value problem for the wave equation

$$\begin{aligned} u_{tt} - u_{xx} &= 0, & 0 < x < L, t > 0, \\ u(0, t) &= u(L, t) = 0, \\ u(x, 0) &= \varphi(x), \\ u_t(x, 0) &= \psi(x), \end{aligned}$$

on the real axis. We introduce a uniform spacetime grid (x_j, t_n) where

$$\begin{aligned} x_j &= hj, \quad j = 0, \dots, N, \\ t_n &= kn, \quad n \in \mathbb{N}, \end{aligned}$$

where $x_0 = 0, x_N = L$ and thus $h = \frac{L}{N}$.

The wave equation is required to hold at the grid points only.

$$(u_{tt})_j^n = (u_{xx})_j^n.$$

The partial derivatives are approximated by centered differences (order h^2, k^2).

$$\frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{k^2} = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2},$$

\Updownarrow

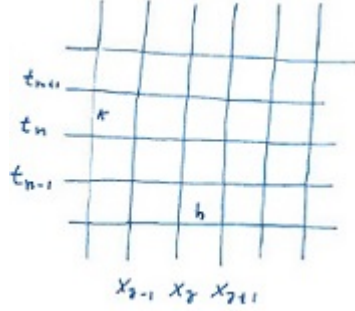


Figure 16

$$u_j^{n+1} = s(u_{j+1}^n + u_{j-1}^n) + 2(1-s)u_j^n - u_j^{n-1}, \quad (50)$$

where $s = \frac{k^2}{h^2}$. The boundary conditions evaluated on the grid simply become $u_0^n = 0, u_N^n = 0$.

Note that the difference equation (50) is of order 2 in n and j . In order to solve the difference equation and compute the solution at time level $n+1$, we need to know the values on time level n and $n-1$. Thus to compute u_j^2 , we need u_j^1 and u_j^0 . These values are provided by the initial conditions

$$\begin{aligned} u(x, 0) &= \varphi(x), \\ u_t(x, 0) &= \psi(x). \end{aligned}$$

Evaluated on the grid these are

$$\begin{aligned} u_j^0 &= \varphi_j, \\ (u_t)_j^0 &= \psi_j. \end{aligned}$$

We now need a difference approximation for u_t at timelevel $n=0$. Since the scheme (50) is of order 2 in t and x , we should use a 2-order approximation for u_t . If we use a 1-order approximation for u_t in the initial condition the whole numerical scheme is of order 1. Using a centered difference in time, we get

$$\frac{u_j^1 - u_j^{-1}}{2k} = \psi_j. \quad (51)$$

We have introduced a “ghost point” t_{-1} in time with associated u value u_j^{-1} . If we insert $n=0$ in (50), we get a second identity involving u_j^{-1}

$$u_j^1 = s(u_{j+1}^0 + u_{j-1}^0) + 2(1-s)u_j^0 - u_j^{-1}. \quad (52)$$

We can use (51) and (52) to eliminate the ghost point and find in conjunction with $u_j^0 = \varphi_j$ that

$$\begin{aligned} u_j^0 &= \varphi_j, \\ u_j^1 &= \frac{1}{2}s(\varphi_{j+1} + \varphi_{j-1}) + (1-s)\varphi_j + k\psi_j. \end{aligned}$$

These are the initial values for the scheme (50). We now test the numerical scheme using initial conditions

$$\begin{aligned}\varphi(x) &= e^{-a(x-\frac{1}{2}L)^2}, \\ \psi(x) &= 0.\end{aligned}$$

The exact solution

$$u(x, t) = \frac{1}{2}(\varphi(x+t) + \varphi(x-t)),$$

to this problem is found using a method that we will discuss later in this class. For the test we use $L = 20$ and chose the parameter a so large that $\varphi(x)$ to a good approximation satisfy the boundary conditions for the wave equation. In figures (17) and (18) we plot the initial condition, the exact solution and the numerical solution on the interval $(0, L)$. The space-time grid parameters are $h = 0.1$, thus $N = 200$, and $k = h\sqrt{s} = 0.0004\sqrt{s}$ for $s = 0.9$ in figure 17 and $s = 1.1$ in figure 18. We use 60 iterations in n .

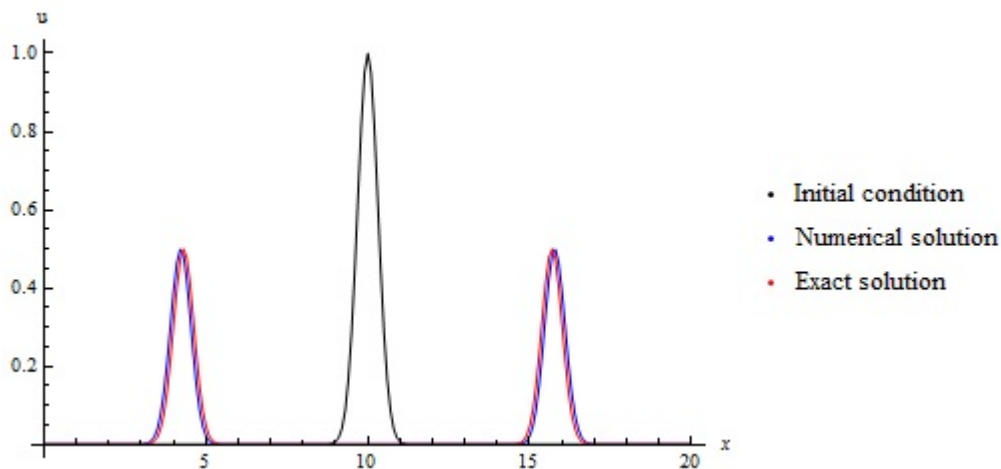


Figure 17

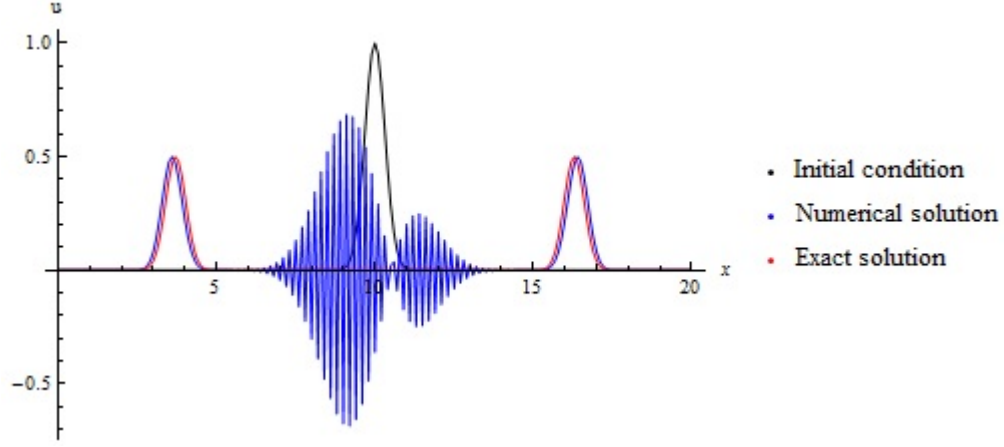


Figure 18

It appears as if the scheme is stable for $s \lesssim 1$ and unstable for $s > 1$. Let us prove this using the Von Neumann Approach.

$$u_j^n = \eta^j \xi^n, \quad \eta = e^{i\theta},$$

where j and n on the right hand side of the first equation are to be interpreted as a power, i.e. $\eta^j = (\eta)^j$, etc. Inserting this into the scheme we get

$$\eta^j \xi^{n+1} = s(\eta^{j+1} \xi^n + \eta^{j-1} \xi^n) + 2(1-s)\eta^j \xi^n - \eta^j \xi^{n-1},$$

$$\Downarrow \quad (\text{divided by } \xi^n \eta^j)$$

$$\xi + \frac{1}{\xi} - 2 = s\left(\eta + \frac{1}{\eta} - 2\right) = 2s(\cos \theta - 1).$$

Define $p = s(\cos \theta - 1)$. Then the equation can be written as

$$\xi^2 - 2(1+p)\xi + 1 = 0.$$

The roots are

$$\xi = 1 + p \pm \sqrt{p^2 + 2p}.$$

Clearly $p \leq 0$. If $p < -2$ then $p^2 + 2p = p(p+2) > 0$ and there are two real roots. One of these must be $< -1 \Rightarrow |\xi| > 1 \Rightarrow$ scheme unstable. If $p > -2$ then $p^2 + 2p < 0$ and we have two complex conjugate roots

$$\xi_{\pm} = 1 + p \pm i\sqrt{-p^2 - 2p}.$$

These two roots satisfy

$$\begin{aligned} |\xi_{\pm}|^2 &= (1+p)^2 + (-p^2 - 2p) \\ &= 1 + 2p + p^2 - p^2 - 2p = 1. \end{aligned}$$

we therefore have stability if $p > -2$. It is easy to verify that the scheme is also stable for $p = -2$. Taking into account the definition of p we thus have stability if

$$s(\cos \theta - 1) \geq -2,$$

$$s \leq \frac{2}{1 - \cos \theta}.$$

This holds for all θ if $s \leq 1$ because clearly $\frac{2}{1 - \cos \theta} \geq 1$. Thus we have stability if

$$s \leq 1.$$

which is what our numerical test indicated.

5.4 Finite difference method for the Laplace equation

Let us consider the problem

$$\begin{aligned} \Delta u &= 0, & (x, y) &\in (0, 1) \times (0, 1) \equiv D, \\ u &= f & (x, y) &\in \partial D. \end{aligned}$$

We solve the problem by introducing a grid

$$(x_i, y_j) \quad x_i = ih, y_j = jp, \quad i = 0, \dots, N, j = 0, \dots, M$$

where $h = \frac{1}{N}, p = \frac{1}{M}$.

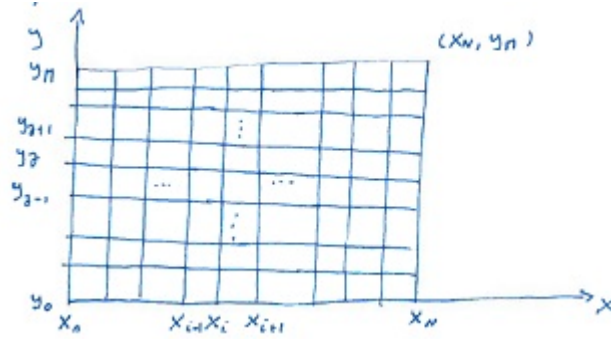


Figure 19

Using center difference for u_{xx} and u_{yy} , we get the scheme

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{p^2} = 0.$$

This scheme is 2-order in h and p . Because of the symmetry there is no reason to choose h and p differently. For the case $h = p$, which corresponds to $M = N$, we get

$$u_{i,j} = \frac{1}{4}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}). \quad (53)$$



Figure 20

Observe that the value in the center is simply the mean of the value of all closest neighbors. The *stencil* of the scheme is shown in figure (20).

There is no iteration process in the scheme (53) (There is no time!). It is a linear system of variables for the $(N-1)(N-1)$ interior gridpoints. We are thus looking at a linear system of equations

$$A\mathbf{x} = \mathbf{b}, \quad (54)$$

where \mathbf{b} comes from the boundary values. Observe however, that the linear system (53) is not actually written in the form (54). In order to apply linear algebra to solve (53), we must write (53) in the form (54). The problem that arises at this point is that the interior gridpoints in (53) are not *ordered*, whereas the vector components in (54) are ordered. We must thus *choose* an ordering of the gridpoints. Each choice will lead to a different *structure* for the matrix A (for instance which elements are zero and which are nonzero). Since the efficiency of iterative methods depends on the structure of the matrix, this is a point where we can be clever (or stupid!).

For the problem at hand it is natural to order the interior gridpoints *lexicographical*.

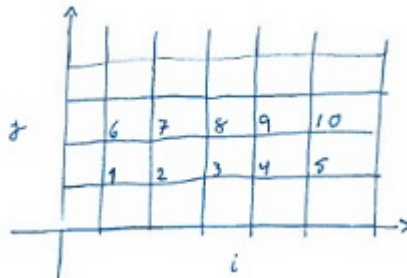


Figure 21: Ordering the grid points.

Thus we define the vector \mathbf{x} as

$$\mathbf{x} = (u_{1,1}, u_{2,1}, \dots, u_{N-1,1}, u_{1,2}, u_{2,2}, \dots, u_{N,1}, \dots, u_{N-1,N-1}).$$

The \mathbf{b} vector is obtained from the values of u at the boundary. For example

when $j = 1$, we have

$$u_{i,1} = \frac{1}{4}(u_{i+1,1} + u_{i-1,1} + u_{i,2} + u_{i,0}), \quad 1 \leq i \leq N-1.$$

The boundary condition gives

$$\begin{aligned} u_{i,0} &= f_{i,0} \equiv f(x_i, y_0), \\ u_{i,N} &= f_{i,N}, \\ u_{0,j} &= f_{0,j}, \\ u_{N,j} &= f_{N,j}, \end{aligned}$$

$i = 1$,

$$\begin{aligned} 4u_{1,1} &= u_{2,1} + u_{0,1} + u_{1,2} + u_{1,0}, \\ &\Downarrow \end{aligned}$$

$$-4u_{1,1} + u_{1,2} + u_{2,1} = -f_{0,1} - f_{1,0},$$

$1 < i < N-1$,

$$\begin{aligned} 4u_{i,1} &= u_{i+1,1} + u_{i-1,1} + u_{i,2} + u_{i,0}, \\ &\Downarrow \end{aligned}$$

$$u_{i-1,1} - 4u_{i,1} + u_{i+1,1} + u_{i,2} = -f_{i,0},$$

$i = N-1$,

$$\begin{aligned} 4u_{N-1,1} &= u_{N,1} + u_{N-2,1} + u_{N-1,2} + u_{N-1,0}, \\ &\Downarrow \end{aligned}$$

$$u_{N-2,1} - 4u_{N-1,1} + u_{N-1,2} = -f_{N,1} - f_{N-1,0}.$$

Thus the matrix for the system has a structure that starts off as

$$\begin{array}{cccccccccccccccc} & & & & & & & N-1 & & & & & & & & (N-1)(N-1) \\ -4 & 1 & 0 & \cdots & \cdots & \cdots & 0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 1 & -4 & 1 & 0 & \cdots & \cdots & 0 & 0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & & & & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 & -4 & 0 & \cdots & \cdots & \cdots & 1 & 0 & \cdots & 0 \end{array}.$$

By writing down a larger part of the linear system we eventually realize that the matrix of the system has the form of a $(N-1) \times (N-1)$ block matrix

$$A = \begin{pmatrix} B & I & 0 & \cdots & \cdots & 0 \\ I & B & I & 0 & \cdots & 0 \\ & & \vdots & \vdots & & \\ 0 & \cdots & \cdots & \cdots & I & B \end{pmatrix},$$

where I is the $(N-1)(N-1)$ identity and where B is the following $(N-1) \times (N-1)$ tri-diagonal matrix

$$B = \begin{pmatrix} -4 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & -4 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -4 & 1 & 0 & \cdots & \cdots \\ & & \vdots & & \vdots & & \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 & -4 \end{pmatrix}.$$

We observe that most elements in the matrices A and B are zero, we say the matrices are *sparse*. The structural placing of the nonzero elements represents the *sparsity structure* of the matrix. Finite difference methods for PDEs usually leads to sparse matrices with a simple sparsity structure.

When N is large, as it often is, the linear system (54) represents an enormous computational challenge.

In 3D with 1000 points in each direction, we get a linear system with 10^9 independent variables and the matrix consists of 10^{18} entries!

There are two broad classes of methods available for solving large linear systems of P equations in P unknowns.

- 1) Direct methods: These are variants of elementary Gaussian elimination. If the matrix is sparse (most elements are zero) direct methods can solve large linear systems fast. For dense matrices, the complexity of Gaussian elimination ($O(P^3)$) makes it impossible to solve more than few thousand equations in reasonable time.
- 2) Iterative methods: These methods try to find approximate solutions to (53) by defining an iteration

$$\mathbf{x}^{n+1} = M\mathbf{x}^n,$$

that converges to a solution of the linear system of interest. Since each multiplication cost at most $O(P^2)$ operations, fast convergence ($n \ll P$) will give an approximate solution to (54) in $O(P^2)$ operations, even for a dense matrix.

The area of iterative solution methods for linear systems is vast and we can not possibly give an overview of the available methods and issues in a introductory class on PDEs. However, it is appropriate to discuss some of the central ideas.

5.4.1 Iterations in \mathbb{R}^n

Let $\mathbf{u}^0 \in \mathbb{R}^n$ be a given vector in \mathbb{R}^n and consider the following iteration

$$\mathbf{u}^{l+1} = M\mathbf{u}^l + \mathbf{c},$$

where M is a $n \times n$ matrix and \mathbf{c} is a fixed vector. Let us assume that $\mathbf{u}^l \rightarrow \mathbf{u}^*$ when $l \rightarrow \infty$. Then we have

$$\mathbf{u}^{l+1} = M\mathbf{u}^l + \mathbf{c},$$

$$\begin{aligned}
& \Downarrow \\
& \mathbf{u}^* = M\mathbf{u}^* + \mathbf{c}, \\
& \Downarrow \\
& (I - M)\mathbf{u}^* = \mathbf{c}.
\end{aligned}$$

Thus \mathbf{u}^* solves a linear system. The idea behind iterative solution of linear system can be described as follows:

Find a matrix M and a vector \mathbf{c} such that

$$\begin{aligned}
\text{i)} \quad & A\mathbf{u} = f, \\
& \Updownarrow \\
& (I - M)\mathbf{u} = \mathbf{c},
\end{aligned}$$

ii) The iteration

$$\mathbf{u}^{l+1} = M\mathbf{u}^l + \mathbf{c},$$

converges.

The condition i) can be realized in many ways and one seeks choices M and \mathbf{c} such that the convergence is *fast*.

In order to get anywhere, it is at this point evident that we need to be more precise about convergence in \mathbb{R}^n . This we do by introducing a vector norm on \mathbb{R}^n that measures distance between vectors. We will also need a corresponding norm on $n \times n$ matrices.

A $n \times n$ matrix is a linear operator on \mathbb{R}^n . Let us assume that \mathbb{R}^n comes with a norm $\|\cdot\|$ defined. There are many such norms in use for \mathbb{R}^n . The following three norms are frequently used

$$\begin{aligned}
\|\mathbf{x}\|_2 &= \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}, \\
\|\mathbf{x}\|_1 &= \sum_{i=1}^n |x_i|, \\
\|\mathbf{x}\|_\infty &= \sup_i \{|x_i|\}.
\end{aligned}$$

Given a norm, $\|\cdot\|$, on \mathbb{R}^n there is a corresponding norm on linear operators (matrices) defined by

$$\|A\| = \sup_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|.$$

For matrix norms corresponding to vector norms we have the following two important inequalities.

$$\begin{aligned}
\|A\mathbf{x}\| &\leq \|A\| \|\mathbf{x}\|, \\
\|AB\| &\leq \|A\| \|B\|.
\end{aligned}$$

For $\|\cdot\|_1$ and $\|\cdot\|_\infty$, one can show that

$$\begin{aligned}\|A\|_\infty &= \max_j \sum_i |a_{i,j}|, \\ \|A\|_1 &= \max_i \sum_j |a_{i,j}|.\end{aligned}$$

For any matrix, define the *spectral radius* $\rho(A)$ by

$$\rho(A) = \max_s |\lambda_s|,$$

where λ_s are the eigenvalues of A . For the $\|\cdot\|_2$ norm we have the following identity

$$\|A\|_2 = \sqrt{\rho(AA^*)},$$

where A^* is the adjoint of A . For the special case when A is selfadjoint, $A = A^*$, we have

Property 1. *Let $A = A^*$. Then*

$$\|A\|_2 = \rho(A).$$

Proof. Let λ be an eigenvalue of A with associated eigenvector u . Then we have

$$A^2(u) = A(Au) = A(\lambda u) = \lambda A(u) = \lambda^2 u,$$

$$\Downarrow$$

$$\text{spectrum}(A^2) = \{\lambda^2 | \lambda \in \text{spectrum}(A)\},$$

$$\Downarrow$$

$$\begin{aligned}\|A\|_2 &= (\rho(AA^*))^{\frac{1}{2}} = (\rho(A^2))^{\frac{1}{2}} = (\max_s |\lambda_s|^2)^{\frac{1}{2}} \\ &= \max_s \sqrt{|\lambda_s|^2} = \max_s |\lambda_s| = \rho(A).\end{aligned}$$

□

Let us now consider the iteration

$$\mathbf{u}^{l+1} = M\mathbf{u}^l + \mathbf{c},$$

and let us use the $\|\cdot\|_2$ norm on \mathbb{R}^n with the associated norm $\|\cdot\|_2$ on $n \times n$ matrices. The iteration will then converge to \mathbf{u}^* in \mathbb{R}^n iff

$$\lim_{l \rightarrow \infty} \|\mathbf{u}^l - \mathbf{u}^*\|_2 = 0.$$

But

$$\begin{aligned}\|\mathbf{u}^l - \mathbf{u}^*\|_2 &= \|M\mathbf{u}^{l-1} + \mathbf{c} - (M\mathbf{u}^* + \mathbf{c})\|_2 \\ &= \|M(\mathbf{u}^{l-1} - \mathbf{u}^*)\|_2 \leq \|M\|_2 \|\mathbf{u}^{l-1} - \mathbf{u}^*\|_2.\end{aligned}$$

By induction we get

$$\|\mathbf{u}^l - \mathbf{u}^*\|_2 \leq (\|M\|_2)^l \|\mathbf{u}^0 - \mathbf{u}^*\|_2.$$

Let us assume that M is symmetric, then

$$\|\mathbf{u}^l - \mathbf{u}^*\|_2 \leq (\rho(M))^l \|\mathbf{u}^0 - \mathbf{u}^*\|_2.$$

This gives us the following sufficient condition for convergence.

Theorem 1. *Let M be symmetric and assume that for all eigenvalues λ for M we have*

$$|\lambda| < 1.$$

Then the iteration

$$\mathbf{u}^{l+1} = M\mathbf{u}^l + \mathbf{c},$$

converges to the unique solution of

$$(I - M)\mathbf{u} = \mathbf{c},$$

for all initial values \mathbf{u}^0 .

Note that even though all initial values converges to the solution, the number of iterations we need to get a predetermined accuracy *will* depend on the choice of \mathbf{u}^0 .

We are now ready to describe the two simplest iteration methods for solving $A\mathbf{u} = \mathbf{b}$.

5.4.2 Gauss-Jacobi

Any matrix can be written in the form

$$A = L + D + U,$$

where L is lower triangle, U is upper triangular and D is diagonal. Using this decomposition we have

$$A\mathbf{u} = \mathbf{b},$$

$$\Downarrow$$

$$L\mathbf{u} + D\mathbf{u} + U\mathbf{u} = \mathbf{b},$$

$$\Downarrow$$

$$D\mathbf{u} = -(L + U)\mathbf{u} + \mathbf{b}.$$

Assume that D is invertible (no zeroes on the diagonal). Then we get

$$\mathbf{u} = M_J\mathbf{u} + \mathbf{c},$$

where $M_J = -D^{-1}(L + U)$, $\mathbf{c} = D^{-1}\mathbf{b}$. This gives the Gauss-Jacobi iteration.

$$\mathbf{u}^{l+1} = M_J\mathbf{u}^l + \mathbf{c}.$$

In component form the Gauss-Jacobi iteration can be written in the form

$$u_i^{l+1} = \frac{1}{a_{ii}}(b_i - \sum_{j \neq i} a_{ij}u_j^l), \quad i = 1, \dots, n \quad .$$

Observe that

$$\begin{aligned} M_J &= -D^{-1}(L + U), \\ &\Downarrow \\ DM_J &= -L - U = -A + D, \\ &\Downarrow \\ M_J &= -D^{-1}A + I, \quad \text{and} \quad D^{-1} = -\frac{1}{4}I, \\ &\Downarrow \\ M_J &= \frac{1}{4}A + I. \end{aligned}$$

Let λ be an eigenvalue of A with corresponding eigenvector v . Then we have

$$M_J \mathbf{v} = \frac{1}{4}Av + Iv = (\frac{1}{4}\lambda + 1)v.$$

This means that

$$\text{spectrum}(M_J) = \left\{ \left(\frac{1}{4}\lambda + 1 \right) \mid \lambda \in \text{spectrum}(A) \right\}. \quad (55)$$

In order to find the spectrum of A , we will start from a property of the spectrum for block-matrices.

Let C be a matrix on the form

$$C = \begin{bmatrix} C_{11} & \cdots & \cdots & C_{1m} \\ \vdots & & & \vdots \\ C_{m1} & \cdots & \cdots & C_{mm} \end{bmatrix}.$$

where C_{ij} are $n \times n$ matrices. Thus C is a $(nm) \times (nm)$ matrix. Let us assume that there is a common set of eigenvectors for the matrices C_{ij} . Denote this set by $\{v^k\}_{k=1}^n$. Thus we have

$$C_{ij}v^k = \lambda_{ij}^k v^k,$$

where λ_{ij}^k is the eigenvalue of the matrix C_{ij} corresponding to the eigenvector v^k .

Let $\alpha_1, \dots, \alpha_m$ be real numbers. We will seek eigenvectors of C on the form

$$w^k = \begin{bmatrix} \alpha_1 v^k \\ \alpha_2 v^k \\ \vdots \\ \alpha_m v^k \end{bmatrix}.$$

Then w^k is an eigenvector of C with associated eigenvalue μ^k only if

$$Aw^k = \mu^k w^k,$$

$$\Updownarrow$$

$$\begin{bmatrix} (\lambda_{11}^k \alpha_1 + \dots + \lambda_{1m}^k \alpha_m) v^k \\ \vdots \\ (\lambda_{m1}^k \alpha_1 + \dots + \lambda_{mm}^k \alpha_m) v^k \end{bmatrix} = \begin{bmatrix} \mu^k \alpha_1 v^k \\ \vdots \\ \mu^k \alpha_m v^k \end{bmatrix}.$$

The eigenvalues of C are thus given by the eigenvalues of the matrix

$$\Lambda = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1m} \\ \vdots & & \vdots \\ \lambda_{m1} & \dots & \lambda_{mm} \end{bmatrix}.$$

Let us now return to the convergence problem for Gauss-Jacobi.

A is a $(N-1) \times (N-1)$ block matrix given by

$$A = \begin{bmatrix} B & I & 0 & \dots & \dots & 0 \\ I & B & I & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \dots & \dots & I & B \end{bmatrix},$$

where B is a $(N-1) \times (N-1)$ tri-diagonal matrix given by

$$B = \begin{bmatrix} -4 & 1 & 0 & \dots & \dots & 0 \\ 1 & -4 & 1 & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & \dots & \dots & 1 & -4 \end{bmatrix}.$$

B obviously commute with I and the zero matrix and they therefore has common eigenvectors. Since all vectors in \mathbb{R}^{N-1} are eigenvectors of the identity matrix and the zero matrix, with eigenvalues 1 and zero, respectively, in order to set up the Λ matrix we only need to find the eigenvalues of B .

Let D be a $n \times n$ tri-diagonal matrix of the form

$$D = \begin{bmatrix} a & b & 0 & \dots & \dots & \dots & 0 \\ b & a & b & 0 & \dots & \dots & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & \dots & \dots & \dots & b & a & b \\ 0 & \dots & \dots & \dots & 0 & b & a \end{bmatrix}.$$

Let λ be an eigenvalue with corresponding eigenvector $\mathbf{v} = (v_1, \dots, v_n)$ Thus

$$Dv = \lambda v,$$

$$\Updownarrow$$

$$bv_{i-1} + av_i + bv_{i+1} = \lambda v_i, \quad i = 1, \dots, n,$$

and $v_0 = v_{n+1} = 0$. This is a boundary-value problem for a difference equation of order 2. We write the equation on the form

$$bv_{i-1} + (a - \lambda)v_i + bv_{i+1} = 0,$$

and look for solutions of the form

$$v_i = r^i \Rightarrow v_{i-1} = r^{-1}v_i, \quad v_{i+1} = rv_i.$$

Thus we get the algebraic condition

$$br^2 + (a - \lambda)r + b = 0, \tag{56}$$

$$\Downarrow$$

$$r_{\pm} = \frac{1}{2b}(-(a - \lambda) \pm ((a - \lambda)^2 - 4b^2)^{\frac{1}{2}}).$$

If $r_+ \neq r_-$, the general solution to the difference equation is

$$v_i = Ar_+^i + Br_-^i,$$

the boundary conditions give

$$v_0 = 0, \quad \Leftrightarrow \quad A + B = 0,$$

$$v_{n+1} = 0, \quad \Leftrightarrow \quad Ar_+^{n+1} + Br_-^{n+1} = 0,$$

nontrivial solutions exist only if

$$r_-^{n+1} - r_+^{n+1} = 0,$$

$$\Updownarrow$$

$$\left(\frac{r_+}{r_-}\right)^{n+1} = 1. \tag{57}$$

But $r_+ r_- = 1 \Rightarrow r_- = \frac{1}{r_+}$ and therefore from equation (57)

$$(r_+^2)^{(n+1)} = 1,$$

$$\Updownarrow$$

$$r_+ = e^{\pi i \frac{k}{n+1}},$$

$$\Downarrow$$

$$r_- = e^{-\pi i \frac{k}{n+1}}.$$

Furthermore, from equation (56) we have

$$\frac{a - \lambda}{b} = -(r_+ + r_-),$$

$$\Updownarrow$$

$$\lambda = a + b(r_+ + r_-),$$

$$\Updownarrow$$

$$\lambda = a + 2b \cos(\pi \frac{k}{n+1}). \tag{58}$$

The eigenvalues of the tri-diagonal matrix D are thus

$$\lambda_k = a + 2b \cos\left(\frac{\pi k}{n+1}\right), \quad k = 1, \dots, n.$$

Returning to our analysis of the matrix A , we can now assemble our matrix Λ^k

$$\Lambda^k = \begin{bmatrix} \lambda_k & 1 & 0 & \dots & \dots & 0 \\ 1 & \lambda_k & 1 & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & \dots & \dots & 1 & \lambda_k \end{bmatrix},$$

where $\lambda_k = -4 + 2 \cos\left(\frac{\pi k}{N}\right)$, $k = 1, \dots, N-1$. Λ^k is also tri-diagonal and the previous theory for tri-diagonal matrices give us the eigenvalues

$$\lambda_{kl} = \lambda_k + 2 \cos\left(\frac{\pi l}{N}\right), \quad l = 1, \dots, N-1,$$

\Downarrow

$$\lambda_{kl} = -4 + 2 \cos\left(\frac{\pi k}{N}\right) + 2 \cos\left(\frac{\pi l}{N}\right),$$

or using standard trigonometric identities

$$\lambda_{kl} = -4 \left(\sin^2\left(\frac{\pi k}{2N}\right) + \sin^2\left(\frac{\pi l}{2N}\right) \right), \quad k = 1, \dots, N-1, \quad l = 1, \dots, N-1,$$

and finally the spectrum of M_J in the Gauss-Jacobi iteration method is from (55)

$$\mu_{kl} = 1 - \sin^2\left(\frac{\pi k}{2N}\right) - \sin^2\left(\frac{\pi l}{2N}\right).$$

Using Theorem 1, the stability condition for Gauss-Jacobi is

$$|\mu_{kl}| < 1,$$

\Updownarrow

$$\left| 1 - \sin^2\left(\frac{1}{2}\pi x_k\right) - \sin^2\left(\frac{1}{2}\pi y_l\right) \right| < 1, \quad x_k = \frac{k}{N}, \quad y_l = \frac{l}{N}.$$

Since both $x_k, y_l \in (0, 1)$, it is evident that the condition holds and that we have convergence. We get the largest eigenvalue by choosing $k = l = 1$. The spectral radius can then for large N be approximated as

$$\rho(M) = 1 - 2 \sin^2\left(\frac{1}{2}\pi h\right) \approx 1 - \frac{1}{2}(\pi h)^2,$$

where we have recall that $h = \frac{1}{N}$ grid parameter for the grid defined in figure (19). The relative error is defined as

$$\frac{\|\mathbf{u}^l - \mathbf{u}^*\|}{\|\mathbf{u}^0 - \mathbf{u}^*\|} \leq (\rho(M))^l \approx \left(1 - \frac{1}{2}\pi^2 h^2\right)^l.$$

The discretization error of our scheme is of order h^2 . It is natural to ask how many iterations we need in order for the relative error to be of this size. There is no point in iterating more than this.

$$\left(1 - \frac{1}{2}\pi^2 h^2\right)^l \approx h^2,$$

$$\Updownarrow$$

$$e^{-\frac{1}{2}l\pi^2 h^2} \approx h^2,$$

$$\Updownarrow$$

$$l \approx \frac{4}{\pi^2} N^2 \ln N.$$

Gauss elimination requires in general $O(N^3)$ and since for large N we have $N^2 \ln N \ll N^3$ we observe that Gauss-Jacobi iteration is much faster than Gauss elimination.

5.4.3 Gauss-Seidel

We use the same decomposition of A as for the Gauss-Jacobi method, $A = L + D + U$

$$(L + D + U)\mathbf{u} = \mathbf{b},$$

$$\Downarrow$$

$$-(L + D)\mathbf{u} = U\mathbf{u} - \mathbf{b},$$

$$\Downarrow$$

$$\mathbf{u} = M_S \mathbf{u} + \mathbf{c},$$

where $M_S = -(L + D)^{-1}U$, $\mathbf{c} = (L + D)^{-1}\mathbf{b}$. In component form we have the iteration

$$u_i^{l+1} = \frac{1}{a_{ii}}(b_i - \sum_{j<i} a_{ij}u_j^{l+1} - \sum_{j>i} a_{ij}u_j^l), \quad i = 1, \dots, n.$$

We will now see that Gauss-Seidel converges faster than Gauss-Jacobi. It also use the memory in a more effective way. For the stability of this method we must find the spectrum of the matrix

$$M_S = -(L + D^{-1})U.$$

Note that the iteration matrix for the Gauss-Seidel method can be rewritten into the form

$$M_S = -(I + D^{-1}L)^{-1}D^{-1}U.$$

Using straight forward matrix manipulations it is now easy to verify that λ_s is an eigenvalue to the matrix M_S iff

$$\det(\lambda_s D + \lambda_s L + U) = 0. \quad (59)$$

Let now α be any number, and define a matrix E by

$$E = \begin{bmatrix} I & 0 & \dots & \dots & 0 \\ 0 & \alpha I & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & \dots & \alpha^{N-2}I \end{bmatrix}.$$

Doing the matrix multiplication, it is easy to verify that we have the identity

$$E(D + L + U)E^{-1} = D + \alpha L + \frac{1}{\alpha}U,$$

from which we immediately can conclude

$$\det\left(D + \alpha L + \frac{1}{\alpha}U\right) = \det(D + L + U). \quad (60)$$

This derivation only require D to be a diagonal matrix, L to be a lower triangular matrix and U to be an upper triangular matrix. Using (59) and (60), we have

$$\begin{aligned} 0 &= \det(\lambda_s I - M_S) \\ &= \det(\lambda_s D + \lambda_s L + U) \\ &= \det\left(\lambda_s D + \alpha \lambda_s L + \frac{1}{\alpha}U\right) \\ &= \det\left(\lambda_s D + \sqrt{\lambda_s}(L + U)\right), \quad \alpha = \frac{1}{\sqrt{\lambda_s}} \\ &= \lambda_s^{\frac{(N-1)}{2}} \det(\mu D + L + U), \quad \mu = \sqrt{\lambda_s} \\ &= \lambda_s^{\frac{(N-1)}{2}} \det(D) \det(\mu + D^{-1}(L + U)) \\ &= \lambda_s^{\frac{(N-1)}{2}} \det(D) \det(\mu I - M_J). \end{aligned}$$

Thus, we have proved that the nonzero eigenvalues λ_s of the Gauss-Seidel iteration matrix are determined by the eigenvalues μ of the Jacobi iteration matrix through the formula

$$\lambda_s = \mu^2 \quad (61)$$

Since all eigenvalues of the Jacobi iteration matrix M_J are inside the unit circle, this formula implies that the same is true for the eigenvalues of the Gauss-Seidel iteration matrix M_S . Thus Gauss-Seidel also converges for all initial

conditions. As for the *rate* of convergence, using the results from the analysis of the convergence of the Gauss-Jacobi method, we find that

$$\rho(M_S) \approx \left(1 - \frac{1}{2}\pi^2 h^2\right)^2 \approx 1 - \pi^2 h^2,$$

and therefore

$$\frac{\|\mathbf{u}^l - \mathbf{u}^*\|}{\|\mathbf{u}^0 - \mathbf{u}^*\|} = h^2 \Leftrightarrow (1 - \pi^2 h^2)^l = h^2 \Leftrightarrow l \approx \frac{2}{\pi^2} N^2 \ln N.$$

For a fixed accuracy the Gauss-Seidel only needs half as many iterations as Gauss-Jacobi.

These two examples have given a taste of what one needs to do in order to define iteration methods for linear systems and study their convergence.

As previously mentioned there is an extensive set of iteration methods available for solving large linear systems of equation, and many has been implemented in numerical libraries available for most common programming languages. The choice of method depends to a large extent on the structure of your matrix. Here are two common choices:

Matrix property	Method name
Sparse, symmetric, positive:	Conjugate gradient
Dense, nonsymmetric:	Generalized minimal residual method CGMRE

The conjugate gradient method has also been extended to more general matrices. A standard reference for iterative methods is “Iterative methods for sparse linear systems” by Yoosef Saad.

6 First order PDEs

So far we have only discussed numerical methods for solving PDEs. It is now time to introduce the so-called *analytical methods*. The aim here is to produce some kind of formula(s) for the solution of PDEs. What counts as a formula is not very precisely defined, but in general it involves infinite series and integrals that can not be solved by quadrature. Thus, if the ultimate aim is to get a numerical answer, and it usually is, numerical methods must be used to approximate the said infinite series and integrals. There is thus no clear-cut line separating analytical and numerical methods.

We started our exposition of the finite difference method by applying it to second order equations. This is because the simplest finite difference methods work best for such equations.

For our exposition of analytical methods we start with first order PDEs because the simplest such methods work best for these types of equations.

6.1 Higher order equations and first order systems

It turns out that higher order PDEs and first order systems of PDEs are not really separate things. Usually one can be converted into the other. Whether this is useful or not depends on the context.

Let us recall the analog situation from the theory of ODEs. The ODE for the Harmonic Oscillator is

$$x'' + w^2 x = 0.$$

Define two new functions by

$$\begin{aligned} x_1(t) &= x(t), \\ x_2(t) &= x'(t). \end{aligned}$$

Then

$$\begin{aligned} x_1' &= x' = x_2, \\ x_2' &= x'' = -w^2 x = -w^2 x_1. \end{aligned}$$

Thus we get a first order system of ODEs.

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -w^2 x_1. \end{aligned}$$

In the theory of ODEs this is a very useful device that can be used to transform any ODE of order n into a first order system of ODEs in \mathbb{R}^n .

Note that the transition from one high order ODE, to a system of first order ODEs is not by any means unique. If we for the Harmonic Oscillator rather define

$$\begin{aligned} x_1 &= x + x', \\ x_2 &= x - x', \end{aligned}$$

we get the first order system

$$\begin{aligned} x_1' &= \alpha x_1 - \beta x_2, \\ x_2' &= \beta x_1 - \alpha x_2, \quad \alpha = \frac{1}{2}(1 - w^2), \beta = \frac{1}{2}(1 + w^2). \end{aligned}$$

Sometimes it is also possible to convert a first order system into a higher order scalar equation. This is however not possible in general but can always be done for linear first order systems of both ODEs and PDEs.

Let us consider the wave equation

$$u_{tt} - \gamma^2 u_{xx} = 0.$$

Observe that

$$\begin{aligned} (\partial_t + \gamma \partial_x)(\partial_t - \gamma \partial_x)u &= (\partial_t + \gamma \partial_x)(u_t - \gamma u_x) \\ &= (\partial_t + \gamma \partial_x)u_t - (\partial_t + \gamma \partial_x)\gamma u_x \\ &= u_{tt} + \gamma u_{tx} - \gamma u_{xt} - \gamma^2 u_{xx} \\ &= u_{tt} - \gamma^2 u_{xx} = 0. \end{aligned}$$

Inspired by this let us define a function

$$v = (\partial_t - \gamma \partial_x)u.$$

Then the wave equation implies

$$(\partial_t + \gamma \partial_x)v = 0.$$

Thus the wave equation can be converted into the system

$$\begin{aligned} u_t - \gamma u_x &= v, \\ v_t + \gamma v_x &= 0. \end{aligned}$$

Observe that the second equation is decoupled from the first. We will later in this class see that this can be used to solve the wave equation.

Writing a higher order PDE as a first order system is not always useful. Observe that

$$(\partial_x + i\partial_y)(\partial_x - i\partial_y)u = \partial_{xx}u + \partial_{yy}u.$$

Thus we might write the Laplace equation as the system

$$\begin{aligned} u_x - iu_y &= v, \\ v_x + iv_y &= 0. \end{aligned}$$

However this is not a very useful system since we most of the time are looking for *real* solutions to the Laplace equation.

If we rather put

$$\begin{aligned} v &= u_x, \\ w &= u_y. \end{aligned}$$

Then using the Laplace equation and assuming equality of mixed partial derivatives we have

$$\begin{aligned} v_y &= u_{xy} = u_{yx} = w_x, \\ v_x &= u_{xx} = -u_{yy} = -w_y. \end{aligned}$$

Thus we get the first order system

$$\begin{aligned} v_y - w_x &= 0, \\ v_x + w_y &= 0. \end{aligned} \tag{62}$$

These are the famous *Cauchy-Riemann equations*. If you have studied complex function theory you know that (62) is the key to the whole subject; they determine the complex differentiability of the function

$$f(z) = v(x, y) + iw(x, y), \quad z = x + iy.$$

(62) is a coupled system and does not on its own make it any simpler to find solutions to the Laplace equation. However, by using the link between complex

function theory and the Laplace equation, great insight into its set of solutions can be gained.

Going from first order systems to higher order equations is not always possible but sometimes *very* useful.

A famous case is the Maxwell equations for electrodynamics, which in a region containing no sources of currents and charge are

$$\begin{aligned}\nabla \times \mathbf{E} + \partial_t \mathbf{B} &= 0, \\ \nabla \times \mathbf{B} - \mu_0 \varepsilon_0 \partial_t \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{B} &= 0,\end{aligned}\tag{63}$$

where $\mathbf{E} = (E_1, E_2, E_3)$ and $\mathbf{B} = (B_1, B_2, B_3)$ are the electric and magnetic fields. We can write this first order system for \mathbf{E} and \mathbf{B} as a second order system for \mathbf{E} alone. We have, using (63), that

$$\begin{aligned}\mu_0 \varepsilon_0 \partial_t \mathbf{E} &= \nabla \times \mathbf{B}, \\ \Downarrow \\ \mu_0 \varepsilon_0 \partial_{tt} \mathbf{E} &= \nabla \times \partial_t \mathbf{B} = \nabla \times (-\nabla \times \mathbf{E}), \\ \Downarrow \\ \mu_0 \varepsilon_0 \partial_{tt} \mathbf{E} + \nabla \times \nabla \times \mathbf{E} &= 0,\end{aligned}$$

and

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{E}) &= \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E}, \\ \Downarrow \\ \varepsilon_0 \mu_0 \partial_{tt} \mathbf{E} - \nabla^2 \mathbf{E} &= 0.\end{aligned}$$

This is the 3D wave equation and it predicts, as we will see later in this course, that there are solutions describing electromagnetic disturbances traveling at a speed $v = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$ in source free regions (for example in a vacuum). The parameters ε_0, μ_0 are constants of nature that was known to Maxwell, and when he calculated v , he found that the speed of these disturbances was very close to the known speed of light.

He thus conjectured that light is an electromagnetic disturbance. This was a monumental discovery whose ramifications we are still exploring today.

6.2 First order linear scalar PDEs

A general first order linear scalar PDE with two independent variables, (x, t) , is of the form

$$a \partial_x v + b \partial_t v = cv + d,\tag{64}$$

where in general, a, b, c and d depends on x and t . The solution method for (64) is based on a geometrical interpretation of the equation.

Observe that the two functions $a(x, t)$ and $b(x, t)$ can be thought of as the two components of a vector field $\mathbf{f}(x, t)$ in the plane.

$$\mathbf{f}(x, t) = (a(x, t), b(x, t)).$$

Let $\gamma(s) = (x(s), t(s))$ be a parametrized curve in the plane whose velocity at each point (x, t) is equal to $\mathbf{f}(x, t)$.

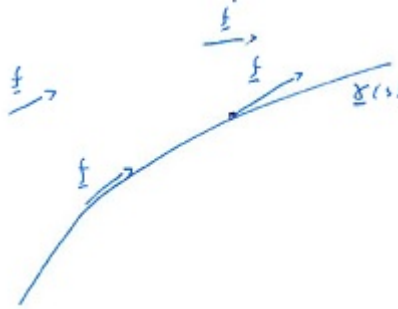


Figure 22

Thus we have

$$\begin{aligned} \frac{d\gamma}{ds} &= \mathbf{f}(\gamma(s)), \\ \Downarrow \\ \frac{dx}{ds} &= a, \\ \frac{dt}{ds} &= b. \end{aligned} \tag{65}$$

Let $v(x, t)$ be a solution of (64) and define

$$v(s) = v(\gamma(s)) = v(x(s), t(s)).$$

Then

$$\begin{aligned} \frac{dv}{ds} &= \partial_x v \frac{dx}{ds} + \partial_t v \frac{dt}{ds} \\ &= a \partial_x v + b \partial_t v = cv + d. \end{aligned}$$

Thus $v(s)$ satisfy an ODE!

$$\frac{dv(s)}{ds} = c(x(s), t(s))v(s) + d(x(s), t(s)). \tag{66}$$

From (65) and (66) we get a coupled system of ODEs

$$\begin{aligned} \frac{dx}{ds} &= a, \\ \frac{dt}{ds} &= b, \\ \frac{dv}{ds} &= cv + d. \end{aligned} \tag{67}$$

The general theory of ODEs tells us that through each point (x_0, t_0, v_0) , there goes exactly one solution curve of (67) if certain regularity conditions on a, b, c and d are satisfied. The family of curves that solves (67) are called *the charac-*

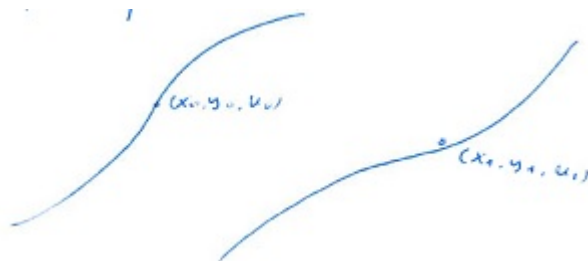


Figure 23: Characteristic curves for the PDE (64).

teristic curves for the PDE (64).

Observe that the system (67) is partly decoupled, the two first equations can be solved independently of the third. The plane curves solving

$$\begin{aligned}\frac{dx}{ds} &= a, \\ \frac{dt}{ds} &= b,\end{aligned}$$

are called *characteristic base curves*. They are also, somewhat imprecisely, called characteristic curves.

Most of the time we are not interested in the general solution of (64), but rather in special solutions that in addition to (64) has given values on some curve C in the (x, t) plane. Any such problem is called an initial value problem

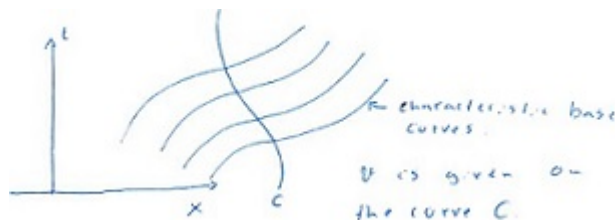


Figure 24: Characteristic base curves for the PDE (64).

for our equation (64).

This is a generalization of the notion of initial value problem for a PDE from what we have discussed earlier. This earlier notion consisted of giving the solution at $t = 0$ (or $t = t_0$)

$$v(x, 0) = \varphi(x),$$

where $\varphi(x)$ is some given function of x .

Let $(x(\tau), t(\tau))$ be a parametrization of C . Then our generalized notion of initial value problem for (64) consists of imposing the condition

$$v(x(\tau), t(\tau)) = \varphi(\tau),$$

where $\varphi(\tau)$ is a given function. Our previous notion of initial value problem corresponds to the choice.

$$(x(\tau), t(\tau)) = (\tau, 0).$$

Since this is a parametrization of the x -axis in the (x, t) plane. Thus in summary, the (generalized) initial value problem for (64) consists of finding functions

$$\begin{aligned} x &= x(s, \tau), \\ t &= t(s, \tau), \\ v &= v(s, \tau), \end{aligned} \tag{68}$$

such that

$$\begin{aligned} \frac{dx}{ds} &= a, & x(0, \tau) &= x(\tau), \\ \frac{dt}{ds} &= b, & t(0, \tau) &= t(\tau), \\ \frac{dv}{ds} &= cv + d, & v(0, \tau) &= \varphi(\tau). \end{aligned} \tag{69}$$

Where $(x(\tau), t(\tau))$ parametrize the initial curve and where $\varphi(\tau)$ are the given values of v on the initial curve. The reader might at this point object and ask how we can possibly consider (68), solving (69), to be a solution to our original problem (64). Where is my $v(x, t)$ you might ask? This is after all what we started out to find.

In fact, the function $v(x, t)$, solving our equation (64), is implicitly contained in (68).

We can in principle use the two first equations in (68) to express s and τ as functions of x and t .

$$\begin{aligned} x &= x(s, \tau), \\ t &= t(s, \tau), \\ &\Downarrow \\ s &= s(x, t), \\ \tau &= \tau(x, t). \end{aligned} \tag{70}$$

It is known that this can be done locally close to a point (s_0, τ_0) if the determinant of the Jacobian matrix at (s_0, τ_0) , this is what is known as the *Jacobian* in calculus, is nonzero.

$$(\partial_s x \partial_\tau t - \partial_s t \partial_\tau x)|_{(s_0, \tau_0)} \neq 0.$$

Thus if the Jacobian is nonzero for all points on the initial curve the initial value problem for (64) has a unique solution for any specification of φ on the initial curve. This solution is

$$v(x, t) = v(s(x, t), \tau(x, t)).$$

Actually finding a formula for the solution $v(x, t)$ using this method, which is called *the method of characteristics*, can be hard. This is because in general it is hard to find explicit solutions to the coupled ODEs (69) and even if you succeed with this, inverting the system (70) analytically might easily be impossible.

However, even if we can not find an explicit solution, the method of characteristics can still give great insight into what the PDE is telling us. Additionally, the method of characteristic is the starting point for a numerical approach to solving PDEs of the type (64) and of related types as well.

In our traffic model, we encountered the equation

$$\partial_t n + c_0 \partial_x n = 0. \quad (71)$$

This equation is of the type (64) and can be solved using the method of characteristic. Let the initial data be given on the x-axis.

$$h(x, 0) = \varphi(x).$$

Our system of ODEs determining the base characteristic curves for the traffic equation (71) is thus

$$\begin{aligned} \frac{dt}{ds} &= 1, & t(0, \tau) &= 0, \\ \frac{dx}{ds} &= c_0, & x(0, \tau) &= \tau, \\ \frac{dn}{ds} &= 0, & n(0, \tau) &= \varphi(\tau). \end{aligned}$$

Observe that the last ODE tell us that the function n is constant along the base characteristic curves.

$$\begin{aligned} \frac{dt}{ds} = 1 &\Rightarrow t = s + t_0, \quad t(0, \tau) = 0 \Rightarrow t_0 = 0 \Rightarrow t(x, \tau) = s, \\ \frac{dx}{ds} = c_0 &\Rightarrow x = c_0 s + x_0, \quad x(0, \tau) = \tau \Rightarrow x_0 = \tau \\ &\Rightarrow x(s, \tau) = c_0 s + \tau, \\ \frac{dn}{ds} = 0 &\Rightarrow n = n_0, \quad n(0, \tau) = \varphi(\tau) \Rightarrow n_0 = \varphi(\tau) \\ &\Rightarrow n(s, \tau) = \varphi(\tau). \end{aligned}$$

Now we must express s and τ in terms of x and t

$$\begin{aligned} t = s &\Rightarrow s = t, \\ x = c_0 s + \tau &\Rightarrow \tau = x - c_0 s = x - c_0 t, \end{aligned}$$

and our solution to the initial value problem is

$$n(x, t) = \varphi(\tau) = \varphi(x - c_0 t).$$

This is the unique solution to our initial value problem. The base characteristic curves are given in parametrized form as

$$\begin{aligned} t &= s, \\ x &= c_0 s + \tau, \quad \tau \in \mathbb{R}, \end{aligned}$$

or in unparametrized form as

$$x - c_0 t = \tau.$$

In either case they are a family of straight lines that is parametrized by τ .

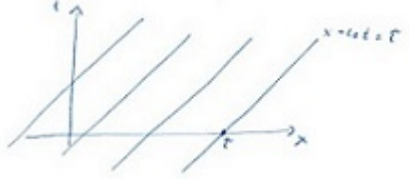


Figure 25: Base characteristic curves for the traffic equation (71).

The solution in general depends on both the choice of initial curve and which values we specify for the solution on this curve. In fact, if we are not careful here, there might be many solutions to the initial value problem or none at all.

Since n is constant along a base characteristic curve, its value is determined by the initial value on the point of the initial curve where the base characteristic curve and the initial curve coincide. If this occurs at more than one point there is in general no solution to the initial value problem.

Let for example C be given through a parametrization

$$x(\tau) = \tau, \tag{72}$$

$$t(\tau) = \tau(1 - \tau). \tag{73}$$

The initial values is some function $\varphi(\tau)$ defined on C . The initial curve is thus a parabola and the base characteristic curves form a family of parallel lines

$$x - c_0 t = \text{const}$$

Since $x - c_0 t = A$ intersect the initial curve at two points corresponding to $x = \tau_1$ and $x = \tau_2$ and since $n = \varphi(\tau)$ is constant on the base characteristic curves, we must have

$$\varphi(\tau_1) = \varphi(\tau_2).$$

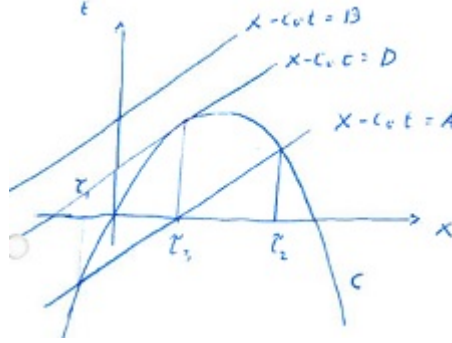


Figure 26: Base characteristic curves and parabolic initial curve for the traffic equation (71).

This only holds for very special initial data. At a certain point $x = \tau_3$, $t = \tau_3(1 - \tau_3)$, the characteristic base curve is tangent to the initial curve.

At this point the equations

$$\begin{aligned} x &= x(s, \tau), \\ t &= t(x, \tau), \end{aligned}$$

can not be inverted, even locally.

In order to see this, we observe that the base characteristic curves for the traffic equation (71), with initial data prescribed on the parabola (72), are given by

$$\begin{aligned} x &= c_0 s + \tau, \\ t &= s + \tau(1 - \tau). \end{aligned}$$

The Jacobian is

$$\Delta = \partial_s x \partial_\tau t - \partial_s t \partial_\tau x = c_0(1 - 2\tau) - 1,$$

and at the point of tangency we have

$$\frac{1}{c_0} = \frac{d}{d\tau}(\tau(1 - \tau)) = (1 - 2\tau),$$

$$\Downarrow$$

$$c_0(1 - 2\tau) = 1,$$

$$\Downarrow$$

$$\Delta = 0.$$

A point where the base characteristic curves and the initial curve are tangent is called a *characteristic point*. The initial curve is said to be *characteristic* at this point. Thus the initial curve C is characteristic at the point (x, t) where

$$\begin{aligned}x &= \frac{1}{2}\left(1 - \frac{1}{c_0}\right), \\t &= \frac{1}{4}\left(1 - \frac{1}{c_0^2}\right).\end{aligned}$$

If the initial curve is characteristic at all of its points, we say that the initial curve is characteristic. It in fact *is* a base characteristic curve. The initial value problem then in general has no solution, and if it does have a solution, the solution is not unique. A characteristic initial value problem for our traffic equation (71) would be

$$\begin{aligned}\partial_t n + c_0 \partial_x n &= 0, \\n(c_0 t, t) &= \varphi(t).\end{aligned}$$

General solution of the equation is

$$n(x, t) = f(x - c_0 t). \quad (74)$$

The initial values require that

$$\begin{aligned}n(c_0 t, t) &= \varphi(t), \\&\Downarrow \\f(0) &= \varphi(t), \quad \forall t.\end{aligned}$$

In order for a solution to exist, $\varphi(t) = A$ must be a constant. And if $\varphi(t) = A$ then (74), with *any* function $f(y)$ where $f(0) = A$, would solve the initial value problem. Thus we have nonuniqueness.

Let us next consider the initial value problem

$$\begin{aligned}t \partial_t v + \partial_x v &= 0, \\v(x, 1) &= \varphi(x).\end{aligned} \quad (75)$$

The initial curve is the horizontal line $t = 1$ which can be parametrized using

$$\begin{aligned}x(\tau) &= \tau, \\t(\tau) &= 1.\end{aligned}$$

Applying the method of characteristics, we get the following system of ODEs

$$\begin{aligned}\frac{dt}{ds} &= t, \quad t(0, \tau) = 1, \\ \frac{dx}{ds} &= 1, \quad x(0, \tau) = \tau, \\ \frac{dv}{ds} &= 0, \quad v(0, \tau) = \varphi(\tau),\end{aligned}$$

and

$$\begin{aligned}\frac{dt}{ds} &= t \Rightarrow t = t_0 e^s, \quad t(0, \tau) = 1 \Rightarrow t_0 = 1 \Rightarrow t = e^s, \\ \frac{dx}{ds} &= 1 \Rightarrow x = s + x_0, \quad x(0, \tau) = \tau \Rightarrow x_0 = \tau \Rightarrow x = s + \tau, \\ \frac{dv}{ds} &= 0 \Rightarrow v = v(\tau), \quad v(0, \tau) = \varphi(\tau) \Rightarrow v = \varphi(\tau).\end{aligned}$$

We now solve for s and τ in terms of x and t

$$\begin{aligned}t &= e^s, \\ x &= s + \tau.\end{aligned}$$

The first equation only have solutions for $t > 0$

$$\begin{aligned}s &= \ln t, \\ \tau &= x - s = x - \ln t, \\ \Downarrow \\ v(x, t) &= \varphi(\tau) = \varphi(x - \ln t).\end{aligned}$$

This solution is constant along the base characteristic curves

$$\begin{aligned}\tau &= x - \ln t, \\ \Updownarrow \\ t &= e^{x-\tau} = e^{-\tau} e^x, \quad \tau \in \mathbb{R}.\end{aligned}$$

The solution we have found exists only for $t > 0$. This is its domain of definition

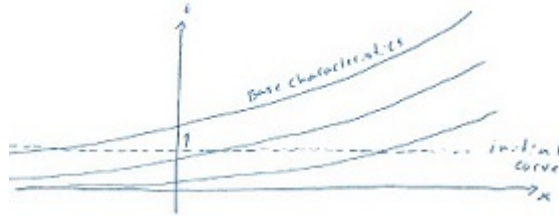


Figure 27: Base characteristic curves and initial curve for the initial value problem (75).

As a final example, let us consider the initial value problem for the wave equation

$$\begin{aligned}v_{tt} - \gamma^2 v_{xx} &= 0, \\ v(x, 0) &= f(x), \\ v_t(x, 0) &= g(x).\end{aligned}\tag{76}$$

We have previously seen that this can be written as a first order system

$$v_t - \gamma v_x = u, \quad (77)$$

$$u_t + \gamma u_x = 0. \quad (78)$$

The initial conditions are

$$\begin{aligned} v(x, 0) &= f(x), \\ u(x, 0) &= (v_t - \gamma v_x)(x, 0) = g(x) - \gamma f'(x). \end{aligned}$$

We start by solving the homogeneous equation (78). We will use the method of characteristics on the problem

$$u_t + \gamma u_x = 0,$$

$$u(x, 0) = g - \gamma f'.$$

The corresponding system of ODEs is

$$\begin{aligned} \frac{dx}{ds} &= \gamma, & x(0, \tau) &= \tau, \\ \frac{dt}{ds} &= 1, & t(0, \tau) &= 0, \\ \frac{du}{ds} &= 0, & u(0, \tau) &= g(\tau) - \gamma f'(\tau). \end{aligned}$$

Solving we find

$$\begin{aligned} \frac{dx}{ds} = \gamma &\Rightarrow x = \gamma s + x_0, \quad x(0, \tau) = \tau \Rightarrow x_0 = \tau \Rightarrow x = \gamma s + \tau, \\ \frac{dt}{ds} = 1 &\Rightarrow t = s + t_0, \quad t(0, \tau) = 0 \Rightarrow t_0 = 0 \Rightarrow t = s, \\ \frac{du}{ds} = 0 &\Rightarrow u = u_0, \quad u(0, \tau) = g(\tau) - \gamma f'(\tau) \Rightarrow u_0 = g(\tau) - \gamma f'(\tau) \\ &\Rightarrow u(s, \tau) = g(\tau) - \gamma f'(\tau). \end{aligned}$$

We now express τ and s as function of x and t

$$\begin{aligned} x &= \gamma s + \tau, & \tau &= x - \gamma t, \\ t &= s, & s &= t, \end{aligned} \quad \Leftrightarrow$$

from which it follows that the function $u(x, t) = u(s(x, t), \tau(x, t))$ is given by

$$u(x, t) = g(x - \gamma t) - \gamma f'(x - \gamma t).$$

Let us next apply the method of characteristic to the initial value problem for equation (77). This problem is

$$v_t - \gamma v_x = g(x - \gamma t) - \gamma f'(x - \gamma t),$$

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

The corresponding system of ODEs is

$$\begin{aligned} \frac{dx}{ds} &= -\gamma, & x(0, \tau) &= \tau, \\ \frac{dt}{ds} &= 1, & t(0, \tau) &= 0, \\ \frac{dv}{ds} &= g(x(s) - \gamma t(s)) - \gamma f'(x(s) - \gamma t(s)), & v(\tau, 0) &= f(\tau). \end{aligned}$$

The first two equations give

$$\begin{aligned} x &= -\gamma s + \tau, \\ t &= s, \end{aligned} \Leftrightarrow \begin{aligned} \tau &= x + \gamma t, \\ s &= t. \end{aligned}$$

The last equation is solved by

$$v(s, \tau) = \int_0^s d\sigma (g(x(\sigma) - \gamma t(\sigma)) - \gamma f'(x(\sigma) - \gamma t(\sigma))) + v_0(\tau),$$

and using the initial data we have

$$v(0, \tau) = f(\tau) \Rightarrow v_0(\tau) = f(\tau),$$

\Downarrow

$$v(s, \tau) = f(\tau) + \int_0^s d\sigma (g(x(\sigma) - \gamma t(\sigma)) - \gamma f'(x(\sigma) - \gamma t(\sigma))).$$

Observing that

$$x - \gamma t = -\gamma s + \tau - \gamma s = -2\gamma s + \tau,$$

the formula for v become

$$v(s, \tau) = f(\tau) + \int_0^s d\sigma (g(-2\gamma\sigma + \tau) - \gamma f'(-2\gamma\sigma + \tau)).$$

Changing variable in the integral using $y = -2\gamma\sigma + \tau$, we get

$$v(s, \tau) = f(\tau) - \frac{1}{2\gamma} \int_{\tau}^{-2\gamma s + \tau} dy (g(y) - \gamma f'(y)),$$

and $s = t$, $\tau = x + \gamma t$ from which it follows that

$$-2\gamma s + \tau = -2\gamma t + x + \gamma t = x - \gamma t.$$

Therefore, the function $v(x, t) = v(s(x, t), \tau(x, t))$ is given by the formula

$$\begin{aligned} v(x, t) &= f(x + \gamma t) - \frac{1}{2\gamma} \int_{x+\gamma t}^{x-\gamma t} dy (g(y) - \gamma f'(y)) \\ &= f(x + \gamma t) + \frac{1}{2} \int_{x+\gamma t}^{x-\gamma t} dy f'(y) + \frac{1}{2\gamma} \int_{x-\gamma t}^{x+\gamma t} dy g(y), \end{aligned}$$

\Downarrow

$$v(x, t) = \frac{1}{2}(f(x + \gamma t) + f(x - \gamma t)) + \frac{1}{2\gamma} \int_{x-\gamma t}^{x+\gamma t} dy g(y).$$

This is the famous D'Alembert formula for the solution of the initial value problem for the wave equation on an infinite domain. It was first derived by the French mathematician D'Alembert in 1747.

Let $G(\xi) = \int g(y)dy$. Then the formula can be rewritten as

$$v(x, t) = \left[\frac{1}{2}f(x + \gamma t) + \frac{1}{2}G(x + \gamma t) \right] + \left[\frac{1}{2}f(x - \gamma t) - \frac{1}{2\gamma}G(x - \gamma t) \right].$$

The solution is clearly a linear superposition of a left moving and a right moving wave.

6.3 First order Quasilinear scalar PDEs

A first order, scalar, quasilinear PDE is an equation of the general form

$$a\partial_x u + b\partial_t u = c, \tag{79}$$

where

$$\begin{aligned} a &= a(x, t, u), \\ b &= b(x, t, u), \\ c &= c(x, t, u). \end{aligned}$$

This equation can also be solved through a geometrical interpretation of the equation.

Let $u(x, t)$ be a solution of (79). The *graph* of u is a surface in (x, t, u) space defined by the equation.

$$F(x, t, u) \equiv u(x, t) - u = 0.$$

This surface is called an *integral surface* for the equation.

From calculus we know that the gradient of F

$$\nabla F = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial t}, -1 \right),$$

is normal to the integral surface F . Observe that the equation (79) can be rewritten as

$$a\partial_x u + b\partial_t u - c = 0,$$

\Updownarrow

$$(a, b, c) \cdot \nabla F(x, t, u) = 0.$$

Thus the vector (a, b, c) must always be in the tangent plane to the integral surface corresponding to a solution $u(x, t)$.

At each point (x, t, u) the vector (a, b, c) determines a direction that is the *characteristic direction* at this point. Thus the vector (a, b, c) determines a *characteristic direction field* in (x, t, u) space.

The characteristic curves are the curves that has (a, b, c) as velocity at each of its points. Thus, if the curve is parametrized by $(x(s), t(s), u(s))$, we must have

$$\begin{aligned}\frac{dx}{ds} &= a(x, t, u), \\ \frac{dt}{ds} &= b(x, t, u), \\ \frac{du}{ds} &= c(x, t, u).\end{aligned}$$

Observe that for this quasilinear case the two first equations do not decouple and the 3 ODEs has to be solved simultaneously. This can obviously be challenging.

For sufficiently smooth a, b and c , there is a solution curve, or *characteristic curve*, passing through any point (x_0, t_0, u_0) in (x, t, u) space.

The initial value problem is specified by giving the value of u on a curve in (x, t) space. This curve together with the initial data determines a curve, C , in (x, t, u) space that we call *the initial curve*. To solve the initial value problem we pass a characteristic curve through every point of the initial curve. If these curves generate a smooth surface, the function, whose graph is equal to the surface, is the solution of the initial value problem.

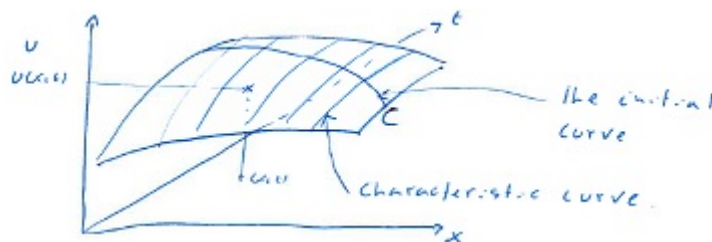


Figure 28

Let the initial curve C be parametrized by

$$(x(\tau), t(\tau), \varphi(\tau)).$$

The characteristic system of ODEs with initial conditions imposed is

$$\begin{aligned}\frac{dx}{ds} &= a, \quad x(0, \tau) = x(\tau), \\ \frac{dt}{ds} &= b, \quad t(0, \tau) = t(\tau), \\ \frac{du}{ds} &= c, \quad u(0, \tau) = \varphi(\tau).\end{aligned}$$

The theory of ODEs gives a unique solution

$$\begin{aligned}x &= x(s, \tau), \\t &= t(s, \tau), \\u &= \varphi(s, \tau).\end{aligned}\tag{80}$$

Observe that

$$\partial_s x = a, \quad \partial_s t = b,$$

so the Jacobian is

$$\Delta(s, \tau) = \begin{vmatrix} \partial_s x & \partial_\tau x \\ \partial_s t & \partial_\tau t \end{vmatrix} = (a\partial_\tau t - b\partial_\tau x)(s, \tau).$$

Let us assume that the Jacobian is nonzero along a smooth initial curve. We also assume that a, b and c are smooth. Then $\Delta(s, \tau)$ is continuous in s and τ , and therefore nonzero in an open set, surrounding the initial curve.

By the inverse function theorem, the equations

$$\begin{aligned}x &= x(s, \tau), \\t &= t(s, \tau),\end{aligned}$$

can be inverted in an open set surrounding the initial curve. Thus we have

$$\begin{aligned}s &= s(x, t), \\\tau &= \tau(x, t).\end{aligned}$$

Then

$$u(x, t) = \varphi(s(x, t), \tau(x, t))$$

is the unique solution to the initial value problem

$$\begin{aligned}a\partial_x u + b\partial_t u &= c, \\u(x(\tau), t(\tau)) &= \varphi(\tau).\end{aligned}$$

In general, going through this program and finding the solution $u(x, t)$ can be very hard. However deploying the method can give great insight into a given PDE.

Let us consider the following quasilinear equation

$$\begin{aligned}u_t + uu_x &= 0, \\u(x, 0) &= \varphi(x).\end{aligned}\tag{81}$$

This equation occurs as a first approximation to many physical phenomena. We have seen how it appears in the modelling of traffic.

It also describe waves on the surface of the ocean. If we assume that the waves moves in one direction only, then (81) appears as a description of the deviation of the ocean surface from a calm, flat state.

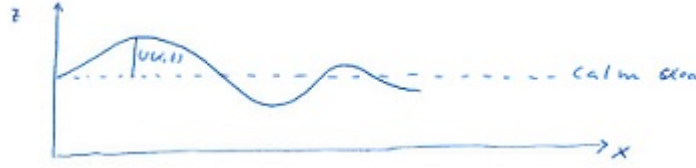


Figure 29: Ocean surface waves.

It is called the (inviscid) Burger's equation. Let us solve Burger's equation using the method of characteristics. Our initial value problem is

$$\begin{aligned} u_t + uu_x &= 0, \\ u(x, 0) &= \varphi(x). \end{aligned}$$

The initial curve is parametrized by

$$x = \tau, \quad t = 0, \quad u = \varphi(\tau).$$

Here

$$\begin{aligned} a(x, t, u) &= u, \\ b(x, t, u) &= 1, \\ c(x, t, u) &= 0. \end{aligned}$$

Let us compute the Jacobian along the initial curve. If the characteristic curves are

$$\begin{aligned} x &= x(s, \tau), \\ t &= t(s, \tau), \\ u &= u(s, \tau), \end{aligned}$$

where

$$\begin{aligned} \frac{dx}{ds} &= u, \quad x(0, \tau) = \tau, \\ \frac{dt}{ds} &= 1, \quad t(0, \tau) = 0, \\ \frac{du}{ds} &= 0, \quad u(0, \tau) = \varphi(\tau). \end{aligned} \tag{82}$$

Then

$$\begin{aligned} \partial_\tau t &= 0, \quad \partial_\tau x = 1 \quad \text{at } (0, \tau), \\ \Downarrow \\ \Delta(0, \tau) &= (a\partial_\tau t - b\partial_\tau x)|_{(0, \tau)} = -1 \neq 0, \end{aligned}$$

so $\Delta \neq 0$ along the entire initial curve. Let us next solve (82). From the last equation in (82) we get

$$u = u_0.$$

Thus u is a constant independent of s . Since u is a constant, the second equation in (82) give us

$$x = u_0 s + x_0,$$

and

$$\begin{aligned} u(0, \tau) = \varphi(\tau) &\Rightarrow u_0 = \varphi(\tau), \\ x(0, \tau) = \tau &\Rightarrow x_0 = \tau, \\ &\Downarrow \\ u &= \varphi(\tau), \\ x &= \varphi(\tau)s + \tau. \end{aligned}$$

Finally for t we get in the usual way

$$t = s.$$

If the conditions for the inverse function theorem holds, the equations

$$\begin{aligned} x &= \varphi(\tau)s + \tau, \\ t &= s, \end{aligned}$$

can be inverted to yield

$$\tau = \tau(x, t), \quad s = t,$$

and the solution to our problem is

$$u(x, t) = \varphi(\tau(x, t)).$$

Without specifying φ , we can not invert the system and find a solution to our initial value problem. We can however write down an *implicit* solution. We have

$$\begin{aligned} u &= \varphi(\tau), \\ x &= \varphi(\tau)s + \tau, \\ t &= s, \\ &\Downarrow \\ \tau &= x - us, \\ &\Downarrow \\ u &= \varphi(x - tu). \end{aligned}$$

Even without precisely specifying φ we can still gain great insight by considering the way the characteristics depends on the initial data. Projection of the characteristics in (x, t, u) space down into (x, t) space give us base characteristics of the form

$$x = \varphi(\tau)t + \tau.$$

This is a family of straight lines in (x, t) space parametrized by the parameter τ . Observe that the slope of the lines depends on τ , thus the lines in the family are *not* parallel in general. This is highly significant. The slope of the base characteristics is

$$\frac{dt}{dx} = \frac{1}{\varphi(\tau)},$$

and is evidently varying as a function of τ . This implies the possibility that the base characteristics will intersect. If they do, we are in trouble because the solution, u , is constant along the base characteristics. Thus if the base characteristics corresponding to parameters τ_1 and τ_2 intersect at a point (x^*, t^*) in (x, t) space then at this point u will have two values

$$\begin{aligned} u_1 &= \varphi(\tau_1), \\ u_2 &= \varphi(\tau_2), \end{aligned}$$

that are different unless it happened to be the case that

$$\varphi(\tau_1) = \varphi(\tau_2).$$

But in this case the two characteristics have the same slope and would not have intersected. The conclusion is that base characteristics can intersect and when that happens u is no longer a valid solution.

There is another way of looking at this phenomenon. Recall that the implicit solution is

$$u = \varphi(x - ut). \quad (83)$$

We know that a formula

$$v = \varphi(x - cv),$$

describes a wave moving at speed c to the right. The formula (83) can then be interpreted to say that each point on the wave has a “local” speed $u(x, t)$ that varies along the wave. Thus the wave does not move as one unit but can deform as it moves.

If $u(x, t) > 0$, the point $(x, u(x, t))$ on the wave moves to the right and if $u(x, t) < 0$, it moves to the left. If $u(x, t)$ is increasing, points on the wave move at greater and greater speed as we move to the right. The wave will stretch out. If $u(x, t)$ is decreasing the wave will compress. In a wave that has the shape of a localized “lump”,



Figure 30: Initial conditions in the form of a symmetric lump.



Figure 31: Steepening of initial lump over time.

there are points to the left that move faster than points further to the right. Thus points to the left can catch up with points to the right and lead to a steepening of the wave

Eventually the steepening can create a wave with a vertical edge. At this point $|u_x| = \infty$ and the model no longer describe what it is supposed to model. We thus have a *breakdown*.

It is evidently important to find out if and when such a breakdown occurs. There are two ways to work this out. The first approach use the implicit solution.

$$u = \varphi(x - ut). \quad (84)$$

We ask when the slope u_x become infinite. From (84) we get

$$\begin{aligned} u_x &= \varphi'(x - ut)(1 - u_x t), \\ &\Downarrow \\ u_x &= \frac{\varphi'(x - ut)}{1 + t\varphi'(x - ut)}. \end{aligned}$$

Clearly $|u_x|$ become infinite if the denominator become zero

$$\begin{aligned} 1 + t\varphi'(x - ut) &= 0, \\ &\Updownarrow \\ t &= -\frac{1}{\varphi'(\tau)}, \end{aligned}$$

where we recall that $\tau = x - ut$. What is important is to find the smallest time for which infinite slope occurs. The solution is not valid for times larger than this. Thus we must determine

$$t^* = \min_{\tau} \left(-\frac{1}{\varphi'(\tau)} \right).$$

When initial data is given, $\varphi(\tau)$ is known and t^* can be found.

The second method involve looking at the geometry of the characteristics. It leads to the same breakdown time t^* .

We can immediately observe that if φ is increasing for all x so that

$$\varphi'(\tau) > 0,$$

then $t^* < 0$ and no breakdown occur. For breakdown to occur we must have $\varphi'(\tau) < 0$ for some τ . Let us look at some specific choices of initial conditions.

$$\begin{aligned}
\text{i) } u(x, 0) = \varphi(x) = -x \\
\Rightarrow t(\tau) = -\frac{1}{\varphi'(\tau)} = 1, \\
\Downarrow \\
t^* = \min_{\tau} t(\tau) = 1.
\end{aligned}$$

Thus we have breakdown at $t = 1$. The solution is

$$\begin{aligned}
u = \varphi(x - ut) = -(x - ut) = -x + ut, \\
\Downarrow \\
u(1 - t) = -x \\
\Downarrow \\
u(x, t) = \frac{x}{t-1}
\end{aligned}$$

The whole graph becomes vertical at the breakdown time $t^* = 1$. This is

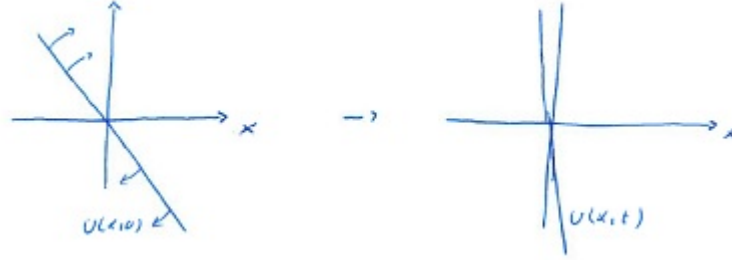


Figure 32: The solution $u(x, t) = \frac{x}{t-1}$ at $t = 0$ and $t = t_1 > 0$.

illustrated in figure 32. Let us describe this same case in terms of base characteristics. They are

$$x + \tau t = \tau.$$

They all intersect the t -axis at $t = 1$ and the base characteristics corresponding to τ intersect the x -axis at $x = \tau$. This case is illustrated in figure 33.

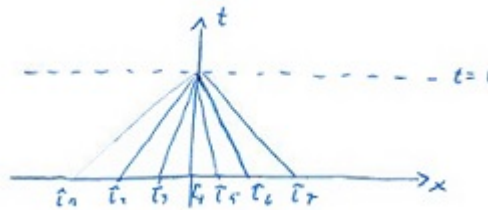


Figure 33: Intersecting base characteristics for the initial condition $u(x, 0) = -x$

ii) $u(x, 0) = \varphi(x) = 1 - x^2$.

$$\Rightarrow t(\tau) = -\frac{1}{\varphi'(\tau)} = \frac{1}{2\tau - 1},$$

$$\Downarrow$$

$$t^* = \min_{\tau} t(\tau) = 0.$$

Thus we have breakdown already at $t = 0$! The implicit formula for the solution now is

$$u = 1 - (x - ut)^2,$$

$$\Downarrow$$

$$u = 1 - x^2 + 2xut - u^2t^2,$$

$$\Downarrow$$

$$u^2t^2 + (1 - 2xt)u + x^2 - 1,$$

$$\Downarrow$$

$$u = \frac{1}{2t^2} \left(-(1 - 2xt) \pm \left((1 - 2xt)^2 - 4t^2(x^2 - 1)^{\frac{1}{2}} \right) \right),$$

$$\Downarrow$$

$$u = \frac{x}{t} - \frac{1 \pm (1 + 4t(t - x))^{\frac{1}{2}}}{2t^2}.$$

The initial condition requires that

$$\lim_{t \rightarrow 0^+} u(x, t) = 1 - x^2,$$

and for small t we have

$$(1 + 4t(t - x))^{\frac{1}{2}} \approx 1 + 2t(t - x) - 2t^2(t - x)^2.$$

Using the ‘-’ solution, we have for small t

$$\begin{aligned} u &\approx \frac{x}{t} - \frac{1 - 1 - 2t(t - x) + 2t^2(t - x)^2}{2t^2}, \\ &= \frac{x}{t} - \frac{1}{2t^2}(-2t^2 + 2tx + 2t^4 - 4t^3x + 2t^2x^2), \\ &= \frac{x}{t} + 1 - \frac{x}{t} - x^2 - 2t^2x - t^2. \end{aligned}$$

Thus

$$u(x, t) \rightarrow 1 - x^2, \quad \text{when } t \rightarrow 0.$$

Our unique solution of the initial value problem is therefore

$$u(x, t) = \frac{x}{t} - \frac{1 - (1 + 4t(t - x))^{\frac{1}{2}}}{2t^2}.$$

We have

$$u_x = \frac{1}{t}(1 - (1 + 4t(t - x))^{-\frac{1}{2}}),$$

and thus infinite slop occurs at all space-time points (x, t) determined by

$$1 + 4t(t - x) = 0,$$

$$\Updownarrow$$

$$x = t + \frac{1}{4t}.$$

This curve is illustarted in figure (34). There clearly are space-time points with arbitrarily small times on this curve and we therefore conclude that the breakdown time is $t^* = 0$.

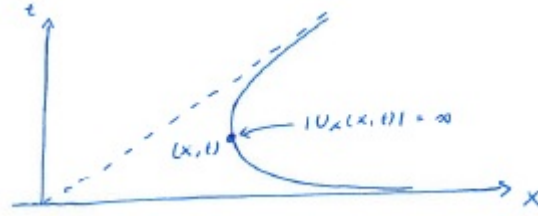


Figure 34: Set of space-time points where $u(x, t)$ has infinite slop.

The actual wave propagation is illustrated in figure 35. The steepening is evi-

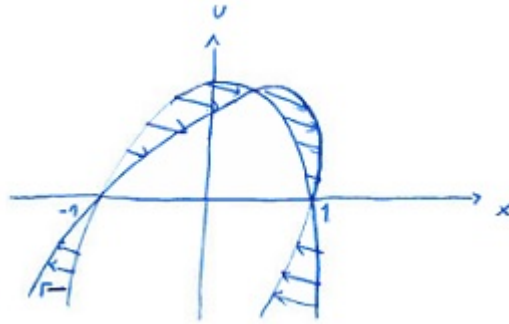


Figure 35: The solution $u(x, t)$ with $u(x, 0) = 1 - x^2$ at $t = 0$ and $t = t_1 > 0$

dent.

iii) $u(x, 0) = \varphi(x) = \sin x.$

The implicit solution is now

$$u = \sin(x - ut)$$

and thus

$$\begin{aligned}\varphi(\tau) &= \sin \tau, \\ \Downarrow \\ \varphi'(\tau) &= \cos \tau, \\ \Downarrow \\ t^* &= \min_{\tau} \left(-\frac{1}{\cos \tau} \right).\end{aligned}$$

From this formula it is evident that

$$t^* = 1,$$

and this minimum breakdown time is achieved for each $\tau_n = (2n + 1)\pi$.

Recall that the characteristic curves are parametrized by τ through

$$x = \sin(\tau)t + \tau.$$

Thus we can conclude that the breakdown occurs at the space-time points (t^*, x_n^*) where

$$x^* = \sin(\tau_n)t^* + \tau_n = (2n + 1)\pi.$$

The breakdown is illustrated in figure 36.

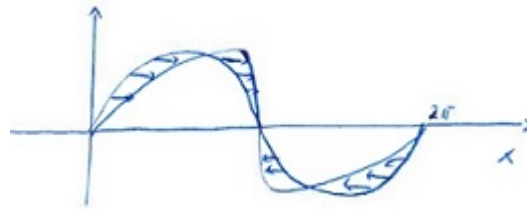


Figure 36: The solution $u(x, t)$ with $u(x, 0) = \sin x$ at $t = 0$ and $t = t_1 > 0$

7 Classification of equations: Characteristics

We will start our investigations by considering second order equations. More specifically we will consider scalar linear second order PDEs with two independent variables. The most general such equation is of the form

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = G, \quad (85)$$

where $u = u(x, y)$, $A = A(x, y)$ etc. The wave equation, diffusion equation and Laplace equation are clearly included in this class.

What separates one equation in this class from another are determined entirely by the functions $A \rightarrow G$. Thus the exact form of these functions determines all properties of the equation and the exact form of the set of solutions

Lectures used B
instead of 2B

(or *space* of solutions as we usually say). In fact if $G = 0$ the space of solutions is a *vectorspace*. This is because the differential operator

$$\mathcal{L} = A\partial_{xx} + 2B\partial_{xy} + C\partial_{yy} + D\partial_x + E\partial_y + F,$$

is a *linear operator*. (Prove it!) Thus even if we at this point knows *nothing* about the functions forming the space of solutions of the given equation, we do know that it has the *abstract property* of being a vectorspace.

Studying abstract properties of the space of solutions of PDEs is a central theme in the modern theory of PDEs. Very rarely will we be able to get an analytical description of the solution space for a PDE.

Knowing that a given PDE has a solution space that has a set of abstract properties is often of paramount importance. This knowledge can both inform us about expected qualitative properties of solutions and act as a guide for selecting which numerical or analytical approximation methods are appropriate for the given equation.

The aim of this section is to describe three abstract properties that the equation (85) might have or not. These are the properties of being *hyperbolic*, *parabolic* or *elliptic*.

Let us then begin: We first assume that $A(x, y)$ is nonzero in the region of interest. This is not a very restrictive assumption and can be achieved for most reasonable A . We can then divide (85) by A

$$u_{xx} + \frac{2B}{A}u_{xy} + \frac{C}{A}u_{yy} + \frac{D}{A}u_x + \frac{E}{A}u_y + \frac{F}{A}u = \frac{G}{A}. \quad (86)$$

We will now focus on the terms in (86) that are of second order in the derivatives. These terms form the *principal part* of the equation. The differential operator defining the principal part is

$$\mathcal{L} = \partial_{xx} + \frac{2B}{A}\partial_{xy} + \frac{C}{A}\partial_{yy}. \quad (87)$$

We now try to factorize the differential operator by writing it in the form

$$\mathcal{L} = (\partial_x - a\partial_y)(\partial_x - b\partial_y) + c\partial_y, \quad (88)$$

where a, b, c are some unknown functions. We find these functions by expanding out the derivatives and comparing with the expression (87)

$$\begin{aligned} \mathcal{L} &= \partial_{xx} - b_x\partial_y - b\partial_{xy} - a\partial_{xy} + ab_y\partial_y + ab\partial_{yy} + c\partial_y \\ &= \partial_{xx} - (a+b)\partial_{xy} + ab\partial_{yy} + (c - b_x + ab_y)\partial_y. \end{aligned}$$

Comparing with (87) we get

$$\begin{aligned} a + b &= -\frac{2B}{A}, \\ ab &= \frac{C}{A}, \quad c = b_x - ab_y. \end{aligned}$$

From the first two equations we find that both a and b solves the equation

$$Az^2 + 2Bz + C = 0. \quad (89)$$

Let us call these solutions ω^\pm so that

$$a = \omega^+, b = \omega^-.$$

We thus have

$$\omega^\pm = \frac{1}{A} \left\{ -B \pm (B^2 - AC)^{\frac{1}{2}} \right\}, \quad (90)$$

and then

$$\begin{aligned} & \partial_{xx} + \frac{2B}{A} \partial_{xy} + \frac{C}{A} \partial_{yy} \\ &= (\partial_x - \omega^+ \partial_y)(\partial_x - \omega^- \partial_y) + (\omega_x^- - \omega^+ \omega_y^-) \partial_y. \end{aligned}$$

We are in general interested in real solutions to our equations and aim to use this factorization to rewrite and simplify our equation. Thus it is of importance to know how many real solutions equation (89) has. From formula (90) we see that this is answered by studying the sign of $B^2(x, y) - A(x, y)C(x, y)$ in the domain of interest. We will here assume that this quantity is of one sign in the region of interest. This can always be achieved by restricting the domain, if necessary.

The equation (85) is called

- i) *Hyperbolic* if $B^2 - AC > 0$,
- ii) *Parabolic* if $B^2 - AC = 0$,
- iii) *Elliptic* if $B^2 - AC < 0$.

For the wave equation

$$u_{tt} - c^2 u_{xx} = 0,$$

we have $A = 1, B = 0, C = -c^2$ ($x \rightarrow t, y \rightarrow x$ in the general form of the equation (85))

$$\Rightarrow B^2 - AC = -1 \cdot (-c^2) = c^2 > 0.$$

Thus the wave equation is *hyperbolic*.

For the diffusion equation we have

$$u_t - D u_{xx} = 0,$$

we have $A = 0, B = 0, C = -D$

$$\Rightarrow B^2 - AC = 0.$$

Thus the diffusion equation is *parabolic*

For the Laplace equation we have

$$u_{xx} + u_{yy} = 0,$$

we have $A = 1, B = 0, C = 1$

$$\Rightarrow B^2 - AC = -1 < 0.$$

Thus the Laplace equation is *elliptic*.

We will now start deriving consequences of the classification into hyperbolic, parabolic and elliptic.

7.1 Canonical forms for equations of hyperbolic type

For this type of equations $\omega^+ \neq \omega^-$ and both ω^+ and ω^- are real. Take ω^+ and write down the following ODE

$$\frac{dy}{dx} = -\omega^+. \quad (91)$$

Let the solution curves of this equation be parametrized by the parameter ξ . Thus the curves can be written as

$$\xi = \xi(x, y).$$

In a similar way we can write the solution curves for the ODE

$$\frac{dy}{dx} = -\omega^-, \quad (92)$$

in the form

$$\eta = \eta(x, y),$$

where η is the parameter parameterizing the solution curves for (92). We now introduce (ξ, η) as new coordinates in the plane. The change of coordinates is thus

$$\begin{aligned} \xi &= \xi(x, y), \\ \eta &= \eta(x, y). \end{aligned} \quad (93)$$

For this to really be a coordinate system the Jacobian has to be nonzero.

$$\Delta = \xi_x \eta_y - \xi_y \eta_x.$$

The solution curves, or *characteristic curve*, of the equation (85) are by definition determined by

$$\begin{aligned} \xi(x, y) &= \text{const}, \\ \eta(x, y) &= \text{const}. \end{aligned}$$

Differentiate implicit with respect to x . This gives, upon using (91) and (92),

$$\begin{aligned} 0 &= \xi_x + y' \xi_y = \xi_x - \omega^+ \xi_y \Rightarrow \xi_x = \omega^+ \xi_y, \\ 0 &= \eta_x + y' \eta_y = \eta_x - \omega^- \eta_y \Rightarrow \eta_x = \omega^- \eta_y. \end{aligned} \quad (94)$$

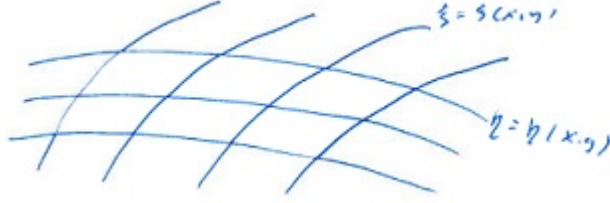


Figure 37

Thus

$$\Delta = (\omega^+ \xi_y) \eta_y - \xi_y (\omega^- \eta_y) = (\omega^+ - \omega^-) \eta_y \xi_y \neq 0,$$

because, by assumption, $\omega^+ \neq \omega^-$ in our domain. Thus the proposed change of coordinates (93) does in fact define a change of coordinates.

Recall that $z = \omega^+$ is a solutions to the quadratic equation (89), thus

$$A(\omega^+)^2 + 2B\omega^+ + C = 0. \quad (95)$$

Inserting $\omega^+ = \frac{\xi_x}{\xi_y}$ from equation (94) we get, after multiplication by ξ_y^2

$$A(x, y) \xi_x^2(x, y) + 2B(x, y) \xi_x(x, y) \xi_y(x, y) + C(x, y) \xi_y^2(x, y) = 0.$$

In a similar way the root $z = \omega^-$ show that the function $\eta(x, y)$ satisfy the equation

$$A(x, y) \eta_x^2(x, y) + 2B(x, y) \eta_x(x, y) \eta_y(x, y) + C(x, y) \eta_y^2(x, y) = 0.$$

Thus both $\xi(x, y)$ and $\eta(x, y)$, that determines the characteristic coordinate system are solutions to the *characteristic PDE* corresponding to our original PDE (85)

$$A(x, y) \varphi_x^2(x, y) + 2B(x, y) \varphi_x(x, y) \varphi_y(x, y) + C(x, y) \varphi_y^2(x, y) = 0.$$

Using the chain rule we have

$$\begin{aligned} u_x &= \xi_x u_\xi + \eta_x u_\eta, \\ u_y &= \xi_y u_\xi + \eta_y u_\eta, \end{aligned}$$

\Downarrow

$$\begin{aligned} u_x - \omega^+ u_y &= \xi_x u_\xi + \eta_x u_\eta - \omega^+ \xi_y u_\xi - \omega^+ \eta_y u_\eta \\ &= (\xi_x - \omega^+ \xi_y) u_\xi + (\eta_x - \omega^+ \eta_y) u_\eta \\ &= (\eta_x - \omega^+ \eta_y) u_\eta = \eta_y (\omega^- - \omega^+) u_\eta, \end{aligned} \quad (96)$$

because

$$\xi_x - \omega^+ \xi_y = \xi_x + y'(x) \xi_y = \frac{d}{dx} \xi(x, y(x)) = 0.$$

Since

$$\xi(x, y) = \text{const},$$

describe the solution curves to the equation (91). In a similar way we find

$$u_x - \omega^- u_y = \xi_y(\omega^+ - \omega^-)u_\xi. \quad (97)$$

Therefore, using equations (94), we have

$$\begin{aligned} \partial_x - \omega^+ \partial_y &= -\eta_y(\omega^+ - \omega^-)\partial_\eta, \\ \partial_x - \omega^- \partial_y &= \xi_y(\omega^+ - \omega^-)\partial_\xi, \end{aligned} \quad (98)$$

and both operators are nonzero because $\omega^+ \neq \omega^-$. But then we have

$$\begin{aligned} (\partial_x - \omega^+ \partial_y)(\partial_x - \omega^- \partial_y) &= [-\eta_y(\omega^+ - \omega^-)\partial_\eta] [\xi_y(\omega^+ - \omega^-)\partial_\xi] \\ &= -\xi_y \eta_y (\omega^+ - \omega^-)^2 \partial_{\xi\eta}^2 - \eta_y(\omega^+ - \omega^-)\partial_\eta(\xi_y(\omega^+ - \omega^-))\partial_\xi. \end{aligned} \quad (99)$$

Substituting this into (85) and expressing all first order derivatives with respect to x and y in terms of ∂_ξ and ∂_η , we get an equation of the form

$$u_{\xi\eta} + au_\xi + bu_\eta + cu = d, \quad (100)$$

where we have divided by the nonzero quantity $\xi_y \eta_y (\omega^+ - \omega^-)^2$ and where now $u = u(\xi, \eta)$, $a = a(\xi, \eta)$ etc. For any particular equation a, b, c and d will be determined explicitly in terms of A, B, C and D .

Since (100) results from (85) by a change of coordinates, (100) and are *equivalent* PDEs. Their spaces of solutions will consist of functions that are related through the change of coordinates (93).

(100) is called *the canonical form* for equations of hyperbolic type.

There is a second common form for hyperbolic equation. We find this by doing the additional transformation

$$\begin{aligned} \xi &= \alpha + \beta, \\ \eta &= \alpha - \beta, \end{aligned}$$

\Downarrow

$$u_{\alpha\alpha} - u_{\beta\beta} + \tilde{a}u_\alpha + \tilde{b}u_\beta + \tilde{c}u = \tilde{d}, \quad (101)$$

for some new functions $\tilde{a} = \tilde{a}(\alpha, \beta)$ etc. This is also called the canonical form for hyperbolic equations. We observe that the wave equation

$$u_{tt} - u_{xx} = 1,$$

is exactly of this form.

7.2 Canonical forms for equations of parabolic type

For parabolic equations we have

$$B^2(x, y) - A(x, y)C(x, y) = 0,$$

in the domain of interest. For this case we have

$$\omega^+ = \omega^- = \omega,$$

and there exists only one family of characteristic curves

$$\xi = \xi(x, y),$$

where the curves

$$\xi(x, y) = \text{const},$$

by definition are solutions to the ODE

$$\frac{dy}{dx} = -\omega.$$

Let

$$\eta = \eta(x, y),$$

be another family of curves chosen such that

$$\begin{aligned}\xi &= \xi(x, y), \\ \eta &= \eta(x, y),\end{aligned}\tag{102}$$

defines a new coordinate system (ξ, η) in the domain of interest. For example if A, B and C are constants we have

$$\xi = y + \omega x,$$

and we can let

$$\eta = \omega y + cx,$$

where $c \neq \omega^2$. Then (102) is a coordinate system because

$$\Delta = \xi_x \eta_y - \xi_y \eta_x = \omega^2 - c \neq 0.$$

We now express all derivatives $u_x, u_{xx}, u_{xy}, u_y, u_{yy}$ in terms of $u_\xi, u_\eta, u_{\xi\eta}, u_{\xi\xi}, u_{\eta\eta}$ using the chain rule. For example we have

$$\begin{aligned}u_x &= \xi_x u_\xi + \eta_x u_\eta, \\ u_y &= \xi_y u_\xi + \eta_y u_\eta, \\ u_{xx} &= \xi_{xx} u_\xi + \eta_{xx} u_\eta + \xi_x^2 u_{\xi\xi} + 2\xi_x \eta_x u_{\xi\eta} + \eta_x^2 u_{\eta\eta},\end{aligned}$$

etc.

Inserting these expressions into the equation we get a principal part of the form

$$\begin{aligned} & [A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2] u_{\xi\xi} + 2[A\xi_x\eta_x + B\xi_x\eta_y + B\xi_y\eta_x + C\xi_y\eta_y] u_{\xi\eta} \\ & + [A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2] u_{\eta\eta}, \end{aligned} \quad (103)$$

Since

$$\xi_x = \omega\xi_y,$$

we have

$$\begin{aligned} A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2 &= A(\omega\xi_y)^2 + 2B(\omega\xi_y)\xi_y + C\xi_y^2 \\ &= \xi_y^2(A\omega^2 + 2B\omega + C) = 0. \end{aligned}$$

Thus, the term $u_{\xi\xi}$ drops out of the equation. Also recall that for the parabolic case we have $\omega = -\frac{B}{A}$. Therefore we have

$$\begin{aligned} A\xi_x\eta_x + B\xi_x\eta_y + B\xi_y\eta_x + C\xi_y\eta_y &= A(-\frac{B}{A}\xi_y)\eta_x + B(-\frac{B}{A}\xi_y)\eta_y + B\xi_y\eta_x + C\xi_y\eta_y \\ &= -B\xi_y\eta_x - \frac{B^2}{A}\xi_y\eta_y + B\xi_y\eta_x + C\xi_y\eta_y \\ &= -\frac{1}{A}\xi_y\eta_y(B^2 - AC) \\ &= 0. \end{aligned}$$

Therefore the term $u_{\xi\eta}$ also drops out of the equation. Also, observe that the term

$$A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2, \quad (104)$$

must be nonzero. This is because if it was zero, $\eta = \eta(x, y)$ would define a second family of characteristic curves different from $\xi = \xi(x, y)$ and we know that no such family exists for the parabolic case.

We can therefore divide the equation by the term (104) and we get

$$u_{\eta\eta} + au_{\xi} + bu_{\eta} + cu = d. \quad (105)$$

This is the canonical form for equations of parabolic type. We observe that the diffusion equation

$$u_t - Du_{xx} = 0,$$

is of this form.

7.3 Canonical forms for equations of elliptic type

For elliptic equations we have

$$B^2 - AC < 0,$$

so both w^+ and w^- are complex and ξ and η become complex coordinates. In [1] all functions are extended into the complex domain, and an argument like the one for the hyperbolic case lead to the canonical form

$$u_{\alpha\alpha} + u_{\beta\beta} + au_{\alpha} + bu_{\beta} + cu = d, \quad (106)$$

where $a = a(\alpha, \beta)$ etc. The same canonical form can be found using only real variables. The key step in the reduction to canonical form is to simplify the principal part of the equation using a change of variables. Let us therefore do a general change of variables

$$\begin{aligned}\xi &= \xi(x, y), \\ \eta &= \eta(x, y).\end{aligned}$$

From equation (103) we have the following expression for the principal part of the transformed equation

$$\begin{aligned}& [A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2] u_{\xi\xi} + 2[A\xi_x\eta_x + B\xi_x\eta_y + B\xi_y\eta_x + C\xi_y\eta_y] u_{\xi\eta} \\ & + [A\eta_x^2 + B\eta_x\eta_y + C\eta_y^2] u_{\eta\eta}.\end{aligned}$$

The functions $\xi(x, y)$ and $\eta(x, y)$ are at this point only constrained by the Jacobi condition

$$\Delta(x, y) = (\xi_x\eta_y - \xi_y\eta_x)(x, y) \neq 0.$$

In the elliptic case, both solutions ξ, η to the characteristic equation are complex. Thus, for the elliptic case the coefficients in front of $u_{\xi\xi}$ and $u_{\eta\eta}$ can *not* be made to vanish. However if we let $\eta(x, y)$ be arbitrary and let $\xi = \xi(x, y)$ be the solution family of the system

$$\frac{dy}{dx} = \frac{B\eta_x + C\eta_y}{A\eta_x + B\eta_y}. \quad (107)$$

Then along the solution curves $\xi(x, y) = \text{const}$, we have

$$\begin{aligned}\xi_x + \frac{dy}{dx}\xi_y &= 0, \\ \Downarrow \\ \xi_x &= -\xi_y \frac{B\eta_x + C\eta_y}{A\eta_x + B\eta_y}.\end{aligned}$$

Thus

$$\begin{aligned}\Delta &= \xi_x\eta_y - \xi_y\eta_x = -\xi_y\eta_y \frac{B\eta_x + C\eta_y}{A\eta_x + B\eta_y} - \xi_y\eta_x \\ &= \frac{-\xi_y}{A\eta_x + B\eta_y} \{ \eta_y(B\eta_x + C\eta_y) + \eta_x(A\eta_x + B\eta_y) \} \\ &= \frac{-\xi_y}{A\eta_x + B\eta_y} \{ B\eta_x\eta_y + C\eta_y^2 + A\eta_x^2 + B\eta_x\eta_y \} \\ &= \frac{-\xi_y}{A\eta_x + B\eta_y} \{ A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2 \} \neq 0,\end{aligned}$$

because the characteristic equation has no real solutions.

Therefore

$$\begin{aligned}\xi &= \xi(x, y), \\ \eta &= \eta(x, y),\end{aligned}$$

is a change of coordinates.

Furthermore

$$A\xi_x\eta_x + B\xi_x\eta_y + B\xi_y\eta_x + C\xi_y\eta_y = (A\eta_x + B\eta_y)\left(\xi_x + \frac{B\eta_x + C\eta_y}{A\eta_x + B\eta_y}\xi_y\right) = 0.$$

Thus, the coefficient in front of $u_{\xi\eta}$ is zero. Since the quantity

$$A\varphi_x^2 + 2B\varphi_x\varphi_y + C\varphi_y^2 \neq 0, \quad \forall \text{ real } \varphi,$$

the coefficients in front of $u_{\xi\xi}$ and $u_{\eta\eta}$ are nonzero and of the same sign. Dividing, we find that the principal part takes the form

$$u_{\xi\xi} + h(\xi, \eta)u_{\eta\eta}, \quad (108)$$

where the function $h(x, y)$ is a positive function.

Let us at this point simplify our considerations by assuming that the coefficients of the equation (85) are in fact constant. If that is the case, we choose the function $\eta(x, y)$ to be linear in x and y . Then the derivatives of η are constant and therefore the function $\xi(x, y)$ defined in (107) is also a linear function. For this particular case, the function $h(\xi, \eta)$ is a positive constant which we denote by h_0 . Thus equation (108) takes the form

$$u_{\xi\xi} + h_0u_{\eta\eta}. \quad (109)$$

Introduce a change of variables defined by

$$\alpha = \xi, \quad (110)$$

$$\beta = \frac{1}{\sqrt{h_0}}\eta. \quad (111)$$

Using the chain rule we find that the principal part of an elliptic equation with constant coefficients, takes the form

$$u_{\alpha\alpha} + u_{\beta\beta}.$$

Thus the canonical form of elliptic equations with constant coefficients is

$$u_{\alpha\alpha} + u_{\beta\beta} + au_{\alpha} + bu_{\beta} + cu = d. \quad (112)$$

One can show that also elliptic equations with variable coefficients can be put into canonical form (112) by a suitable change of coordinates. This argument however either relies on complex analysis, or some rather intricate real variable manipulations.

7.4 Characteristic curves

One consequence of the classification of equations into hyperbolic, parabolic and elliptic, is the possibility to reduce each class into a canonical form where the principal part has the form

$$\begin{array}{ll} u_{\xi\xi} - u_{\eta\eta} & \text{hyperbolic,} \\ u_{\xi\xi} & \text{parabolic,} \\ u_{\xi\xi} + u_{\eta\eta} & \text{elliptic.} \end{array}$$

In these reductions, certain special curve families played an important role. We will now see that these special curve families reappear when we consider the possibility of solutions that contain singularities. These are solutions whose existence, or not, is a consequence of which class our equation belongs to. This will give us a new window into what the abstract classification of equations into the three classes hyperbolic, parabolic and elliptic actually means.

Let our domain, D , be separated into two parts D_1 and D_2 by a curve.

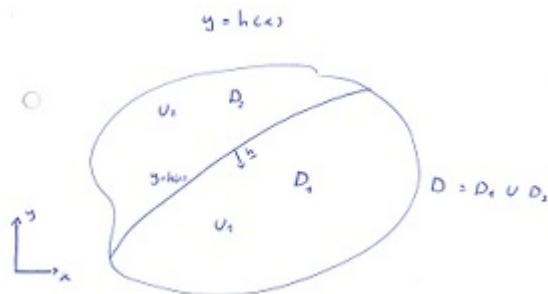


Figure 38

Let u_1 and u_2 be solutions of the equation

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = G, \quad (113)$$

in D_1 and D_2 . Define a function U on the whole domain D by

$$u(x, y) = \begin{cases} u_1(x, y) & (x, y) \in D_1, \\ u_2(x, y) & (x, y) \in D_2. \end{cases}$$

We will assume that u, u_x, u_y are continuous on D but that u_{xx}, u_{xy}, u_{yy} can be discontinuous across the curve separating D_1 and D_2 . Thus u is not a solution in the usual sense to (113) on the whole domain D .

The curve is defined by

$$\xi(x, y) = y - h(x) = 0$$

Introduce a curve family

$$\xi(x, y) = \xi, \quad \xi \in \mathbb{R}$$

Our curve $y = h(x)$ corresponds to $\xi = 0$. Introduce a second curve family determined $\eta = \eta(x, y)$ in such a way that

$$\begin{aligned}\xi &= \varphi(x, y), \\ \eta &= \psi(x, y),\end{aligned}$$

define a change of coordinates. For this to be the case the Jacobian must be nonzero everywhere.

We now transform our equation into the new coordinate system in the same way as on page 79, and get

$$\begin{aligned}(A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2)u_{\xi\xi} + 2(A\xi_x\eta_x + B(\xi_x\eta_y + \xi_y\eta_x) \\ + C\xi_y\eta_y)u_{\xi\eta} + (A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2)u_{\eta\eta} + \dots = G,\end{aligned}\tag{114}$$

where we only have displayed the principal part of the equation. What we precisely mean when we say that u_{xx}, u_{xy} and u_{yy} can be discontinuous across $y = h(x)$ is that

$$u, u_\xi, u_\eta, u_{\xi\eta}, u_{\eta\eta},$$

are continuous but that $u_{\xi\xi}$ can be discontinuous. Evaluating (114) above and below the curve $\xi = 0 \Leftrightarrow y = h(x)$, taking the difference and going to the limit as the points approach $y = h(x)$ from above and below we get

$$(A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2)[u_{\xi\xi}]|_{\xi=0} = 0,$$

where $[u_{\xi\xi}]|_{\xi=0}$ is the jump in $u_{\xi\xi}$ across $y = h(x)$. Since $[u_{\xi\xi}]|_{\xi=0} \neq 0$, we *must* have

$$A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2 = 0.$$

Thus $\xi(x, y)$ must satisfy the characteristic PDE and consequently

$$\xi = 0 \Leftrightarrow y = h(x),$$

must be a *characteristic curve* for the PDE (113).

Thus, discontinuities in the second derivatives can only occur on characteristic curves.

Let us see how this works out for the wave equation, diffusion equation and Laplace equation. We have seen that these three equations are respectively of hyperbolic type, parabolic type and elliptic type.

The wave equation is

$$u_{tt} - \gamma^2 u_{xx} = 0,$$

$$A = 1, B = 0, C = \gamma.$$

The characteristic PDE is

$$\varphi_t^2 - \gamma^2 \varphi_x^2 = (\varphi_t + \gamma \varphi_x)(\varphi_t - \gamma \varphi_x) = 0,$$

and the characteristic curves are

i) $\varphi_t - \gamma\varphi_x = 0 \Leftrightarrow \varphi(x, t) = x + \gamma t,$

ii) $\varphi_t + \gamma\varphi_x = 0 \Leftrightarrow \varphi(x, t) = x - \gamma t.$

The two families of characteristic curves are

$$x + \gamma t = \text{const},$$

$$x - \gamma t = \text{const}.$$

Singularities, that is discontinuities in u_{xx}, u_{xy}, u_{yy} , can only occur along these

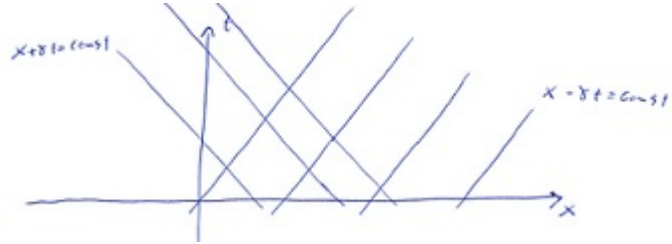


Figure 39

curves. Singularities given in initial data at $t = 0$ is propagated along the characteristic curves and persists forever. This behavior is typical for all hyperbolic equations.

The diffusion equation is

$$u_t = Du_{xx},$$

$$\Rightarrow A = D, B = C = 0.$$

So the characteristic PDE is

$$\varphi_x^2 = 0, \Rightarrow \varphi = \varphi(t),$$

and characteristic curves $\varphi(x, t) = \text{const}$ must be of the form

$$t = \text{const}.$$



Figure 40

Singularities in the initial data given on the x-axis can not leave the x-axis and will not be seen at any point (x, t) for $t > 0$. Thus, all singularities in initial data

are washed out instantly for $t > 0$. This behaviour is typical for all parabolic equations.

Finally, the Laplace equation is

$$u_{xx} + u_{yy} = 0,$$

and the characteristic equation is

$$\varphi_x^2 + \varphi_y^2 = 0, \Leftrightarrow \varphi_x = \varphi_y = 0, \Leftrightarrow \varphi(x, y) = \text{const.}$$

There are no characteristic curves that can support singularities. The solutions of elliptic equations are very smooth. This is typical for all elliptic equations.

7.5 Characteristic curves again: A different point of view

We first discussed characteristic curves in the context of linear and quasilinear first order scalar PDEs in two independent variables. There they played the star role in a method for solving the initial value problem for such equations. This was the *method of characteristics*. Next, we met them in the context of second order linear scalar PDEs in two independent variables. Here we have seen that they played a dual role. On the one hand they were instrumental in the reduction of hyperbolic and parabolic equations to canonical form, and on the other hand they were curves along which solutions to the PDEs could have discontinuities in their second derivatives. This second role is also played by the characteristic of first order linear scalar PDEs.

We will now discuss a third way in which characteristic curves play a role.

Let us start by considering the case of first order linear scalar PDEs in two independent variables. The initial value problem is assumed to be given in terms of an initial curve that is the graph of a function $h(x)$

$$y = h(x).$$

On this curve the solution u is assumed to be given by some function $f(x)$. Thus we have

$$au_x + bu_y = cu + d, \tag{115}$$

$$u(x, h(x)) = f(x). \tag{116}$$

Differentiating (116) we get

$$u_x + h'u_y = f', \tag{117}$$

and evaluating (115) along the initial curve we get

$$au_x + bu_y = cf + d, \tag{118}$$

(117) and (118) is a linear system of equations for determining u_x and u_y

$$\begin{bmatrix} 1 & h' \\ a & b \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} f' \\ cf + d \end{bmatrix}. \tag{119}$$

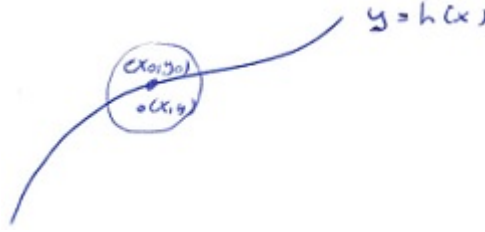


Figure 41

For points close to the initial curve we can express the solution of the initial value problem (115),(116) as a Taylor series.

$$u(x, y) = u(x_0, y_0) + u_x(x_0, y_0)(x - x_0) + u_y(x_0, y_0)(y - y_0) + \dots$$

Clearly if the solution $u(x, y)$ is to exist and be unique, then at least $u_x(x_0, y_0)$ and $u_y(x_0, y_0)$ must exist as a unique solution of the system (119).

This is ensured if

$$\det \begin{bmatrix} 1 & h' \\ a & b \end{bmatrix} \bigg|_{(x_0, y_0)} \neq 0.$$

Even if this is true it is not certain that the initial value problem is uniquely solvable, there is the question of the coefficients of the rest of the Taylor series and of course there is the matter of convergence for such series

However, if on the other hand,

$$\det \begin{bmatrix} 1 & h' \\ a & b \end{bmatrix} \bigg|_{(x_0, y_0)} = 0, \quad (120)$$

$u_x(x_0, y_0), u_y(x_0, y_0)$ either does not exist and if they do exist they are not unique. This follows from the elementary theory of linear systems of equations. So if (120) holds we can know for sure that we do not have existence and uniqueness for the initial value problem (115),(116).

Condition (120) can be written as

$$\begin{aligned} b - ah' &= 0, \\ \Updownarrow \\ h' &= \frac{b}{a}. \end{aligned} \quad (121)$$

This equation tells us that $y = h(x)$ must be a characteristic curve! This follows, because the family of base characteristic curves $(x(s), y(s))$ are determined by being solutions to

$$\frac{dx}{ds} = a, \frac{dy}{ds} = b. \quad (122)$$

If the curves in question are graphs of functions we can parametrise them using x . So that they are given by a function $y(x)$. For this function (122) together with the chain rule gives

$$\frac{dy}{dx} = \frac{dy}{ds} \cdot \frac{ds}{dx} = \frac{b}{a}.$$

and (121) say that $h(x)$ solves this equation so that $y = h(x)$ is a base characteristic curve.

Let us next consider the case of second order linear scalar PDEs in two independent variables.

The characteristic curves have preciously been found to satisfy

$$A\varphi_x^2 + 2B\varphi_x\varphi_y + C\varphi_y^2 = 0. \quad (123)$$

The curves on which the PDE could have singularities, $y = h(x)$, were related to $\varphi(x, y)$ through

$$\varphi(x, y) = y - h(x) = 0,$$

and (123) then give the equation

$$Ah'^2 - 2Bh' + C = 0 \quad (124)$$

The solutions of this nonlinear ODE gives the characteristic curves for PDEs of the form

$$Au_{xx} + Bu_{xy} + Cu_{yy} = F(x, y, u, u_x, u_y), \quad (125)$$

where

$$F(x, y, u, u_x, u_y) = Du_x + Eu_y + Fu + G.$$

Recall that all parameters are functions of x and y .

Let $y = h(x)$ be an initial curve where we have given u, u_x and u_y . Thus

$$\begin{aligned} u(x, h(x)) &= p(x), \\ u_x(x, h(x)) &= q(x), \\ u_y(x, h(x)) &= r(x). \end{aligned} \quad (126)$$

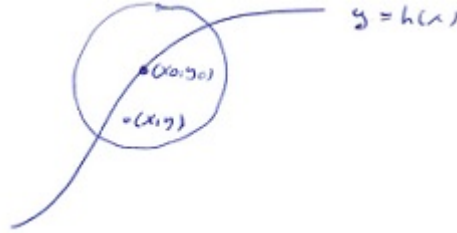


Figure 42

For points close to the initial curve the solution can be expand as a Taylor series.

$$\begin{aligned} u(x, y) &= u(x_0, y_0) + u_x(x_0, y_0)(x - x_0) + u_y(x_0, y_0)(y - y_0) \\ &+ \frac{1}{2}u_{xx}(x_0, y_0)(x - x_0)^2 + u_{xy}(x_0, y_0)(x - x_0)(y - y_0) \\ &+ \frac{1}{2}u_{yy}(x_0, y_0)(y - y_0)^2. \end{aligned}$$

Differentiating the last two equations in the initial data (126) with respect to x , we get

$$u_{xx} + h'u_{xy} = q', \quad (127)$$

$$u_{xy} + h'u_{yy} = r'. \quad (128)$$

Evaluating (125) along $y = h(x)$ we get the equation

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} = F(x, h(x), p, q, r). \quad (129)$$

Equations (127), (128) and (129) define a linear system of equations for u_{xx} , u_{xy} and u_{yy}

$$\begin{bmatrix} 1 & h' & 0 \\ 0 & 1 & h' \\ A & 2B & C \end{bmatrix} \begin{bmatrix} u_{xx} \\ u_{xy} \\ u_{yy} \end{bmatrix} = \begin{bmatrix} q' \\ r' \\ F(x, h(x), p, q, r) \end{bmatrix}. \quad (130)$$

The Taylor expansion show that the the problem (125) with initial data (126) can have a unique solution only if the determinant of the matrix in the linear system (130) is nonzero.

If, on the other hand, the determinant is zero, the problem is sure to either have no solution or to have a nonunique solution. This condition is

$$\det \begin{bmatrix} 1 & h' & 0 \\ 0 & 1 & h' \\ A & 2B & C \end{bmatrix} = 0, \quad (131)$$

$$\Downarrow \\ C - 2Bh' + Ah'^2 = 0, \quad (132)$$

and this is exactly the equation for the characteristic curves (124).

Thus characteristic curves *are* curves for which the (generalized) initial value problem either can not be solved or for which the solution is non-unique.

There is something awkward about the presentation given and equation (132). Since we are considering the general initial value problem as giving u , u_x and u_y along some curve in the (x, y) plane, there is really no point in restricting to curves that is the graph of a function $y = h(s)$. Why not consider curves that are graph of a function $x = k(y)$ or indeed a curve that is neither a graph of a function of x or a function of y ?

Let us consider case iii). Let $\varphi(x, y)$ be a function such that C is a level curve of φ

$$C : \varphi(x, y) = 0.$$

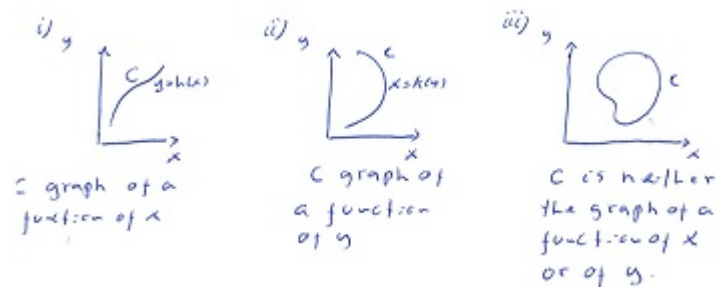


Figure 43

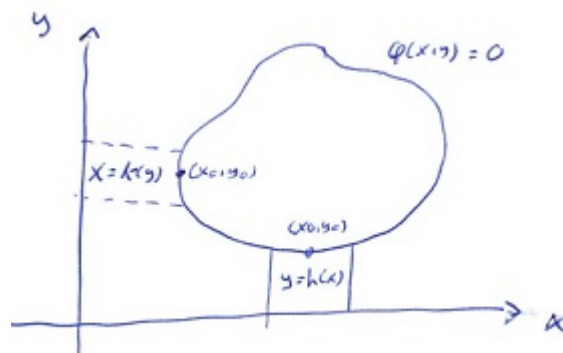


Figure 44

Pick any point (x_0, y_0) on C . Then the piece of the curve close to (x_0, y_0) will either be the graph of a function of x or the graph of a function of y

Let us assume, without loss of generality, that the piece of C close to (x_0, y_0) is the graph of a function of x , $y = h(x)$. But then we must have

$$\varphi(x, h(x)) = 0,$$

$$\Downarrow$$

$$\varphi_x + h' \varphi_y = 0,$$

$$\Downarrow$$

$$h' = -\frac{\varphi_x}{\varphi_y}.$$

Inserting this into equation (132) gives

$$C + 2B \frac{\varphi_x}{\varphi_y} + A \left(\frac{\varphi_x}{\varphi_y} \right)^2 = 0,$$

$$\Updownarrow$$

$$A\varphi_x^2 + 2B\varphi_x\varphi_y + C\varphi_y^2 = 0. \quad (133)$$

The conclusion of this is that curves, C , for which the generalized initial value problem (125), (126) either has no solution or has a nonunique solution, must be level curves

$$\varphi(x, y) = 0,$$

where φ is a solution of (133). And all solutions of (133) determines such special initial curves.

We have met equation (133) before on page 77 of these notes where it was called the characteristic equation for (125).

Thus the special initial curves discussed in this section are in fact *characteristic curves* as defined previously.

The characteristic curves are seen to play three different roles for PDEs of type (125).

- i) Level curves for special coordinate system used to reduce the PDE to canonical form.
- ii) Curves on which solutions of the PDE can have singularities.
- iii) Curves for which the initial value problem for the PDE either have no solution or have a non-unique solution.

So far the statements i), ii) and iii) only applies to linear second order scalar PDEs in two independent variables. Can we extend these results to more general equations?

The answer is yes we can! Many of these ideas can be extended far beyond the class of equations we have discussed so far. In [1] classification and associated characteristic curves and surfaces for large classes of equations are discussed on pages 137-151.

In this course we will not discuss all these various cases but will rather restrict to linear scalar second order equations in n independent variables where n can be any positive integer.

Equations of this type takes the form

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij} \partial_{x_i} \partial_{x_j} u + \sum_{i=1}^n b_i \partial_{x_i} u + cu + d = 0, \quad (134)$$

where all parameters $\{a_{ij}\}$ and $\{b_i\}, c, d$ are functions of (x_1, \dots, x_n)

Remark on notation: Here and in many other areas of applied mathematics and theoretical physics it is common to use the so-called *Einstein summation convention*.

In this convention repeated indices in any expression are implicitly summed over. Thus the expressions

$$x_i y_i,$$

$$a_{ij} x_j,$$

$$x_i a_{ij} y_j,$$

means

$$\begin{aligned} & \sum_{i=1}^n x_i y_i, \\ & \sum_{j=1}^n a_{ij} x_j, \\ & \sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} y_j. \end{aligned}$$

Using this convention, equation (134) is written as

$$a_{ij} \partial_{x_i} \partial_{x_j} u + b_i \partial_{x_i} u + cu + d = 0, \quad (135)$$

Observe that

$$\begin{aligned} a_{ij} \partial_{x_i} \partial_{x_j} u &= \frac{1}{2} a_{ij} \partial_{x_i} \partial_{x_j} u + \frac{1}{2} a_{ij} \partial_{x_i} \partial_{x_j} u \\ &= \frac{1}{2} a_{ij} \partial_{x_i} \partial_{x_j} u + \frac{1}{2} a_{ji} \partial_{x_j} \partial_{x_i} u \\ &= \frac{1}{2} a_{ij} \partial_{x_i} \partial_{x_j} u + \frac{1}{2} a_{ji} \partial_{x_i} \partial_{x_j} u \\ &= \left(\frac{1}{2} a_{ij} + \frac{1}{2} a_{ji} \right) \partial_{x_i} \partial_{x_j} u \\ &= \tilde{a}_{ij} \partial_{x_i} \partial_{x_j} u, \end{aligned}$$

where $\tilde{a}_{ij} = \frac{1}{2}(a_{ij} + a_{ji})$.

\tilde{a}_{ij} are evidently the components of a symmetric matrix

$$\tilde{a}_{ij} = \tilde{a}_{ji}.$$

Thus, we can without loss of generality assume that the matrix $A = (a_{ij})$ in equation (135) is a symmetric matrix.

Let us first assume that all a_{ij} are constants. Let λ_k be the eigenvalues of A and $R = (r_{ki})$ be the orthogonal matrix whose rows are the corresponding eigenvectors. In terms of matrices we have

$$RAR^t = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

This can be written, using the Einstein convention, in the form

$$r_{pi} a_{ij} r_{qj} = \Lambda_{pq},$$

where Λ_{pq} is defined by

$$\Lambda_{pq} = \begin{cases} \lambda_p, & p = q \\ 0, & p \neq q \end{cases}.$$

Introduce new coordinates ξ_1, \dots, ξ_n through

$$\xi_p = r_{pk}x_k.$$

This is evidently a linear change of coordinates. Observe that

$$\begin{aligned} \partial_{x_i} &= \partial_{x_i} \xi_p \partial_{\xi_p} = r_{pi} \partial_{\xi_p}, \\ \partial_{x_j} &= \partial_{x_j} \xi_q \partial_{\xi_q} = r_{qj} \partial_{\xi_q}, \\ &\Downarrow \end{aligned}$$

$$\begin{aligned} a_{ij} \partial_{x_i} \partial_{x_j} u &= a_{ij} r_{pi} r_{qj} \partial_{\xi_p} \partial_{\xi_q} u \\ &= \Lambda_{pq} \partial_{\xi_p} \partial_{\xi_q} u \\ &= \lambda_p \partial_{\xi_p} \partial_{\xi_p} u. \end{aligned}$$

Thus the second order mixed partial derivatives all vanish. This is the canonical form for (135). If we return to the $n = 2$ case we observe that

$$\begin{aligned} \lambda_1, \lambda_2 \text{ same sign} &- \text{elliptic case,} \\ \lambda_1, \lambda_2 \text{ opposite sign} &- \text{hyperbolic case,} \\ \text{one of } \lambda_1, \lambda_2 \text{ is zero} &- \text{parabolic case.} \end{aligned}$$

Based on this we now make the following classification of equations of type (135).

$$\begin{aligned} &\begin{cases} \lambda_i > 0, & \forall i \\ \text{or} \\ \lambda_i < 0, & \forall i \end{cases} && - \text{elliptic type,} \\ &\begin{cases} \text{one of } \lambda_i > 0 \text{ or } \lambda_i < 0 \\ \text{All other } \lambda_i \text{ opposite sign} \end{cases} && - \text{hyperbolic type,} \\ &\text{one or more } \lambda_i = 0 && - \text{parabolic type.} \end{aligned}$$

For the elliptic case we can make the further change of coordinates

$$\alpha_i = \frac{1}{|\lambda_i|} \xi_i, \tag{136}$$

leading to a principal part of the form

$$\pm \partial_{\alpha_i} \partial_{\alpha_i} u. \tag{137}$$

Assuming $\lambda_1 > 0$ and $\lambda_2 \dots \lambda_n < 0$, the hyperbolic case is reduced, using (136), to the form

$$\partial_{x_1}^2 - \sum_{i=2}^n \partial_{x_i}^2. \quad (138)$$

Similar reductions are possible for the parabolic case. Some $n = 3, 4$ basic examples of elliptic, hyperbolic and parabolic equations are

$$\begin{aligned} \nabla^2 u &= u_{xx} + u_{yy} + u_{zz} = 0 && \text{elliptic,} \\ u_{tt} - \nabla^2 u &= u_{tt} - u_{xx} - u_{yy} - u_{zz} = 0 && \text{hyperbolic,} \\ u_t - \nabla^2 u &= u_t - u_{xx} - u_{yy} - u_{zz} = 0 && \text{parabolic.} \end{aligned}$$

The classification of equations of type (135) into elliptic, hyperbolic and parabolic based on eigenvalues of the matrix A is still possible when a_{ij} depends on (x_1, \dots, x_n) and are thus not constant. The eigenvalues and the classification will however in general depend on which point we are at. Also, it is in general not possible to reduce it to a canonical form where all terms with mixed partial derivatives vanish. This is seen to be true by a simple counting argument.

Let

$$\begin{aligned} \xi_1 &= \xi_1(x_1, \dots, x_n), \\ &\vdots \\ \xi_n &= \xi_n(x_n, \dots, x_n), \end{aligned} \quad (139)$$

be a change of coordinates in \mathbb{R}^n . Using the chain rule the principal part of equation (135) can be written as

$$\sum_{ij} F_{ij}(\xi_1, \dots, \xi_n) \partial_{\xi_i} \partial_{\xi_j} u, \quad (140)$$

where F_{ij} is some expression containing derivatives of the functions ξ_j . Let us try to eliminate all mixed partials. Then for this to happen, we must have

$$F_{ij}(\xi_1, \dots, \xi_n) = 0, \quad \forall i \neq j. \quad (141)$$

There are exactly $\frac{1}{2}n(n-1)$ terms with mixed partial derivatives. Thus there are $\frac{1}{2}n(n-1)$ equations in (141). On the other hand we only have n unknown functions ξ_1, \dots, ξ_n . If

$$\frac{1}{2}n(n-1) > n,$$

the system (141) in general has no solution. This happened for $n \geq 4$.

For equations of type (135), surfaces supporting singularities of solutions, and on which given initial data does not give a unique solution for the initial value problem, are given as solutions to

$$a_{ij} \varphi_{x_i} \varphi_{x_j} = 0. \quad (142)$$

These surfaces are called *characteristic* surfaces and equation (142) is called the *characteristic equation* for equation (135). For example the 3D wave equation

$$u_{tt} - u_{xx} - u_{yy} - u_{zz} = 0,$$

has a characteristic equation

$$\varphi_t^2 - \varphi_x^2 - \varphi_y^2 - \varphi_z^2 = 0.$$

One important type of solution is

$$\varphi = \omega t - k_1 x - k_2 y - k_3 z + c,$$

where $\omega^2 = k_1^2 + k_2^2 + k_3^2$. For each value of t the characteristic surface

$$\varphi = 0,$$

$$\Updownarrow$$

$$k_1 x + k_2 y + k_3 z = \omega t + c,$$

is a plane normal to the vector $\mathbf{k} = (k_1, k_2, k_3)$. Such surfaces play a pivotal role in the theory of wave equations where they appear through special solutions called *plane waves*.

7.6 Stability theory, energy conservation and dispersion

So far we have used the highest order terms, the so-called principal part of the equation, to classify our equations. The lower order terms however also play an important role.

In this section of the notes we will investigate the effects of lower order terms by considering the class of equations of the form

$$Au_{xx} + 2Bu_{xt} + Cu_{tt} + Du_x + Eu_t + Fu = 0, \quad (143)$$

where A, B, C, D, E and F are real constants. Certain aspects of our conclusions for this class can be extended to a wider class of equations[1].

We start our investigations by looking for special solutions to (143) of exponential form

$$u(x, t) = a(k)e^{(ikx + \lambda(k)t)}. \quad (144)$$

This is a complex solution but both the real and imaginary part of u are real solutions of (143). (Verify this). We will in this section interpret the variable t as time and study the evolution of the solutions (144) for $t > 0$. The special solutions (144) are called *normal modes* for (143). By the principle of linear super-position we can add together several normal modes corresponding to different values of k to create more general solutions. Taking an infinite set of discrete values $\{k_j\}$ we get a solution of the form

$$u(x, t) = \sum_{j=-\infty}^{\infty} a(k_j)e^{(ik_j x + \lambda(k_j)t)}. \quad (145)$$

This is a *Fourier series*, and the sense in which such series solves (143) will be discussed extensively later in the class. There we will also discuss a much more general notion of normal mode that applies to a much wider class of equations. We will see that the normal modes (144) can be combined in a “continuous sum” by which we mean an integral

$$u(x, t) = \int_{-\infty}^{\infty} dk a(k) e^{(ikx + \lambda(k)t)}, \quad (146)$$

to generate even more general solutions to (143). We recognize (146) as a *Fourier transform*. More general transforms are associated with other types of normal modes. These are discussed in chapter 5 in [1].

At this point, the only thing to note is that through the general representations (145) and (146) the behavior of the normal modes (144) will determine what any solution to equation (143) will do for $t > 0$.

So, let us now start analyzing the normal modes. From (144) we have

$$\begin{aligned} u(x, t) &= a(k) e^{(ikx + \lambda(k)t)}, \\ &\Downarrow \\ u_t &= \lambda(k)u, \quad u_x = ik u, \quad u_{tt} = \lambda^2(k)u, \\ u_{xt} &= ik\lambda(k)u, \quad u_{xx} = -k^2 u. \end{aligned}$$

Inserting (147) into (143) we get

$$(-Ak^2 + 2iBk\lambda + C\lambda^2 + iDk + E\lambda + F)u = 0.$$

A solution of this type thus only exists if

$$C\lambda^2 + (2ikB + E)\lambda + (-Ak^2 + iDk + F) = 0. \quad (147)$$

This equation has two complex solutions. For each of these we can write

$$\lambda(k) = \text{Re}[\lambda(k)] + i\text{Im}[\lambda(k)], \quad (148)$$

and the corresponding normal mode can be written as

$$u(x, t) = a(k) e^{\text{Re}[\lambda(k)]t} e^{i(kx + \text{Im}[\lambda(k)]t)}. \quad (149)$$

The magnitude of $u(x, t)$ is

$$|u(x, t)| = |a(k)| e^{\text{Re}[\lambda(k)]t}. \quad (150)$$

We will assume that the expression $|a(k)|$ is bounded for all k . The growth of $|u(x, t)|$ as a function of k for $t > 0$ is then determined by $\text{Re}[\lambda(k)]$.

There are now two distinct possibilities. Either $\text{Re}[\lambda(k)]$ is bounded above for all $k \in \mathbb{R}$, or it is not.

Define

$$\Omega = \sup_{k \in \mathbb{R}} \text{Re}[\lambda(k)]. \quad (151)$$

If $Re[\lambda(k)]$ is unbounded we let $\Omega = +\infty$. Let us first consider the case $\Omega = +\infty$. Previously in this class we introduced the notion of being well-posed for initial and or boundary value problems for PDEs. In order to be well-posed, three criteria had to be satisfied

- 1) The problem must have at least one solution.
- 2) The problem must have at most one solution.
- 3) The solution must depend in a continuous way on the data of the problem: Small data must produce a small solution.

Data here means parameters occurring in the PDEs, description of the boundary curves or surfaces, if any, and functions determining boundary and/or initial conditions

For the case $\Omega = +\infty$, let $a(k) = \frac{1}{\lambda(k)^2}$ and consider the initial value problem for the normal mode with data at $t = 0$. The data are from (144)

$$u(x, 0) = \frac{1}{\lambda^2(k)} e^{ikx},$$

$$u_t(x, 0) = \frac{1}{\lambda(k)} e^{ikx}.$$

Since $Re[\lambda(k)]$ are unbounded ($\Omega = +\infty$) there are values of k such that $|u(x, 0)|$ and $|u_t(x, 0)|$ are as small as we like. Thus the data in this initial value problem can be made arbitrarily small. However

$$|u(x, t)| = \frac{1}{\lambda^2(k)} e^{Re[\lambda(k)]t},$$

grows exponentially as a function of k and can for any fixed $t > 0$ be made arbitrarily large. This is because an algebraic decay $\frac{1}{\lambda^2(k)}$ is always overwhelmed by an exponential growth $e^{Re[\lambda(k)]t}$ (Basic calculus). Thus arbitrarily small data can be made arbitrarily large for any $t > 0$. Condition 3) on page 52 is not satisfied and consequently the problem is *not* well posed. If the case $\Omega = +\infty$ is realized, it is back to the drawing board, something is seriously wrong with the model! If $\Omega < +\infty$ one can show that the problem is well posed.

Let us look at some specific examples.

The equation

$$u_{xx} + u_{tt} + \rho u = 0,$$

where ρ is constant, is an elliptic equation. Our equation for $\lambda(k)$ for this case is

$$\lambda^2 + (-k^2 + \rho) = 0,$$

$$\Downarrow$$

$$\lambda(k) = \pm \sqrt{k^2 - \rho}.$$

$\lambda(k)$ is clearly unbounded when k ranges over \mathbb{R} . Thus $\Omega = +\infty$ and the Cauchy problem for the equation with data at $t = 0$ is not well posed.

The equation

$$\rho u_{xx} + u_t = 0,$$

where ρ is constant is a parabolic equation. In fact, we can also write it as

$$u_t = -\rho u_{xx}.$$

For $\rho < 0$ it is the heat equation. For $\rho > 0$ it is called *the reverse heat equation*. We get the reverse heat equation from the heat equation by making the transformation $t' = -t$. Increasing values of t' then corresponds to decreasing values of t . Solving the reverse heat equation thus corresponds to solving the heat equation backward in time.

We now get the following equation for $\lambda(k)$

$$-\rho k^2 + \lambda = 0 \Rightarrow \lambda = \rho k^2. \quad (152)$$

Clearly $\lambda(k)$ is unbounded for $\rho > 0 \Rightarrow \Omega = +\infty$ which implies that the reverse heat equation is not well posed. This says something profound about the process of heat diffusion or any other process described by the diffusion equation. Such processes are *irreversible*, they have a built-in direction of time. It is a well known fact that no fundamental physical theory is irreversible, they have no direction of time. So how come the diffusion equation has one? Where and how did the direction of time enter into the logical chain connecting microscopic reversible physics to macroscopic irreversible diffusion processes? This question has been discussed for more than a hundred years and no consensus has been reached. Nobody really knows.

For $\rho < 0$, $\lambda(k)$ is bounded above by zero $\Rightarrow \Omega = 0$ and the Cauchy problem for the heat equation is well-posed.

The equation

$$u_{tt} - u_{xx} + \rho u = 0,$$

where ρ is a constant, is a hyperbolic equation. The equation for $\lambda(k)$ is

$$\lambda^2 + k^2 + \rho = 0,$$

$$\Downarrow$$

$$\lambda = \pm i\sqrt{k^2 + \rho}.$$

For $\rho < 0$ the real part of λ vanishes for $|k|$ large enough and for $\rho > 0$ the real part vanishes for all k . Thus in either case $\Omega < +\infty$ and the Cauchy problem for this equation is well-posed.

The constant Ω is called the *stability index* for the equation. If $\Omega > 0$ there is some set or range of k values whose corresponding normal modes experience exponential growth. If this is the case the equation is said to be *unstable*. If $\Omega < +\infty$ the Cauchy problem is well posed, but the unbounded growth can still invalidate the PDE as a mathematical model for some $t > 0$. The reason for

this is that PDE models are usually derived using truncated Taylor expansions or expansions of some other sort. In order for these truncations to remain valid the solutions must remain small. This assumption will break down in the case of unbounded growth. If $\Omega < 0$ all normal modes decay exponentially. If this is the case the PDE is *strictly stable*. If $\Omega = 0$ there may be some modes that does not decay and will continue to play a role for all $t > 0$. If this is the case the PDE is *neutrally stable*.

Consider the parabolic equation

$$u_t - c^2 u_{xx} + au_x + bu = 0.$$

For this equation the equation for $\lambda(k)$ is

$$\lambda + c^2 k^2 + iak + b = 0,$$

$$\Downarrow$$

$$\lambda(k) = -c^2 k^2 - b - iak,$$

$$\text{Re}[\lambda(k)] = -b - c^2 k^2.$$

Clearly $\Omega = -b$. It is stable if $b > 0$ and unstable if $b < 0$ and neutrally stable for $b = 0$.

The equation

$$u_{xx} + 2u_{xt} + u_{tt} = 0,$$

is parabolic and has an equation for λ given by

$$-k^2 + 2ik\lambda + \lambda^2 = 0,$$

$$\Updownarrow$$

$$(\lambda + ik)^2 = 0,$$

$$\Updownarrow$$

$$\lambda(k) = -ik,$$

and thus $\Omega = 0$. The equation is therefore neutrally stable. It has a normal mode solution

$$u_1(x, t) = ae^{ik(x-t)}.$$

It is straight forward to verify that it also have a solution

$$u_2(x, t) = bte^{ik(x-t)}.$$

By linear superposition it has a solution

$$u(x, t) = ae^{ik(x-t)} + bte^{ik(x-t)}.$$

The solution to the Cauchy problem with start data

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

$$u_t(x, 0) = 1,$$

is then easily shown to be

$$u(x, t) = te^{-ikx}.$$

Thus, even if the data for this case is bounded, the solution grows without bound as t approach infinity. However, the growth is algebraic rather than exponential, which was the case for an unstable equation. This kind of behavior is generally true for the neutrally stable case; if there is an instability, then the solution grows algebraically rather than exponentially.

A special situation occur when $Re[\lambda(k)] = 0, \forall k$. Then

$$\begin{aligned} |u(x, t)| &= \left| a(k)e^{i(kx + Im[\lambda(k)]t)} \right| \\ &= |a(k)| = |u(x, 0)|. \end{aligned}$$

The amplitude of the normal mode is preserved for all time. In many cases of importance, $|u|^2$ is a measure of the *energy* of the normal mode and we have that the energy in the system is conserved in time.

In general, an equation where $Re[\lambda(k)] = 0$, is called *conservative*. For the conservative case we *define* the positive quantity

$$\omega = |Im[\lambda(k)]|,$$

to be the *frequency* of the mode. Given this definition we have

$$\lambda(k) = \pm i\omega(k),$$

depending on the sign of the imaginary part of the complex root $\lambda(k)$ of equation (147).

The normal mode then takes the form

$$u(x, t) = a(k)e^{i(kx \pm \omega(k)t)}. \quad (153)$$

The equation

$$\omega = \omega(k), \quad (154)$$

is called the *dispersion relation* for the equation. If

$$\omega''(k) \neq 0, \quad (155)$$

and thus $\omega(k)$ is not a linear function of k , the equation is said to be of *dispersive type*. The normal mode can, for the conservative case, be written as

$$u(x, t) = a(k)e^{i\theta}, \quad (156)$$

where $\theta = kx \pm \omega(k)t$ is called the *phase* of the normal mode. It is clear that for $k > 0$, (156) represents a wave moving from right to left (+) or from left to right (-), with speed

$$v_f(k) = \frac{\omega(k)}{k}. \quad (157)$$

The quantity v_f is called the *phase speed* of the wave. If $\omega''(k) = 0$ then

$$\omega(k) = ck$$

where c is some constant. For this case the phase speed is independent of k

$$v_f = \frac{\omega(k)}{k} = \frac{ck}{k} = c.$$

Thus all normal modes moves at the same speed.

For the case when $\omega''(k) \neq 0$, each normal mode moves at its own speed, and the linear superpositions (145) and (146) represents spreading or *dispersing* waves. For such solutions the phase speed is not the relevant quantity to consider. We will see later that for such waves the *group velocity*.

$$v_g = \frac{d\omega}{dk}, \quad (158)$$

is the relevant quantity.

If the stability index $\Omega \leq 0$ and $Re[\lambda(k)] < 0$ for all, except a finite number of k -values, the equation is said to be of *dissipative type*.

Let us consider the so-called Telegrapher's equation

$$u_{tt} - \gamma^2 u_{xx} + 2\hat{\lambda}u_t = 0, \quad \hat{\lambda} > 0.$$

This is a hyperbolic equation and the polynomial determining $\lambda(k)$ is

$$\lambda^2 + \gamma^2 k^2 + 2\hat{\lambda}\lambda = 0,$$

$$\Downarrow$$

$$\lambda(k) = -\hat{\lambda} \pm \sqrt{\hat{\lambda}^2 - \gamma^2 k^2}.$$

Since $\hat{\lambda} > 0$, $Re[\lambda(k)] < 0$, $\forall k \neq 0$. For $k = 0$, either $\lambda(0) = 0$ or $\lambda(0) = -2\hat{\lambda}$ so that

$$\Omega = 0.$$

The telegraphers equation is thus neutrally stable. It is also of dissipative type. When $\hat{\lambda} = 0$ we get the regular wave equation which is hyperbolic and conservation. For this reason the telegrapher's equation is also called the *damped wave equation*.

The *Klein-Gordon* equation takes the form

$$u_{tt} - \gamma^2 u_{xx} + c^2 u = 0.$$

The equation for $\lambda(k)$ is

$$\lambda^2 + \gamma^2 k^2 + c^2 = 0,$$

$$\Downarrow$$

$$\lambda(k) = \pm i\sqrt{\gamma^2 k^2 + c^2}.$$

The real part of $\lambda(k)$ is zero so the equation is of conservative type, also $\Omega = 0$ so it is neutrally stable. The dispersion relation is

$$\omega = \sqrt{\gamma^2 k^2 + c^2},$$

and clearly $\omega''(k) \neq 0$ so the Klein-Gordon equation is of dispersive type. If $\gamma = 0$ we get

$$\omega = ck.$$

and for this case the equation is not of dispersive type.

The Klein-Gordon equation was proposed as a relativistic extension of the Schrodinger equation of quantum mechanics. It was first found by Schrodinger but was rediscovered by Klein and Gordon. In both cases it was rejected as an equation for a relativistic electron on physical grounds. It predicted negative probability! Later it however reappeared in quantum field theory and is now one of the fundamental equations of nature. It for example describes the Higgs Boson.

8 Adjoint differential operators

Let us consider the differential operator

$$L = \frac{d^2}{dx^2}. \quad (159)$$

For any pair of functions u, w we have

$$\begin{aligned} uLw - wLu &= uw'' - wu'' \\ &= uw'' + u'w' - wu'' - u'w' \\ &= (uw' - wu')'. \end{aligned} \quad (160)$$

Thus $uLw - wLu$ is a total derivative. Identities like (160) are of outmost importance in the theory of PDEs. Here is another example. Let

$$L = \frac{d}{dx}, \quad (161)$$

and define a operator L^* by

$$L^* = -\frac{d}{dx}. \quad (162)$$

Then

$$uLw - wL^*u = uw' + wu' = (uw)'.$$

Thus we get a total derivative. Next consider an operator

$$L = \nabla^2. \quad (163)$$

Then from vector calculus we have

$$\begin{aligned} uLw - wLu &= u\nabla^2 w - w\nabla^2 u \\ &= \nabla \cdot (u\nabla w - w\nabla u). \end{aligned}$$

Thus we get a total derivative. In general if L is a differential operator and if we have, for some vector field \mathbf{F} , the identity

$$uLw - wL^*u = \nabla \cdot \mathbf{F},$$

then L^* is the *formal adjoint* of L . If $L = L^*$ the operator is *formally self-adjoint*. The operators (159), (163) are formally self-adjoint. The operator $L^* = -\frac{d}{dx}$ is the formal adjoint of $L = \frac{d}{dx}$.

Let D be some domain in \mathbb{R}^n and let L be a differential operator defined for functions whose domain include D . Let L^* be the formal adjoint of L . Then by the divergence theorem we have

$$\int_D dv \{uLw - wL^*u\} = \int_D dv \nabla \cdot \mathbf{F} = \int_S dA \mathbf{n} \cdot \mathbf{F}. \quad (164)$$

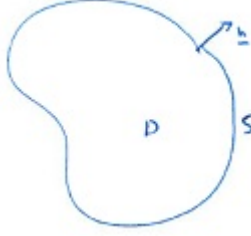


Figure 45

It is the possibility of reducing an integral over a domain to an integral over the boundary of the domain that is the reason why the formal adjoint of differential operators are important. We can use this possibility in many important ways. We will return to this topic many times in this and the next class (Mat-3202) in the spring.

For now observe the following:

Let $L = \nabla^2 \Rightarrow L^* = \nabla^2$. Then (164) is

$$\begin{aligned} \int_D dv \{uLw - wLu\} &= \int_S dA \mathbf{n} \cdot \{u\nabla w - w\nabla u\} \\ &= \int_S dA \{u\partial_{\mathbf{n}}w - w\partial_{\mathbf{n}}u\}, \end{aligned} \quad (165)$$

where $\partial_{\mathbf{n}} = \mathbf{n} \cdot \nabla$ is the normal derivative on the surface S . Let us assume that

the domain of L consists only of functions that satisfy the boundary conditions

$$\begin{aligned} u &= 0 \text{ on } S, \\ \text{or} \\ \partial_{\mathbf{n}} u &= 0 \text{ on } S. \end{aligned} \tag{166}$$

Then (165) implies that

$$\int_D dv \{uLw - wLu\} = 0, \tag{167}$$

which can be rewritten into the form

$$\int_D dv uLw = \int_D wLu. \tag{168}$$

Let V be the vector space of smooth functions defined on D and that satisfy the boundary conditions (166) (Prove that this is a vector space). On the vector space V , we can define an inner product using the formula

$$(u, w) = \int_D dv uw.$$

Then (168) show that

$$(u, Lw) = (Lu, w). \tag{169}$$

So, L is a self-adjoint operator on V . If we choose different boundary conditions so that the right hand side of (165) does not vanish then (169) will not hold and the operator is not self-adjoint in the sense of linear algebra. It is however still formally self-adjoint according to the definition on the previous page.

We know from linear algebra that self-adjoint operators have many good properties, we will discuss many of these in the next section of these lecture notes.

An important class of differential operators in the theory of PDEs consists of operators of the form

$$Lu = a_{ij}\partial_{x_i}\partial_{x_j}u + b_i\partial_{x_i}u + cu, \tag{170}$$

where we use the Einstein convention.

The formal adjoint of this operator is

$$L^*u = \partial_{x_i}\partial_{x_j}(a_{ij}u) - \partial_{x_i}(b_iu) + cu. \tag{171}$$

To see why this is so take the single term $wa_{11}\partial_{x_1}\partial_{x_1}u$. We have

$$\begin{aligned} wa_{11}\partial_{x_1}\partial_{x_1}u &= \partial_{x_1}(wa_{11}\partial_{x_1}u) - \partial_{x_1}(wa_{11})\partial_{x_1}u \\ &= \partial_{x_1}(wa_{11}\partial_{x_1}u) - \partial_{x_1}(u\partial_{x_1}(wa_{11})) + u\partial_{x_1}\partial_{x_1}(wa_{11}), \\ &\quad \Downarrow \\ wa_{11}\partial_{x_1}\partial_{x_1}u - \partial_{x_1}\partial_{x_1}(wa_{11})u &= \partial_{x_1}(wa_{11}\partial_{x_1}u - u\partial_{x_1}(wa_{11})). \end{aligned}$$

The other terms are treated in the same way and (171) appears.
 By carrying out all differentiations in (171) and comparing with (170) it is easy to verify that L is formally self-adjoint only if

$$b_i = \partial_{x_j} a_{ij}.$$

If this is the case we have

$$\begin{aligned} & a_{ij} \partial_{x_i} \partial_{x_j} u + b_i \partial_{x_i} u \\ &= a_{ij} \partial_{x_i} \partial_{x_j} u + \partial_{x_j} a_{ij} \partial_{x_i} u \\ &= a_{ij} \partial_{x_i} \partial_{x_j} u + \partial_{x_i} a_{ji} \partial_{x_j} u \\ &= a_{ij} \partial_{x_i} \partial_{x_j} u + \partial_{x_i} a_{ij} \partial_{x_j} u \\ &= \partial_{x_i} (a_{ij} \partial_{x_j} u), \end{aligned} \tag{172}$$

where we have used the fact that a_{ij} is a symmetric matrix. From (172) we see that a self-adjoint operator of the form (170) can be rewritten as

$$Lu = \partial_{x_i} (a_{ij} \partial_{x_j} u) + cu. \tag{173}$$

If the coefficients are constants (173) imply that a self-adjoint operator can not have any first derivative terms.

Let us look at some examples.

The operator

$$Lu = -\nabla \cdot (p \nabla u) + qu, \quad p > 0,$$

is an elliptic operator (Show this).

Observe that

$$\begin{aligned} wLu - uLw &= -w \nabla \cdot (p \nabla u) + quw + u \nabla \cdot (p \nabla w) - quw \\ &= \nabla \cdot (pu \nabla w - pw \nabla u), \end{aligned}$$

so L is formally self-adjoint.

The operator

$$\tilde{L}u = \rho u_{tt} + Lu,$$

is a hyperbolic operator (Show this).

Assuming that ρ is independent of time we observe that

$$\begin{aligned} w\tilde{L}u - u\tilde{L}w &= \rho w u_{tt} + wLu - \rho w u_{tt} - uLw \\ &= \partial_t (\rho w u_t - \rho u w_t) + \nabla \cdot (pu \nabla w - pw \nabla u) \\ &= \tilde{\nabla} \cdot (pu \nabla w - pw \nabla u, \rho w u_t - \rho u w_t), \end{aligned}$$

where $\tilde{\nabla} = (\nabla \cdot, \partial_t)$ is a 4D divergence operator. Thus \tilde{L} is formally self-adjoint.

Finally define the operator

$$\hat{L}u = \rho u_t + Lu,$$

where we still assume that ρ is independent of time. This operator is parabolic (Show this). Define another parabolic operator

$$\widehat{L}^* u = -\rho u_t + Lu.$$

Then

$$\begin{aligned} w\widehat{L}u - u\widehat{L}^*w &= \rho w u_t + wLu + \rho w u_t - uLw \\ &= \partial_t(\rho w) + \nabla \cdot (\rho u \nabla w - \rho w \nabla u) \\ &= \widetilde{\nabla} \cdot (\rho u \nabla w - \rho w \nabla u, \rho w). \end{aligned}$$

Thus \widehat{L}^* is the formal adjoint of \widehat{L} . Since $\widehat{L}^* \neq \widehat{L}$ the operator \widehat{L} is not formally self-adjoint.

The notion of a formal adjoint can be extended to a much larger class of operators[1] than (170).

9 Initial and boundary value problems in bounded domains

In this section we will introduce the standard approach for solving PDEs in bounded domains. This approach consists of expanding solutions into infinite series of eigenfunctions. The eigenfunctions are determined by the PDE and the boundary values. Eigenfunctions are generalizations of the normal modes introduced in the previous section. In fact eigenfunctions are often called normal modes.

This approach is very general and work for a large range of linear PDEs. We will see, at the end of this section, that they are also very useful when we investigate nonlinear PDEs.

In these lecture notes, we choose a certain level of generality. Our first restriction is to consider only scalar PDEs. Within the class of scalar PDEs we chose a subclass that is quite general and cover situations that occur in many applications in science and technology. In chapter 8 of [1] a much larger class of PDEs is investigated using similar methods to the ones we discuss here.

Let L be a second order differential operator defined by

$$Lu = -\nabla \cdot (p(\mathbf{x}) \nabla u) + q(\mathbf{x})u, \quad (174)$$

where $\mathbf{x} \in \mathbb{R}^2$ or \mathbb{R}^3 . For one space-dimension we define L by

$$Lu = -\partial_x(p(x)\partial_x u) + q(x)u. \quad (175)$$

We will in general assume that $p > 0$ and $q \geq 0$. The type of equations we want to consider are

Elliptic case:

$$Lu = \rho(x)F(\mathbf{x}), \quad \rho > 0. \quad (176)$$

Parabolic case:

$$\rho(\mathbf{x})\partial_t u + Lu = \rho(\mathbf{x})F(\mathbf{x}, t), \quad \rho > 0. \quad (177)$$

Hyperbolic case:

$$\rho(\mathbf{x})\partial_{tt} u + Lu = \rho(\mathbf{x})F(\mathbf{x}, t), \quad \rho > 0. \quad (178)$$

It is an important point here that the functions p, ρ and q does not depend on time.

These equations, of elliptic, parabolic and hyperbolic type appear as the first approximation to a large set of mathematical modelling problems in science and technology.

Let us consider the case of mass flow. Let D be some domain with boundary S .

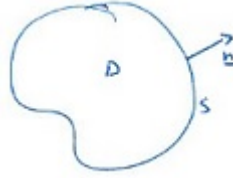


Figure 46

In D we have a concentration of some material substance given by $u(\mathbf{x}, t)$. The total mass of the material inside D is

$$M(t) = \int_D dv u(\mathbf{x}, t)\rho, \quad (179)$$

where ρ have the interpretation of mass density.

Mass of substance = concentration of substance \times mass density

This requirement explains the assumption $\rho > 0$ in equations (176),(177) and (178).

We have the balance equation for mass

$$\frac{dM}{dt} = - \int_S dS \mathbf{j} \cdot \mathbf{n} + \int_D dV H. \quad (180)$$

where \mathbf{j} is the mass flux and H is a source of mass. In general, $\mathbf{j} = \mathbf{j}(\mathbf{x}, u, \nabla u)$, but for low concentration and small gradients we can Taylor expand and find

$$\mathbf{j} \approx -p(\mathbf{x})\nabla u, \quad p > 0. \quad (181)$$

Here a constant term and a term only depending on u has been dropped on physical grounds: All mass flux is driven by gradients, and we have introduced a minus sign and assumed $p > 0$ in order to accommodate the fact that mass



Figure 47

always flow from high to low concentration. In general $H = H(\mathbf{x}, t, u)$ but for low concentration we can also here Taylor expand, and find

$$H = -q(\mathbf{x})u + \tilde{F}(\mathbf{x}, t), \quad q \geq 0. \quad (182)$$

For notational convenience we introduce F by

$$\tilde{F} = \rho F. \quad (183)$$

The minus sign in the first term is introduced to model a situation where high concentration leads to high loss. Imagine a sink hole where the amount leaving increase in proportion to the mass present.

Using (181), (182) and (183) in (180) we get

$$\begin{aligned} \int_V dv \rho \partial_t u &= \int_S dS p \nabla u \cdot \mathbf{n} + \int_D dV \{-qu + \rho F\}, \\ &\Downarrow \text{divergence theorem} \\ \int_V dV \{\rho \partial_t u - \nabla \cdot (p \nabla u) + qu - \rho F\} &= 0. \end{aligned}$$

This is assumed to hold for all domains D . Therefore we must have

$$\rho \partial_t u + Lu = \rho F,$$

where L is the operator defined in (174). In a similar way, using conservation of momentum, equation (178) can be shown to model vibrations in homogeneous solids, membranes and bars.

Our basic model equations must be subject to boundary conditions since the domains in this section are assumed to be bounded. In two and three spatial dimensions we pose the conditions

$$\alpha(\mathbf{x})u + \beta(\mathbf{x})\partial_{\mathbf{n}}u|_{\partial G} = B(\mathbf{x}, t), \quad (184)$$

where α, β and B are given functions defined on the boundary, ∂G , of G . In the elliptic case, (176), there is no time dependence and $B = B(\mathbf{x})$. For the parabolic and hyperbolic case G is a space-time domain of the form $G = G_{\mathbf{x}} \times [t_0, t_1]$



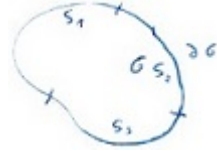
Figure 48

For most common situations occurring in applications, both α and β are positive. They can not both be zero at the same point since then the boundary condition (184) would not actually be a constraint!

There are 3 special cases of (184).

- i) $\alpha \neq 0, \beta = 0$ –Dirichelet condition or condition of the first kind.
- ii) $\alpha = 0, \beta \neq 0$ –Neuman condition or condition of the second kind.
- iii) $\alpha \neq 0, \beta \neq 0$ –Robin condition or condition of the third kind.

It is not uncommon that conditions of the first, second and third kind are specified at different parts of the boundary.



Conditions of type j are specified on $S_j, j = 1, 2, 3$.

Figure 49

For heat conduction problems, part of the boundary could have a given temperature (first kind) whereas other parts could be insulated (second kind). For the one dimensional case, ∂G consists of the two end points of the interval $G = (0, l)$. So $\partial G = \{0, l\}$. For the hyperbolic and parabolic case the general conditions (184) are

$$\begin{aligned} \alpha_1 u(0, t) - \beta_1 u_x(0, t) &= g_1(t), \\ \alpha_2 u(l, t) + \beta_2 u_x(l, t) &= g_2(t), \quad \alpha_j, \beta_j > 0. \end{aligned} \quad (185)$$

The minus sign at $x = 0$ occur because the unit normal at the point $x = 0$ is given by the number (-1).

If F and B in (176), (177), (178) and (184) are zero the corresponding boundary value problem is *homogeneous*.

9.1 Separation of variables

We will now apply the method of separation of variables to the equations and boundary conditions introduced in the previous section.

For the *hyperbolic case* we consider the homogeneous PDE

$$\rho(\mathbf{x})\partial_{tt}u + Lu = 0, \quad \mathbf{x} \in G, t > 0, \quad (186)$$

and homogeneous boundary conditions

$$\alpha(\mathbf{x})u + \beta(\mathbf{x})\partial_{\mathbf{n}}u|_{\partial G} = 0, \quad t > 0, \quad (187)$$

in 2D and 3D.

In 1D, the homogeneous condition is

$$\begin{aligned} \alpha_1 u(0, t) - \beta_1 u_x(0, t) &= 0, \\ \alpha_2 u(l, t) + \beta_2 u_x(l, t) &= 0, \quad t > 0. \end{aligned} \quad (188)$$

In order to get a completely determined problem we also must specify initial conditions

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad u_t(\mathbf{x}, 0) = g(\mathbf{x}) \quad \mathbf{x} \in G. \quad (189)$$

For the *parabolic case* we consider the homogeneous equation

$$\rho\partial_t u + Lu = 0, \quad \mathbf{x} \in G, t > 0, \quad (190)$$

with the boundary conditions (187) or (188) and the single initial condition

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad \mathbf{x} \in G. \quad (191)$$

In order to give a uniform treatment of the hyperbolic, parabolic and elliptic case, we introduce for the elliptic case a function $u = u(\mathbf{x}, y)$ where now \mathbf{x} is in \mathbb{R} or \mathbb{R}^2 . The variable y is scalar and restricted to the domain $0 < y < \widehat{l}$. The elliptic PDE can in terms of these variables be written as

$$\rho(\mathbf{x})\partial_{yy}u - Lu = 0, \quad \mathbf{x} \in G, 0 < y < \widehat{l}, \quad (192)$$

where ρ, p, q are functions of \mathbf{x} alone. The equation (192) is supplied with the boundary conditions (187) for the 2D case or (188) for the 1D case. The analog of the initial conditions in the time dependent case are the following conditions at $y = 0$ and $y = l$.

$$u(\mathbf{x}, 0) = f(\mathbf{x}), u(\mathbf{x}, \widehat{l}) = g(\mathbf{x}), \mathbf{x} \in G. \quad (193)$$

This assumed form for the elliptic equation certainly does not cover all possibilities that occur in applications. We observe for example that the domain always is a generalized cylinder

However, the assumed form makes a unified presentation for the hyperbolic, parabolic and elliptic case possible.

After all this preparatory work, let us start separating variables!

The method of separation of variables starts by looking for special solutions

$$u(\mathbf{x}, t) = M(\mathbf{x})N(t), \quad (194)$$

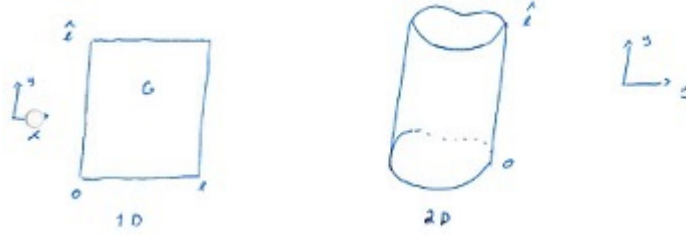


Figure 50

for (186) and (187) and

$$u(x, y) = M(\mathbf{x})N(y), \quad (195)$$

for (192). Inserting (194) and (195) into (186), (190) and (192) gives, upon division by NM

$$\frac{N''(t)}{N(t)} = -\frac{LM(\mathbf{x})}{\rho(\mathbf{x})M(\mathbf{x})}, \quad \text{hyperbolic case,} \quad (196)$$

$$\frac{N'(t)}{N(t)} = -\frac{LM(\mathbf{x})}{\rho(\mathbf{x})M(\mathbf{x})}, \quad \text{parabolic case,} \quad (197)$$

$$-\frac{N''(y)}{N(y)} = -\frac{LM(\mathbf{x})}{\rho(\mathbf{x})M(\mathbf{x})}, \quad \text{elliptic case.} \quad (198)$$

Since the two sides of equations (196) \rightarrow (198) depends on different variables, both sides must be equal to the same constant. This constant we write, by convention, in the form $-\lambda$. This give us in all three cases the following equation for M .

$$LM(\mathbf{x}) = \lambda\rho(\mathbf{x})M(\mathbf{x}), \quad (199)$$

and

$$N''(t) + \lambda N(t) = 0, \quad \text{hyperbolic case,} \quad (200)$$

$$N'(t) + \lambda N(t) = 0, \quad \text{parabolic case,} \quad (201)$$

$$N''(t) - \lambda N(y) = 0, \quad \text{elliptic case.} \quad (202)$$

The function $M(\mathbf{x})$ in equation (199) is subject to boundary conditions (187) or (188). Since both conditions are homogeneous, as are equation (199), $M(\mathbf{x}) = 0$ is a solution. This solution, also called the *trivial solution*, gives $u = 0$ and thus plays no role.

Defining $A = \frac{1}{\rho}L$, we observe that the structure of equation (199) is

$$AM = \lambda M$$

Thus M is an eigenvector of A and λ the corresponding eigenvalue. The vector space is the set of smooth functions defined on G and satisfying the homogeneous boundary conditions (187) or (188) on the boundary of G . (Show that this is a

vector space!) In this vector space, vectors are functions, so solutions of (199) are often called *eigenfunctions*. The big difference between the eigenvalue problems that are studied in elementary linear algebra and the current one, is that here, the vector space is infinite dimensional. This makes the problem much harder, much more interesting and also much more useful.

9.2 Self-adjoint and positive operators

The operator

$$L = -\nabla \cdot (p(\mathbf{x})\nabla u) + q(\mathbf{x})u, \quad (203)$$

occurring in equation (199) and determining the functions $M(\mathbf{x})$ in the separation

$$u = NM,$$

for the hyperbolic case (186), the parabolic case (190) and the elliptic case (192), is very special. We will now determine some of the properties of the *spectrum* of L . The spectrum of a linear operator is the collection of eigenvalues and corresponding eigenspaces of the operator. Recall that the eigenspace corresponding to a given eigenvalue consists of all eigenvectors of the operator that has the same eigenvalue.

Using vector calculus we observe that

$$\begin{aligned} w\nabla \cdot (p\nabla u) &= \nabla \cdot (pw\nabla u) - p\nabla w \cdot \nabla u, \\ u\nabla \cdot (p\nabla w) &= \nabla \cdot (pu\nabla w) - p\nabla u \cdot \nabla w, \\ \Rightarrow w\nabla \cdot (p\nabla u) - u\nabla \cdot (p\nabla w) &= \nabla \cdot (pw\nabla u - pu\nabla w), \\ &\Downarrow \\ wLu - uLw &= \nabla \cdot (pu\nabla w - pw\nabla u). \end{aligned} \quad (204)$$

Thus L is formally self-adjoint.

Let G be a domain in \mathbb{R}^2 or \mathbb{R}^3 . Integrating (204) over G gives

$$\int_G dV \{wLu - uLw\} = \int_{\partial G} dS p(u\partial_{\mathbf{n}}w - w\partial_{\mathbf{n}}u). \quad (205)$$

Let us assume that both u and w satisfy the boundary conditions (184) with $B = 0$. Thus

$$\begin{aligned} \alpha(\mathbf{x})u + \beta(\mathbf{x})\partial_{\mathbf{n}}u|_{\partial G} &= 0, \\ \alpha(\mathbf{x})w + \beta(\mathbf{x})\partial_{\mathbf{n}}w|_{\partial G} &= 0. \end{aligned} \quad (206)$$

For a given u, w that satisfy (206) we can think of (206) as a linear system for α and β

$$\begin{bmatrix} u & \partial_{\mathbf{n}}u \\ w & \partial_{\mathbf{n}}w \end{bmatrix} \Big|_{\partial G} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 0. \quad (207)$$

But our assumptions on the functions α and β imply that they can not both be zero at any point. Thus, the system (207) must have a nonzero solution. This imply that the determinant of the matrix defining the linear system (207), must be zero or

$$(u \partial_{\mathbf{n}} w - w \partial_{\mathbf{n}} u)|_{\partial G} = 0. \quad (208)$$

Thus, the right hand side of (205) is zero

$$\Rightarrow \int_G dV \{wLu - uLw\} = 0. \quad (209)$$

On the vector space, V , consisting of real valued functions defined on G and satisfying the homogeneous boundary conditions (206), we define an inner product

$$(f, g) = \int_G dV \rho(\mathbf{x}) f(\mathbf{x}) g(\mathbf{x}). \quad (210)$$

Given this inner product we define the *norm* of functions in V as

$$\|f\| = ((f, f))^{1/2} = \sqrt{\int_G dV \rho(\mathbf{x}) f^2(\mathbf{x})}. \quad (211)$$

Assuming f is continuous we have

$$\|f\| = 0 \Leftrightarrow f = 0. \quad (212)$$

Thus, what we have defined is a norm in the usual sense of linear algebra. Using this inner product (209) can be written as

$$(w, \frac{1}{\rho} Lu) = (\frac{1}{\rho} Lw, u).$$

Thus the operator $\tilde{L} = \frac{1}{\rho} L$ is self-adjoint in the sense of linear algebra. In the most common cases when $\rho = 1$, L itself is self-adjoint.

Let now M_m, M_l be eigenfunctions of L corresponding to *different* eigenvalues, $\lambda_k \neq \lambda_l$. Since both M_k and M_l satisfy the homogeneous boundary conditions, (206), we have

$$\int_G dV \{M_k L M_l - M_l L M_k\} = 0,$$

but

$$\begin{aligned} LM_k &= \rho \lambda_k M_k, \\ LM_l &= \rho \lambda_l M_l. \end{aligned}$$

Thus

$$\begin{aligned} 0 &= \int_G dV \{M_k L M_l - M_l L M_k\} = \int_G dV \rho(\lambda_l - \lambda_k) M_k M_l \\ &= (\lambda_l - \lambda_k)(M_k, M_l). \end{aligned}$$

But $\lambda_k \neq \lambda_l \Rightarrow \lambda_k - \lambda_l \neq 0$, and therefore we must have $(M_k, M_l) = 0$.

"Eigenfunctions corresponding to different eigenvalues are orthogonal."

Actually we have in our choice of inner product assumed that the eigenfunctions are real valued and the spectrum is real. This is in fact true:

Let φ be a complex valued function. Then, redoing the argument leading up to (205) using $w = u^*$ (the complex conjugate of u), we get

$$\int_G dV \{u(Lu)^* - (Lu)u^*\} = \int_{\partial G} dS \rho \{u^* \partial_{\mathbf{n}} u - u \partial_{\mathbf{n}} u^*\}, \quad (213)$$

and (208) becomes

$$\{u^* \partial_{\mathbf{n}} u - u \partial_{\mathbf{n}} u^*\}|_{\partial G} = 0,$$

\Downarrow

$$\int_G dV \{uLu^* - Luu^*\} = 0.$$

Let now λ be a complex eigenvalue and M the corresponding complex valued eigenfunction. Then (9.2) imply that

$$\begin{aligned} 0 &= \int_G dV \{M(\rho\lambda M)^* - \rho\lambda M M^*\} \\ &= \int_G dV \{\rho\lambda^* M M^* - \rho\lambda M M^*\} \\ &= (\lambda^* - \lambda) \int_G dV \rho M M^*. \end{aligned}$$

But

$$\int_G dV \rho |M|^2 = 0 \Rightarrow M = 0,$$

so we must have

$$\lambda^* = \lambda,$$

which means that the eigenvalues are *real*. The eigenfunctions can then always be chosen to be real.

Let now u be a (real valued) function that satisfies the boundary condition (184) with $B = 0$. Thus $u \in V$, as defined on page 112. Then we have

$$\begin{aligned} \int_G dV uLu &= - \int_G dV \{ \nabla \cdot (pu \nabla u) - p(\nabla u)^2 \} + \int_G dV qu^2 \\ &= \int_G dV \{ p(\nabla u)^2 + qu^2 \} - \int_{\partial G} dS pu \partial_{\mathbf{n}} u. \end{aligned}$$

Split the boundary ∂G into three pieces S_1, S_2 and S_3 , where on S_i the boundary conditions are of kind i , where $i = 1, 2, 3$ (see page 108). Then on $S_1 \cup S_2, pu \partial_{\mathbf{n}} u = 0$ and on $S_3, pu \partial_{\mathbf{n}} u = -p \frac{\alpha}{\beta} u^2$

$$\Rightarrow \int_G dV uLu = \int_G dV \{ p(\nabla u)^2 + qu^2 \} + \int_{S_3} dS p \frac{\alpha}{\beta} u^2. \quad (214)$$

The assumptions we have put on p, α, β and q earlier implies that the right hand side of (214) is positive. Thus, using the inner product defined on page 112, we have found that

$$\left(\frac{1}{\rho} Lu, u \right) \geq 0. \quad (215)$$

This means, by definition, that $\frac{1}{\rho} L$ is a *positive* operator.

Let now M_k be an eigenfunction corresponding to the eigenvalue λ_k . From (215) we have

$$\left(M_k, \frac{1}{\rho} LM_k \right) = \int_G dV M_k LM_k = \int_G dV M_k \rho \lambda_k M_k = \lambda_k \int_G dV \rho M_k^2 \geq 0,$$

and

$$\int_G dV \rho M_k^2 > 0.$$

Thus we must have

$$\lambda_k \geq 0.$$

Therefore, the spectrum of L is not only real but non-negative.

One can prove[1] that there are countably infinite many eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ and that ∞ is the only limit point of the sequence of eigenvalues. Recall that a limit point for a sequence $\{a_n\}$ is a point $a^* \in \mathbb{R}$ such that $\forall \epsilon > 0$ there exists infinitely many n such that

$$|a_n - a^*| < \epsilon.$$

Thus the eigenvalues only "accumulate" around $\lambda = \infty$.

Let now $\{\lambda_k\}$ be the countably infinite set of eigenvalues. One can also prove that there are at most a finite set of linearly independent eigenvectors corresponding to each λ_k . We can assume that these has been orthogonalized using the Gram-Schmidt orthogonalization process.

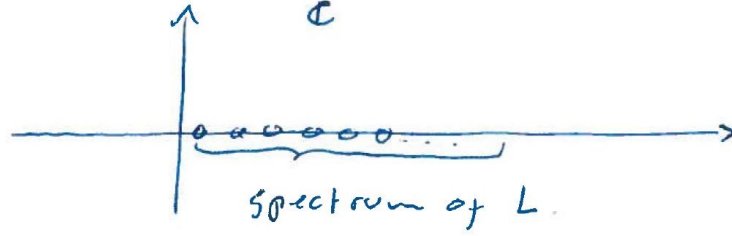


Figure 51

Let us in the following for simplicity only consider the case when each eigenspace is of dimension one and where all k , $\lambda_k > 0$. Thus $\lambda = 0$ is not an eigenvalue. Then the equations (200-202) for the function N gives upon solution the following:

$$N_k(t) = a_k \cos \sqrt{\lambda_k} t + b_k \sin \sqrt{\lambda_k} t, \quad \text{hyperbolic case,} \quad (216)$$

$$N_k(t) = a_k e^{\lambda_k t}, \quad \text{parabolic case,} \quad (217)$$

$$N_k(y) = a_k e^{\sqrt{\lambda_k} y} + b_k e^{-\sqrt{\lambda_k} y}, \quad \text{elliptic case.} \quad (218)$$

Here, a_k, b_k are arbitrary constants, one for each eigenvalue $\lambda_1, \lambda_2, \dots$. We now have infinitely many solutions

$$u_k = N_k M_k,$$

that solves equations (186), (190) or (192) and the boundary conditions (187) or (188). These functions will however not in general satisfy the initial conditions. In order to also satisfy the initial conditions (183) for the hyperbolic case, (191) for the parabolic case and (193) for the elliptic case we consider formal superposition

$$u = \sum_{k=1}^{\infty} u_k = \sum_{k=1}^{\infty} M_k N_k. \quad (219)$$

For the hyperbolic case we get the conditions

$$u(\mathbf{x}, 0) = \sum_{k=1}^{\infty} M_k(\mathbf{x}) N_k(0) = \sum_{k=1}^{\infty} a_k M_k(\mathbf{x}) = f(\mathbf{x}), \quad (220)$$

$$u_t(\mathbf{x}, 0) = \sum_{k=1}^{\infty} M_k(\mathbf{x}) N'_k(0) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} b_k M_k(\mathbf{x}) = g(\mathbf{x}). \quad (221)$$

(220) and (221) require us to find *eigenfunction expansions* for the functions f and g in terms of the eigenfunctions $\{M_k\}_{k=1}^{\infty}$. In what sense f and g can be represented by their eigenfunction expansions is an important, and in general very hard, question that will be discussed more fully only for the 1D case in this class.

You might recall from elementary calculus that for infinite series of functions, several different notions of convergence are available. The most important ones are pointwise convergence, uniform convergence and mean-square convergence. We will see that the last type of convergence will hold for very mild conditions on f and g . However, mean-square convergence is not enough to ensure that (219) is an actual, honest to God, solution to our initial-boundary value problem (186),(187),(189). For this to be the case we need a stronger convergence than mean-square. This stronger form of convergence will allow us to differentiate the formal superposition series (219) term by term. If this is possible, we have a *classical* solution. If it is not possible, (219) is a *generalized solution* that might still carry important information about the problem.

In order to express the coefficients $\{a_k\}$ and $\{b_k\}$ in terms of f and g , we use the fact that the eigenfunctions $\{M_k\}$ are orthogonal for $k \neq k'$

$$\begin{aligned} (f, M_j) &= \sum_{k=1}^{\infty} a_k (M_k, M_j) = a_j (M_j, M_j), \\ \Rightarrow a_j &= \frac{(f, M_j)}{(M_j, M_j)}. \end{aligned} \quad (222)$$

In the same way we find

$$b_j = \frac{(g, M_j)}{\sqrt{\lambda_j} (M_j, M_j)}. \quad (223)$$

The *formal* solution to the hyperbolic initial-boundary value problem (186-188) is then

$$u(\mathbf{x}, t) = \sum_{k=1}^{\infty} (a_k \cos \sqrt{\lambda_k} t + b_k \sin \sqrt{\lambda_k} t) M_k(\mathbf{x}), \quad (224)$$

where $\{a_k\}$ and $\{b_k\}$ are given by (222) and (223).

In a similar way, the solution of the initial-boundary problem for the parabolic case is

$$u(\mathbf{x}, t) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t} M_k(\mathbf{x}), \quad (225)$$

where $\{a_k\}$ is given by (222).

For the elliptic case we get the conditions

$$\begin{aligned} u(\mathbf{x}, 0) &= \sum_{k=1}^{\infty} (a_k + b_k) M_k(\mathbf{x}) = f(\mathbf{x}), \\ u(\mathbf{x}, \tilde{l}) &= \sum_{k=1}^{\infty} (a_k e^{\sqrt{\lambda_k} \tilde{l}} + b_k e^{-\sqrt{\lambda_k} \tilde{l}}) M_k(\mathbf{x}) = g(\mathbf{x}). \end{aligned}$$

Proceeding as for the hyperbolic case we get

$$\begin{aligned} a_k + b_k &= (f, M_k), \\ e^{\sqrt{\lambda_k} \tilde{l}} a_k + e^{-\sqrt{\lambda_k} \tilde{l}} b_k &= (g, M_k). \end{aligned} \quad (226)$$

These equations can be solved to give unique values for $\{a_k\}$ and $\{b_k\}$. Our formal solution for the elliptic case is then

$$u(\mathbf{x}, y) = \sum_{k=1}^{\infty} (a_k e^{\sqrt{\lambda_k} y} + b_k e^{-\sqrt{\lambda_k} y}) M_k(\mathbf{x}). \quad (227)$$

We will now take a closer look at the formal solutions (224), (225) and (227) for the 1D case.

10 The Sturm-Liouville problem and Fourier series

The 1D version of the eigenvalue problem (199) + (187) is

$$L[u(x)] = -\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] + q(x)u(x) = \lambda \rho(x)u(x), \quad (228)$$

for $0 < x < l$ and boundary conditions

$$\begin{aligned} \alpha_1 u(0) - \beta_1 u'(0) &= 0, \\ \alpha_2 u(l) + \beta_2 u'(l) &= 0. \end{aligned} \quad (229)$$

The problem defined by (228) and (229) is known as the *Sturm-Liouville problem*. If we require the following additional conditions

$$\begin{aligned} i) \quad & p(x) > 0, \rho(x) > 0, q(x) \geq 0, \quad 0 \leq x \leq l, \\ ii) \quad & p(x), p'(x), \rho(x), q(x) \text{ are continuous for } 0 \leq x \leq l, \\ iii) \quad & \alpha_i \geq 0, \beta_i \geq 0, \alpha_i + \beta_i > 0, \end{aligned} \quad (230)$$

to hold, (228), (230) define a *regular Sturm-Liouville* problem. If we relax one or more condition on the coefficient functions p, q and ρ we get more generalized versions of the Sturm-Liouville problem. For example, one can allow the functions p and/or ρ to vanish at the endpoints $x = 0, x = l$. We then get a *singular Sturm-Liouville* problem. This generalization do occur in important applications of this theory.

Before we start developing the theory, let us introduce some important constructions. These constructions have already been introduced in the more general setting discussed earlier, but will be repeated here in the 1D context.

For real valued functions on $0 \leq x \leq l$ which are bounded and integrable, we introduce the *inner product*

$$(\varphi, \psi) = \int_0^l \rho(x) \varphi(x) \psi(x) dx. \quad (231)$$

Observe that (231) is evidently bilinear and symmetric. For continuous functions we have

$$\int_0^l \rho(x) \varphi(x)^2 dx = 0, \Rightarrow \varphi(x) = 0 \quad \forall x.$$

Thus, (231) really defines an inner product in the sense of linear algebra on the linear vector space consisting of continuous functions on $0 \leq x \leq l$.

Introduce the norm of a function by

$$\|\varphi\| = \left(\int_0^l \rho(x) \varphi^2(x) dx \right)^{1/2}. \quad (232)$$

A function such that $\|\varphi\| < \infty$ is said to be *square-integrable*. Furthermore, a square-integrable function is *normalized to unity* if

$$\|\varphi\| = 1. \quad (233)$$

If φ is any square-integrable function we can *normalize* this function by defining a new function

$$\begin{aligned} \hat{\varphi} &= \frac{1}{\|\varphi\|} \varphi, \\ &\Downarrow \\ \|\hat{\varphi}\| &= \left\| \frac{1}{\|\varphi\|} \varphi \right\| = \frac{1}{\|\varphi\|} \|\varphi\| = 1. \end{aligned}$$

Two functions φ and ψ are *orthogonal* on the interval $0 < x < l$ if

$$(\varphi, \psi) = 0. \quad (234)$$

A set of functions $\{\varphi_k(x)\}$, $k = 1, 2, \dots$ is said to be an *orthogonal set* if

$$(\varphi_k, \varphi_j) = 0, \quad k \neq j. \quad (235)$$

If also $\|\varphi_k\| = 1$ for all k , the set is *orthonormal*. An orthogonal set can always be turned into an orthonormal set by using the approach described above.

If we allow the functions to assume complex values, we replace the inner product (231) by

$$(\varphi, \psi) = \int_0^l \rho(x) \varphi(x) \overline{\psi(x)} dx, \quad (236)$$

where for any complex number z , \bar{z} is its complex conjugate. Observe that we have bilinearity and

$$(\varphi, \psi) = \overline{(\psi, \varphi)} \quad (237)$$

Furthermore for any φ

$$(\varphi, \varphi) = \int_0^l \rho(x) |\varphi(x)|^2 dx \geq 0, \quad (238)$$

and if φ is, say, continuous then

$$(\varphi, \varphi) = 0 \Rightarrow \varphi(x) = 0, \quad \forall x. \quad (239)$$

Thus (236) defines a *Hermitian* inner product in the sense of linear algebra.

Let us return to the real valued case. Let an orthonormal set of functions $\{\varphi_k\}$ be given and let φ be a square-integrable function on $0 < x < l$. Then the infinite set of numbers (φ, φ_k) are called the *Fourier coefficients* of φ with respect to $\{\varphi_k\}$. The formal series

$$\sum_{k=1}^{\infty} (\varphi, \varphi_k) \varphi_k, \quad (240)$$

is called the *Fourier series* of φ with respect to the set $\{\varphi_k\}$. Note that we use the summation sign in (240) in a formal sense - we have at this point not assumed anything about the convergence of the infinite series of functions (240).

The name, Fourier series, is taken from the special case when $\{\varphi_k\}$ is a set of trigonometric functions. Often in the literature the term Fourier series refer *only* to this trigonometric case, but we will use the term Fourier series for all formal series (240).

So, in what sense does (240) converge? And if it converge, does it converge to the function φ that we used to construct the formal series?

Let us start by considering a finite sum,

$$S_N = \sum_{k=1}^N (\varphi, \varphi_k) \varphi_k.$$

For this sum we have

$$\begin{aligned} \|\varphi - \sum_{k=1}^N (\varphi, \varphi_k) \varphi_k\|^2 &= (\varphi - \sum_{k=1}^N (\varphi, \varphi_k) \varphi_k, \varphi - \sum_{j=1}^N (\varphi, \varphi_j) \varphi_j) \\ &= (\varphi, \varphi) - \sum_{k=1}^N (\varphi, \varphi_k)^2 - \sum_{j=1}^N (\varphi_j, \varphi_j)^2 + \sum_{k,j=1}^N (\varphi, \varphi_k)(\varphi, \varphi_j)(\varphi_k, \varphi_j) \\ &= \|\varphi\|^2 - 2 \sum_{k=1}^N (\varphi, \varphi_k)^2 + \sum_{k=1}^N (\varphi, \varphi_k)^2 \\ &= \|\varphi\|^2 - \sum_{k=1}^N (\varphi, \varphi_k)^2. \end{aligned}$$

But the norm of any function is non-negative. Thus we must have

$$\begin{aligned}\|\varphi\|^2 - \sum_{k=1}^N (\varphi, \varphi_k) &\geq 0, \\ \Downarrow \\ \sum_{k=1}^N (\varphi, \varphi_k) &\leq \|\varphi\|^2.\end{aligned}$$

This inequality says that *all* partial sums of the *numerical series* $\sum_{k=1}^N (\varphi, \varphi_k)^2$, consisting entirely of positive terms, are bounded by $\|\varphi\|^2$. It then follows, from the fact that all increasing bounded sequences are convergent that the series $\sum_{k=1}^N (\varphi, \varphi_k)^2$ is *convergent* and that

$$\sum_{k=1}^{\infty} (\varphi, \varphi_k)^2 \leq \|\varphi\|^2, \quad (241)$$

holds. (241) is known as *Bessel's inequality*.

A sequence of square-integrable functions $\{S_N\}_{N=1}^{\infty}$ is said to converge to a function $\varphi(x)$ *in the mean* if

$$\lim_{N \rightarrow \infty} \|\varphi - S_N\| = 0.$$

This kind of convergence does not in general imply pointwise convergence. It is called *mean square* convergence. Observe that if we have equality in Bessel's inequality (241),

$$\sum_{k=1}^{\infty} (\varphi, \varphi_k)^2 = \|\varphi\|^2, \quad (242)$$

then we have

$$\begin{aligned}\lim_{N \rightarrow \infty} \|\varphi - S_N\|^2 &= \lim_{N \rightarrow \infty} \left\| \varphi - \sum_{k=1}^N (\varphi, \varphi_k) \varphi_k \right\|^2 \\ &= \lim_{N \rightarrow \infty} \left\{ \|\varphi\|^2 - \sum_{k=1}^N (\varphi, \varphi_k)^2 \right\} = 0.\end{aligned}$$

(242) is called *Parseval's equality*. The argument we just did shows that if Parseval's equality holds, for a given φ and given orthonormal set $\{\varphi_k\}$, then the Fourier series (240) converges to φ in the mean.

An orthonormal set $\{\varphi_k\}_{k=1}^{\infty}$ is said to be *complete* if the Parseval equality (242) holds for all square integrable functions. We have proved that if an orthonormal set of functions is complete then the formal Fourier series (240) for any square integrable functions φ , converge in the mean square sense to the function φ .

10.1 Properties of eigenvalues and eigenfunctions

The properties of the eigenvalues and eigenspaces are the same as for the 2D and 3D case discussed earlier. After all, they follow from structural properties of L and the boundary conditions that are completely analogous in 1D, 2D and 3D. Let us just record these properties here:

- i) Eigenfunctions corresponding to different eigenvalues are orthogonal.
- ii) The eigenvalues are real and non-negative and the eigenfunctions can be chosen to be real.
- iii) Each eigenvalue is simple. This means that each eigenspace is of dimension one.
- iv) There is a countable infinity of eigenvalues having a limit point at infinity. The set of eigenvalues can be arranged as

$$0 \leq \lambda_1 < \lambda_2 < \lambda_3 < \dots,$$

with $\lim_{k \rightarrow \infty} \lambda_k = \infty$.

- v) The set of eigenfunctions $\{u_k(x)\}_{k=1}^{\infty}$ form a complete orthonormal set of square integrable functions. Thus, for any square integrable function on $0 < x < l$, the Fourier series

$$u = \sum_{k=1}^{\infty} (u, u_k) u_k,$$

converge to u in the mean.

- vi) If $u(x)$ is continuous on $0 \leq x \leq l$ and has a piecewise continuous first derivative on $0 \leq x \leq l$ and $u(x)$ satisfies the boundary conditions (229) then the Fourier series (240) converge absolutely and uniformly to $u(x)$ on $0 \leq x \leq l$.
- vii) If $u(x)$ has some jump discontinuity at a point x_0 in the interior of the interval $0 \leq x \leq l$ then (240) converge to $\frac{1}{2}(u(x_0^-) + u(x_0^+))$ at this point.

Here

$$u(x_0^-) = \lim_{x \rightarrow x_0^-} u(x), \quad (\text{approaching from the left}),$$

$$u(x_0^+) = \lim_{x \rightarrow x_0^+} u(x), \quad (\text{approaching from the right}),$$

so the only thing that remains to do is to actually find the eigenvalues and eigenfunctions for particular cases of interest. We will do so for a few important cases.

All regular Sturm-Liouville problems can be simplified if we use the following uniform approach:

Let $v(x; \lambda)$, $w(x; \lambda)$ be solutions of the initial value problem

$$-\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] + q(x)u(x) = \lambda + \rho(x)u(x), \quad (243)$$

with initial conditions

$$\begin{aligned} v(0; \lambda) &= 1, & w(0; \lambda) &= 0, \\ v'(0; \lambda) &= 0, & w'(0; \lambda) &= 1. \end{aligned} \quad (244)$$

Then it is easy to verify that the function

$$u(x; \lambda) = \beta_1 v(x; \lambda) + \alpha_1 w(x; \lambda), \quad (245)$$

is a solution to (243) and satisfies the boundary condition

$$\alpha_1 u(0; \lambda) - \beta_1 u'(0; \lambda) = 0.$$

In order to also satisfy the boundary condition at $x = l$ we must have

$$\alpha_2 \beta_1 v(l; \lambda) + \alpha_2 \alpha_1 w(l; \lambda) + \beta_2 \beta_1 v'(l; \lambda) + \beta_2 \alpha_1 w'(l; \lambda) = 0. \quad (246)$$

The eigenvalues are solutions of equation (246) and the corresponding eigenfunctions are determined by (245).

We will now apply the uniform approach to several cases leading to various trigonometric eigenfunctions and then turn to a couple of important examples of singular Sturm-Liouville problems.

10.1.1 Trigonometric eigenfunctions

For this case we choose $p(x) = \rho(x) = 1$ and $q(x) = 0$. Equation (226) in the Sturm-Liouville problem for this case reduces to

$$-u''(x) = \lambda u(x). \quad (247)$$

It is easy to verify that the functions

$$v(x; \lambda) = \cos(\sqrt{\lambda}x), \quad (248)$$

$$w(x; \lambda) = \frac{\sin(\sqrt{\lambda}x)}{\sqrt{\lambda}}, \quad (249)$$

satisfy (247) and the initial values (244). Here we assume $\lambda > 0$. We leave the case $\lambda = 0$ as an exercise for the reader. The eigenfunction for $\lambda > 0$ is found from (245) and is

$$u(x) = \beta_1 \cos(\sqrt{\lambda}x) + \alpha_1 \frac{\sin(\sqrt{\lambda}x)}{\sqrt{\lambda}}. \quad (250)$$

The eigenvalues are determined from the equation (229), which for this case simplifies into

$$\sqrt{\lambda}(\alpha_1\beta_2 + \beta_1\alpha_2) \cos(\sqrt{\lambda}l) + (\alpha_2 + \alpha_1 - \lambda\beta_2\beta_1) \sin(\sqrt{\lambda}l) = 0. \quad (251)$$

We will get different sets of eigenvalues and eigenfunctions depending on which boundary conditions (227) we assume for the Sturm-Liouville problem (247).

Fourier Sine series Assume $\alpha_1 = \alpha_2 = 1, \beta_1 = \beta_2 = 0$. Thus the boundary conditions are

$$u(0) = u(l) = 0. \quad (252)$$

Our equation (251) then reduces to

$$\sin(\sqrt{\lambda}l) = 0,$$

which gives us the following nonzero eigenvalues

$$\lambda_k = \left(\frac{\pi k}{l}\right)^2, \quad k = 1, 2, 3, \dots, \quad (253)$$

and the corresponding normalized eigenfunctions are from (250)

$$u_k(x) = \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l}x\right), \quad k = 1, 2, \dots. \quad (254)$$

For this case $\lambda = 0$ is not an eigenvalue because for $\lambda = 0$ equation (247) reduces to

$$\begin{aligned} -u''(x) &= 0, \\ \Downarrow \\ u(x) &= Ax + B, \end{aligned}$$

$$\left. \begin{aligned} u(0) = 0 &\Rightarrow B = 0 \\ u(l) = 0 &\Rightarrow A = 0 \end{aligned} \right\} \Rightarrow u(x) = 0.$$

The expansion of a function in a series

$$u = \sum_{k=1}^{\infty} (u, u_k) u_k,$$

is called the *Fourier Sine series*. Using explicit expressions for the eigenfunction and the inner product we get the following formulas for the Fourier Sine series

$$u(x) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k \sin \frac{\pi k}{l} x,$$

where

$$a_k = \sqrt{\frac{2}{l}} \int_0^l dx u(x) \sin \frac{\pi k}{l} x.$$

Fourier Cosine series Assume $\alpha_1 = \alpha_2 = 0$, $\beta_1 = \beta_2 = 1$. Thus the boundary conditions are

$$u'(0) = u'(l) = 0. \quad (255)$$

Our equation (251) now reduces to

$$\lambda \sin \sqrt{\lambda} l = 0.$$

$\lambda_0 = 0$ is clearly a solution. For $\lambda_0 = 0$ our equation reduces to

$$\begin{aligned} -u''(x) &= 0, \\ \Downarrow \\ u(x) &= Ax + B, \end{aligned}$$

$$\left. \begin{aligned} u'(0) = 0 &\Rightarrow A = 0 \\ u'(l) = 0 &\Rightarrow A = 0 \end{aligned} \right\} \Rightarrow u(x) = B.$$

Thus, $\lambda = 0$ is an eigenvalue for this case. The corresponding normalized eigenfunction is

$$u_0(x) = \frac{1}{\sqrt{l}}. \quad (256)$$

For $\lambda > 0$ we can divide the eigenvalue equation by λ and get

$$\sin(\sqrt{\lambda} l) = 0, \quad (257)$$

$$\Downarrow \quad (258)$$

$$\lambda_k = \left(\frac{\pi k}{l} \right)^2, \quad k = 1, 2, \dots \quad (259)$$

The corresponding normalized eigenfunctions are

$$u_k(x) = \sqrt{\frac{2}{l}} \cos \left(\frac{\pi k}{l} x \right), \quad k = 1, 2, \dots \quad (260)$$

The series expansion

$$u = (u, u_0)u_0 + \sum_{k=1}^{\infty} (u, u_k) u_k,$$

is the *Fourier Cosine* series. Using explicit expression for the eigenfunctions and the inner product we get

$$u(x) = \sqrt{\frac{1}{l}}a_0 + \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k \cos \frac{\pi k}{l}x,$$

where

$$a_0 = \sqrt{\frac{1}{l}} \int_0^l dx u(x),$$

$$a_k = \sqrt{\frac{2}{l}} \int_0^l dx u(x) \cos \frac{\pi k}{l}x.$$

Note that all the general properties of eigenvalues and eigenfunctions for Sturm-Liouville problems can be verified to hold for the eigenvalues and eigenfunctions we have found for the Fourier Sine and Fourier Cosine cases.

Fourier series In this case we consider a slightly generalized Sturm-Liouville problem. We assume that the domain is $(-l, l)$ instead of $(0, l)$. We still make the assumptions $p(x) = \rho(x) = 1$, $q(x) = 0$, so the eigenvalue equation is

$$-u''(x) = \lambda u(x), \quad -l < x < l. \quad (261)$$

Instead of using the boundary conditions (227) we use *periodic* boundary conditions

$$u(-l) = u(l), \quad (262)$$

$$u'(-l) = u'(l). \quad (263)$$

The general solution of (261) is for $\lambda > 0$

$$u(x) = a \cos \sqrt{\lambda}x + b \sin \sqrt{\lambda}x, \quad (264)$$

$$\Downarrow \quad (265)$$

$$u'(x) = -a\sqrt{\lambda} \sin \sqrt{\lambda}x + b\sqrt{\lambda} \cos \sqrt{\lambda}x,$$

and

$$u(-l) = u(l) \Rightarrow a \cos \sqrt{\lambda}l - b \sin \sqrt{\lambda}l = a \cos \sqrt{\lambda}l + b \sin \sqrt{\lambda}l,$$

$$\Downarrow$$

$$2b \sin \sqrt{\lambda}l = 0. \quad (266)$$

$$\begin{aligned}
u'(-l) = u'(l) &\Rightarrow a\sqrt{\lambda} \sin \sqrt{\lambda}l + b\sqrt{\lambda} \cos \sqrt{\lambda}l = -a\sqrt{\lambda} \sin \sqrt{\lambda}l + b\sqrt{\lambda} \cos \sqrt{\lambda}l, \\
&\Downarrow \\
2a \sin \sqrt{\lambda}l &= 0.
\end{aligned} \tag{267}$$

The eigenvalues $\lambda > 0$ are then clearly

$$\lambda_k = \left(\frac{\pi k}{l} \right)^2, \quad k = 1, 2, \dots$$

For each such eigenvalue a and b are arbitrary. Thus (264) gives us an eigenspace of dimension 2 spanned by $\cos \sqrt{\lambda_k}x$ and $\sin \sqrt{\lambda_k}x$. Introducing an inner product and corresponding norm

$$\begin{aligned}
(u, v) &= \int_{-l}^l dx u(x)v(x), \\
\|u\|^2 &= \int_{-l}^l dx u^2(x),
\end{aligned}$$

we can normalize the eigenfunctions and find

$$\begin{aligned}
u_k(x) &= \frac{1}{\sqrt{l}} \sin \frac{\pi k}{l} x, \quad k = 1, 2, \dots \\
\hat{u}_k(x) &= \frac{1}{\sqrt{l}} \cos \frac{\pi k}{l} x.
\end{aligned} \tag{268}$$

One can show that there are no negative eigenvalues and that $\lambda_0 = 0$ is an eigenvalue with corresponding normalized eigenfunction

$$\hat{u}_0(x) = \frac{1}{\sqrt{2l}}. \tag{269}$$

Using these functions we get a *Fourier* series

$$u = (u, \hat{u}_0)\hat{u}_0 + \sum_{k=1}^{\infty} [(u, \hat{u}_k)\hat{u}_k + (u, u_k)u_k.] \tag{270}$$

Using expressions for the eigenfunctions and inner product we get

$$u(x) = \frac{1}{\sqrt{2l}}a_0 + \frac{1}{\sqrt{l}} \sum_{k=1}^{\infty} a_k \cos \frac{\pi k}{l} x + b_k \sin \frac{\pi k}{l} x,$$

where

$$\begin{aligned} a_0 &= \frac{1}{\sqrt{2l}} \int_{-l}^l dx u(x), \\ a_k &= \frac{1}{\sqrt{l}} \int_{-l}^l dx u(x) \cos \frac{\pi k}{l} x, \\ b_k &= \frac{1}{\sqrt{l}} \int_{-l}^l dx u(x) \sin \frac{\pi k}{l} x. \end{aligned}$$

If u is continuous, has a piecewise continuous derivative in the interval $[-l, l]$ and satisfies the boundary conditions (262), then the series (270) converges absolutely and uniformly to $u(x)$.

10.1.2 Bessel eigenfunctions

We are now going to consider a singular Sturm-Liouville problem that arise from the eigenvalue problem for Bessel's equation of order n

$$L(u(x)) = -\frac{d}{dx} \left(x \frac{du}{dx} \right) + \frac{n^2}{x} u(x) = \lambda x u(x), \quad 0 < x < l, \quad (271)$$

where $n \geq 0$ is an integer. This is a Sturm-Liouville problem with $p(x) = \rho(x) = x$, $q(x) = \frac{n^2}{x}$. The boundary conditions of interest for (271) are

$$\begin{aligned} u(x) &\text{ bounded at } x=0, \\ u(l) &= 0. \end{aligned} \quad (272)$$

Requiring that $u(x)$ is bounded at $x = 0$ is a significant condition here because (271) is a singular Sturm-Liouville problem. We observe that p , ρ and q are positive on $0 < x < l$ but that p and ρ vanish at $x = 0$ and that q is singular at $x = 0$. We might therefore in general expect there to be solutions that are singular at $x = 0$ and (272) eliminate such solutions as acceptable eigenfunctions. One might worry that *all* solutions are singular at $x = 0$. Then we will get no eigenvalues and eigenfunctions at all! By using the *method of Frobenius* from the theory of ordinary differential equations one can show that (271) has solutions that are bounded at $x = 0$. These solutions can be expressed in terms of *Bessel functions* of order n , J_n , in the following simple way

$$u(x) = J_n(\sqrt{\lambda}x). \quad (273)$$

We find the eigenvalues by solving the equation

$$J_n(\sqrt{\lambda}l) = 0. \quad (274)$$

One can prove that this equation has infinitely many solutions λ_{kn} , $k = 1, 2, 3, \dots$. They are all positive and $\lambda_{kn} \rightarrow \infty$ as $k \rightarrow \infty$. A large number of zeros for

the function J_n has been found using numerical methods. Let these zeros be α_{kn} , $k = 1, 2, 3, \dots$. In terms of these zeros the eigenvalues are

$$\lambda_{kn} = \left(\frac{\alpha_{kn}}{l} \right)^2, \quad k = 1, 2, \dots \quad (275)$$

The corresponding eigenfunctions are

$$\hat{u}_{kn}(x) = J_n(\sqrt{\lambda_{kn}}x), \quad k = 1, 2, \dots \quad (276)$$

Using the inner product

$$(u, v) = \int_0^l dx \, x u(x) v(x),$$

the eigenfunctions can be normalized

$$u_{kn} = \frac{\sqrt{2}}{l} \frac{J_n(\sqrt{\lambda_{kn}}x)}{|J_{n+1}(\sqrt{\lambda_{kn}}l)|}, \quad k = 1, 2, \dots, \quad (277)$$

where we have used the formulas

$$\|J_n(\sqrt{\lambda_{kn}}x)\|^2 = \int_0^l x J_n^2(\sqrt{\lambda_{kn}}x) dx = \frac{l^2}{2} J_{n+1}^2(\sqrt{\lambda_{kn}}l). \quad (278)$$

As you have seen in calculus there are a large number of formulas linking elementary functions like $\sin(x)$, $\cos(x)$, $\ln(x)$, e^x etc. and their derivatives and anti-derivatives. The Bessel functions J_n is an extension of this machinery. There are a large set of formulas linking the J_n and their derivatives and anti-derivatives to each other and other functions like $\sin(x)$, $\cos(x)$, $\ln(x)$ etc. The functions in calculus ($\sin(x)$, $\cos(x)$, \dots) are called *elementary functions*. The function J_n is our first example of a much larger machinery of functions called *special functions*. There are many handbooks that list some of the important relations and formulas involving special functions. Symbolic systems like mathematica, maple, etc. embody most of these formulas and often produce answers to problems formulated in terms of such functions.

Note that the boundary conditions (227) are not of the type (229) discussed in the general theory of Sturm-Liouville theory. Thus, there is no guarantee at this point that eigenfunctions corresponding to different eigenvalues are orthogonal. This is typical for singular Sturm-Liouville problems; the properties of the eigenvalues and eigenfunctions must be investigated in each separate case.

Observe that, in general, with L defined as in (228) we have

$$\begin{aligned}
\int_0^l dx uLv &= \int_0^l dx u[-(pv')' + qv] \\
&= - \int_0^l dx u(pv')' + \int_0^l dx uqv \\
&= -u(pv')|_0^l + \int_0^l dx u'pv' + \int_0^l dx quv \\
&= -u(pv')|_0^l + (u'p)v|_0^l - \int_0^l dx (pu')'v + \int_0^l dx uqv \\
&= \{pu'v - puv'\}|_0^l + \int_0^l dx Luv \\
&\Rightarrow \int_0^l dx \{uLv - vLu\} = \{pu'v - pv'u\}|_0^l \tag{279}
\end{aligned}$$

Recall that for our current singular Sturm-Liouville problem, the space of functions we need to expand in terms of the eigenfunctions (277), consists of functions such that

$$\begin{aligned}
u(l) &= 0, \\
\lim_{x \rightarrow 0^+} u(x) &\text{ bounded.}
\end{aligned} \tag{280}$$

Thus on the right hand side of (279) the evaluation at $x = 0$ must be interpreted as a limit as x approach 0 from inside the interval $(0, l)$.

Let us apply (279) to a pair of eigenfunctions u_i, u_j corresponding to different eigenvalues $\lambda_i \neq \lambda_j$. Thus

$$\begin{aligned}
Lu_i &= \rho\lambda_i u_i, \\
Lu_j &= \rho\lambda_j u_j.
\end{aligned}$$

But then (279) gives us

$$\begin{aligned}
(\lambda_i - \lambda_j)(u_i, u_j) &= (\lambda_i u_i, u_j) - (u_i, \lambda_j u_j) \\
&= \left(\frac{1}{\rho} Lu_i, u_j\right) - \left(u_i, \frac{1}{\rho} Lu_j\right) \\
&= \int_0^l dx \{Lu_i u_j - u_i Lu_j\} = [pu_j u_i' - pu_i u_j']|_0^l.
\end{aligned} \tag{281}$$

In order for eigenvectors corresponding to different eigenvalues to be orthogonal, the right hand side (281) must be zero. The evaluation at $x = l$ is zero since the eigenfunctions are zero at this point. In order for the right hand side of (281) to

vanish, it is in general *not* enough to require that the eigenfunctions are bounded at $x = 0$. The derivative of the eigenfunctions might still be unbounded at $x = 0$ in such a way that the right hand side of (281) is nonzero and thus ruining the orthogonality of the eigenfunctions. However, if we also require that

$$p(x)u'_i(x) \rightarrow 0 \text{ when } x \rightarrow 0^+, \quad (282)$$

everything is okay and the eigenfunctions form an orthogonal set. Since $p(x) = x$ we thus have the requirement

$$\lim_{x \rightarrow 0^+} x u'_i(x) = 0.$$

From the theory of Bessel functions we have

$$\left. \begin{aligned} J_0(\sqrt{\lambda}x) &= 1 + \mathcal{O}(x^2), \\ J_n(\sqrt{\lambda}x) &= \mathcal{O}(x^n), \quad n > 0 \end{aligned} \right\} \text{ as } x \rightarrow 0^+.$$

Thus the functions are bounded at $x = 0$. Also

$$\left. \begin{aligned} J'_0(\sqrt{\lambda}x) &= \mathcal{O}(x), \\ J'_n(\sqrt{\lambda}x) &= \mathcal{O}(x^{n-1}), \quad n > 0 \end{aligned} \right\} \text{ as } x \rightarrow 0^+,$$

but then

$$\begin{aligned} \lim_{x \rightarrow 0^+} x J'_0(\sqrt{\lambda}x) &= 0, \\ \lim_{x \rightarrow 0^+} x J'_n(\sqrt{\lambda}x) &= 0, \quad n > 0. \end{aligned}$$

Thus, by the argument given, the eigenfunctions

$$u_{kn}(x) = \alpha_{kn} J_n(\sqrt{\lambda_{kn}}x), \text{ with } \alpha_{kn} = \frac{\sqrt{2}}{l |J_{n+1}(\sqrt{\lambda_{kn}}l)|},$$

form an orthonormal set.

One can show that these eigenfunctions form a complete set. Any smooth function defined on the interval $0 \leq x \leq l$ that satisfies the boundary condition $u(l) = 0$ can be expanded in a convergent series

$$u = \sum_{k=1}^{\infty} (u, u_{kn}) u_{kn}. \quad (283)$$

10.1.3 Legendre polynomial eigenfunctions

Another singular Sturm-Liouville problem that occur frequently is

$$L(u(x)) = -\frac{d}{dx} \left[(1-x^2) \frac{du}{dx} \right] = \lambda u(x), \quad -1 < x < 1, \quad (284)$$

with boundary conditions

$$u(x) \text{ bounded at } x = \pm 1. \quad (285)$$

The function $p(x) = 1 - x^2$ vanishes at both endpoints $x = \pm 1$ so we indeed have a singular Sturm-Liouville problem.

One can use the method of Frobenius, alluded to earlier, to show that the only bounded solutions to (284) are in fact polynomials and that these exist only for

$$\lambda_k = k(k+1), \quad k = 0, 1, 2, \dots \quad (286)$$

The normalized eigenfunctions, P_k , are called the *Legendre polynomials* of degree k

$$P_k(x) = \left(\frac{1}{2^k k!} \right) \frac{d^k}{dx^k} (x^2 - 1)^k. \quad (287)$$

It is very easy to verify that

$$\lim_{x \rightarrow \pm 1} (1 - x^2) P'_k(x) = 0, \quad (288)$$

so this singular Sturm-Liouville problem has an orthogonal set of eigenfunctions according to previous arguments. One can show that

$$\|P_k(x)\|^2 = \int_{-1}^1 P_k^2(x) dx = \frac{2}{2k+1}, \quad k = 0, 1, 2, \dots \quad (289)$$

Thus we get an orthonormal set of eigenfunctions

$$u_k(x) = \sqrt{\frac{2k+1}{2}} P_k(x), \quad k = 0, 1, 2, \dots \quad (290)$$

One can show that this set is complete and that any smooth function defined over the interval $-1 \leq x \leq 1$ can be expanded in a convergent series

$$u = \sum_{k=0}^{\infty} (u, u_k) u_k. \quad (291)$$

10.2 Series solutions to initial boundary value problems

It is now time for returning to the solution of initial boundary value problems. We will consider a few specific examples and do detailed calculations for these cases. However, before we do that let us recapitulate the separation of variable strategy for the hyperbolic case. As we have seen, the strategy for parabolic and elliptic equations is very similar.

i) Look for functions of the form

$$u(\mathbf{x}, t) = M(\mathbf{x})N(t) \quad (292)$$

that solves the PDE (186) *and* the boundary conditions. These could be of the form (187), but as we have seen in the previous section on singular Sturm-Liouville problems, other possibilities exist.

- ii) If the boundary condition has been chosen in an appropriate way we find that M must be one of a complete set of eigenfunctions $\{M_k\}_{k=1}^{\infty}$ of a self-adjoint eigenvalue problem with eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ corresponding to the eigenfunctions $\{M_k\}_{k=1}^{\infty}$.
- iii) The function $N(t)$ is found to be one of

$$N_k(t) = a_k \cos(\sqrt{\lambda}t) + b_k \sin(\sqrt{\lambda}t), \quad (293)$$

where $\{a_k\}$, $\{b_k\}$ are constants that can be freely specified.

- iv) A formal solution to the PDE and the boundary conditions is

$$u(\mathbf{x}, t) = \sum_{k=1}^{\infty} \{a_k \cos(\sqrt{\lambda}t) + b_k \sin(\sqrt{\lambda}t)\} M_k(\mathbf{x}). \quad (294)$$

- v) Let a_k and b_k be the Fourier coefficients

$$a_k = \frac{(f, M_k)}{(M_k, M_k)}, \quad b_k = \frac{(g, M_k)}{\sqrt{\lambda_k}(M_k, M_k)}, \quad (295)$$

where f and g are the initial data. Then

$$\begin{aligned} u(\mathbf{x}, 0) &= \sum_{k=1}^{\infty} a_k M_k(\mathbf{x}) = f(\mathbf{x}), \\ u_t(\mathbf{x}, 0) &= \sum_{k=1}^{\infty} \sqrt{\lambda_k} b_k M_k(\mathbf{x}) = g(\mathbf{x}) \end{aligned} \quad (296)$$

and thus the formal solution (294) satisfies the PDE, the boundary conditions *and* the initial conditions.

- vi) Investigate the convergence of the series (294). If the initial data f and g are smooth enough the series will converge so fast that it can be differentiated twice termwise. In this case the formal series solution is a classical (honest to God!) solution to our initial/boundary value problem. For less smooth f and g the convergence is weaker and will lead to generalized (non-classical) solutions.

10.2.1 Vibrations of a string with fixed endpoints

We have seen earlier in this class that a vibrating string with fixed endpoints can in certain cases be modelled using the equation

$$u_{tt}(x, t) - c^2 u_{xx}(x, t) = 0, \quad 0 < x < l, t > 0, \quad (297)$$

$$u(0, t) = 0, u(l, t) = 0, \quad (298)$$

where $c = \sqrt{\frac{T}{\rho}}$, and where T and ρ are the, assumed constant, tension and mass density of the string. In order to get a unique motion of the string, we must specify the displacement and movement of the string at $t = 0$

$$\begin{aligned} u(x, 0) &= f(x), \quad 0 < x < l, \\ u_t(x, 0) &= g(x). \end{aligned} \quad (299)$$

These are the initial values. We now separate variables, $u(x, t) = M(x)N(t)$, and get the following regular Sturm-Liouville problem for M

$$\begin{aligned} -M''(x) &= \lambda M(x), \\ M(0) &= M(l) = 0. \end{aligned} \quad (300)$$

This is the problem we investigated on page 122. We found eigenvalues and normalized eigenfunctions of the form

$$\left. \begin{aligned} \lambda_k &= \left(\frac{\pi k}{l}\right)^2 \\ M_k(x) &= \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l}x\right) \end{aligned} \right\} \quad k = 1, 2, \dots \quad (301)$$

For the function $N(t)$ we find

$$N_k(t) = a_k \cos\left(\frac{\pi k c}{l}t\right) + b_k \sin\left(\frac{\pi k c}{l}t\right), \quad k = 1, 2, \dots \quad (302)$$

and the corresponding separated solutions $u_k(x, t)$ are

$$u_k(x, t) = \left[a_k \cos\left(\frac{\pi k c}{l}t\right) + b_k \sin\left(\frac{\pi k c}{l}t\right) \right] \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l}x\right), \quad \text{for } k = 1, 2, \dots \quad (303)$$

In order to satisfy the initial conditions we write down the formal solution

$$u(x, t) = \sum_{k=1}^{\infty} u_k(x, t). \quad (304)$$

Upon applying the initial conditions (299) we get

$$u(x, 0) = \sum_{k=1}^{\infty} u_k(x, 0) = \sum_{k=1}^{\infty} a_k \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right) = f(x),$$

$$u_t(x, 0) = \sum_{k=1}^{\infty} \partial_t u_k(x, 0) = \sum_{k=1}^{\infty} \left(\frac{\pi k c}{l}\right) b_k \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right) = g(x).$$

These are formally satisfied if

$$a_k = (f, M_k) = \sqrt{\frac{2}{l}} \int_0^l dx f(x) \sin\left(\frac{\pi k}{l} x\right), \quad (305)$$

$$b_k = \frac{l}{\pi k c} (g, M_k) = \frac{\sqrt{2l}}{\pi k c} \int_0^l dx g(x) \sin\left(\frac{\pi k}{l} x\right).$$

With these values for a_k and b_k , (304), is the formal solution of the initial value problem for the vibrating string fixed at both ends.

The separated solutions (303) are called *normal modes* for the vibrating string. The formal solution (304) is a sum, or *linear superposition*, of normal modes. The notion of normal modes is a concept that goes way beyond their use for the vibrating string. Separated solutions satisfying a hyperbolic PDE and boundary conditions are in general called normal modes.

Normal modes is a physical concept that has very wide applicability and deep roots. Physicists in general are not very familiar with eigenvalue problems for linear differential operators and their associated eigenfunctions, but they all know what normal modes are!

In quantum theory applied to fields, normal modes is a key idea. In this setting, they are in fact the mathematical embodiment of elementary particles. The newly discovered Higgs Boson is a normal mode for a certain field equation closely related to the 3D wave equation. In the quantum theory of materials, so-called many-body theory, the interaction between material normal modes (phonons) and electronic normal modes is responsible for the phenomenon of superconductivity.

Now then, when can we expect the formal solution (304) to be an actual, honest to God, classical solution? In general, finding sharp conditions on the initial data $f(\mathbf{x})$ and $g(\mathbf{x})$ that will imply the existence of a classical solution is a hard mathematical problem that is best left to specialists in mathematical analysis. In fact, sometimes the problem is so hard that entirely new mathematical ideas have to be developed in order to solve this problem.

You might know that all of modern mathematics is formulated in terms of, and gets its consistency from, *set theory*. You might not know that set theory was invented by George Cantor in the 1880's while studying the convergence of trigonometric Fourier series!

So what does mathematical analysis tell us about the convergence of the formal solution (304)? The following conditions are *sufficient* to ensure the existence of a classical solution for the vibrating string:

- i) f must have two continuous derivatives and a third piecewise continuous derivative,
- ii) $f(0) = f(l) = 0, \quad f''(0) = f''(l) = 0,$
- iii) g must have one continuous derivative and a second piecewise continuous derivative,
- iv) $g(0) = g(l) = 0.$

By using standard trigonometric formulas, the normal modes (303) can be rewritten as

$$u_k(x, t) = \alpha_k \cos\left(\frac{\pi k c}{l} t + \delta_k\right) \sin\left(\frac{\pi k}{l} x\right), \quad k = 1, 2, \dots, \quad (306)$$

where

$$\left. \begin{aligned} \alpha_k &= \sqrt{\frac{2}{l}} \sqrt{a_k^2 + b_k^2} \\ \delta_k &= -tg^{-1}\left(\frac{b_k}{a_k}\right) \end{aligned} \right\} \quad k = 1, 2, \dots \quad (307)$$

The formula (306) shows that for a normal mode each point x_0 perform a harmonic vibration with frequency

$$w_k = \frac{\pi k c}{l} \quad (308)$$

and amplitude

$$A_k = \alpha_k \sin\left(\frac{\pi k}{l} x_0\right) \quad (309)$$

Solutions to the wave equation are in general called *waves*. The normal mode (303) then is a *standing wave*; each point on the string oscillate in place as t varies.

Some points on the string remain fixed during the oscillation. These are the points

$$x_m = \frac{ml}{k}, \quad m = 1, 2, \dots, k-1$$

which are called *nodes* for the standing wave. The standing wave has maximum amplitude at the points

$$x_n = \frac{(2n+1)l}{2k}, \quad n = 0, 1, 2, \dots, k-1,$$

which are called *antinodes*.

Since each normal mode performs a harmonic oscillation, $u_k(x, t)$ is often called the k^{th} harmonic. Since a vibrating string makes sound, the normal mode u_k will produce a pure tone at frequency ω_k . Thus the decomposition of

the string into normal modes corresponds to the decomposition of sound into pure tones. If the tone-spectrum is pleasing we call it music!

Recall that the energy of a small piece of a vibrating string located between x and $x + dx$ is

$$E(x, t) = \underbrace{\{\rho(\partial_t u(x, t))^2\}}_{\text{kinetic energy}} + \underbrace{T(\partial_x u(x, t))^2}_{\text{potential energy}} dx.$$

Thus, the total energy of the vibrating string is

$$E(t) = \frac{1}{2} \int_0^l dx \{ \rho(\partial_t u(x, t))^2 + T(\partial_x u(x, t))^2 \}.$$

Writing

$$u_k(x, t) = N_k(t)M_k(x),$$

where

$$\begin{aligned} N_k(t) &= \alpha_k \cos(\omega_k t + \delta_k), \\ M_k(x) &= \sin\left(\frac{\pi k}{l}x\right), \end{aligned}$$

we observe that

$$\begin{aligned} \int_0^l dx \{(\partial_t u(x, t))^2\} &= (\partial_t u, \partial_t u) \\ &= \left(\sum_k N'_k(t)M_k, \sum_{k'} N'_{k'}(t)M_{k'} \right) \\ &= \sum_{kk'} N'_k(t)N'_{k'}(t)(M_k, M_{k'}) \\ &= \sum_k (N'_k(t))^2(M_k, M_k) = \sum_k (N'_k(t)M_k, N'_k(t)M_k) \\ &= \sum_k \int_0^l dx (\partial_t u_k(x, t))^2. \end{aligned}$$

In a similar way we find that

$$\int_0^l dx (\partial_x u(x, t))^2 = \sum_k \int_0^l dx (\partial_x u_k(x, t))^2.$$

Therefore

$$E(t) = \sum_k E_k(t),$$

where

$$\begin{aligned}
E_k(t) &= \frac{1}{2} \int_0^l dx \{ \rho (\partial_t u_k(x, t))^2 + T (\partial_x u_k(x, t))^2 \} \\
&= \frac{\omega_k^2 m (a_k^2 + b_k^2)}{2l},
\end{aligned}$$

where $m = \rho l =$ mass of string. (Do this calculation!)

Thus each E_k is independent of t and so is the total energy

$$E(t) = \frac{m}{2l} \sum_{k=1}^{\infty} \omega_k^2 (a_k^2 + b_k^2),$$

which is the sum of the energies, E_k , of each normal mode.

The decomposition of the total energy of a system into the energy of each normal mode plays a key role in the quantum theory of fields and in many-body theory.

Using trigonometric addition formulas one can show that

$$\begin{aligned}
2 \cos \frac{A-B}{2} \sin \frac{A+B}{2} &= \sin A + \sin B \\
-2 \sin \frac{A-B}{2} \sin \frac{A+B}{2} &= \cos A - \cos B
\end{aligned} \tag{310}$$

Using these formulas on the expression (303) for the normal modes we find

$$u_k(x, t) = \frac{1}{\sqrt{2l}} (a_k \sin \left(\frac{\pi k}{l} (x + ct) \right) + a_k \sin \left(\frac{\pi k}{l} (x - ct) \right)) \tag{311}$$

$$+ \frac{1}{\sqrt{2l}} (b_k \cos \left(\frac{\pi k}{l} (x - ct) \right) - b_k \cos \left(\frac{\pi k}{l} (x + ct) \right)). \tag{312}$$

Thus the normal mode can be written as a sum of left and right traveling waves. This is not surprising since we know that *all* solutions of the wave equation can be written as a sum of left and right traveling waves. What is special for the normal modes is that their traveling wave components are arranged in such a way that we get a standing wave.

10.2.2 Heat conduction in a finite rod

As we have seen previously in this class, heat conduction in a finite rod, where there are no heat sources and where the endpoints of the rod are kept at zero temperature, is in certain cases modeled by the boundary value problem

$$u_t(x, t) - c^2 u_{xx}(x, t) = 0, \quad 0 < x < l, t > 0, \tag{313}$$

$$u(0, t) = u(l, t) = 0, \tag{314}$$

where $u(x, t)$ is the temperature of the rod as a function of time, and where c^2 is the coefficient of heat conductivity. In order to get a unique solution we must specify the initial temperature distribution

$$u(x, 0) = f(x), \quad 0 < x < l. \quad (315)$$

The separation of variable approach leads to the same eigenvalue problem as for the string. The corresponding functions $N_k(t)$ now become

$$N_k(t) = a_k e^{-(\frac{\pi k c}{l})^2 t}, \quad k = 1, 2, \dots \quad (316)$$

and we get a formal solution

$$u(x, t) = \sum_{k=1}^{\infty} N_k(t) M_k(x) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k e^{-(\frac{\pi k c}{l})^2 t} \sin\left(\frac{\pi k}{l} x\right). \quad (317)$$

The initial value (313) implies that the constants a_k must be

$$a_k = \sqrt{\frac{2}{l}} \int_0^l dx f(x) \sin\left(\frac{\pi k}{l} x\right). \quad (318)$$

(317) with the constants given by (318) is a formal solution to our initial boundary value problem.

Because of the exponential term in the solution for $t > 0$, the series (317) converges so strongly that the series can be differentiated termwise as many times as we like. Thus (317) satisfies the PDE and boundary conditions if we merely assume that $f(x)$ is a bounded function on $0 \leq x \leq l$. However, as $t \rightarrow 0$ the exponential term vanishes, so in order to ensure that (317) also satisfies the initial condition, we need stronger conditions on f . Our friendly mathematical analyst tells us that the following conditions are sufficient:

- i) f must be continuous and f' piecewise continuous in $0 \leq x \leq l$
- ii) $f(0) = f(l) = 0$

For the wave equation we know that discontinuities in the initial data are preserved by the equation and propagated along the characteristic curves $x \pm ct = \text{const}$. For the heat conduction problem, the exponential factor will ensure that the solution is infinitely many times differentiable for any $t > 0$ even if the initial data is discontinuous!

The heat conduction equation can not support any kind of discontinuities for $t > 0$. We have seen this same result appear previously from the fact that the only characteristic curves for the heat conduction equation are of the form $t = \text{const}$.

10.2.3 The Laplace equation in a rectangle

Consider the boundary value problem

$$\begin{aligned} u_{xx}(x, y) + u_{yy}(x, y) &= 0, \quad 0 < x < l, 0 < y < \hat{l}, \\ u(x, 0) &= f(x), \quad u(0, y) = 0, \\ u(x, \hat{l}) &= g(x), \quad u(l, y) = 0. \end{aligned} \quad (319)$$

The equation, which is the Laplace equation, and boundary conditions determine the steady state displacement of a stretched membrane on a rectangular domain. The functions f and g specify the stretching of the membrane on the boundary curves $y = 0$ and $y = \hat{l}$.

The separation of variable method leads to the same eigenvalue problem as before but now the corresponding functions N_k are

$$N_k(y) = \hat{a}_k e^{\frac{\pi k}{l} y} + \hat{b}_k e^{-\frac{\pi k}{l} y}. \quad (320)$$

The separated solutions then are

$$u_k(x, y) = [a_k \sinh\left(\frac{\pi k}{l} y\right) + b_k \sinh\left(\frac{\pi k}{l} (y - \hat{l})\right)] \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right), \quad (321)$$

where we have introduced hyperbolic functions instead of exponential ones.

The formal solution to our equation and homogeneous boundary conditions are then

$$u(x, y) = \sum_{k=1}^{\infty} u_k(x, y). \quad (322)$$

The boundary conditions at $y = 0$ and $y = \hat{l}$ imply that

$$u(x, 0) = \sum_{k=1}^{\infty} b_k \sinh\left(-\frac{\pi k}{l} \hat{l}\right) \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right) = f(x), \quad (323)$$

$$u(x, \hat{l}) = \sum_{k=1}^{\infty} a_k \sinh\left(\frac{\pi k}{l} \hat{l}\right) \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right) = g(x). \quad (324)$$

These are formally satisfied if

$$\left. \begin{aligned} a_k &= \frac{(g, M_k)}{\sinh\left(\frac{\pi k \hat{l}}{l}\right)} \\ b_k &= \frac{(f, M_k)}{\sinh\left(-\frac{\pi k \hat{l}}{l}\right)} \end{aligned} \right\} \quad k = 1, 2, \dots \quad (325)$$

(322) with a_k and b_k given by (325) determines the formal solution to the problem (319).

Assume that

$$\left. \begin{aligned} \int_0^l dx |f(x)| &\leq m \\ \int_0^l dx |g(x)| &\leq m \end{aligned} \right\} \text{ for some } m > 0. \quad (326)$$

Then

$$\begin{aligned} |b_k| &= \frac{1}{|\sinh\left(-\frac{\pi k \hat{l}}{l}\right)|} \left| \int_0^l dx f(x) \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right) \right| \\ &\leq \frac{m \sqrt{\frac{2}{l}}}{|\sinh\left(-\frac{\pi k \hat{l}}{l}\right)|} = \frac{m \sqrt{\frac{2}{l}}}{\frac{1}{2} e^{\frac{\pi k \hat{l}}{l}} (1 - e^{-\frac{2\pi k \hat{l}}{l}})} \end{aligned} \quad (327)$$

and using the fact that for $z \geq 0$

$$\sinh z = \frac{1}{2}(e^z - e^{-z}) \leq \frac{1}{2}e^z \quad (328)$$

we get the estimate

$$\begin{aligned} |b_k \sinh\left(\frac{\pi k}{l}(y - \hat{l})\right)| &\leq \frac{m \sqrt{\frac{2}{l}} e^{\frac{\pi k}{l}(l-y)}}{e^{\frac{\pi k \hat{l}}{l}} (1 - e^{-\frac{2\pi k \hat{l}}{l}})} \\ &= \frac{m \sqrt{\frac{2}{l}} e^{-\frac{\pi k}{l} y}}{(1 - e^{-\frac{2\pi k \hat{l}}{l}})} \end{aligned}$$

Thus, this term is bounded by an exponential as $k \rightarrow \infty$, $\forall 0 < y < \hat{l}$. In a similar way $|a_k \sinh\left(\frac{\pi k}{l} y\right)|$ is bounded by a decaying exponential as $k \rightarrow \infty$. Therefore the terms in the formal solution (322) decay exponentially as $k \rightarrow \infty$ for all $0 < x < l, 0 < y < \hat{l}$. The formal series (322) therefore converges and defines a function in the inside of the rectangle $[0, l] \times [0, \hat{l}]$ whose series representation can be differentiated termwise as many times as we like. Thus the formal series (322) defines a classical solution to the Laplace equation inside a closed rectangle $[0, l] \times [0, \hat{l}]$. It, by construction of the separated solutions, satisfies the boundary conditions on the edges $x = 0, x = l, 0 \leq y \leq \hat{l}$. As we approach $y = 0$ or $y = \hat{l}$, the exponential decay is weakened and in order to ensure that the formal solution satisfies the boundary conditions on $y = 0, \hat{l}, 0 \leq x \leq l$ our friendly mathematical analysis tells us that it is sufficient that:

- i) f, g are continuous and f', g' piecewise continuous in the interval $0 \leq x \leq l$

$$\text{ii) } f(0) = f(l) = g(0) = g(l) = 0$$

The fact that solutions to the Laplace equations are smooth in the interior of the rectangle even if the boundary data has discontinuities is, as we have seen, also a consequence of having no real characteristics at all.

10.3 Inhomogenous equations - Duhamel's principle

So far, we have only considered the separation of variable method for homogeneous equations. This means physically that we are modeling source-free problems. It is now time to consider the presence of sources. We will here focus on the hyperbolic and parabolic case:

$$\begin{aligned} \rho(\mathbf{x})u_{tt}(\mathbf{x}, t) + L[u(\mathbf{x}, t)] &= g(\mathbf{x}, t), & \text{Hyperbolic case,} \\ \rho(\mathbf{x})u_t(\mathbf{x}, t) + L[u(\mathbf{x}, t)] &= g(\mathbf{x}, t), & \text{Parabolic case.} \end{aligned} \quad (329)$$

The equations (329) are supplied with homogeneous boundary conditions and also homogeneous initial conditions

$$\begin{aligned} u(\mathbf{x}, 0) = 0, \quad u_t(\mathbf{x}, 0) &= 0, & \text{Hyperbolic case,} \\ u(\mathbf{x}, 0) &= 0, & \text{Parabolic case.} \end{aligned} \quad (330)$$

The initial/boundary value problem (329),(330) is solved using *Duhamel's principle*. We proceed as follows:

Let $v(\mathbf{x}, t; \tau)$ be a function that satisfies

$$\begin{aligned} \rho(\mathbf{x})v_{tt}(\mathbf{x}, t) + L[v(\mathbf{x}, t)] &= 0, & \text{Hyperbolic case,} \\ \rho(\mathbf{x})v_t(\mathbf{x}, t) + L[v(\mathbf{x}, t)] &= 0, & \text{Parabolic case.} \end{aligned}$$

with initial conditions

$$\begin{aligned} v(\mathbf{x}, \tau; \tau) = 0, \quad v_t(\mathbf{x}, \tau; \tau) &= \frac{g(\mathbf{x}, \tau)}{\rho(\mathbf{x})}, & \text{Hyperbolic case,} \\ v(\mathbf{x}, \tau; \tau) &= \frac{g(\mathbf{x}, \tau)}{\rho(\mathbf{x})}, & \text{Parabolic case.} \end{aligned} \quad (331)$$

We assume this solution has been found using separation of variables or some other method.

Define a function $u(\mathbf{x}, t)$ by

$$u(\mathbf{x}, t) = \int_0^t d\tau v(\mathbf{x}, t; \tau). \quad (332)$$

Then we have by the rules of calculus

$$\begin{aligned}
u_t(\mathbf{x}, t) &= v(\mathbf{x}, t; t) + \int_0^t dt v_t(\mathbf{x}, t; \tau), \\
u_{tt} &= \partial_t(v(\mathbf{x}, t; t)) + v_t(\mathbf{x}, t; t) + \int_0^t d\tau v_{tt}(x, t; \tau), \\
L[u(\mathbf{x}, t)] &= \int_0^t d\tau L[v(\mathbf{x}, t; \tau)]
\end{aligned} \tag{333}$$

and from the initial conditions

$$\begin{aligned}
v(\mathbf{x}, t; t) = 0 &\Rightarrow \partial_t(v(\mathbf{x}, t; t)) = 0, \quad \text{hyperbolic}, \\
v(\mathbf{x}, t; t) &= \frac{g(\mathbf{x}, t)}{\rho(\mathbf{x})}, \quad \text{parabolic}
\end{aligned} \tag{334}$$

Thus we have for the hyperbolic case

$$\begin{aligned}
\rho u_{tt}(\mathbf{x}, t) + L[u(\mathbf{x}, t)] &= \rho(\mathbf{x}) \left(\frac{g(\mathbf{x}, t)}{\rho(\mathbf{x})} \right) \\
+ \int_0^t d\tau \{ \rho v_{tt}(\mathbf{x}, t; \tau) + L[v(\mathbf{x}, t; \tau)] \} &= g(\mathbf{x}, t).
\end{aligned} \tag{335}$$

(335) show that the $u(\mathbf{x}, t)$ constructed in (332) solves the inhomogeneous problem for the hyperbolic case. In a similar way we can show that the function $v(\mathbf{x}, t; \tau)$ for the parabolic case also solves the inhomogeneous equation for that case. (332) is Duhamel's principle.

10.3.1 The inhomogeneous wave equation

Let the following initial/boundary value problem be given

$$\begin{aligned}
u_{tt}(x, t) - c^2 u_{xx}(x, t) &= g(x, t), \quad -\infty < x < \infty, t > 0, \\
u(x, 0) = u_t(x, 0) &= 0, \quad -\infty < x < \infty.
\end{aligned} \tag{336}$$

According to Duhamel's principle we should construct a function $v(x, t; \tau)$ which solves

$$\begin{aligned}
v_{tt}(x, t; \tau) - c^2 v_{xx}(x, t; \tau) &= 0, \quad -\infty < x < \infty, t > \tau, \\
v(x, \tau; \tau) &= 0, \\
v_t(x, \tau; \tau) &= g(x, \tau).
\end{aligned} \tag{337}$$

We solve (337) using the d'Alembert formula

$$u(x, t; \tau) = \frac{1}{2c} \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\theta g(\theta, \tau). \tag{338}$$

The solution to (336), which we denote by u_p , is then according to Duhamel's principle

$$u_p(x, t) = \frac{1}{2c} \int_0^t d\tau \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\theta g(\theta, \tau). \quad (339)$$

Let $u_h(x, t)$ be the solution to the homogeneous wave equation with arbitrary initial data

$$\begin{aligned} u(x, 0) &= F(x), \quad -\infty < x < \infty, \\ u_t(x, 0) &= G(x). \end{aligned} \quad (340)$$

According to the d'Alembert formula the solution with initial data (340) is

$$u_h(x, t) = \frac{1}{2} \{F(x+ct) + F(x-ct)\} + \frac{1}{2c} \int_{x-ct}^{x+ct} d\theta G(\theta). \quad (341)$$

Let now

$$u(x, t) = u_h(x, t) + u_p(x, t). \quad (342)$$

Verify that $u(x, t)$ satisfies the inhomogeneous wave equation with initial data (340). Inserting the formulas for u_p and u_h we get

$$\begin{aligned} u(x, t) &= \frac{1}{2} \{F(x+ct) + F(x-ct)\} + \frac{1}{2c} \int_{x-ct}^{x+ct} d\theta G(\theta) \\ &\quad + \frac{1}{2c} \int_0^t d\tau \int_{x-c(t-\tau)}^{x+c(t-\tau)} d\theta g(\theta, \tau). \end{aligned} \quad (343)$$

From this formula it is clear that the value of u at a point (x_0, t_0) , with $t_0 > 0$ depends only on the data at points in a *characteristic triangle*.

The characteristic triangle is called *the domain of dependence* of the solution at the point (x_0, t_0) .

Similarly, we define the *domain of influence* of a point (x_0, t_0) to be the set of space-time points (x, t) such that the domain of dependence of the solution at (x, t) includes the point (x_0, t_0) .

The point (x_0, t_0) will have no causal connection to points beyond its domain of influence. These notions play a fundamental role in the theory of relativity, where the domain of influence of a space-time point (x, t) is called a *light cone*. All material particles and force mediating particles must move within the light cone. This is strictly true only for classical physics. In quantum theory, both material particles and force mediators can, in the form of virtual particles, move beyond the light cone.

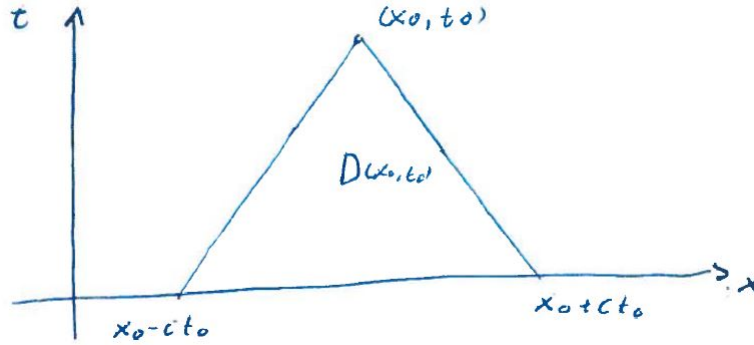


Figure 52

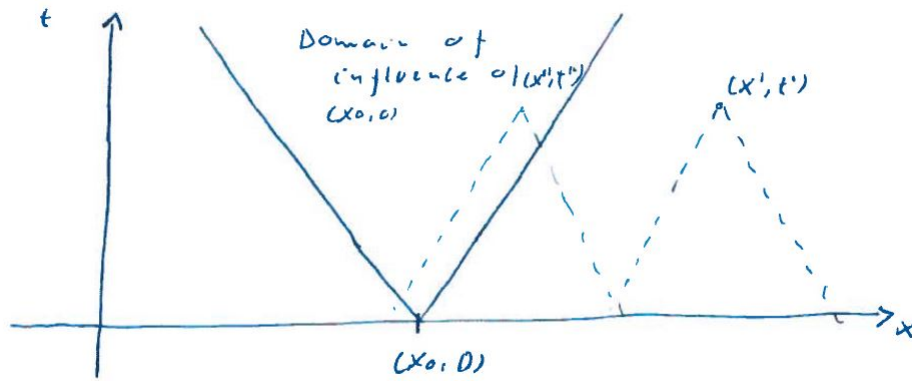


Figure 53

10.3.2 The inhomogeneous heat equation

Let us consider the 1D heat equation on a finite interval, $0 < x < l$, driven by a heat source $g(x, t)$ located in the interval. Assuming the endpoints are kept at zero temperature, our problem is

$$\begin{aligned} u_t - c^2 u_{xx} &= g(x), \quad 0 < x < l, t > 0, \\ u(x, 0) &= 0, \\ u(0, t) &= u(l, t) = 0. \end{aligned} \tag{344}$$

The function $v(x, t; \tau)$ appearing in Duhamel's principle satisfies (344) with g replaced by zero and the initial condition

$$v(x, \tau; \tau) = g(x). \tag{345}$$

Separation of variables for this problem gives (page 72)

$$v(x, t; \tau) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k(\tau) e^{-\left(\frac{\pi k c}{l}\right)^2 (t-\tau)} \sin\left(\frac{\pi k}{l} x\right), \quad (346)$$

where

$$a_k(\tau) = \sqrt{\frac{2}{l}} \int_0^l dx g(x, \tau) \sin\left(\frac{\pi k}{l} x\right), \quad k = 1, 2, \dots \quad (347)$$

Duhamel's principle now gives the solution to (344) in the form

$$u(x, t; \tau) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} \left\{ \int_0^t d\tau a_k(\tau) e^{-\left(\frac{\pi k c}{l}\right)^2 (t-\tau)} \right\} \sin\left(\frac{\pi k}{l} x\right), \quad (348)$$

Let us consider the special case

$$g(x, t) = \sin\left(\frac{\pi}{l} x\right). \quad (349)$$

It is easy to verify that for this choice

$$\begin{aligned} a_1 &= \sqrt{\frac{l}{2}}, \\ a_k &= 0, \quad k > 1, \end{aligned}$$

$$\Rightarrow u(x, t) = \left(\frac{l}{\pi c}\right)^2 \left\{ 1 - e^{-\left(\frac{\pi c}{l}\right)^2 t} \right\} \sin\left(\frac{\pi}{l} x\right). \quad (350)$$

Observe that

$$u(x, t) \rightarrow \left(\frac{l}{\pi c}\right)^2 \sin\left(\frac{\pi}{l} x\right) \equiv u(x), \quad (351)$$

when $t \rightarrow +\infty$. By direct differentiation we find that the limiting function, $u(x)$, is a solution to

$$\begin{aligned} -c^2 u_{xx}(x) &= \sin\left(\frac{\pi}{l} x\right), \\ u(x) &= u(l) = 0, \end{aligned} \quad (352)$$

which shows that the function $u(x)$ is a time independent, or *stationary* solution to the heat equation.

In fact, if $g(x, t) = G(x)$ then from (348) we get

$$u(x, t) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} \left\{ \left(\frac{l}{\pi k c}\right)^2 a_k \left(1 - e^{-\left(\frac{\pi k c}{l}\right)^2 t}\right) \sin\left(\frac{\pi k}{l} x\right) \right\}. \quad (353)$$

For this case the limiting function $u(x)$ is

$$u(x) = \lim_{t \rightarrow +\infty} u(x, t) \quad (354)$$

$$= \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} \left(\frac{l}{\pi k c} \right)^2 a_k \sin \left(\frac{\pi k}{l} x \right). \quad (355)$$

It is easy to verify that $u(x)$ is in fact a stationary solution to the driven heat equation (344)

$$\begin{aligned} -c^2 u_{xx}(x) &= G(x), \\ u(x) &= u(l) = 0. \end{aligned} \quad (356)$$

10.4 The finite Fourier transform

Duhamel's principle can, as we have seen, be used to construct solutions to inhomogeneous equations. However, the solutions are expressed as an integral over the parameter τ and are rarely explicit solutions. In order to get actual numbers out we typically have to do numerical integration. This can quickly become awkward and expensive to do in realistic situations.

There is however another approach that is more general and which is better behaved numerically. This is the *finite Fourier transform*. In this method we expand the solutions in terms of eigenfunctions of the spatial part of the differential operator defining the equation. The expansion coefficients are time dependent and the PDE is reduced to an infinite set of coupled ODEs. We have thus transformed the PDE into a system of ODEs. This process is called the finite Fourier transform. The infinite series of eigenfunctions expressing the solution of the PDE in terms of the time dependent coefficients is called the *inverse finite Fourier transform*.

We have thus the following solution strategy:

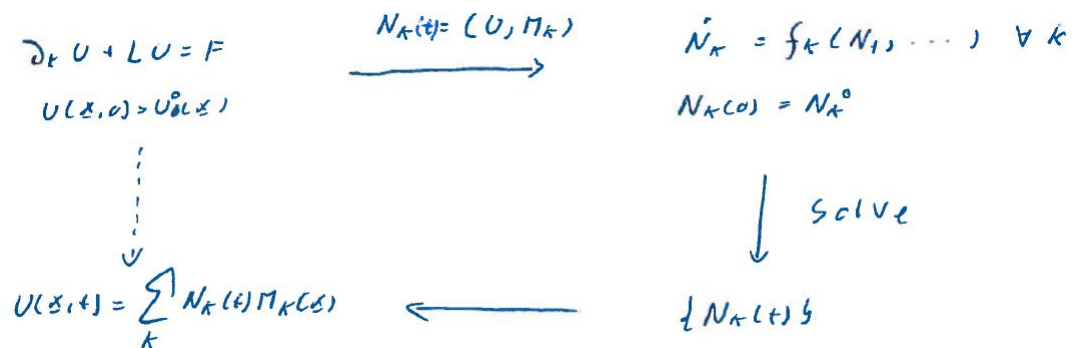


Figure 54

This *transform strategy* is an important approach to solving complex problems in applied mathematics and theoretical physics.

We will stick to the restricted class discussed previously in these notes. For that class we have an operator, L , given by

$$Lu = -\nabla \cdot (p\nabla u) + qu. \quad (357)$$

We can give a unified presentation of the finite Fourier transform if we introduce an operator K which is

$$\begin{aligned} K = \partial_t & \quad - \text{parabolic case,} \\ K = \partial_{tt} & \quad - \text{hyperbolic case,} \end{aligned} \quad (358)$$

$$K = -\partial_{yy} \quad - \text{elliptic case.} \quad (359)$$

$$(360)$$

In fact, if we also allow $K = 0$, the general 3D elliptic problem can be included. With this definition of K our problem of interest is

$$\rho(\mathbf{x})Ku + Lu = \rho(\mathbf{x})F. \quad (361)$$

Let $\{M_k\}$ be the complete set of eigenfunctions for the eigenvalue problem

$$LM_k = \lambda_k \rho M_k, \quad (362)$$

where as usual we apply the spatial boundary conditions. We will assume that the eigenfunctions have been normalized so that

$$(M_k, M_{k'}) = \delta_{kk'}. \quad (363)$$

We now expand the solutions of (361) in terms of the eigenfunctions M_k

$$u = \sum_k N_k M_k, \quad (364)$$

where

$$N_k = (u, M_k). \quad (365)$$

The coefficients N_k are functions of t for the hyperbolic and parabolic case and a function of y for the elliptic case with (generalized) cylinder symmetry. For the general elliptic case the N_k 's are constants.

The formula (365) defines the finite Fourier transform and (364) the inverse finite Fourier transform.

In order to derive an equation for the coefficients N_k we multiply (361) by $M_k(\mathbf{x})$ and integrate over the domain for the variable \mathbf{x} . We let this domain be denoted by G .

This integration gives us the equation

$$\int_G dV \rho(\mathbf{x}) M_k(\mathbf{x}) K u(\mathbf{x}, t) + \int_G dV M_k(\mathbf{x}) L u(\mathbf{x}, t) = \int_G dV \rho(\mathbf{x}) M_k(\mathbf{x}) F. \quad (366)$$

Using the definition of the finite Fourier transform and the inner product, each term in this equation can be reformulated into a more convenient form.

For the first term we have

$$\begin{aligned} \int_G dV \rho(\mathbf{x}) M_k(\mathbf{x}) K u(\mathbf{x}, t) &= K \int_G dV \rho(\mathbf{x}) u(\mathbf{x}, t) M_k(\mathbf{x}) \\ &= K(u, M_k(\mathbf{x})) = K N_k. \end{aligned} \quad (367)$$

For the second term in (366) we have

$$\begin{aligned} \int_G dV M_k(\mathbf{x}) L u(\mathbf{x}, t) &= \int_G dV u(\mathbf{x}, t) L M_k(\mathbf{x}) \\ &\quad - \int_{\partial G} dS p(\mathbf{x}) \{ M_k(\mathbf{x}) \partial_{\mathbf{n}} u(\mathbf{x}, t) - u(\mathbf{x}, t) \partial_{\mathbf{n}} M_k(\mathbf{x}) \} \\ &= \lambda_k N_k - \int_{\partial G} dS p(\mathbf{x}) \{ M_k(\mathbf{x}) \partial_{\mathbf{n}} u(\mathbf{x}, t) - u(\mathbf{x}, t) \partial_{\mathbf{n}} M_k(\mathbf{x}) \}. \end{aligned} \quad (368)$$

The second term in (368) can be further simplified using the boundary conditions.

Recall that the boundary conditions on u and M_k are

$$\begin{aligned} \alpha M_k + \beta \partial_{\mathbf{n}} M_k &= 0, \\ \alpha u + \beta \partial_{\mathbf{n}} u &= B, \end{aligned} \quad (369)$$

where u and B are functions of \mathbf{x} and t for the hyperbolic and parabolic case, functions of \mathbf{x} and y for the restricted elliptic case and a function of \mathbf{x} for the general elliptic case. As usual $\alpha = \alpha(\mathbf{x})$ and $\beta = \beta(\mathbf{x})$.

Using the classification of the boundary conditions into type 1, 2, 3 introduced on page 102 we have for the second term in (368)

$$\begin{aligned}
& \int_{\partial G} dS p \{M_k \partial_{\mathbf{n}} u - u \partial_{\mathbf{n}} M_k\}. \\
&= \int_{\partial G} dS p M_k \partial_{\mathbf{n}} u - \int_{\partial G} dS p \partial_{\mathbf{n}} M_k u \\
&= \int_{S_2 \cup S_3} dS p M_k \partial_{\mathbf{n}} u - \int_{S_1} dS \frac{p}{\alpha} \partial_{\mathbf{n}} M_k B \\
&\quad - \int_{S_2 \cup S_3} dS p \partial_{\mathbf{n}} M_k u \\
&= \int_{S_2 \cup S_3} dS \frac{p}{\beta} M_k \beta \partial_{\mathbf{n}} u + \int_{S_2 \cup S_3} dS \frac{p}{\beta} M_k \alpha u \\
&\quad - \int_{S_1} dS \frac{p}{\alpha} \partial_{\mathbf{n}} M_k B \\
&= \int_{S_2 \cup S_3} dS \frac{p}{\beta} M_k B - \int_{S_1} dS \frac{p}{\alpha} \partial_{\mathbf{n}} M_k B \\
&\equiv B_k.
\end{aligned} \tag{370}$$

Substituting (367), (368) and (370) into (366) gives us the system

$$KN_k + \lambda_k N_k = F_k + B_k, \quad k = 1, 2, \dots, \tag{371}$$

where

$$F_k = (F, M_k).$$

For the general elliptic case $K = 0$ so (371) are algebraic equations for the N_k . For the restricted elliptic case we have a system of ordinary differential equations

$$-\partial_{yy} N_k(y) + \lambda_k N_k(y) = F_k(y) + B_k(y). \tag{372}$$

This system of ODEs must be solved as a boundary value problem. The boundary conditions are

$$\begin{aligned}
N_k(0) &= (u(0), M_k) = (f, M_k), \\
N_k(\hat{l}) &= (u(\hat{l}), M_k) = (g, M_k),
\end{aligned} \tag{373}$$

where we have used the boundary conditions and where $u(0) \equiv u(\mathbf{x}, 0)$ etc.

For the parabolic case $K = \partial_t$ we have a system of ODEs

$$\partial_t N_k(t) + \lambda_k N_k(t) = F_k(t) + B_k(t), \quad k = 1, 2, \dots, \quad (374)$$

subject to the initial condition

$$N_k(0) = (u(0), M_k) = (f, M_k). \quad (375)$$

Observe that (374) is an infinite system of ODEs, but they are *uncoupled*. This is a result of the linearity of the underlying PDEs and the fact that the M_k 's are eigenfunctions for the operator L defining the equations. The finite Fourier transform can also be used on nonlinear PDEs, as we will see, and for this case the ODE's for the $N_k(t)$ will be *coupled*. The ODEs (374) are of a particular simple type. They are linear first order equations. Such equations can be solved using an integrating factor. The solution is found to be

$$N_k(t) = N_k(0)e^{-\lambda_k t} + \int_0^t d\tau [F_k(\tau) + B_k(\tau)]e^{-\lambda_k(t-\tau)}. \quad (376)$$

This formula is only useful if we can get some exact or approximate analytical solution of the integral. If a numerical solution is required it might be more efficient to solve the ODEs (374) numerically, using for example Runge-Kutta.

For the hyperbolic case, $K = \partial_{tt}$ and we get a system of second order ODEs

$$\partial_{tt} N_k(t) + \lambda_k N_k(t) = F_k(t) + B_k(t), \quad (377)$$

subject to the initial conditions

$$\begin{aligned} N_k(0) &= (u(0), M_k) = (f, M_k), \\ \partial_t N_k(0) &= (u_t(0), M_k) = (g, M_k), \end{aligned} \quad (378)$$

This second order equation is also of a simple type which is solvable by the method of variation of parameters. We get the formula

$$\begin{aligned} N_k(t) &= N_k(0) \cos \sqrt{\lambda_k} t + \frac{1}{\sqrt{\lambda_k}} \partial_t N_k(0) \sin \sqrt{\lambda_k} t \\ &+ \frac{1}{\sqrt{\lambda_k}} \int_0^t d\tau [F_k(\tau) + B_k(\tau)] \sin \sqrt{\lambda_k}(t - \tau). \end{aligned} \quad (379)$$

Just as for the parabolic case the formula (379) is only useful if some exact or approximate analytical solution can be found. If a numerical solution is required it is more efficient to solve the ODEs (377) using a numerical ODE solver.

Recall that when the N_k s have been computed, the solution to the PDE is constructed using the inverse finite Fourier transform

$$u = \sum_k N_k M_k. \quad (380)$$

The representation (380) is problematic when we approach the boundary since all the M_k s satisfy homogeneous boundary conditions, whereas u satisfies inhomogeneous boundary conditions. If these conditions are of, say, type 1 *and* if (380) is a classical solution (series converges pointwise) then (380) implies that u is zero on the boundary. But we know that u in general satisfies nonzero boundary conditions. Thus the series (380) can not converge pointwise as we approach the boundary. Therefore, (380) is in general not a classical solution, but rather some form of generalized solution.

The series might still converge pointwise away from the boundary, but the convergence will typically be slow since the expansion coefficients, N_k , "know about" the function over the whole domain and this leads to slow convergence everywhere.

There is however a way, which we will discuss shortly, to avoid this boundary problem and get series that converge fast enough to get a classical solution.

Let us now look at a couple of examples.

10.4.1 Hyperbolic equations

We consider the following hyperbolic PDE

$$\rho(\mathbf{x})u_{tt}(\mathbf{x}, t) + Lu(\mathbf{x}, t) = \rho(\mathbf{x})M_i(\mathbf{x}) \sin \omega t, \quad (381)$$

where $M_i(\mathbf{x})$ are one of the eigenfunctions of L . We assume that initial and boundary conditions are homogeneous

$$\begin{aligned} f(\mathbf{x}) &= 0, \\ g(\mathbf{x}) &= 0, \\ B(\mathbf{x}, t) &= 0. \end{aligned} \quad (382)$$

For this case we have

$$F_k(t) = (F(t), M_k) = \sin \omega t (M_i, M_k) = \delta_{ik} \sin \omega t, \quad (383)$$

and because the initial conditions for the PDE (381) are homogeneous, we have

$$\begin{aligned} N_k(0) &= \partial_t N_k(0) = 0, \\ B_k(t) &= 0. \end{aligned} \quad (384)$$

Thus our system of ODEs simplifies into

$$\partial_{tt} N_k(t) + \lambda_k N_k(t) = \delta_{ik} \sin \omega t, \quad (385)$$

with initial conditions

$$\begin{aligned} N_k(0) &= 0, \\ \partial_t N_k(0) &= 0. \end{aligned}$$

We get $N_k = 0$, $k \neq i$ and from (379)

$$\begin{aligned} N_i(t) &= \frac{1}{\sqrt{\lambda_i}} \int_0^t d\tau \sin \omega \tau \sin \sqrt{\lambda_i}(t - \tau) \\ &= \frac{\omega \sin(\sqrt{\lambda_i}t) - \sqrt{\lambda_i} \sin \omega t}{\sqrt{\lambda_i}(\omega^2 - \lambda_i)}. \end{aligned} \quad (386)$$

The inverse finite Fourier transform gives us

$$u(\mathbf{x}, t) = \frac{1}{\omega^2 - \lambda_i} \left[\frac{\omega}{\sqrt{\lambda_i}} \sin \sqrt{\lambda_i}t - \sin \omega t \right] M_i(\mathbf{x}). \quad (387)$$

This is now the exact solution to this problem. From (387) it is clear that as $\omega \rightarrow \sqrt{\lambda_i}$ the amplitude for the solution (387) becomes larger. We can not evaluate (387) at $\omega = \sqrt{\lambda_i}$, since the denominator is then zero, but if we take the limit of (387) as $\omega \rightarrow \sqrt{\lambda_i}$, using L'Hopitals rule on the way, we find

$$u(\mathbf{x}, t) = \frac{1}{2\sqrt{\lambda_i}} \left[\frac{\sin \sqrt{\lambda_i}t}{\sqrt{\lambda_i}} - t \cos \sqrt{\lambda_i}t \right] M_i(\mathbf{x}). \quad (388)$$

The formulas (387) and (388) tell the following story: The general solution of the homogeneous equation

$$\rho(\mathbf{x})u_{tt}(\mathbf{x}, t) + Lu(\mathbf{x}, t) = 0, \quad (389)$$

is a sum of normal modes $u_k(\mathbf{x}, t)$, whose oscillation frequencies are $\omega_k = \sqrt{\lambda_k}$. If we drive (386) with a source that oscillates at a frequency ω and whose spatial profile is an eigenfunction, M_i , the solution is a linear superposition of terms oscillating at ω and ω_i , with a common spatial profile $M_i(\mathbf{x})$. As ω approaches ω_i , the amplitude of the oscillation grows and in the limit $\omega \rightarrow \omega_i$ the system displays linear unbounded growth in time.

What we have here is the well known phenomena of *resonance* from particle mechanics. Driven continuum system modeled by PDEs can display resonance. Understanding, predicting and preventing resonance is of major importance in applied mathematics and physics. After all we don't want to end up in a plane whose jet engines fall off because resonance has weakened the bolts tying the engine to the plane fuselage!

10.4.2 The Poisson equation in a disk

We consider the boundary value problem for Poisson's equation

$$\begin{aligned} \nabla^2 u(x, y) &= -F(x, y), \quad x^2 + y^2 < R^2, \\ u|_{x^2+y^2=R^2} &= f. \end{aligned} \quad (390)$$

For this geometry it makes sense to use polar coordinates

$$\begin{aligned}x &= r \cos \theta, \\y &= r \sin \theta.\end{aligned}\tag{391}$$

Then

$$\begin{aligned}\partial_r &= \partial_r x \partial_x + \partial_r y \partial_y = \cos \theta \partial_x + \sin \theta \partial_y, \\ \partial_\theta &= \partial_\theta x \partial_x + \partial_\theta y \partial_y = -r \sin \theta \partial_x + r \cos \theta \partial_y, \\ &\Downarrow \\ \partial_x &= \cos \theta \partial_r - \frac{1}{r} \sin \theta \partial_\theta, \\ \partial_y &= \sin \theta \partial_r + \frac{1}{r} \cos \theta \partial_\theta, \\ &\Downarrow \\ \partial_{xx} &= (\cos \theta \partial_r - \frac{1}{r} \sin \theta \partial_\theta)(\cos \theta \partial_r - \frac{1}{r} \sin \theta \partial_\theta) \\ &= \cos^2 \theta \partial_{rr} + \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta - \frac{1}{r} \sin \theta \cos \theta \partial_{r\theta} \\ &\quad + \frac{1}{r} \sin^2 \theta \partial_r - \frac{1}{r} \sin \theta \cos \theta \partial_{r\theta} + \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta \\ &\quad + \frac{1}{r^2} \sin^2 \theta \partial_{\theta\theta}, \\ \partial_{yy} &= (\sin \theta \partial_r + \frac{1}{r} \cos \theta \partial_\theta)(\sin \theta \partial_r + \frac{1}{r} \cos \theta \partial_\theta) \\ &= \sin^2 \theta \partial_{rr} - \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta + \frac{1}{r} \sin \theta \cos \theta \partial_{r\theta} \\ &\quad + \frac{1}{r} \cos^2 \theta \partial_r + \frac{1}{r} \sin \theta \cos \theta \partial_{\theta r} - \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta \\ &\quad + \frac{1}{r^2} \cos^2 \theta \partial_{\theta\theta}, \\ &\Downarrow \\ \nabla^2 &= \cos^2 \theta \partial_{rr} + \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta - \frac{1}{r} \sin \theta \cos \theta \partial_{r\theta} \\ &\quad + \frac{1}{r} \sin^2 \theta \partial_r - \frac{1}{r} \sin \theta \cos \theta \partial_{r\theta} + \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta \\ &\quad + \frac{1}{r^2} \sin^2 \theta \partial_{\theta\theta} + \sin^2 \theta \partial_{rr} - \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta \\ &\quad + \frac{1}{r} \sin \theta \cos \theta \partial_{r\theta} + \frac{1}{r} \cos^2 \theta \partial_r + \frac{1}{r} \sin \theta \cos \theta \partial_{\theta r} \\ &\quad - \frac{1}{r^2} \sin \theta \cos \theta \partial_\theta + \frac{1}{r^2} \cos^2 \theta \partial_{\theta\theta} \\ &= \partial_{rr} + \frac{1}{r} \partial_r + \frac{1}{r^2} \partial_{\theta\theta}.\end{aligned}$$

Thus in polar coordinates (390) can be written as

$$\begin{aligned}\partial_{rr}u(r, \theta) + \frac{1}{r}\partial_r u(r, \theta) + \frac{1}{r^2}\partial_{\theta\theta}u(r, \theta) &= F(r, \theta) \\ u(R, \theta) &= f(\theta).\end{aligned}\tag{392}$$

By finding the eigenfunctions, $M_k(r, \theta)$, of the 2D elliptic operator

$$L = \partial_{rr} + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta},\tag{393}$$

we can express the solution u as a Fourier series

$$u(r, \theta) = \sum_k a_k M_k(r, \theta).\tag{394}$$

This is done in chapter 8 of our text-book. Here we will solve this problem using the finite Fourier transform. Introduce the operators

$$K = -r^2\partial_{rr} - r\partial_r,\tag{395}$$

$$\hat{L} = -\partial_{\theta\theta}.\tag{396}$$

Then (392) can be written as

$$Ku(r, \theta) + \hat{L}u(r, \theta) = r^2 F(r, \theta).\tag{397}$$

The idea is now to introduce a finite Fourier transform based on the eigenfunctions of the operator \hat{L} . Observe that since θ is the polar angle, running from zero to 2π , all functions are periodic in θ with period 2π . Thus we are lead to the eigenvalue problem

$$-\partial_{\theta\theta}M(\theta) = \lambda M(\theta), \quad 0 < \theta < 2\pi,\tag{398}$$

$$M(\theta + 2\pi) = M(\theta).\tag{399}$$

This problem is easiest to solve using complex exponentials. The general solution of the equation (398) is for $\lambda > 0$

$$M(\theta) = Ae^{i\sqrt{\lambda}\theta} + Be^{-i\sqrt{\lambda}\theta},\tag{400}$$

and the periodicity condition (399) implies that

$$\begin{aligned}Ae^{i\sqrt{\lambda}\theta} + Be^{-i\sqrt{\lambda}\theta} &= Ae^{i\sqrt{\lambda}(\theta+2\pi)} + Be^{-i\sqrt{\lambda}(\theta+2\pi)}, \\ \Updownarrow \\ e^{i\sqrt{\lambda}\theta} \left(1 - e^{i\sqrt{\lambda}2\pi}\right) A + e^{-i\sqrt{\lambda}\theta} \left(1 - e^{-i\sqrt{\lambda}2\pi}\right) B &= 0.\end{aligned}\tag{401}$$

The functions $e^{\pm i\sqrt{\lambda}\theta}$ are linearly independent (verify this), so we must conclude from (401) that

$$\begin{aligned} A(1 - e^{i\sqrt{\lambda}2\pi}) &= 0, \\ B(1 - e^{-i\sqrt{\lambda}2\pi}) &= 0, \Leftrightarrow B(1 - e^{i\sqrt{\lambda}2\pi}) = 0. \end{aligned} \quad (402)$$

If $(1 - e^{-i\sqrt{\lambda}2\pi}) \neq 0$ both A and B must be zero. This gives $M(\theta) = 0$, so such λ s are not eigenvalues. The other possibility is that

$$e^{-i\sqrt{\lambda}2\pi} = 1, \quad (403)$$

and then A, B are arbitrary. The equation (403) is only satisfied if

$$\begin{aligned} 2\pi\sqrt{\lambda_k} &= 2\pi k, \quad k = 1, 2, \dots, \\ &\Updownarrow \\ \lambda_k &= k^2. \end{aligned} \quad (404)$$

These are the positive eigenvalues. The corresponding eigenspaces are two-dimensional and spanned by the normalized eigenfunctions

$$\frac{1}{\sqrt{\pi}} \cos k\theta, \quad \frac{1}{\sqrt{\pi}} \sin k\theta. \quad (405)$$

In order to find out if $\lambda = 0$ is an eigenvalue we put $\lambda = 0$ in (398) and get the equation

$$\partial_{\theta\theta} M(\theta) = 0, \quad (406)$$

whose general solution is

$$M(\theta) = A\theta + B. \quad (407)$$

The boundary condition now gives

$$\begin{aligned} M(\theta + 2\pi) &= M(\theta), \\ &\Updownarrow \\ A(\theta + 2\pi) + B &= A\theta + B, \\ &\Updownarrow \\ 2\pi A &= 0 \Rightarrow A = 0, \end{aligned}$$

and B is arbitrary, so the eigenspace corresponding to $\lambda = 0$ is one-dimensional and spanned by the normalized eigenfunction

$$\frac{1}{\sqrt{2\pi}}. \quad (408)$$

Verify that there are no negative eigenvalues.

Using (405) and (408) we define a finite Fourier transform for functions $u(r, \theta)$ by

$$a_0(r) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} d\theta u(r, \theta), \quad (409)$$

$$\left. \begin{aligned} a_k(r) &= \frac{1}{\sqrt{\pi}} \int_0^{2\pi} d\theta u(r, \theta) \cos k\theta \\ b_k(r) &= \frac{1}{\sqrt{\pi}} \int_0^{2\pi} d\theta u(r, \theta) \sin k\theta \end{aligned} \right\} \quad k = 1, 2, \dots \quad (410)$$

The corresponding inverse Finite Fourier transform is

$$\begin{aligned} u(r, \theta) &= \frac{1}{\sqrt{2\pi}} a_0(r) \\ &+ \sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi}} \{a_k(r) \cos k\theta + b_k(r) \sin k\theta\}. \end{aligned} \quad (411)$$

The ODEs for the coefficients a_k and b_k are now found by multiplying (397) by the eigenfunctions and integrating over the interval $(0, 2\pi)$. Verify that we get

$$\begin{aligned} a_k''(r) + \frac{1}{r} a_k'(r) - \frac{k^2}{r^2} a_k(r) &= -A_k(r), \quad k = 0, 1, 2, \dots, \\ b_k''(r) + \frac{1}{r} b_k'(r) - \frac{k^2}{r^2} b_k(r) &= -B_k(r), \quad k = 0, 1, 2, \dots, \end{aligned} \quad (412)$$

where A_k and B_k are the Fourier coefficients of F ,

$$A_0(r) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} d\theta F(r, \theta), \quad (413)$$

$$\left. \begin{aligned} A_k(r) &= \frac{1}{\sqrt{\pi}} \int_0^{2\pi} d\theta F(r, \theta) \cos k\theta \\ B_k(r) &= \frac{1}{\sqrt{\pi}} \int_0^{2\pi} d\theta F(r, \theta) \sin k\theta \end{aligned} \right\} \quad k = 1, 2, \dots \quad (414)$$

Note that only solutions of the singular equations (412) that are bounded as $r \rightarrow 0$ are acceptable. So our boundary condition at $r = 0$ is

$$\lim_{r \rightarrow 0} \{a_k(r), b_k(r)\} \text{ exists.} \quad (415)$$

At $r = R$ the boundary condition is

$$u(R, \theta) = f(\theta).$$

Taking the Finite Fourier transform of the boundary data $f(\theta)$, we find the the boundary conditions for the functions $a_k(r), b_k(r)$ at $r = R$ are

$$a_0(R) = \alpha_0 \equiv \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} d\theta f(\theta), \quad (416)$$

$$\left. \begin{aligned} a_k(R) &= \alpha_k \equiv \frac{1}{\sqrt{\pi}} \int_0^{2\pi} d\theta f(\theta) \cos k\theta \\ b_k(R) &= \beta_k \equiv \frac{1}{\sqrt{\pi}} \int_0^{2\pi} d\theta f(\theta) \sin k\theta \end{aligned} \right\} \quad k = 1, 2, \dots \quad (417)$$

Note that both of the equations in (412) are inhomogeneous versions of the homogeneous *Euler equation*

$$C_k''(r) + \frac{1}{r} C_k'(r) - \frac{k^2}{r^2} C_k(r) = 0. \quad (418)$$

A basis for the solution space of such equations can be found using standard methods. We have

$$C_0(r) = \begin{cases} \text{constant} \\ \ln r \end{cases}, \quad C_k(r) = \begin{cases} r^k \\ r^{-k} \end{cases} \quad k = 1, 2, \dots \quad (419)$$

The solutions of (412) can now be found using variation of parameters. We have

$$\begin{aligned} a_0(r) &= \int_0^r dt \ln \left(\frac{R}{r} \right) A_0(t)t + \int_r^R dt \ln \left(\frac{R}{t} \right) A_0(t)t + \alpha_k(R), \\ a_k(r) &= \frac{1}{2k} \int_0^r dt \left[\left(\frac{R}{r} \right)^k - \left(\frac{r}{R} \right)^k \right] \left(\frac{t}{R} \right)^k A_k(t)t \end{aligned} \quad (420)$$

$$\begin{aligned} &+ \frac{1}{2k} \int_r^R dt \left[\left(\frac{R}{t} \right)^k - \left(\frac{t}{R} \right)^k \right] \left(\frac{r}{R} \right)^k A_k(t)t + \alpha_k(R) \left(\frac{r}{R} \right)^k, \\ b_k(r) &= \frac{1}{2k} \int_0^r dt \left[\left(\frac{R}{r} \right)^k - \left(\frac{r}{R} \right)^k \right] \left(\frac{t}{R} \right)^k B_k(t)t \\ &+ \frac{1}{2k} \int_r^R dt \left[\left(\frac{R}{t} \right)^k - \left(\frac{t}{R} \right)^k \right] \left(\frac{r}{R} \right)^k B_k(t)t + \beta_k(R) \left(\frac{r}{R} \right)^k. \end{aligned} \quad (421)$$

This completes the formal solution of (390).

Let us now return to the general scheme. As we have indicated, the solutions found using the finite Fourier transform will in general have slow convergence because the eigenfunctions satisfy the homogeneous boundary conditions, whereas in general the solution might satisfies inhomogeneous boundary conditions.

However, the problem of slow convergence can in general be solved using an approach which we will now explain.

Let us first consider the case when the inhomogeneous terms in the equation and boundary conditions are independent of time (we are here thinking about the hyperbolic and the parabolic cases).

Thus we have

$$\begin{aligned}\rho(\mathbf{x})Ku(\mathbf{x},t) + Lu(\mathbf{x},t) &= \rho(\mathbf{x})F(\mathbf{x}) \quad \text{in } G, \text{ for } t > 0, \\ \alpha(\mathbf{x})u(\mathbf{x},t) + \beta(\mathbf{x})\partial_{\mathbf{n}}u(\mathbf{x},t) &= B(\mathbf{x}) \quad \text{on } \partial G, \text{ for } t > 0.\end{aligned}\tag{422}$$

Let $v(\mathbf{x})$ be a time independent solution to (422). Thus

$$\begin{aligned}Lv(\mathbf{x}) &= \rho(\mathbf{x})F(\mathbf{x}) \quad \text{in } G, \\ \alpha(\mathbf{x})v(\mathbf{x}) + \beta(\mathbf{x})\partial_{\mathbf{n}}v(\mathbf{x}) &= B(\mathbf{x}) \quad \text{on } \partial G.\end{aligned}\tag{423}$$

We will assume that we can solve (423). Define $w(\mathbf{x},t) = u(\mathbf{x},t) - v(\mathbf{x})$. Then $w(\mathbf{x})$ solves

$$\rho(\mathbf{x})Kw(\mathbf{x},t) + Lw(\mathbf{x},t)\tag{424}$$

$$\begin{aligned}&= \rho(\mathbf{x})Ku(\mathbf{x},t) - \rho(\mathbf{x})Kv(\mathbf{x}) + Lu(\mathbf{x},t) - Lv(\mathbf{x}) \\ &= \rho(\mathbf{x})Ku(\mathbf{x},t) + Lu(\mathbf{x},t) - Lv(\mathbf{x}) \\ &= \rho(\mathbf{x})(F(\mathbf{x}) - F(\mathbf{x})) = 0,\end{aligned}\tag{425}$$

and has boundary condition

$$\alpha(\mathbf{x})w(\mathbf{x},t) + \beta(\mathbf{x})\partial_{\mathbf{n}}w(\mathbf{x},t)\tag{426}$$

$$\begin{aligned}&= \alpha(\mathbf{x})u(\mathbf{x},t) - \alpha(\mathbf{x})v(\mathbf{x}) \\ &+ \beta(\mathbf{x})\partial_{\mathbf{n}}u(\mathbf{x},t) - \beta(\mathbf{x})\partial_{\mathbf{n}}v(\mathbf{x}) \\ &= B(\mathbf{x}) - B(\mathbf{x}) = 0.\end{aligned}\tag{427}$$

Thus $w(\mathbf{x},t)$ satisfies a homogeneous equation and homogeneous boundary conditions. The formal series for $w(\mathbf{x},t)$ will therefore have better convergence properties than the formal series for $u(\mathbf{x},t)$. If $v(\mathbf{x})$ can be found without using eigenfunction expansions, the solution $u(\mathbf{x},t)$ will be represented by a Fourier series with good convergence properties.

Finding an expression for $v(\mathbf{x})$ might however not be easy.

There is however a related approach that can be used and that do not rely on finding a stationary solution.

Let us assume that we can find a function $v(\mathbf{x}, t)$ such that

$$\alpha(\mathbf{x})v(\mathbf{x}, t) + \beta(\mathbf{x})\partial_{\mathbf{n}}v(\mathbf{x}, t) = B(\mathbf{x}, t). \quad (428)$$

The function v must be sufficiently smooth so that it is possible to substitute into the equation

$$\rho(\mathbf{x})Ku(\mathbf{x}, t) + Lu(\mathbf{x}, t) = \rho(\mathbf{x})F(\mathbf{x}, t). \quad (429)$$

Define a function $w(\mathbf{x}, t) = u(\mathbf{x}, t) - v(\mathbf{x}, t)$. Then

$$\begin{aligned} & \alpha(\mathbf{x})w(\mathbf{x}, t) + \beta(\mathbf{x})\partial_{\mathbf{n}}w(\mathbf{x}, t) \\ &= \alpha(\mathbf{x})u(\mathbf{x}, t) - \alpha(\mathbf{x})v(\mathbf{x}, t) \\ &+ \beta(\mathbf{x})\partial_{\mathbf{n}}u(\mathbf{x}, t) - \beta(\mathbf{x})\partial_{\mathbf{n}}v(\mathbf{x}, t) = B(\mathbf{x}, t) - B(\mathbf{x}, t) = 0, \end{aligned} \quad (430)$$

and

$$\begin{aligned} & \rho(\mathbf{x})Kw(\mathbf{x}, t) + Lw(\mathbf{x}, t) \\ &= \rho(\mathbf{x})Ku(\mathbf{x}, t) - \rho(\mathbf{x})Kv(\mathbf{x}, t) + Lu(\mathbf{x}, t) - Lv(\mathbf{x}, t) \\ &= \rho(\mathbf{x})F(\mathbf{x}, t) - \rho(\mathbf{x})Kv(\mathbf{x}, t) - Lv(\mathbf{x}, t) \\ &\equiv \rho(\mathbf{x})\hat{F}(\mathbf{x}, t), \end{aligned} \quad (431)$$

where

$$\hat{F}(\mathbf{x}, t) = F(\mathbf{x}, t) - Kv(\mathbf{x}, t) - \frac{Lv(\mathbf{x}, t)}{\rho(\mathbf{x})}.$$

Thus $w(\mathbf{x}, t)$ solves an inhomogeneous equation with homogeneous boundary conditions. The solution to this problem can be found using Duhamel's principle or the finite Fourier transform. Convergence of the series should now be much faster and thus lead to more efficient numerical calculations.

Let us illustrate this method for a 1D inhomogeneous heat equation.

$$\begin{aligned} & u_t(x, t) - c^2u_{xx}(x, t) = g(x, t), \quad 0 < x < l, \quad t > 0, \\ & u(x, 0) = f(x), \quad 0 < x < l, \\ & u(0, t) = g_1(t), \quad t > 0, \\ & u(l, t) = g_2(t), \quad t > 0. \end{aligned} \quad (432)$$

It is easy to see, using linear interpolation that the function

$$v(x, t) = \frac{1}{l}(xg_2(t) + (l - x)g_1(t)), \quad (434)$$

satisfy the inhomogeneous boundary conditions. Defining $w(x, t) = u(x, t) - v(x, t)$, we find immediately that

$$w_t(x, t) - c^2 w_{xx}(x, t) = g(x, t) - \frac{1}{l}(xg_2'(t) + (l-x)g_1'(t)), \quad (435)$$

for $t > 0$ and $0 < x < l$. The boundary condition are by construction homogeneous

$$\begin{aligned} w(0, t) &= 0, \\ w(l, t) &= 0, \end{aligned} \quad (436)$$

and the initial value is

$$\begin{aligned} w(x, 0) &= u(x, 0) - v(x, 0) \\ &= f(x) - \frac{1}{l}(xg_2(0) - (l-x)g_1(0)). \end{aligned} \quad (437)$$

This problem can now be solved by, for example series expansion, and the convergence of the series should be good.

10.5 Nonlinear stability theory

The main focus of this class is the theory of *linear* partial differential equations. The reason for this is two-fold. Firstly, many systems of scientific and technological interest can be described by such equation. We have seen several modeling examples where this turns out to be the case. Secondly, there is no theory of nonlinear PDEs with a generality anywhere close to the case for linear PDEs. Nonlinear PDEs in general require an individual treatment. In fact some people spend their whole working career studying one or a few such nonlinear equations. This is, however, if one is looking for exact solutions, precise statements about solution spaces etc. If one is satisfied with approximate answers, analytic or numeric, a lot more can be achieved.

The reason for this is that in many situations the nonlinear terms in the PDEs are much smaller than the linear ones. For example, we saw at the beginning of this class that the PDE

$$\partial_t n + a(n) \partial_x n = 0, \quad (438)$$

is a simple model of traffic flow on a one-lane road. Assuming n is small, thus the density of cars is low, we can Taylor expand $a(n)$ and truncate at first order

$$a(n) \approx a_0 + a_1 n. \quad (439)$$

Inserting (439) into (438) and rearranging the terms we get the nonlinear equation

$$\partial_t n + a_0 \partial_x n = -a_1 n \partial_x n. \quad (440)$$

By assumption (439), the nonlinear term in this equation is smaller than the linear ones.

The use of truncated expansions like (439) are very common when modeling physical systems by PDEs. We saw this for example for the vibrating string.

The refractive index, n , is a material property that determines the optical response of dielectrics and gases. For inhomogeneous materials, like corrugated glass for example, it will depend on position $n = n(\mathbf{x})$.

An important equation in the theoretical description of light is

$$\frac{1}{c^2} \partial_{tt}(n^2 E) - \nabla^2 E = 0. \quad (441)$$

This is an equation which models light under normal conditions well. It is clearly a linear equation. For light under more extreme conditions, for example light generated by a laser, one finds that the refractive index depends on the light intensity $I = |E|^2$. We thus in general have $n = n(\mathbf{x}, I)$. This function has been measured for many materials and one finds that in many cases a good approximation is

$$n = n_0 + n_2 I, \quad (442)$$

where both n_0 and n_2 in general depend on position. Inserting (442) into (441) and keeping only terms that are linear in the intensity I , we get the equation

$$\frac{n_0^2}{c^2} \partial_{tt} E - \nabla^2 E = -\frac{2n_0 n_2}{c^2} \partial_{tt} (|E|^2 E). \quad (443)$$

This is now a nonlinear equation. The factor n_2 is for most materials very small, so light of very high intensity $I = |E|^2$ is required in order for the nonlinearity to have any effect. From your basic physics class you might recall that light tends to bend towards areas of higher refractive index; this property is the basis for Snell's law in optics. Thus the intensity dependent index (442) will tend to focus light into areas of high light intensity. This increases the intensity, creating an even larger index, which focuses light even more strongly. This is a runaway effect that quickly creates a local light intensity high enough to destroy the material the light is traveling through.

The study of nonlinear properties of high-intensity light is a large and very active field of research which has given us amazing technologies, like the laser, optical fibers, perfect lenses of microscopic size and even invisibility cloaks! (Well almost anyway)

For some physical systems there are no good linear approximations. One such area is fluid dynamics. The most important equation in this field of science is the *Navier-Stokes* equation.

$$\rho_0 (\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}) = -\nabla p + \mu \nabla^2 \mathbf{u}, \quad \nabla \cdot \mathbf{u} = 0, \quad (444)$$

where ρ_0 is the mass density, $\mathbf{u}(\mathbf{x}, t)$ the fluid velocity field, $p(\mathbf{x}, t)$ the pressure in the fluid and μ the viscosity (friction). The second term on the left hand side of (444) is clearly nonlinear and in many important applications not small. The

essential nonlinearity of the Navier-Stokes equation gives rise to the phenomenon of *turbulence* and other physical phenomena that are extremely hard to describe and predict. In fact, it is not even known, after almost 200 years, if the Cauchy problem for (444) is well posed. There is a possibility that solutions of (444) develop singularities in finite time. If you can solve this problem, the Clay mathematics institute will pay you one million dollars!

For systems where the nonlinear terms are small, one would expect that methods introduced for solving linear PDEs should be useful, somehow. They are, and we will illustrate this using a nonlinear heat equation

$$u_t - u_{xx} = \hat{\lambda}u(1 - u^2). \quad (445)$$

The parameter $\hat{\lambda}$ measures the strength of the nonlinearity. The equation can model situations where we have a nonlinear heat source. The equation (445) clearly has a solution

$$u_0(x, t) = 0. \quad (446)$$

This corresponds to a situation where the system has a uniform temperature equal to zero. We know that if the nonlinear heat source is not present, any initial disturbance will decay, and the system will return to a uniform zero temperature state. We have seen this happening using normal modes and also Fourier series solutions. The question we want to ask is if $u_0(x, t) = 0$ is stable also in the presence of the nonlinear heat source. We will investigate this problem both for an infinite system, $-\infty < x < \infty$, and a finite one $(0, l)$. We will see that the stability condition is different in the two cases. This is a general fact: Finite and infinite systems display different stability behavior.

Let us first set the problem up. Since we will be concerned with solutions close to $u_0(x, t) = 0$, we introduce a small parameter $0 < \epsilon < 1$ and write the initial condition as

$$u(x, 0) = \epsilon h(x), \quad (447)$$

where $h(x)$ is of normal size, $h(x) = \mathcal{O}(1)$. We are similarly introducing a $\mathcal{O}(1)$ function $w(x, t)$ through

$$u(x, t) = \epsilon w(x, t). \quad (448)$$

The equation (445) now gives

$$w_t - w_{xx} = \hat{\lambda}w - \epsilon^2 \hat{\lambda}w^3. \quad (449)$$

The first step is to do a *linear stability analysis*. For this we recognize that, if $w = \mathcal{O}(1)$, the last term in (449) is much smaller than the other terms because $\epsilon \ll 1$. We therefore drop the small term and are left with the equation

$$w_t - w_{xx} = \hat{\lambda}w. \quad (450)$$

(450) is an arbitrary good approximation to (449) if ϵ is small enough and $w = \mathcal{O}(1)$. Going from (449) to (450) is called *linearization*. Linearization is an idea of outmost importance in applied mathematics.

We will now investigate the linearized equation (450) for the infinite and finite case separately.

10.5.1 The infinite case

In the infinite case $-\infty < x < \infty$, we can analyze the stability of equation (450) using normal modes as discussed previously in this class. Inserting a normal mode solution

$$w(x, t) = a(k)e^{ikx + \lambda(k)t}, \quad (451)$$

into (450), we find the algebraic equation

$$\lambda(k) + k^2 = \hat{\lambda}, \quad (452)$$

$$\Updownarrow$$

$$\lambda(k) = \hat{\lambda} - k^2. \quad (453)$$

From this we see that the stability index for the equation (450) is

$$\Omega = \sup_k \Re[\lambda(k)] = \hat{\lambda}. \quad (454)$$

The conclusion is then that

- i) u_0 is stable if $\hat{\lambda} < 0$,
- ii) u_0 is unstable if $\hat{\lambda} > 0$.

In the unstable case the normal mode (451) experience unbounded exponential growth. However, when the growth makes

$$\hat{\lambda}w \sim \epsilon^2 \hat{\lambda}w^3, \quad (455)$$

the nonlinear term is of the same size as the linear one and the linearized equation is not valid anymore. From (455) this happend when

$$\begin{aligned} \epsilon^2 w^2 &\sim 1, \\ \Updownarrow \\ w &\sim \frac{1}{\epsilon}, \end{aligned} \quad (456)$$

which gives us, using (451), a break-down time

$$t^* = \frac{|\log \epsilon|}{\hat{\lambda}}. \quad (457)$$

So what happens after the break-down time? Observe that the nonlinear equation (445) has another simple uniform solution

$$u_1(x, t) = 1. \quad (458)$$

We can now ask when this solution is stable. We introduce a small parameter ϵ and a $\mathcal{O}(1)$ function $v(x, t)$ through

$$u(x, t) = 1 + \epsilon v(x, t). \quad (459)$$

Inserting this into (445) we find

$$\begin{aligned} \epsilon(v_t - v_{xx}) &= \hat{\lambda}(1 + \epsilon v)(1 - (1 + \epsilon v)^2) \\ &= \hat{\lambda}(1 + \epsilon v)(1 - 1 - 2\epsilon v - \epsilon^2 v^2) \\ &= -\epsilon \hat{\lambda}(1 + \epsilon v)(2v + \epsilon v^2) \\ &= -2\epsilon \hat{\lambda}v - 2\epsilon^2 \hat{\lambda}v^2 - \epsilon^2 \hat{\lambda}v^2 - \epsilon^2 \hat{\lambda}v^3, \\ &\Downarrow \\ v_t - v_{xx} &= -2\hat{\lambda}v - 3\epsilon \hat{\lambda}v^2 - \epsilon^2 \hat{\lambda}v^3. \end{aligned}$$

Linearizing, we find the equation

$$v_t - v_{xx} = -2\hat{\lambda}v. \quad (460)$$

Inserting a normal mode now gives

$$\lambda = -2\hat{\lambda} - k^2, \quad (461)$$

and we can conclude

- i) u_1 is stable if $\hat{\lambda} > 0$,
- ii) u_1 is unstable if $\hat{\lambda} < 0$.

Thus, the solution u_1 is stable when u_0 is unstable. A reasonable guess now is that a small perturbation to the solution u_0 in the unstable regime ($\hat{\lambda} > 0$) grows and eventually approaches the stable solution u_1 . Solving equation (445) numerically with a broad selection of small initial data $u(x, 0) = \epsilon h(x)$ is needed to support our conjecture. Such situations can not *prove*, in the mathematical sense, that our conjecture is true. However, in real-life applied mathematics, proofs are few and far between, and we must use the tools at our disposal to do as much as we can to dispel uncertainty as to the validity of our conclusions.

Another way to support our conjecture is to find solutions for special initial data. Let us consider the special initial data

$$u(x, 0) = \epsilon \quad (\text{independent of } x). \quad (462)$$

Let us look for a solution to (445) that is independent of x , $u = u(t)$. Such a solution must satisfy

$$\begin{aligned} u'(t) &= \hat{\lambda}u(t)(1 - u^2(t)), \\ u(0) &= \epsilon. \end{aligned} \tag{463}$$

This first order ODE that can be solved exactly

$$u(x, t) = u(t) = \frac{\epsilon e^{\hat{\lambda}t}}{(1 + \epsilon^2(e^{2\hat{\lambda}t} - 1))^{\frac{1}{2}}}. \tag{464}$$

It is easy to verify that

$$\lim_{t \rightarrow +\infty} u(x, t) = 1 = u_1(x, t), \tag{465}$$

for the case $\hat{\lambda} > 0$. Thus this special solution lends further support to our conjecture.

10.5.2 The finite case

Let us consider the nonlinear heat equation (445) on a finite interval $0 < x < \pi$ with zero boundary conditions

$$u(0, t) = u(\pi, t) = 0. \tag{466}$$

Recall that linearizing (445) around the simple solution $u_0(x, t) = 0$ gives us an equation of the form

$$w_t - w_{xx} = \hat{\lambda}w, \quad 0 < x < \pi, \tag{467}$$

$$w(0, t) = w(\pi, t) = 0. \tag{468}$$

We can solve this problem using the finite Fourier transform. Introduce the operator $L = -\partial_{xx}$. The relevant eigenfunctions satisfy

$$\begin{aligned} LM_k(x) &= \lambda M_k(x) \\ M_k(0) &= M_k(\pi) = 0 \end{aligned} \tag{469}$$

We have analyzed this eigenvalue problem before. The eigenvalues are

$$\lambda_k = -k^2, \quad k = 1, 2, \dots$$

and the normalized eigenfunctions are

$$M_k(x) = \sqrt{\frac{2}{\pi}} \sin kx. \tag{470}$$

The associated finite Fourier transform of a given function $f(x)$ is

$$N_k = (f, M_k), \quad k = 1, 2, \dots \quad (471)$$

and the inverse finite Fourier transform is

$$f(x) = \sum_k N_k M_k(x). \quad (472)$$

We now express an arbitrary solution to (467) in terms of an inverse finite Fourier transform

$$w(x, t) = \sum_k N_k(t) M_k(x). \quad (473)$$

The evolution equations for the $N_k(t)$ are found from (467) if we multiply by $M_k(x)$ and integrate

$$\begin{aligned} \int_0^\pi dx w_t(x, t) M_k(x) - \int_0^\pi dx w_{xx}(x, t) M_k(x) \\ = \int_0^\pi dx \hat{\lambda} w(x, t) M_k(x). \end{aligned} \quad (474)$$

For each of the three terms in this equation we have

$$\int_0^\pi dx w_t(x, t) M_k(x) = \partial_t \int_0^\pi dx w(x, t) M_k(x) = N'_k(t), \quad (475)$$

$$\int_0^\pi dx w_{xx}(x, t) M_k(x) = w_x(x, t) M_k(x) \Big|_0^\pi - \int_0^\pi dx w_x(x, t) M_{kx}(x) \quad (476)$$

$$\begin{aligned} - w(x, t) M_{kx}(x) \Big|_0^\pi + \int_0^\pi dx w(x, t) M_{kxx}(x) \\ = \int_0^\pi dx w(x, t) (-k^2) M_k(x) = -k^2 N_k(t), \end{aligned} \quad (477)$$

$$\int_0^\pi dx \hat{\lambda} w(x, t) M_k(x) = \hat{\lambda} N_k(t) \quad (478)$$

Inserting (466), (476) and (478) into (474) gives the following uncoupled system of ODEs for the unknown functions $N_k(t)$

$$N'_k(t) + k^2 N_k(t) = \hat{\lambda} N_k(t), \quad (479)$$

$$\Updownarrow$$

$$N'_k(t) = (\hat{\lambda} - k^2) N_k(t),$$

$$\Updownarrow$$

$$N_k(t) = a_k e^{(\hat{\lambda} - k^2)t}.$$

The general solution to (467) can thus be written as

$$w(x, t) = \sum_{k=1}^{\infty} w_k(x, t), \quad (480)$$

where the *normal modes*, $w_k(x, t)$, are

$$w_k(x, t) = a_k e^{(\hat{\lambda} - k^2)t} \sin kx, \quad k = 1, 2, \dots \quad (481)$$

We see that

- i) $\hat{\lambda} < 1 \Rightarrow$ all normal modes decay $\Rightarrow u_0(x, t) = 0$ is stable.
- ii) $\hat{\lambda} > 1 \Rightarrow$ at least one normal mode grows exponentially $\Rightarrow u_0(x, t)$ is unstable.

For the infinite domain case the stability condition was $\hat{\lambda} < 0$. Thus the stability condition is not the same for the case of infinite and finite domain. This is in general true.

For the unstable case, $\hat{\lambda} > 1$, one or more normal modes experience exponential growth, which after a finite time invalidate the linearized equation (467). What is the system going to do after this time? This is a questing of *nonlinear stability*.

Nonlinear stability is a much harder problem than linear stability but powerful techniques exist in the form of asymptotic expansions which in many important cases can provide answer to the question of nonlinear stability.

Here we will not dive into this more sophisticated approach but will rather introduce an approach based on the finite Fourier transform.

Observe that the finite Fourier transform can also be applied to the nonlinear equation

$$w_t - w_{xx} = \hat{\lambda}w(1 - \epsilon^2 w^2) = \hat{\lambda}w - \hat{\lambda}\epsilon^2 w^3. \quad (482)$$

Representing the solution using the Finite Fourier transform we have

$$w(x, t) = \sum_{k=1}^{\infty} N_k(t) M_k(x). \quad (483)$$

The evolution equations for the $N_k(t)$ s are found as for the linearized case. We have

$$\begin{aligned} & \int_0^{\pi} dx w_t(x, t) M_k(x) - \int_0^{\pi} dx w_{xx}(x, t) M_k(x) \\ &= \int_0^{\pi} dx \hat{\lambda} w(x, t) M_k(x) - \int_0^{\pi} dx \hat{\lambda} \epsilon^2 w^3(x, t) M_k(x). \end{aligned} \quad (484)$$

The only new term as compared with the linearized case is the second term on the right hand side of (484). For this term we have

$$\begin{aligned}
& \int_0^\pi dx \hat{\lambda} \epsilon^2 w^3(x, t) M_k(x) \\
&= \hat{\lambda} \epsilon^2 \sum_{i,j,l=1}^\infty \int_0^\pi dx N_i(t) N_j(t) N_l(t) M_i(x) M_j(x) M_l(x) M_k(x) \\
&= \hat{\lambda} \epsilon^2 \sum_{i,j,l=1}^\infty a_k^{ijl} N_i(t) N_j(t) N_l(t),
\end{aligned} \tag{485}$$

where

$$a_k^{ijl} = \int_0^\pi dx M_i(x) M_j(x) M_k(x) M_l(x). \tag{486}$$

Using (475), (476), (478) and (485) in (484) gives us

$$N'_k(t) + (k^2 - \hat{\lambda}) N_k(t) = -\hat{\lambda} \epsilon^2 \sum_{i,j,l=1}^\infty a_k^{ijl} N_i(t) N_j(t) N_l(t). \tag{487}$$

(487) is an infinite system of nonlinear coupled ODEs. This is evidently a very hard problem to solve in any exact sense. (487) shows the complexity of nonlinear PDEs. The very compact notation from calculus makes the nonlinear PDE (482) to *appear* simple, but as (487) shows the problem of solving (482) is actually horrendously difficult.

In order to get anywhere, we must restrict to situations where some of the complexity of (487) is reduced. Finding such situations is what good applied mathematicians are trained to do.

Let us assume that $\hat{\lambda} > 1$ so we are in the unstable regime, but that $\hat{\lambda} < 4$ so that only the normal mode $w_1(x, t)$ is unstable. All the other ones decay exponentially. For the $N_k(t)$ s this means that

$$N_1(t) \gg N_k(t), \quad \forall k \geq 2. \tag{488}$$

Thus, on the right hand side of (487) the term

$$-\hat{\lambda} \epsilon^2 a_1^{111} N_1^3(t), \tag{489}$$

is much larger than the other terms. A good approximation to (487) therefore appears to be

$$N'_1(t) + (1 - \hat{\lambda}) N_1(t) + \hat{\lambda} \epsilon^2 a_1^{111} N_1^3(t) = 0. \tag{490}$$

The coefficient a_1^{111} can be calculated using (486)

$$a_1^{111} = \int_0^\pi dx (M_1(x))^4 = \frac{4}{\pi} \int_0^\pi dx \sin^4 x = \frac{3}{2\pi}. \tag{491}$$

$$\Rightarrow N_1'(t) + (1 - \hat{\lambda})N_1(t) + \frac{3\hat{\lambda}}{2\pi}\epsilon^2 N_1^3(t) = 0 \quad (492)$$

If we multiply this equation by $N_1(t)$ we obtain

$$\frac{1}{2}y' + (1 - \hat{\lambda})y + \frac{3\hat{\lambda}\epsilon^2}{2\pi}y^2 = 0, \quad (493)$$

where $y(t) = N_1^2(t)$. (493) is a so-called Ricatti equation and can be solved exactly. There is a general way to solve such equations.

A general Ricatti equation is of the form

$$y' = q_0(x) + q_1(x)y + q_2(x)y^2. \quad (494)$$

First introduce a function $v(x)$ by

$$v(x) = q_2(x)y(x). \quad (495)$$

Then we get the following equation for $v(x)$

$$v' = v^2 + R(x)v + S(x), \quad (496)$$

where

$$S(x) = q_2(x)q_0(x), \quad (497)$$

$$R(x) = q_1(x) + \frac{q_2'(x)}{q_2(x)}.$$

Next, introduce the function $u(x)$ by

$$v = -\frac{u'}{u}. \quad (498)$$

Then we get the following *linear* equation for u

$$u''(x) - R(x)u'(x) + S(x)u = 0. \quad (499)$$

Using (498) and (495), any solution of (499) gives a solution of the Ricatti equation (494). Since the general solution to (499) can be found explicitly in some cases, the general solution (494) can also be found for those cases.

For our equation (493) we have

$$q_0(x) = 0 \quad (500)$$

$$q_1(x) = 2(\hat{\lambda} - 1)$$

$$q_2(x) = -\frac{3\hat{\lambda}\epsilon^2}{\pi}$$

We can find the general solution to this equation which leads to the general solution (493). With initial condition given as in (447) we get

$$N_1(t) = \frac{(h, M_1)e^{(\hat{\lambda}-1)t}}{\left[1 + \left[3\hat{\lambda}\epsilon^2(h, M_1)^2/(2\pi(\hat{\lambda}-1))\right] \left(e^{2(\hat{\lambda}-1)t} - 1\right)\right]^{\frac{1}{2}}} \quad (501)$$

This solution tends to a finite stationary value as $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} N_1(t) = \frac{(h, M_1)}{|(h, M_1)|} \sqrt{\frac{2\pi}{3}} \left(\frac{\hat{\lambda}-1}{\hat{\lambda}\epsilon^2}\right)^{\frac{1}{2}} \quad (502)$$

rather than growing unbounded as the linearized equation predicts.

Note that the current method can be extended to study more general non-linear terms $\hat{\lambda}f(u)$.

11 Integral transforms

Integral transforms are used to solve PDEs on unbounded domains. They have very wide applicability and play a fundamental role in applied mathematics and theoretical physics.

In these notes I will concentrate on equations of the form

$$\rho(\mathbf{x})Ku(\mathbf{x}, t) + Lu(\mathbf{x}, t) = \rho(\mathbf{x})F(\mathbf{x}, t), \quad (503)$$

where

$$Lu(\mathbf{x}, t) = -\nabla \cdot (p\nabla u(\mathbf{x}, t)) + q(\mathbf{x})u(\mathbf{x}, t), \quad (504)$$

and where

$$K = \begin{cases} \partial_{tt} & \text{-- hyperbolic case,} \\ \partial_t & \text{-- parabolic case,} \\ \partial_{yy} & \text{-- elliptic case, cylinder symmetry,} \\ 0 & \text{-- general elliptic case.} \end{cases} \quad (505)$$

Recall that the basic idea in the separation of variables is to look for special solution to (503) of the form

$$u(\mathbf{x}, t) = N(t)M(\mathbf{x}), \quad (506)$$

($t \rightarrow y$ in the restricted elliptic case). We found that a separation constant λ appeared and that (506) is a solution to (503) only if

$$LM(\mathbf{x}) = \lambda^2 \rho(\mathbf{x})M(\mathbf{x}), \quad (507)$$

and

$$\begin{cases} N''(t) + \lambda^2 N(t) = 0 & - \text{hyperbolic case,} \\ N'(t) + \lambda^2 N(t) = 0 & - \text{parabolic case,} \\ N'(y) - \lambda^2 N(y) = 0 & - \text{restricted elliptic case.} \end{cases} \quad (508)$$

Here we have written the separation constant as λ^2 for convenience. Equation (507) must be supplied with boundary conditions. They must of course be inherited from the boundary conditions on equation (503). The most common ones are that the solution is bounded or that it vanishes as $\|x\| \rightarrow \infty$. Thus we will require that $M(\mathbf{x})$ is bounded or vanishes at infinity.

As for the case of bounded domains, (507) with boundary conditions will not have solutions for all λ . The λ s that give a solution are part of the *spectrum* of L . Things look very similar to the case of bounded domains, so far. There is however a big difference as compared with the bounded case.

In order to illustrate this difference, let us consider the boundary value problem

$$\begin{aligned} -M''(x) &= \lambda^2 M(\mathbf{x}) \quad -l < x < l, \\ M(l) &= M(-l) = 0. \end{aligned} \quad (509)$$

The general solution to the equation is for $\lambda > 0$

$$M(x) = Ae^{i\lambda x} + Be^{-i\lambda x}. \quad (510)$$

The boundary conditions imply that

$$\begin{aligned} Ae^{-i\lambda l} + Be^{i\lambda l} &= 0, \\ Ae^{i\lambda l} + Be^{-i\lambda l} &= 0, \end{aligned} \quad (511)$$

$$\begin{bmatrix} e^{-i\lambda l} & e^{i\lambda l} \\ e^{i\lambda l} & e^{-i\lambda l} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = 0. \quad (512)$$

In order for (512) to have a nontrivial solution, and thus for λ to be in spectrum of $L = -\partial_{xx}$, the matrix in (512) must have zero determinant

$$\det \begin{bmatrix} e^{-i\lambda l} & e^{i\lambda l} \\ e^{i\lambda l} & e^{-i\lambda l} \end{bmatrix} = 0, \quad (513)$$

$$\Updownarrow$$

$$e^{4i\lambda l} = 1,$$

$$\Updownarrow$$

$$4l\lambda_k = 2\pi k,$$

$$\lambda_k = \frac{\pi k}{2l}, \quad k = 1, 2, \dots \quad (514)$$

There are no eigenvalues for $\lambda \leq 0$. Thus (514) defines the spectrum of l .

Observe that the distance $\Delta\lambda$ between two spectral points is

$$\Delta\lambda = \frac{\pi}{2l}. \quad (515)$$

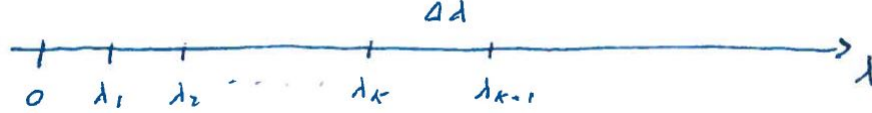


Figure 55

When the size of the domain increases without bound we observe that the distance between spectral points go to zero

$$\lim_{l \rightarrow \infty} \Delta\lambda = 0. \quad (516)$$

Thus in this limit the spectrum of L will form a continuum

$$0 < \lambda < \infty. \quad (517)$$

This is in general the case for unbounded domains; the spectrum of the operators form a continuum.

For the discrete case, with eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ and corresponding eigenfunctions $\{M_k\}_{k=1}^{\infty}$, the basic idea was to look for solutions to the PDE in the form of an infinite sum

$$u = \sum_{k=1}^{\infty} N_k M_k. \quad (518)$$

For the unbounded case we have eigenfunctions M_{λ} and corresponding N_{λ} , where the λ s form a continuum that we denote by D . By analogy with (512) we seek solutions to the PDE in the form of "sums over the continuum". Such sums are of course integrals. Thus

$$u = \int_D d\lambda N_{\lambda} M_{\lambda}. \quad (519)$$

We recall that in the case of finite domains, the fitting of initial conditions to our solution consisted in finding constants $\{a_k\}_{k=1}^{\infty}$ such that

$$f(\mathbf{x}) = \sum_{k=1}^{\infty} a_k M_k(\mathbf{x}). \quad (520)$$

Using orthonormality of the eigenfunctions we found that

$$a_k = (M_k, f) = \int_G dV \rho(\mathbf{x}) f(\mathbf{x}) M_k(\mathbf{x}). \quad (521)$$

Fitting initial conditions for the case of unbounded domains will by analogy involve finding coefficients a_λ , $\lambda \in D$ such that

$$f(\mathbf{x}) = \int_D d\lambda a_\lambda M_\lambda(\mathbf{x}). \quad (522)$$

In order to determine a_λ we need a continuum analog of the orthonormality condition from the discrete case

$$(M_k, M'_k) = \delta_{kk'}, \quad (523)$$

where we recall that $\delta_{kk'}$ is the Kronecker delta.

If we had something like a continuum version of the Kronecker delta one would guess at

$$(M_\lambda, M_{\lambda'}) = \delta_{\lambda\lambda'}, \quad (524)$$

This is in fact what *Paul Dirac* did early in the previous century. Dirac was an English genius that was one of the founding fathers of quantum theory. He came to this idea while trying to solve the Schrodinger equation from quantum mechanics. He found that $\delta_{\lambda\lambda'}$ only depended on $\lambda - \lambda'$ and he wrote it as

$$\delta_{\lambda\lambda'} = \delta(\lambda - \lambda'). \quad (525)$$

$\delta(\lambda)$ is called the *Dirac delta function* after him. In order to do its intended job he found that it would have to have the following two properties

$$\delta(\lambda) = 0, \quad \lambda \neq 0, \quad (526)$$

$$\int_D d\lambda \delta(\lambda) = 1. \quad (527)$$

It is a trivial exercise to prove that no function exists that have the two properties (526) and 527). Mathematicians were quick to point this out but Dirac didn't care, he did not know much mathematics, and he did not have to know it either, because he was in fact a genius and people like him move to a different drum.

Eventually mathematicians did make sense of the Dirac delta function and in fact discovered a whole new mathematical universe called *the theory of distributions*. We will discuss elements of this theory next semester. In this semester we will use objects like the Dirac delta sparingly and when we do only in a purely formal way like Dirac himself.

There is an important moral to the story of the Dirac delta. Pure mathematics is important because it gives us abstract frameworks to help organize

our mathematical methods. Pure mathematics can also give us more efficient ways of doing our calculations and make precise under which conditions our manipulations will keep us within the boundaries of the known mathematical universe.

However the framework and bounds imposed by pure mathematics should not stop us from sometimes disregarding the rules and play with the formulas, like Dirac did. In the final analysis, virtually all of mathematics originate from such rule-defying play.

It is however also appropriate to give a warning; disregarding the rules and playing with the formulas should not be a cover for sloppy thinking. This is unfortunately the case quite frequently in applied science.

We can't all be a Dirac. Disregarding the rules imposed by pure mathematics will, by most people, most of the time, produce useless nonsense.

In these notes I will sometimes break the rules a la Dirac and I will try to remember telling you when I do. All my rule breaking can be made sense of within the universe of distributions.

All finite Fourier transforms will in the limit of unbounded domains give rise to integral transforms. There is thus a large set of such transforms available, each tailored to a particular class of problems. In these notes I will focus on the Sine-, Cosine-, Fourier- and Laplace transforms.

11.1 The 1D Fourier transform

Let us consider the eigenvalue problem

$$M''(x) + \lambda^2 M(x) = 0, \quad -\infty < x < +\infty. \quad (528)$$

The boundary condition is that $M(x)$ is bounded when $|x| \rightarrow \infty$. The general solution of the equation (528) is

$$M_\lambda(x) = a(\lambda)e^{i\lambda x} + b(\lambda)e^{-i\lambda x}, \quad (529)$$

where $\lambda \in \mathbb{C}$. Writing $\lambda = \alpha + i\beta$ we have

$$\begin{aligned} e^{\pm i\lambda x} &= e^{\mp \beta x} e^{\pm i\alpha x}, \\ &\Downarrow \\ |e^{\pm i\lambda x}| &= e^{\mp \beta x}, \end{aligned}$$

and the functions $e^{\mp \beta x}$ are unbounded on $-\infty < x < \infty$ as long as $\beta \neq 0$. Thus to satisfy the boundary conditions we must restrict λ to the real axis $-\infty < x < \infty$. On the real axis impose the conditions

$$\begin{aligned} |a(\lambda)| &< M_a, \\ |b(\lambda)| &< M_b, \end{aligned} \quad \forall \lambda, \quad (530)$$

then the boundedness of M_λ follows from the triangle inequality

$$\begin{aligned}
|M_\lambda(x)| &= |a(\lambda)e^{i\lambda x} + b(\lambda)e^{-i\lambda x}| \\
&\leq |a(\lambda)||e^{i\lambda x}| + |b(\lambda)||e^{-i\lambda x}| \\
&= |a(\lambda)| + |b(\lambda)| \leq M_a + M_b.
\end{aligned} \tag{531}$$

Therefore the spectrum in this case consists of the whole real axis. We can now use these eigenfunctions to represent functions $f(x)$ in the following way

$$\begin{aligned}
f(x) &= \int_{-\infty}^{\infty} d\lambda \{a(\lambda)e^{i\lambda x} + b(\lambda)e^{-i\lambda x}\} \\
&= \int_{-\infty}^{\infty} d\lambda a(\lambda)e^{i\lambda x} + \int_{-\infty}^{\infty} d\lambda b(\lambda)e^{-i\lambda x} \\
&= \int_{-\infty}^{\infty} d\lambda a(-\lambda)e^{-i\lambda x} + \int_{-\infty}^{\infty} d\lambda b(\lambda)e^{-i\lambda x} \\
&= \int_{-\infty}^{\infty} d\lambda c(\lambda)e^{-i\lambda x},
\end{aligned} \tag{532}$$

where $c(\lambda) = a(-\lambda) + b(\lambda)$. Note that in (532), $f(x)$ is represented by $e^{i\lambda x}$ in terms of a standard improper integral. In the theory of integral transforms, other, less standard, improper integrals appear. We will discuss these when they arise.

The formula (532) is by definition the inverse Fourier transform. The key question is how to find the $c(\lambda)$ corresponding to a given function $f(x)$.

For this purpose, let us introduce a small number $\epsilon > 0$ and define

$$\epsilon(x) = \begin{cases} -i\epsilon, & x > 0, \\ i\epsilon, & x < 0. \end{cases} \tag{533}$$

Observe that for all real λ we have

$$\int_{-\infty}^{\infty} dx e^{-i(\lambda + \epsilon(x))x} = \lim_{L \rightarrow \infty} \int_{-L}^L dx e^{-i(\lambda + \epsilon(x))x} \tag{534}$$

$$= \lim_{L \rightarrow \infty} \left\{ \int_{-L}^0 dx e^{-i(\lambda + i\epsilon)x} + \int_0^L dx e^{-i(\lambda - i\epsilon)x} \right\} \tag{535}$$

$$= \lim_{L \rightarrow \infty} \left\{ \frac{1}{-i(\lambda + i\epsilon)} e^{-i(\lambda + i\epsilon)x} \Big|_{-L}^0 + \frac{1}{-i(\lambda - i\epsilon)} e^{-i(\lambda - i\epsilon)x} \Big|_0^L \right\} \tag{536}$$

$$\begin{aligned}
&= \lim_{L \rightarrow \infty} \left\{ \frac{1}{-i(\lambda + i\epsilon)} + \frac{1}{i(\lambda - i\epsilon)} \right. \\
&\quad \left. + \frac{1}{i(\lambda + i\epsilon)} e^{-\epsilon L} e^{i\lambda L} - \frac{1}{i(\lambda - i\epsilon)} e^{-\epsilon L} e^{-i\lambda L} \right\} \\
&= \frac{1}{i(\lambda - i\epsilon)} - \frac{1}{-i(\lambda + i\epsilon)} = \frac{2\epsilon}{\epsilon^2 + \lambda^2}.
\end{aligned} \tag{537}$$

Thus

$$\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} dx e^{-i(\lambda + \epsilon(x))x} = \begin{cases} 0, & \lambda \neq 0 \\ \infty, & \lambda = 0 \end{cases}. \quad (538)$$

This result makes the following formula plausible (I am playing with formulas now)

$$\int_{-\infty}^{\infty} dx e^{-i\lambda x} \propto \delta(\lambda). \quad (539)$$

In the theory of distributions it becomes clear that the constant of proportionality is 2π . So we have the formula

$$\delta(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{-i\lambda x} \quad (540)$$

Since the Dirac delta function is zero for $\lambda \neq 0$ it appears that for all reasonable functions, $f(\lambda)$, we have

$$\int_{-\infty}^{\infty} d\lambda f(\lambda) \delta(\lambda - \hat{\lambda}) = f(\hat{\lambda}). \quad (541)$$

(I am playing now.)

Let us now return to our original question of how to find the $c(\lambda)$ corresponding to a given function $f(x)$ in formula (532)

$$f(x) = \int_{-\infty}^{\infty} d\lambda c(\lambda) e^{-i\lambda x}. \quad (542)$$

Multiplying (542) by $e^{i\hat{\lambda}x}$ and integrate with respect to x over $-\infty < x < \infty$ gives

$$\int_{-\infty}^{\infty} dx f(x) e^{i\hat{\lambda}x} = \int_{-\infty}^{\infty} dx e^{i\hat{\lambda}x} \int_{-\infty}^{\infty} d\lambda c(\lambda) e^{-i\lambda x} \quad (543)$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} d\lambda c(\lambda) \int_{-\infty}^{\infty} dx e^{-i(\lambda - \hat{\lambda})x} \\ &= \int_{-\infty}^{\infty} d\lambda c(\lambda) 2\pi \delta(\lambda - \hat{\lambda}) = 2\pi c(\hat{\lambda}). \end{aligned} \quad (544)$$

So the formula we are looking for is

$$c(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx f(x) e^{i\lambda x}. \quad (545)$$

This formula defines the Fourier transform. Note that there are several conventions about where to place the factor of 2π . A more symmetric choice[1], is

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f(x) e^{i\lambda x}, \quad (546)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx F(x) e^{-i\lambda x}. \quad (547)$$

Note that not only are there different conventions on where to place the factor of 2π but also in which exponent to place the minus sign.

In this class (546) defines the Fourier transform and (547) defines the inverse Fourier transform.

Pure mathematicians have made extensive investigations into for which pair of functions $f(x)$ and $F(\lambda)$ the formulas (546) and (547) are valid. For example if $f(x)$ is piecewise continuously differentiable on each finite interval and if additionally

$$\int_{-\infty}^{\infty} dx |f(x)| < \infty, \quad (548)$$

formula (546) and (547) holds pointwise. This result and related ones are important but we will use (546) and (547) in a much wider context than this.

Let $f(x) = \delta(x)$. Then (546) gives

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \delta(x) e^{i\lambda x} = \frac{1}{\sqrt{2\pi}}, \quad (549)$$

and (547) then gives

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{1}{\sqrt{2\pi}} e^{-i\lambda x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} = \delta(x), \quad (550)$$

where we have used the formula (540) which we argued for previously. More generally, the pair of formulas (546) and (547) holds for any distribution, not just the Delta distribution. More about this next semester. In this semester we will basically use (546) and (547) to take Fourier transform and inverse Fourier transform of anything we want!

The Fourier transform has many powerful properties. You can find them in handbooks or on the web whenever you need them. For example if $F(\lambda)$ and $G(\lambda)$ are Fourier transforms of $f(x)$ and $g(x)$ then $aF(\lambda) + bG(\lambda)$ is the Fourier transform of $af(x) + bg(x)$. A less trivial, and extremely useful for us, property involves products of Fourier transforms $F(\lambda)G(\lambda)$ of functions $f(x)$ and $g(x)$. Such products often occur when we use Fourier transforms to solve ODEs and PDEs.

Playing with formulas we get the following

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F(\lambda) G(\lambda) e^{-i\lambda x} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda G(\lambda) e^{-i\lambda x} \int_{-\infty}^{\infty} dt f(t) e^{i\lambda t} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dt f(t) \int_{-\infty}^{\infty} d\lambda G(\lambda) e^{-i\lambda(x-t)} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt f(t) g(x-t). \tag{551}
\end{aligned}$$

The integral on the right hand side of (551) is known as a *convolution integral*. Canceling the factor $\frac{1}{\sqrt{2\pi}}$ we get the *convolution integral theorem* for Fourier transforms

$$\int_{-\infty}^{\infty} dt f(t) g(x-t) = \int_{-\infty}^{\infty} d\lambda F(\lambda) G(\lambda) e^{-i\lambda x}. \tag{552}$$

Evaluating at $x = 0$ we get

$$\int_{-\infty}^{\infty} dt f(t) g(-t) = \int_{-\infty}^{\infty} d\lambda F(\lambda) G(\lambda). \tag{553}$$

Now observe that

$$\begin{aligned}
F(\lambda) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt f(t) e^{i\lambda t}, \\
&\Downarrow \\
\bar{F}(\lambda) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt \bar{f}(t) e^{-i\lambda t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt \bar{f}(-t) e^{i\lambda t}, \tag{554}
\end{aligned}$$

where $\bar{F}(\lambda)$ is the complex conjugate of $F(\lambda)$.

Thus the fourier transform of $\bar{f}(-t)$ is $\bar{F}(\lambda)$. Putting $g(t) = \bar{f}(-t)$ into (553) we get

$$\int_{-\infty}^{\infty} dt |f(t)|^2 = \int_{-\infty}^{\infty} dt f(t) \bar{f}(t) = \int_{-\infty}^{\infty} dt f(t) g(-t) = \int_{-\infty}^{\infty} d\lambda |F(\lambda)|^2. \tag{555}$$

The identity

$$\int_{-\infty}^{\infty} dt |f(t)|^2 = \int_{-\infty}^{\infty} d\lambda |F(\lambda)|^2. \tag{556}$$

is called the Parseval equation and is clearly a generalization of the Parseval equation for generalized Fourier series to continuous spectrum.

Another important property of Fourier transforms involve derivatives. Observe that

$$\begin{aligned}\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f'(x) e^{i\lambda x} &= \frac{1}{\sqrt{2\pi}} f(x) e^{i\lambda x} \Big|_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f(x) (i\lambda) e^{i\lambda x} \\ &= -i\lambda \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f(x) e^{i\lambda x} = -i\lambda F(\lambda).\end{aligned}\quad (557)$$

Thus, derivatives are converted into multiplication by the Fourier transform. (557) can be easily generalized to

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f^{(n)}(x) e^{i\lambda x} = (-i\lambda)^n F(\lambda). \quad (558)$$

Formulas (557) and (558) are the main sources for the great utility of Fourier transforms in the theory of PDEs and ODEs.

11.1.1 A boundary value problem for an ODE

Let us consider the following boundary value problem

$$y''(x) - k^2 y(x) = -f(x), \quad -\infty < x < \infty, \quad (559)$$

where k is a fixed constant and f a given function. The boundary conditions attached to (559) are

$$y(x), y'(x) \rightarrow 0, \quad \text{when } |x| \rightarrow \infty. \quad (560)$$

Let $F(\lambda)$ and $Y(\lambda)$ be the Fourier transforms of $f(x)$ and $y(x)$. Then the transform of the equation (559) is, upon using (558), given by

$$-(\lambda^2 + k^2)Y(\lambda) = -F(\lambda) \quad (561)$$

This is a simple algebraic equation whose solution is

$$Y(\lambda) = \frac{F(\lambda)}{\lambda^2 + k^2}. \quad (562)$$

Using the inverse Fourier transform and the convolution theorem we get the following representation of $y(x)$:

$$\begin{aligned}y(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{1}{\lambda^2 + k^2} F(\lambda) e^{-i\lambda x} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt g(x-t) f(t),\end{aligned}\quad (563)$$

where

$$g(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{e^{-i\lambda \xi}}{\lambda^2 + k^2}. \quad (564)$$

Using a table of transforms or residue calculus from complex function theory we have

$$g(\xi) = \frac{\sqrt{2\pi}}{2k} e^{-k|\xi|}, \quad (565)$$

so that the solution $y(x)$ is

$$y(x) = \frac{1}{2k} \int_{-\infty}^{\infty} dt e^{-k|x-t|} f(t). \quad (566)$$

Because of the exponentially decaying factor, $y(x)$ and $y'(x)$ will both go to zero as $|x| \rightarrow \infty$ if $f(t)$ decays fast enough at infinity.

Actually, computing $y(x)$ for a given $f(t)$ using formula (566) is no simple matter; tables of integrals, complex function theory, numerical methods and asymptotic methods are all of help here.

For some simple choices for $f(t)$ the integral in (566) can be evaluated easily. Let us first choose $f(t) = 1$. Formula (566) implies

$$\begin{aligned} y(x) &= \frac{1}{2k} \int_{-\infty}^{\infty} dt e^{-k|x-t|} \\ &= \frac{1}{2k} \left[\int_{-\infty}^x dt e^{-k(x-t)} + \int_x^{\infty} dt e^{k(x-t)} \right] \\ &= \frac{1}{2k} \left[e^{-kx} \left(\frac{1}{k} e^{kt} \right) \Big|_{-\infty}^x + e^{kx} \left(-\frac{1}{k} e^{-kt} \right) \Big|_x^{\infty} \right] \\ &= \frac{1}{2k} \left[\frac{1}{k} + \frac{1}{k} \right] = \frac{1}{k^2}. \end{aligned} \quad (567)$$

Observe that

$$y''(x) - k^2 y(x) = -k^2 \left(\frac{1}{k^2} \right) = -1 = -f(x), \quad (568)$$

so $y(x)$ defined in (567) is a solution to the equation (559), but it does clearly not satisfy the boundary conditions (560). The reason why this occurs is that in deriving the formula (566) we assumed that $f(x)$ had a Fourier transform $F(\lambda)$. However $f(x) = 1$ leads to the divergent integral

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{i\lambda x} \quad (569)$$

for $F(\lambda)$. Using (550) we conclude that $F(\lambda)$ is not a function but a distribution

$$F(\lambda) = \sqrt{2\pi} \delta(\lambda). \quad (570)$$

Calculations like the one leading up to the "solution" (567) are very common whenever transforms are used to solve PDEs. They are *ideal* calculations that is meant to represent the *limit* of a sequence of problems, each producing an

honest to God solution to the boundary value problem (559),(560). Such ideal calculations are often possible to do in closed form and does often lead to some insight into the general case, but in what sense the result of the calculation is a limit of a sequence of actual solution is for the most part left unsaid. This will be clarified when we discuss distributions next semester. However, let us preview next semester by saying exactly what we mean by a limit here.

Define a sequence of functions $\{f_N(x)\}$

$$f_N(x) = \begin{cases} 1, & |x| \leq N, \\ 0, & |x| > N. \end{cases} \quad (571)$$

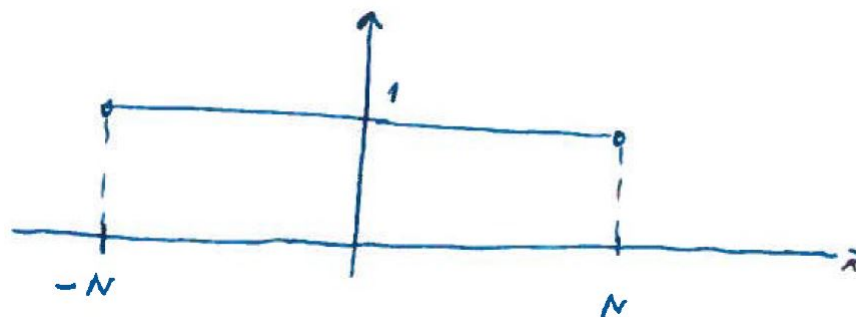


Figure 56

This function has a Fourier transform for each N and our formula (566) should give a solution to the boundary value problem (559),(560)

$$\begin{aligned} y_N(x) &= \frac{1}{2k} \int_{-\infty}^{\infty} dt e^{-k|x-t|} f_N(t) \\ &= \frac{1}{2k} \int_{-N}^N dt e^{-k|x-t|}. \end{aligned} \quad (572)$$

Let us first assume that $|x| \leq N$

$$\begin{aligned}
\Rightarrow y_N(x) &= \frac{1}{2k} \left[\int_{-N}^x dt e^{-k(x-t)} + \int_x^N dt e^{k(x-t)} \right] \quad (573) \\
&= \frac{1}{2k} \left[e^{-kx} \left(\frac{1}{k} e^{kt} \right) \Big|_{-N}^x + e^{kx} \left(-\frac{1}{k} e^{-kt} \right) \Big|_x^N \right] \\
&= \frac{1}{2k^2} [e^{-kx} (e^{kx} - e^{-kN}) - e^{kx} (e^{-kN} - e^{-kx})] \\
&= \frac{1}{2k^2} [1 - e^{-kN} e^{-kx} - e^{-kN} e^{kx} + 1] \\
&= \frac{1}{k^2} [1 - e^{-kN} \cosh kx] . \quad (574)
\end{aligned}$$

The calculations for $|x| > N$ is completed in a similar way. This gives us

$$y_N(x) = \begin{cases} \frac{1}{k^2} (1 - e^{-kN} \cosh kx), & |x| \leq N, \\ \frac{1}{k^2} (e^{-k|x|} \sinh kN), & |x| \geq N. \end{cases} \quad (575)$$

$y_N(x)$ is a solution to (559) and it is easy to verify that $y_N(x), y'_N(x) \rightarrow 0$ when $|x| \rightarrow \infty$, so we have an infinite sequence of solutions $\{y_N(x)\}$ to the boundary value problem defined by (559), (560).

For any x there evidently exists a N so large that $N \geq |x|$. Therefore for such an x we have $f_N(x) = f(x)$ for all $N > |x|$ and consequently the sequence of functions $\{f_N\}_N^\infty$ converge pointwise to the functions $f(x) = 1$. This convergence is however *not* uniform.

Observe that for any x there obviously exists infinitely many N such that $N \geq |x|$. For such N we have from (575) that

$$y_N(x) = \frac{1}{k^2} (1 - e^{-kN} \cosh kx) \rightarrow \frac{1}{k^2} = y(x), \quad (576)$$

when $N \rightarrow \infty$

Thus, this convergence, like the one for the source $f(x)$, is also pointwise but not uniform. This non-uniformity of the limit is why the limiting function, $y(x)$, does not satisfy the boundary condition (560) even though all $y_N(x)$ does.

As a second example of an ideal calculation let us choose $f(t) = \delta(t)$. Then formula (566) gives

$$y(x) = \frac{1}{2k} \int_{-\infty}^{\infty} dt e^{-k|x-t|} \delta(t) = \frac{1}{2k} e^{-k|x|}. \quad (577)$$

This function satisfies the boundary condition and the equation for all $x \neq 0$. It is however *not* differentiable at $x = 0$.

Introduce a sequence of functions $\{g_N(x)\}$ by

$$g_N(x) = \begin{cases} \frac{N}{2}, & |x| \leq \frac{1}{N}, \\ 0, & |x| > \frac{1}{N}. \end{cases} \quad (578)$$

Using formula (566) with $f(x) = g_N(x)$, it is straight forward to show that the corresponding solution is

$$y_N(x) = \begin{cases} \frac{N}{2k^2} \left(1 - e^{-\frac{k}{N}} \cosh(kx)\right), & |x| \leq \frac{1}{N}, \\ \frac{N}{2k^2} e^{-|x|} \sinh\left(\frac{k}{N}\right), & |x| > \frac{1}{N}. \end{cases} \quad (579)$$

For any x there exist infinitely many N such that $|x| \geq \frac{1}{N}$, and for such N we have from (579) that

$$y_N(x) = \frac{N}{2k^2} e^{-|x|} \sinh\left(\frac{k}{N}\right) \rightarrow \frac{1}{2k} e^{-|x|},$$

when N approach infinity. Furthermore, for $x = 0$ we have from (579) that

$$y_N(0) = \frac{N}{2k^2} \left(1 - e^{-\frac{k}{N}}\right) \rightarrow \frac{1}{2k},$$

when N approach infinity.

We have thus established that the infinite sequence of solutions (579) of equation (559) corresponding to the sequence $\{g_N\}$ converge pointwise to the solution $y(x) = \frac{1}{2k} e^{-k|x|}$ corresponding to the Dirac delta function $f(x) = \delta(x)$. The sequence (579) does not converge uniformly to the function $y(x)$ and this explains the fact that all functions from the sequence (579) are differentiable at $x = 0$ whereas the limiting function $y(x) = \frac{1}{2k} e^{-k|x|}$ is not.

In the previous example the sequence $\{f_N(x)\}$ converged pointwise to the limiting function $f(x) = 1$ that generated the solution $y(x) = \frac{1}{k^2}$. Is it similarly true that in the current example the sequence $\{g_N(x)\}$ converge to the Dirac delta function that generated the solution $f(x) = \frac{1}{2k} e^{-k|x|}$?

In order to answer this question we start by observing that for all N

$$\int_{-\infty}^{\infty} dx g_N(x) = \int_{-\frac{1}{N}}^{\frac{1}{N}} dx \frac{N}{2} = \frac{N}{2} \left(\frac{1}{N} + \frac{1}{N}\right) = 1, \quad (580)$$

so $g_N(x)$ satisfies the same integral identity (527) as the Dirac delta function (Here $D = (-\infty, \infty)$). Furthermore for any $x \neq 0$ there exists a N_x such that

$$|x| > \frac{1}{N_x}. \quad (581)$$

Therefore for all $N \geq N_x$, we have $g_N(x) = 0$. Therefore

$$\lim_{N \rightarrow \infty} g_N(x) = 0, \quad (582)$$

so the limit of the sequence g_N when $N \rightarrow \infty$ also satisfies the identity (526). So in fact

$$\lim_{N \rightarrow \infty} g_N(x) = \delta(x). \quad (583)$$

I am playing with formulas here; the limit in (583) does not mean the usual thing. After all, $\delta(x)$ is not a function but something different, namely a distribution. It would perhaps be more fair to say that the limit of the sequence $\{g_N(x)\}$, when $N \rightarrow \infty$ does *not* exist, since there can be *no* function satisfying (526) and (527).

What you see here is an example of a very subtle and profound mathematical idea called *completion*. You have actually used this idea constantly since grade school, but you may not have been told.

When we learn children arithmetic we start with whole numbers and eventually move on to fractions or *rational numbers* as they are called. The universe of numbers are *rational numbers* at grade school level. At some point the children are confronted with the equation

$$x^2 = 2. \quad (584)$$

If the teacher is any good he will by trial and error convince the children that (584) has no solutions. So what is there to do? He will say that (584) actually has a solution that we denote by $\sqrt{2}$, and he will show, perhaps using a calculator, how to find a sequence of rational numbers that more and more closely approximate $\sqrt{2}$. In a similar way the teacher introduces $\sqrt{3}$, $\sqrt{5}$ etc. and he develops rules for calculations with such things.

But the children has been tricked. When the teacher argues that calculated sequences of rational numbers approximate $\sqrt{2}$ better and better and in the limit equals $\sqrt{2}$, he has been pulling wool over their eyes. What he has done is to *complete* the set of rational numbers, in the process *creating* the *real numbers*. There is *no* sequence of rational numbers converging to $\sqrt{2}$, in reality $\sqrt{2}$ is defined to *be* a certain set of sequences of rational numbers. The existence of $\sqrt{2}$ is *postulated*, pure and simple, through the completion process. After the completion process has lead to the set of real numbers, pure and applied mathematics can give them additional intuitive meaning by mapping them onto other physical and mathematical systems. In this way real numbers have come to be interpreted as geometrical points on a line etc.

The relation between functions and distributions is exactly the same as the one between rational and real numbers. For the case of distributions Dirac is the teacher and the rest of us the children. Next semester we will map distributions onto a system derived from linear algebra and in this way gain more insight into the nature of distributions.

The particular "solution" (577) found by choosing $f(x) = \delta(x)$ is an example of what is called a *Green's function*. Writing

$$g(x - y) = \frac{1}{2k} e^{-k|x-y|}. \quad (585)$$

We have seen that $g(x - y)$ is a solution to the equation

$$g'' - k^2 g = -\delta(x - y), \quad (586)$$

where we are taking the derivative with respect to the variable x , the variable y is a parameter here. Green's functions are of *outmost* importance in applied mathematics and theoretical physics and will be discussed extensively next semester. As you can see from (586), Green's functions are solutions to differential equations formulated in the universe of distributions. For this reason Green's functions are not actually functions but distributions.

Let us now proceed to consider a slightly different ODE

$$y''(x) + k^2 y(x) = -f(x), \quad -\infty < x < \infty. \quad (587)$$

For this equation the boundary conditions can not be that $y(x)$ and $y'(x) \rightarrow 0$ as $|x| \rightarrow \infty$. The reason is that (587) does not have solutions decaying at infinity. In fact for $f(x) = 0$ the solution consists of $\sin kx$ and $\cos kx$, which are bounded but does not vanish. We will therefore only require that y and y' are bounded at infinity. This implies, by an argument identical to the one on page 174, that the spectrum is real. We apply the Fourier transform like for the previous example and find

$$Y(\lambda) = \frac{F(\lambda)}{\lambda^2 - k^2}. \quad (588)$$

We observe that, compared to the previous case, a problem has appeared: $Y(\lambda)$ is not defined at $\lambda = \pm k$. The proper way to handle this is to use complex function theory and for more complicated situations complex function theory becomes a necessity. Here we will proceed in a more heuristic way by observing that (587) becomes (559) if we let $k \rightarrow ik$ or $k \rightarrow -ik$. Making the same replacement in the solution formula (566) we get two solutions

$$y_{\pm}(x) = \pm \frac{i}{2k} \int_{-\infty}^{\infty} dt e^{\pm ik|x-t|} f(t). \quad (589)$$

Both of $y_{\pm}(x)$ satisfy (587) and are bounded if $f(t)$ is absolutely integrable $\left(\int_{-\infty}^{\infty} dt |f(t)| < \infty\right)$. Thus the boundary value problem (587) does not have a unique solution. Why is that?

Let us start with the wave equation

$$-\frac{1}{c^2} \varphi_{tt} + \varphi_{xx} = -f(x), \quad -\infty < x < \infty, \quad (590)$$

and let the source, $f(x)$, be supported in some bounded domain G . Outside this domain G , (590) is the homogeneous wave equation whose general solution consists of left- and right going waves.

We assume that the source is the originator of all waves. Thus, no source, no waves. Then it is evident that there can be no left moving waves to the right of G and no right-moving waves to the left of G . This is a physical condition

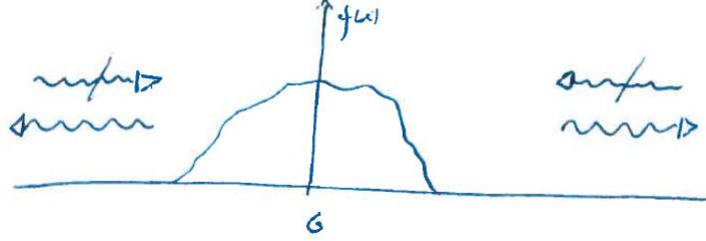


Figure 57

satisfied by all known wave phenomena; waves move *away* from the source that creates them. This is in fact part of the definition of what a physical source is!

Let us separate variables in (590)

$$\varphi(x, t) = e^{-i\omega t} y(x) \quad (591)$$

$$\Rightarrow y''(x) + \left(\frac{\omega}{c}\right)^2 y(x) = -f(x). \quad (592)$$

Defining $k = \frac{\omega}{c}$ we find our equation (587). The homogeneous equation (592) has a solution proportional to $e^{\pm ikx}$. Taking these solutions together with (591) we observe that

$$\begin{aligned} e^{ikx} &\rightarrow e^{i(kx - \omega t)}, & \text{right travelling wave,} \\ e^{-ikx} &\rightarrow e^{-i(kx + \omega t)}, & \text{left travelling wave.} \end{aligned} \quad (593)$$

The physical boundary condition thus consists of killing terms containing e^{-ikx} to the right of the support of f and terms containing e^{ikx} to the left of the support of f . This can be expressed by saying that

$$\begin{aligned} \partial_x y(x) &= ik y(x) & x > \text{Supp}(f) \\ \partial_x y(x) &= -ik y(x) & x < \text{Supp}(f). \end{aligned} \quad (594)$$

where $\text{Supp}(f) = \{x; f(x) \neq 0\}$. Note that $\text{Supp}(f)$ by definition includes its boundary points and is thus a closed set.

We have here assumed that the support of f is bounded. However it is enough that $f(x) \rightarrow 0$, $|x| \rightarrow \infty$. For this more general case we pose the conditions

$$\begin{aligned} \lim_{x \rightarrow \infty} (\partial_x y(x) - ik y(x)) &= 0, \\ \lim_{x \rightarrow -\infty} (\partial_x y(x) + ik y(x)) &= 0. \end{aligned} \quad (595)$$

(594) or more generally (595) is called *the radiation condition at infinity*, or *outgoing waves at infinity* and is the correct boundary condition for equations describing wavelike phenomena. (595) picks out the unique solution $y_+(x)$ to our boundary value problem (587),(595). Choosing $f(x) = \delta(x)$ gives us the following Green's function for this case

$$G(x - \xi) = \frac{i}{2k} e^{ik|x-\xi|}. \quad (596)$$

11.1.2 The Cauchy problem for the heat equation

Let us next apply the Fourier transform to the Cauchy problem for the heat equation

$$\begin{aligned} u_t(x, t) - c^2 u_{xx}(x, t) &= 0, & -\infty < x < \infty, t > 0, \\ u(x, 0) &= f(x), & -\infty < x < \infty. \end{aligned} \quad (597)$$

In order to apply the Fourier transform we must assume that f has a Fourier transform and that u, u_x goes to zero at $x = \pm\infty$.

Taking the Fourier transform with respect to x of the heat equation (597) we get

$$\begin{aligned} \partial_t U(\lambda, t) + c^2 \lambda^2 U(\lambda, t) &= 0, \\ U(\lambda, 0) &= F(\lambda), \end{aligned} \quad (598)$$

where

$$U(\lambda, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{i\lambda x} u(x, t), \quad (599)$$

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{i\lambda x} f(x). \quad (600)$$

The equations (598) are uncoupled for different values of λ . It is the linearity of the heat equation that produce an *uncoupled* system of ODEs for the set of functions $U(\lambda, t)$ indexed by the spectral parameter λ . Note how similar this is to what we got when we applied the finite Fourier transform to linear PDEs on a finite domain.

The initial value problems (598) are easy to solve. We get

$$U(\lambda, t) = F(\lambda) e^{-c^2 \lambda^2 t}. \quad (601)$$

Using the inverse Fourier transform we get the solution

$$\begin{aligned}
u(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} U(\lambda, t) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} F(\lambda) e^{-c^2 \lambda^2 t} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx' e^{-i\lambda x} e^{i\lambda x'} f(x') e^{-c^2 \lambda^2 t} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx' \left[\int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x-x')-c^2 \lambda^2 t} \right] f(x'),
\end{aligned} \tag{602}$$

where we have changed order of integration. Defining

$$g(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda \alpha - c^2 \lambda^2 t}, \tag{603}$$

the solution can be written in the form

$$u(x, t) = \int_{-\infty}^{\infty} dx' g(x - x') f(x'). \tag{604}$$

Note that we could have gotten this expression more directly by using the convolution theorem. Observe that

$$\begin{aligned}
g(\alpha) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda \alpha - c^2 \lambda^2 t} \\
&= \frac{1}{2\pi} \int_{-\infty}^0 d\lambda e^{-i\lambda \alpha - c^2 \lambda^2 t} + \frac{1}{2\pi} \int_0^{\infty} d\lambda e^{-i\lambda \alpha - c^2 \lambda^2 t} \\
&= \frac{1}{2\pi} \int_0^{\infty} d\lambda e^{i\lambda \alpha} e^{-c^2 \lambda^2 t} + \frac{1}{2\pi} \int_0^{\infty} d\lambda e^{-i\lambda \alpha} e^{-c^2 \lambda^2 t} \\
&= \frac{1}{\pi} \int_0^{\infty} d\lambda \cos \lambda \alpha e^{-c^2 \lambda^2 t}.
\end{aligned} \tag{605}$$

This integral, because of the exponential factor present for $t > 0$, converges very fast, so we can differentiate under the integral sign

$$\begin{aligned}
\frac{dg}{d\alpha}(\alpha) &= -\frac{1}{\pi} \int_0^{\infty} d\lambda \lambda \sin \lambda \alpha e^{-c^2 \lambda^2 t} \\
&= \frac{1}{2\pi c^2 t} \sin \lambda \alpha e^{-c^2 \lambda^2 t} \Big|_{\lambda=0}^{\lambda=\infty} - \frac{\alpha}{2\pi c^2 t} \int_0^{\infty} d\lambda \cos \lambda \alpha e^{-c^2 \lambda^2 t},
\end{aligned} \tag{606}$$

so

$$\frac{dg}{d\alpha} = -\frac{\alpha}{2c^2 t} g. \tag{607}$$

This ODE is simple to solve. The solution is

$$g(\alpha) = g(0)e^{-\frac{\alpha^2}{4c^2t}}, \quad (608)$$

and

$$g(0) = \frac{1}{\pi} \int_0^\infty d\lambda e^{-c^2\lambda^2t} = \frac{1}{\sqrt{4\pi c^2t}}. \quad (609)$$

(This is a Gaussian integral.) Thus

$$g(\alpha) = \frac{1}{\sqrt{4\pi c^2t}} e^{-\frac{\alpha^2}{4c^2t}}. \quad (610)$$

The formula for $u(x, t)$ thus becomes

$$u(x, t) = \frac{1}{\sqrt{4\pi c^2t}} \int_{-\infty}^\infty dx' e^{-\frac{(x-x')^2}{4c^2t}} f(x'). \quad (611)$$

If we formally use the initial condition

$$f(x) = \delta(x - \xi), \quad -\infty < x < \infty, \quad (612)$$

we get the solution

$$u(x, t; \xi) \equiv G(x - \xi, t) = \frac{1}{\sqrt{4\pi c^2t}} e^{-\frac{(x-\xi)^2}{4c^2t}}. \quad (613)$$

G is yet another example of a Green's function. In this context it is called the *fundamental solution* to the heat equation. The fundamental solution is the temperature profile generated by a point-like heat source located at $x = \xi$ for $t = 0$.

Note that it is an actual classical solution (verify!) for all $t > 0$ even if the initial condition is extremely singular. This is yet another example illustrating the extreme smoothening properties of the heat equation.

From the solution formula (611) it is evident that even if the initial heat profile is zero beyond a finite interval $[-L, L]$, the solution is nonzero for *any* point $x \in \mathbb{R}$ for any $t > 0$, however small t is.

The heat equation clearly does not respect causality.



Figure 58

Also note that for the fundamental solution we have

$$\int_{-\infty}^{\infty} dx' G(x-x', t)g(x') \rightarrow g(x), \quad t \rightarrow 0^+. \quad (614)$$

Thus

$$G(x-x', t) \rightarrow \delta(x-x'), \quad t \rightarrow 0^+. \quad (615)$$

The fundamental solution is thus a sequence of smooth functions converging to the Dirac delta distribution. Which is exactly what it needs to do in order to satisfy the initial condition!

We can argue for the correctness of (614) as follows. Observe that as $t \rightarrow 0^+$ the function $G(x-\xi, t)$ becomes more and more localized around $x = \xi$. Thus for t close enough to zero and for all reasonable functions $g(x)$ we have

$$\int_{-\infty}^{\infty} dx' G(x-x', t)g(x') \approx g(x) \int_{-\infty}^{\infty} dx' G(x-x', t) = g(x), \quad (616)$$

because

$$\int_{-\infty}^{\infty} dx' G(x-x', t) = 1. \quad (617)$$

Verify this!

11.1.3 The Cauchy problem for the wave equation

Let us consider the problem

$$\begin{aligned} u_{tt}(x, t) - c^2 u_{xx}(x, t) &= 0, \quad -\infty < x < \infty, \quad t > 0, \\ u(x, 0) &= f(x), \\ u_t(x, 0) &= g(x). \end{aligned} \quad (618)$$

Taking the Fourier transform of this problem, proceeding like for the heat equation, we get

$$\begin{aligned} \partial_{tt}U(\lambda, t) - c^2 \lambda^2 U(\lambda, t) &= 0, \quad t > 0, \\ U(\lambda, 0) &= F(\lambda), \\ \partial_t U(\lambda, 0) &= G(\lambda). \end{aligned} \quad (619)$$

The initial value problem (619) is entirely standard and the solution is

$$\begin{aligned} U(\lambda, t) &= \left(\frac{1}{2}F(\lambda) + \frac{1}{2i\lambda c}G(\lambda) \right) e^{i\lambda ct} \\ &+ \left(\frac{1}{2}F(\lambda) - \frac{1}{2i\lambda c}G(\lambda) \right) e^{-i\lambda ct}. \end{aligned} \quad (620)$$

Taking the inverse transform we get

$$u(x, t) = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x-ct)} F(\lambda) + \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x+ct)} F(\lambda) \quad (621)$$

$$+ \frac{1}{2c\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x-ct)} \frac{G(\lambda)}{i\lambda} - \frac{1}{2c\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x+ct)} \frac{G(\lambda)}{i\lambda}.$$

By the definition of the inverse Fourier transform we have

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} G(\lambda). \quad (622)$$

But then

$$\int_0^x ds g(s) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} \frac{G(\lambda)}{i\lambda} + D, \quad (623)$$

where D is a constant. Given (623), (621) can be written as

$$u(x, t) = \frac{1}{2} (f(x - ct) + f(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} ds g(s). \quad (624)$$

This is the d'Alembert formula, now derived using the Fourier transform.

There is however something rather fishy with what we did. Let us say that we did not realize that the identity (623) holds. Then, for given $F(\lambda)$, $G(\lambda)$, we would have to find $u(x, t)$ by solving the integrals in (621) analytically or numerically. We should however worry about the last two integrals. They are of the general form

$$I(\alpha) = \int_{-\infty}^{\infty} d\lambda e^{-i\lambda\alpha} \frac{H(\lambda)}{i\lambda}, \quad \alpha \in \mathbb{R}, \quad (625)$$

where in general $H(0) \neq 0$. The integral is singular at $\lambda = 0$ and the question how the integral should be understood must be raised.

Away from $\lambda = 0$ there are no problems, so the question we need to ask is if we can make sense of integrals of the type

$$\int_{-L}^L d\lambda \frac{1}{\lambda}, \quad (626)$$

where L is arbitrary but fixed. In calculus we have seen that at least some integrals with a singular integrand in the domain of integration can be made sense of by considering it to be an improper integral. Let us try to use this approach on (626):

$$\begin{aligned}
\int_{-L}^L d\lambda \frac{1}{\lambda} &= \int_{-L}^0 d\lambda \frac{1}{\lambda} + \int_0^L d\lambda \frac{1}{\lambda} \\
&= \lim_{\epsilon \rightarrow 0} \int_{-L}^{-\epsilon} d\lambda \frac{1}{\lambda} + \lim_{\delta \rightarrow 0} \int_{\delta}^L d\lambda \frac{1}{\lambda} \\
&= \lim_{\epsilon \rightarrow 0} \int_L^{\epsilon} d\mu \frac{1}{\mu} + \lim_{\delta \rightarrow 0} \int_{\delta}^L d\lambda \frac{1}{\lambda} \\
&= -\lim_{\epsilon \rightarrow 0} \int_{\epsilon}^L d\mu \frac{1}{\mu} + \lim_{\delta \rightarrow 0} \int_{\delta}^L d\lambda \frac{1}{\lambda} \\
&= -\lim_{\epsilon \rightarrow 0} (\ln L - \ln \epsilon) + \lim_{\delta \rightarrow 0} (\ln L - \ln \delta) \\
&= \lim_{\epsilon \rightarrow 0} \ln \epsilon - \lim_{\delta \rightarrow 0} \ln \delta.
\end{aligned} \tag{627}$$

This limit does not exist. Thus (626) can *not* be given meaning as an improper integral. Observe however that if we in (627) let $\epsilon = \delta$ and take the difference *before* we take the limit we get zero, which is a finite answer. Thus what we are proposing is to understand (626) in the following way

$$\begin{aligned}
\int_{-L}^L d\lambda \frac{1}{\lambda} &= \lim_{\epsilon \rightarrow 0} \left\{ \int_{-L}^{-\epsilon} d\lambda \frac{1}{\lambda} + \int_{\epsilon}^L d\lambda \frac{1}{\lambda} \right\} \\
&= \lim_{\epsilon \rightarrow 0} \left\{ \int_L^{\epsilon} d\mu \frac{1}{\mu} + \int_{\epsilon}^L d\lambda \frac{1}{\lambda} \right\} \\
&= \lim_{\epsilon \rightarrow 0} \{ \ln \epsilon - \ln L + \ln L - \ln \epsilon \} = 0.
\end{aligned} \tag{628}$$

When we use the definition (628) we understand the integral as a *Cauchy principal value* integral. This is now a new way for you to calculate integrals with a singular integrand. In general for a function $f(x)$, with a singularity at $x = x_0$, the definition is

$$PV \int_{-L}^L dx f(x) = \lim_{\epsilon \rightarrow 0} \left\{ \int_{-L}^{x_0 - \epsilon} dx f(x) + \int_{x_0 + \epsilon}^L dx f(x) \right\}, \tag{629}$$

where *PV* signifies that we are calculating the principal value.

The principal value is finite only if the singularity in f is of the right type. For example for $f(x) = \frac{1}{x^2}$ we get

$$\begin{aligned}
PV \int_{-1}^1 dx \frac{1}{x^2} &= \lim_{\epsilon \rightarrow 0} \left\{ \int_{-1}^{-\epsilon} dx \frac{1}{x^2} + \int_{\epsilon}^1 dx \frac{1}{x^2} \right\} \\
&= \lim_{\epsilon \rightarrow 0} \left\{ \left(-\frac{1}{x} \right) \Big|_{-1}^{-\epsilon} + \left(-\frac{1}{x} \right) \Big|_{\epsilon}^1 \right\} \\
&= \lim_{\epsilon \rightarrow 0} \left\{ \frac{1}{\epsilon} + 1 - 1 + \frac{1}{\epsilon} \right\} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} = \infty!
\end{aligned} \tag{630}$$

If we interpret the two integrals containing $\frac{1}{\lambda}$ in (615) as Cauchy principal value integrals, we can use the formula to calculate the solution to the wave equation.

However, there is still a weak fishy smell here; why is interpreting the singular integrals in (615) as Cauchy principal value integrals the right thing to do? How do we know that this will give the same solution as, say, the d'Alembert formula? Could one not imagine some other funny way of getting a finite number out of integrals like (626)? And would this not perhaps produce a different solution; which should not be possible because the Cauchy problem for the wave equation has a unique solution? What is going on?

Answering this question, which is an important one, would take us too far from the main focus of the course, so I will refrain from giving an answer. Here I will only ensure you that this *is* the right way to understand the integral (626) and it *does* give the same solution as the d'Alembert formula. It is also worth noting that questions like the one we have been considering frequently turn up when we use transforms to solve PDEs, and solving them can often be a deep and vexing problem.

11.1.4 The Laplace equation in a half-plane

Let us consider the Laplace equation in the upper halfplane:

$$\begin{aligned} u_{xx}(x, y) + u_{yy}(x, y) &= 0, & -\infty < x < \infty, y > 0, \\ u(x, 0) &= f(x), \\ u(x, y) &\text{ is bounded as } y \rightarrow \infty. \end{aligned} \quad (631)$$

We introduce the Fourier transform in x

$$U(\lambda, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{i\lambda x} u(x, y). \quad (632)$$

Using this Fourier transform on (631) gives us the following infinite system of uncoupled ODEs:

$$\begin{aligned} \partial_{yy} U(\lambda, y) - \lambda^2 U(\lambda, y) &= 0, & 0 < y < \infty, \\ U(\lambda, 0) &= F(\lambda), \\ U(\lambda, y) &\text{ is bounded as } y \rightarrow \infty. \end{aligned} \quad (633)$$

This problem is standard and the solution is

$$U(\lambda, y) = F(\lambda) e^{-|\lambda|y}. \quad (634)$$

The inverse transform gives us the solution

$$\begin{aligned}
u(x, y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} F(\lambda) e^{-|\lambda|y} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{-i\lambda x} e^{-|\lambda|y} \int_{-\infty}^{\infty} dx' e^{i\lambda x'} f(x') \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx' f(x') \left[\int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x-x')-|\lambda|y} \right].
\end{aligned} \tag{635}$$

Furthermore, observe that

$$\begin{aligned}
&\int_{-\infty}^{\infty} d\lambda e^{-i\lambda(x-x')-|\lambda|y} \\
&= 2 \int_0^{\infty} d\lambda \cos \lambda(x-x') e^{-\lambda y} \\
&= \frac{2y}{(x-x')^2 + y^2},
\end{aligned} \tag{636}$$

and therefore we have the following expression for the solution of the Laplace equation in a half-space

$$u(x, y) = \frac{y}{\pi} \int_{-\infty}^{\infty} dx' \frac{f(x')}{(x-x')^2 + y^2}. \tag{637}$$

By direct substitution it can be verified that (637) is a solution to the Laplace equation for $y > 0$.

When $y \rightarrow 0$ the factor $\frac{y}{\pi}$ goes to zero, but the integrand approaches

$$\frac{f(x')}{(x-x')^2}, \tag{638}$$

which is singular at $x' = x$. A careful study of the limit shows that the limit of (637) when $y \rightarrow 0$ is actually $f(x)$. We will not do this here.

For simple choices of $f(x)$ we can solve the integral (637). For example if $f(x) = H(x)$ where $H(x)$ is the Heaviside step function

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases} \tag{639}$$

we find

$$\begin{aligned}
u(x, y) &= \frac{y}{\pi} \int_{-\infty}^{\infty} dx' \frac{H(x')}{(x-x')^2 + y^2} \\
&= \frac{y}{\pi} \int_0^{\infty} dx' \frac{1}{(x-x')^2 + y^2} \\
&= \frac{1}{\pi y} \int_0^{\infty} dx' \frac{1}{1 + \left(\frac{x-x'}{y}\right)^2}.
\end{aligned} \tag{640}$$

Substitute $u = \frac{x' - x}{y}$ and $du = \frac{1}{y}dx'$ in this integral. This gives us the following explicit expression for $u(x, y)$

$$\begin{aligned} u(x, y) &= \frac{1}{\pi} \int_{-\frac{x}{y}}^{\infty} du \frac{1}{1+u^2} = \frac{1}{\pi} tg^{-1}(u) \Big|_{-\frac{x}{y}}^{\infty} \\ &= \frac{1}{2} + \frac{1}{\pi} tg^{-1} \left(\frac{x}{y} \right). \end{aligned} \quad (641)$$

This function satisfies the Laplace equation and

$$\lim_{y \rightarrow 0} u(x, y) = \frac{1}{2} + \frac{1}{\pi} \lim_{y \rightarrow 0} tg^{-1} \left(\frac{x}{y} \right) = \begin{cases} \frac{1}{2} + \frac{1}{\pi} \cdot \frac{\pi}{2} = 1, & x > 0, \\ \frac{1}{2} + \frac{1}{\pi} \cdot \left(-\frac{\pi}{2} \right) = 0, & x < 0, \end{cases} \quad (642)$$

so $u(x, y)$ satisfies the boundary condition at $y = 0$. Also

$$\lim_{y \rightarrow \infty} u(x, y) = \frac{1}{2}, \quad (643)$$

so $u(x, y)$ is bounded at $y = +\infty$.

However, $u(x, y)$ does not go to zero as $|x| \rightarrow \infty$ and so does not actually have a Fourier transform, which we assumed. The reason for this is that our boundary function $f(x) = H(x)$ does not have a Fourier transform in the conventional sense. Observe that $u(x, y)$ is infinitely smooth for $y > 0$ even if it is discontinuous in x for $y = 0$. We are rediscovering the fact that the Laplace equation can not support solutions with singularities. As we have seen, this is a general feature of elliptic equations.

Let us say that we were rather to solve the Laplace equation with Neumann conditions

$$\begin{aligned} u_{xx}(x, y) + u_{yy}(x, y) &= 0, \quad -\infty < x < \infty, y > 0, \\ u_y(x, 0) &= g(x). \end{aligned} \quad (644)$$

We can reduce this to a problem of the previous type using a trick, invented by an English mathematician, called *Stoke's rule*.

Introduce $v(x, y) = u_y(x, y)$. Differentiating (644) with respect to y gives

$$\begin{aligned} v_{xx} + v_{yy} &= 0, \quad -\infty < x < \infty, y > 0, \\ v(x, 0) &= g(x), \quad -\infty < x < \infty. \end{aligned} \quad (645)$$

The solution to this problem is given by (637) and since

$$u(x, y) = \int_0^y ds v(x, s), \quad (646)$$

we get

$$\begin{aligned}
u(x, y) &= \frac{1}{\pi} \int_0^y ds s \int_{-\infty}^{\infty} dx' \frac{f(x')}{(x - x')^2 + s^2} \\
&= \frac{1}{\pi} \int_{-\infty}^{\infty} dx' f(x') \int_0^y ds \frac{s}{(x - x')^2 + s^2} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx' \ln [(x - x')^2 + y^2] f(x'). \tag{647}
\end{aligned}$$

Observe that $u(x, y)$ defined by (647) is not the only solution to (644).

For any solution $v(x, y)$ to (645) we can add an arbitrary function $C(x)$ to the right hand side of (646) and in this way define a whole infinite family of functions $u(x, y)$

$$u(x, y) = \int_0^y ds v(x, s) + C(x). \tag{648}$$

All such functions satisfy the boundary condition $u_y(x, 0) = g(x)$, but in order to satisfy the Laplace equation (644) the added function $C(x)$ must be linear. Furthermore, in order for the solution to be bounded in x this linear function must be a constant.

11.2 The Fourier sine- and cosine transform

Let us consider the following eigenvalue problem

$$M''(x) + \lambda^2 M(x) = 0, \quad 0 < x < \infty. \tag{649}$$

on a semi-infinite 1D domain. We will consider two sets of boundary conditions

$$M(0) = 0, \quad M(x) \text{ bounded as } x \rightarrow +\infty, \tag{650}$$

$$M'(0) = 0, \quad M(x) \text{ bounded as } x \rightarrow +\infty. \tag{651}$$

Let us first consider (649) with boundary conditions (650).

We will only consider (649) for real λ . For $\lambda = 0$ the general solution of (649) is

$$M(x) = Ax + B, \tag{652}$$

and

$$M(0) = 0 \Rightarrow B = 0 \Rightarrow M(x) = Ax,$$

$$M(x) \text{ bounded} \Rightarrow A = 0 \Rightarrow M(x) = 0.$$

Thus $\lambda = 0$ is not an eigenvalue. For $\lambda \neq 0$ the general solution to (649) is

$$M(x) = A \cos \lambda x + B \sin \lambda x, \tag{653}$$

and

$$M(0) = 0 \Rightarrow A = 0 \Rightarrow M(x) = B \sin \lambda x.$$

This function is clearly bounded for all real λ and therefore all such λ are in the spectrum. Observe however that

$$M_{-\lambda}(x) = -M_{\lambda}(x). \quad (654)$$

Therefore, the eigenfunctions corresponding to λ and $-\lambda$ are linearly dependent. We therefore lose nothing by restricting to the range

$$0 < \lambda < \infty, \quad (655)$$

and eigenfunctions

$$M_{\lambda}(x) = \sin \lambda x. \quad (656)$$

For (649) with the boundary conditions (651) we find in a similar way the spectrum

$$0 \leq \lambda < \infty, \quad (657)$$

and eigenfunctions

$$M_{\lambda}(x) = \cos \lambda x. \quad (658)$$

In order to find the transform corresponding to (656) and (658) it is useful to consider the finite domain

$$0 < x < l, \quad (659)$$

and then let l approach infinity. On the bounded domain (659), the boundary condition at $x = +\infty$ is replaced by the condition $M(l) = 0$. This gives us the following normalized eigenfunctions

$$M_k(x) = \sqrt{\frac{2}{l}} \sin \lambda_k x, \quad k = 1, 2, \dots, \quad (660)$$

where

$$\lambda_k = \frac{k\pi}{l}. \quad (661)$$

A function $\tilde{f}(x)$ can now be expanded in a Fourier series with respect to (660) according to

$$\tilde{f}(x) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} F_k \sin \lambda_k x, \quad (662)$$

$$F_k = \sqrt{\frac{2}{l}} \int_0^l dx \sin \lambda_k x \tilde{f}(x). \quad (663)$$

Observe that the distance between the points λ_k on the λ -axis is

$$\Delta\lambda = \lambda_{k+1} - \lambda_k = \frac{\pi}{l}. \quad (664)$$

We use this and rewrite (662), (663) in the following way

$$\tilde{f}(x) = \frac{\sqrt{2}}{\pi} \sum_{k=1}^{\infty} \Delta\lambda \sin \lambda_k x \tilde{F}_k, \quad (665)$$

$$\tilde{F}_k(x) = \sqrt{2} \int_0^l dx \sin \lambda_k x \tilde{f}(x), \quad (666)$$

where $\tilde{F}_k = \sqrt{l} F_k$. If we formally let $l \rightarrow \infty$ in (665) and (666) we get

$$\tilde{f}(x) = \frac{\sqrt{2}}{\pi} \int_0^{\infty} d\lambda \sin \lambda x \tilde{F}(\lambda), \quad (667)$$

$$\tilde{F}(\lambda) = \sqrt{2} \int_0^{\infty} dx \sin \lambda x \tilde{f}(x). \quad (668)$$

In order to get a more symmetric and pleasing form of the transform (667), (668) we introduce scalings

$$\tilde{F} = \alpha F, \quad \tilde{f} = \beta f, \quad (669)$$

$$\Rightarrow f(x) = \frac{\sqrt{2}\alpha}{\pi\beta} \int_0^{\infty} d\lambda \sin \lambda x F(\lambda), \quad (670)$$

$$F(\lambda) = \frac{\sqrt{2}\beta}{\alpha} \int_0^{\infty} dx \sin \lambda x f(x). \quad (671)$$

Choose $\alpha = \sqrt{\pi}$, $\beta = 1$. Then we get the *Fourier sine transform*

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} d\lambda \sin \lambda x F_s(\lambda), \quad (672)$$

$$F_s(\lambda) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} dx \sin \lambda x f(x). \quad (673)$$

Our friendly analysis can tell us for which pair of nice functions the improper integrals converge and for which (672) and (673) holds.

However we will apply it to much more general things and will proceed formally and hope that in the end everything will be ok in the universe of distributions. Here it is only worth noting that even though $\sin \lambda x$ is zero for $x = 0$, $f(x)$ in (672) might not approach zero as x approach zero. This is because in general the convergence of the improper integral will not allow us to interchange integral and limit.

A similar derivation for the boundary condition (651) gives us the *Fourier cosine transform*

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty d\lambda \cos \lambda x F_c(\lambda), \quad (674)$$

$$F_c(\lambda) = \sqrt{\frac{2}{\pi}} \int_0^\infty dx \cos \lambda x f(x). \quad (675)$$

The Fourier sine- and cosine transforms have several interesting and useful properties[1]. Here we will discuss how they behave with respect to derivatives. Observe that

$$\int_0^\infty dx \sin \lambda x f'(x) = \sin \lambda x f(x) \Big|_0^\infty - \lambda \int_0^\infty dx \cos \lambda x f(x). \quad (676)$$

If we assume that $f(x)$ vanishes at $x = +\infty$ we get the identity

$$\sqrt{\frac{2}{\pi}} \int_0^\infty dx \sin \lambda x f'(x) = -\lambda F_c(\lambda). \quad (677)$$

In a similar way we find, assuming that also $f'(x)$ vanishes at infinity,

$$\sqrt{\frac{2}{\pi}} \int_0^\infty dx \sin \lambda x f''(x) = \lambda \sqrt{\frac{2}{\pi}} f(0) - \lambda^2 F_s(\lambda). \quad (678)$$

Thus the Fourier sine transform of f' is expressed in terms of the Fourier cosine transform. The Fourier sine transform of f'' is expressed in terms of the Fourier sine transform of f and a term depending on $f(0)$.

For the Fourier cosine transform we get in a similar way

$$\sqrt{\frac{2}{\pi}} \int_0^\infty dx \cos \lambda x f'(x) = -\sqrt{\frac{2}{\pi}} f(0) + \lambda F_s(\lambda), \quad (679)$$

$$\sqrt{\frac{2}{\pi}} \int_0^\infty dx \cos \lambda x f''(x) = -\lambda \sqrt{\frac{2}{\pi}} f'(0) - \lambda^2 F_c(\lambda). \quad (680)$$

We observe that in order to apply the Fourier sine- and cosine transforms to solve differential equations on a semi infinite domain the boundary conditions must be of a particular type. For second order equations the Fourier sine transform requires Dirichlet condition and the cosine transform requires Neumann condition. Since one typically solve time dependent problems on a semi-infinite domain $t > 0$, one might think that the Fourier Sine- and Cosine transforms should be useful for this, but they are in fact not.

If the equation contains a first derivative with respect to time we will find that the Fourier Sine transform of u_t is expressed in terms of the Fourier Cosine transform for u and vice versa. If the equation contains two time derivatives, u_{tt} , there are two data given at $t = 0$, u and u_t . But both the Fourier Sine and

the Fourier Cosine transform can only include one condition at $t = 0$, u for the Sine transform and u_t for the Cosine transform.

We will see that for initial value problems, the *Laplace transform* is the appropriate one.

Another useful property of the Fourier Sine- and Cosine transform is

$$\sqrt{\frac{2}{\pi}} \int_0^\infty dx x \sin \lambda x f(x) = -\frac{\partial F_c}{\partial \lambda}, \quad (681)$$

$$\sqrt{\frac{2}{\pi}} \int_0^\infty dx x \cos \lambda x f(x) = \frac{\partial F_s}{\partial \lambda}. \quad (682)$$

11.2.1 The heat equation on a semi-infinite interval

We consider the equation

$$u_t(x, t) - c^2 u_{xx}(x, t) = 0, \quad 0 < x < \infty, t > 0, \quad (683)$$

with initial condition

$$u(x, 0) = f(x), \quad 0 < x < \infty, \quad (684)$$

and boundary condition

$$u(0, t) = g(t), \quad t > 0. \quad (685)$$

Because of the Dirichelet boundary condition at $x = 0$ it is appropriate to use the Fourier sine transform on the x -variable.

Taking the Fourier Sine transform of (683) and make use properties (678) and (685), we get

$$\partial_t U_s(\lambda, t) + (\lambda c)^2 U_s(\lambda, t) = \lambda c^2 \sqrt{\frac{2}{\pi}} g(t), \quad (686)$$

with initial condition

$$U_s(\lambda, 0) = F_s(\lambda), \quad (687)$$

where

$$U_s(\lambda, t) = \sqrt{\frac{2}{\pi}} \int_0^\infty dx \sin \lambda x u(x, t), \quad (688)$$

$$F_s(\lambda) = \sqrt{\frac{2}{\pi}} \int_0^\infty dx \sin \lambda x f(x). \quad (689)$$

The initial value problem (686), (687) is standard with solution

$$\begin{aligned}
U_s(\lambda, t) &= F_s(\lambda) e^{-\lambda^2 c^2 t}, \\
&+ \lambda c^2 \sqrt{\frac{2}{\pi}} \int_0^t d\tau e^{-\lambda^2 c^2 (t-\tau)} g(\tau).
\end{aligned} \tag{690}$$

Taking the inverse Fourier Sine transform (672) we get the solution in the form

$$\begin{aligned}
u(x, t) &= \sqrt{\frac{2}{\pi}} \int_0^\infty d\lambda \sin \lambda x F_s(\lambda) e^{-\lambda^2 c^2 t} \\
&+ \frac{2c^2}{\pi} \int_0^\infty d\lambda \sin \lambda x \lambda \int_0^t d\tau e^{-\lambda^2 c^2 (t-\tau)} g(\tau) \\
&= \frac{2}{\pi} \int_0^\infty d\lambda \sin \lambda x \int_0^\infty ds \sin \lambda s f(s) e^{-\lambda^2 c^2 t} \\
&+ \frac{2c^2}{\pi} \int_0^\infty d\lambda \sin \lambda x \lambda \int_0^t d\tau e^{-\lambda^2 c^2 (t-\tau)} g(\tau) \\
&= \frac{2}{\pi} \int_0^\infty ds \left[\int_0^\infty d\lambda \sin \lambda x \sin \lambda s e^{-\lambda^2 c^2 t} \right] f(s) \\
&+ \frac{2c^2}{\pi} \int_0^t d\tau \left[\int_0^\infty d\lambda \lambda \sin \lambda x e^{-\lambda^2 c^2 (t-\tau)} \right] g(\tau).
\end{aligned} \tag{691}$$

Using the addition formulas for cosine we have

$$2 \sin \lambda x \sin \lambda s = \cos \lambda(x-s) - \cos \lambda(x+s). \tag{692}$$

But then

$$\begin{aligned}
&\frac{2}{\pi} \int_0^\infty d\lambda \sin \lambda x \sin \lambda s e^{-\lambda^2 c^2 t} \\
&= \frac{1}{\pi} \int_0^\infty d\lambda \{ \cos \lambda(x-s) - \cos \lambda(x+s) \} e^{-\lambda^2 c^2 t} \\
&= \frac{1}{\pi} \int_0^\infty d\lambda \cos \lambda(x-s) e^{-\lambda^2 c^2 t} - \frac{1}{\pi} \int_0^\infty d\lambda \cos \lambda(x+s) e^{-\lambda^2 c^2 t} \\
&= \frac{1}{\sqrt{4\pi c^2 t}} e^{-\frac{(x-s)^2}{4c^2 t}} - \frac{1}{\sqrt{4\pi c^2 t}} e^{-\frac{(x+s)^2}{4c^2 t}} \\
&= G(x-s, t) - G(x+s, t),
\end{aligned} \tag{693}$$

where we have used quantities defined in (605) and (610). We also have

$$\begin{aligned}
&\frac{1}{\pi} \int_0^\infty d\lambda \lambda \sin \lambda x e^{-\lambda^2 c^2 (t-\tau)} \\
&= -\frac{1}{\pi} \partial_x \int_0^\infty d\lambda \cos \lambda x e^{-\lambda^2 c^2 (t-\tau)} \\
&= -\partial_x G(x, t-\tau),
\end{aligned} \tag{694}$$

so the solution can compactly be written as

$$u(x, t) = \int_0^\infty ds \{G(x - s, t) - G(x + s, t)\} f(s) \quad (695)$$

$$- 2c^2 \int_0^t d\tau \partial_x G(x, t - \tau) g(\tau). \quad (696)$$

Let us look at a couple of simple cases.

For the first case we assume $g(t) = 0$, so the boundary condition is homogeneous, and that the initial temperature profile is uniform $u(x, 0) = f(x) = u_0$. Then the second term in (695) vanishes and

$$\begin{aligned} u(x, t) &= \int_0^\infty ds \{G(x - s, t) - G(x + s, t)\} u_0 \quad (697) \\ &= \frac{u_0}{\sqrt{4\pi c^2 t}} \left\{ \int_0^\infty ds e^{-\frac{(x-s)^2}{4c^2 t}} - \int_0^\infty ds e^{-\frac{(x+s)^2}{4c^2 t}} \right\} \\ &\quad r = \frac{s-x}{2c\sqrt{t}} \quad r = \frac{s+x}{2c\sqrt{t}} \\ &= \frac{u_0}{\sqrt{\pi}} \left\{ \int_{-\frac{x}{2c\sqrt{t}}}^\infty dr e^{-r^2} - \int_{\frac{x}{2c\sqrt{t}}}^\infty dr e^{-r^2} \right\} \\ &= \frac{2u_0}{\sqrt{\pi}} \int_0^{\frac{x}{2c\sqrt{t}}} dr e^{-r^2} = u_0 \operatorname{erf} \left(\frac{x}{2c\sqrt{t}} \right), \end{aligned}$$

where $\operatorname{erf}(\xi)$ is the *error function*

$$\operatorname{erf}(\xi) = \frac{2}{\sqrt{\pi}} \int_0^\xi dr e^{-r^2}. \quad (698)$$

We observe that

$$\operatorname{erf}(0) = 0, \quad \operatorname{erf}(\infty) = 1. \quad (699)$$

Using the properties of the error function we find that

$$u(x, 0) = u_0 \operatorname{erf}(\infty) = u_0, \quad (700)$$

$$u(0, t) = u_0 \operatorname{erf}(0) = 0, \quad (701)$$

$$u(x, t) \rightarrow 0 \text{ as } t \rightarrow \infty. \quad (702)$$

Thus $u(x, t)$ satisfies *all* requirements of our problem even if $u(x, 0) = u_0$ does not have a Fourier sine transform. One can show that the solution (697) is the limit of a sequence of problems that all have initial data for which the transform exist.

For the second case we don't assume that $f(x)$ is constant, but merely that it is bounded

$$|f(x)| \leq M < \infty, \quad \text{for all } x > 0. \quad (703)$$

We let the boundary data at $x = 0$ be given by an arbitrary time dependent function $g(t)$. Observe that for this case, the first term in the solution formula (695) approach zero for large t

$$\begin{aligned} & \left| \int_0^\infty ds \{G(x-s, t) - G(x+s, t)\} f(s) \right| \\ & \leq M \int_0^\infty ds \{G(x-s, t) - G(x+s, t)\} = M \operatorname{erf} \left(\frac{x}{2c\sqrt{t}} \right) \rightarrow 0. \end{aligned}$$

Therefore as $t \rightarrow \infty$ the solution $u(x, t)$ approaches

$$u(x, t) \approx -2c^2 \int_0^t d\tau \partial_x G(x, t - \tau) g(\tau). \quad (704)$$

This solution is called a *steady state*. It is not time invariant but represent the solution for times so large that the initial data has been dissipated. For such large times the solution only depends on the boundary condition, or *drive at the boundary*. Recall that an analogy behavior is displayed by a driven, damped harmonic oscillator.

11.3 The Laplace transform

The last transform we will discuss in this class is the *Laplace transform*. It is, as we will see, well suited for solving initial value problems. We will introduce the Laplace transform using the Fourier transform.

The Laplace transform is concerned with functions $f(t)$ which are zero for $t < 0$. The function $f(t)$ can be quite wild for $t > 0$, but we will set a bound on how fast it can grow when $t \rightarrow \infty$.

We will therefore require $f(t)$ to be of *exponential order* as $t \rightarrow \infty$. This means that there are constants $c > 0$, k and M such that

$$|f(t)| \leq M e^{kt}, \quad \text{for } t > c. \quad (705)$$

Let now α be a constant such that $\alpha > k$. Then the Fourier transform of $e^{-\alpha t} f(t)$ will exist and is given by

$$\hat{F}(\hat{\lambda}) = \frac{1}{\sqrt{2\pi}} \int_0^\infty dt e^{i\hat{\lambda}t} e^{-\alpha t} f(t), \quad (706)$$

where the fact that $f(t) = 0$ for $t \leq 0$ has been taken into account. For example if $f(t) = 1$ we get

$$\hat{F}(\hat{\lambda}) = \frac{1}{\sqrt{2\pi}} \int_0^\infty dt e^{(i\hat{\lambda} - \alpha)t} = \frac{-1}{\sqrt{2\pi}(i\hat{\lambda} - \alpha)}. \quad (707)$$

In the Fourier transform we know that $\hat{\lambda}$ is a real number related to the spectrum of a certain differential operator. Observe, however, that the right hand side of (707) can be evaluated for complex $\hat{\lambda}$ also. For example

$$\hat{F}(i) = \frac{-1}{\sqrt{2\pi}(i^2 - \alpha)} = \frac{1}{\sqrt{2\pi}(1 + \alpha)}. \quad (708)$$

Also observe that (707) can not be evaluated for *all* complex numbers, in fact for $\hat{\lambda} = -i\alpha$ the denominator in (707) is zero. However, this is the only problematic point and $\hat{F}(\hat{\lambda})$ can be extended to the whole complex plane otherwise

$$\hat{F}(\hat{\lambda}) = \frac{-1}{\sqrt{2\pi}(i\hat{\lambda} - \alpha)}, \quad \hat{\lambda} \in \mathbb{C} - \{-i\alpha\}. \quad (709)$$

In a similar way we will now prove that $\hat{F}(\lambda)$ can be extended to complex values for any $f(t)$ of exponential order.

Evaluating (706) at $\hat{\lambda} = \hat{\lambda}_r + i\hat{\lambda}_i \in \mathbb{C}$ gives us

$$\hat{F}(\hat{\lambda}) = \frac{1}{\sqrt{2\pi}} \int_0^\infty dt e^{(i\hat{\lambda} - \alpha)t} f(t). \quad (710)$$

For this to make sense the integral has to converge. But

$$\begin{aligned} |\hat{F}(\hat{\lambda})| &\leq \frac{1}{\sqrt{2\pi}} \int_0^\infty dt e^{-(\hat{\lambda}_i + \alpha)t} |f(t)| \\ &\leq \frac{1}{\sqrt{2\pi}} \int_0^c dt e^{-(\hat{\lambda}_i + \alpha)t} |f(t)| \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_c^\infty dt M e^{(k - \hat{\lambda}_i - \alpha)t} \end{aligned} \quad (711)$$

and the last integral is finite only if

$$\hat{\lambda}_i > -(\alpha - k) \quad (\text{Recall that } \alpha - k > 0). \quad (712)$$

So $\hat{F}(\lambda)$ can not be defined in the whole complex plane but only in a domain defined by (712).

Below the horizontal line $\hat{\lambda} = -i(\alpha - k)$ in the complex plane, we expect the function $\hat{F}(\hat{\lambda})$ to have singularities. For the special choice

$$f(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0, \end{cases} \quad (713)$$

this was exactly what occurred.

Change variables in the complex plane to λ through

$$\hat{\lambda} = i(\lambda - \alpha) \quad (714)$$

and define a function $F(\lambda)$ by

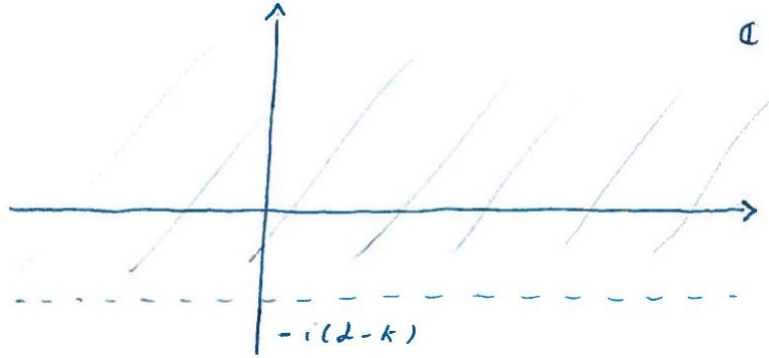


Figure 59

$$F(\lambda) = \sqrt{2\pi} \hat{F}(i(\lambda - \alpha)). \quad (715)$$

From (710) we then get

$$F(\lambda) = \int_0^\infty dt e^{-\lambda t} f(t). \quad (716)$$

For a function of exponential order k , $F(\lambda)$ is defined for complex arguments $\lambda = \lambda_r + i\lambda_i$, with

$$\lambda_r > k. \quad (717)$$

$F(\lambda)$ is the *Laplace transform* of $f(t)$. Using the Fourier inversion formula we have

$$e^{-\alpha t} f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\hat{\lambda} e^{-i\hat{\lambda} t} \hat{F}(\hat{\lambda}). \quad (718)$$

Introducing $\hat{\lambda} = i(\lambda - \alpha)$ gives us $d\hat{\lambda} = i d\lambda$ so that (718) implies

$$f(t) = \frac{1}{2\pi i} \int_L d\lambda e^{\lambda t} F(\lambda), \quad (719)$$

where L is a parameterized curve

$$L: \quad \lambda = \alpha + iu, \quad -\infty < u < \infty. \quad (720)$$

The curve L is indicated in Figure 60.

The formula (719) is the *inversion formula* for the Laplace transform. The inversion formula is a very effective and flexible way to calculate the inverse Laplace transform, but using it requires *residue calculus* from complex analysis, so we will not make further use of formula (719) in this class. Inverse Laplace

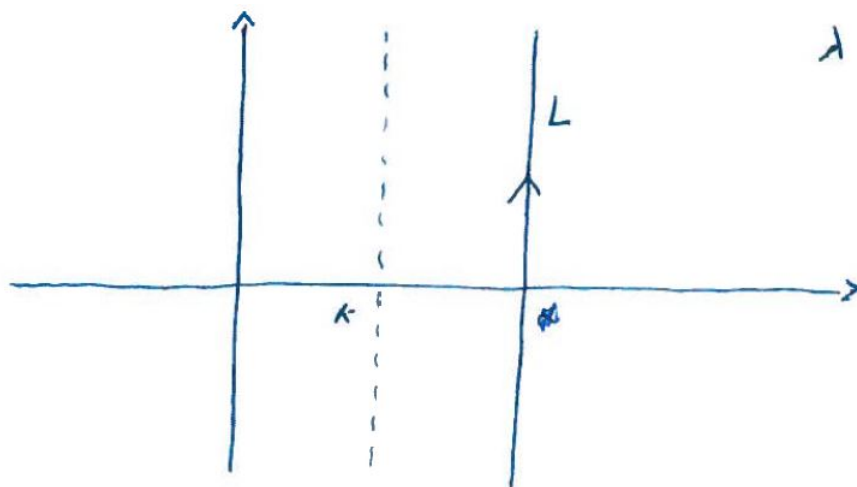


Figure 60

transforms can in many common situations be found using a combination of formal properties of the inverse Laplace transform and tables of inverse transforms. Lists of such formal properties can be found in mathematical tables or on the web. Here we will only display the most important property we need when we solve initial value problems for differential equations using the Laplace transform

$$\int_0^\infty dt e^{-\lambda t} f^{(n)}(t) = \lambda^n \int_0^\infty dt e^{-\lambda t} f(t) - \lambda^{n-1} f(0) - \dots f^{(n-1)}(0). \quad (721)$$

This formula is easy to prove using integration by parts. From (721) it is clear that the n initial conditions for a n^{th} order differential equation can be included in a natural way. The Laplace transform is therefore very useful for solving initial value problem for both ODEs and PDEs. Here we will apply it to initial boundary value problems for PDEs.

We will now apply the Laplace transform to two initial value problems for PDEs.

11.3.1 The wave equation on a half interval

$$\begin{aligned} u_{tt} - c^2 u_{xx} &= 0, & 0 < x < \infty, t > 0, \\ u(x, 0) &= u_t(x, 0) = 0, \\ u(0, t) &= f(t), \\ u(x, t) &\rightarrow 0, & x \rightarrow \infty. \end{aligned} \quad (722)$$

Taking the Laplace transform of the equation we get

$$\lambda^2 U(x, \lambda) - \lambda u(x, 0) - \lambda^2 u_t(x, 0) - c^2 U_{xx}(x, \lambda) = 0 \quad (723)$$

Using the initial conditions we get

$$\begin{aligned} \lambda^2 U(x, \lambda) - c^2 U_{xx}(x, \lambda) &= 0, \\ U(0, \lambda) &= F(\lambda). \end{aligned} \quad (724)$$

The general solution to the family of ordinary differential equations (724) is

$$U(x, \lambda) = a(\lambda)e^{-\frac{\lambda}{c}x} + b(\lambda)e^{\frac{\lambda}{c}x}. \quad (725)$$

Recall that we evaluate (725) in the right halfspace only, so $\lambda_r > 0$. In order to satisfy the boundary condition

$$u(x, t) \rightarrow 0, x \rightarrow \infty \Leftrightarrow U(x, \lambda) \rightarrow 0, x \rightarrow \infty. \quad (726)$$

We must clearly assume that

$$b(\lambda) = 0. \quad (727)$$

Finally applying the boundary condition at $x = 0$ we get

$$U(x, \lambda) = F(\lambda)e^{-\frac{\lambda}{c}x}. \quad (728)$$

We now must take the inverse Laplace transform of (728).

Observe that if $H(\xi)$ is the Heaviside function

$$H(\xi) = \begin{cases} 1, & \xi \geq 0, \\ 0, & \xi < 0. \end{cases} \quad (729)$$

Then

$$\begin{aligned} & \int_0^\infty dt H(t-b) f(t-b) e^{-\lambda t} \\ &= \int_b^\infty dt f(t-b) e^{-\lambda t}, \quad \xi = t-b, d\xi = dt, \\ &= \int_0^\infty d\xi f(\xi) e^{-\lambda(\xi+b)} = e^{-\lambda b} F(\lambda). \end{aligned} \quad (730)$$

This is equal to (728) with $b = \frac{x}{c}$. Thus the inverse Laplace transform of (728) must be

$$u(x, t) = H\left(t - \frac{x}{c}\right) f\left(t - \frac{x}{c}\right) \quad (731)$$

and we have solved the initial value problem.

11.3.2 The heat equation on a finite interval

Consider the problem

$$u_t(x, t) - c^2 u_{xx}(x, t) = 0, \quad 0 < x < l, \quad t > 0, \quad (732)$$

$$u(0, t) = u(l, t) = 0, \quad t > 0,$$

$$u(x, 0) = f(x), \quad 0 < x < l. \quad (733)$$

Taking the Laplace transform of (732) we get

$$\lambda U(x, \lambda) - f(x) - c^2 U_{xx}(x, \lambda) = 0, \quad 0 < x < l, \quad (734)$$

$$U(0, \lambda) = U(l, \lambda) = 0.$$

This is now a boundary value problem. Equation (734) can be solved using variation of parameters from the theory of ODEs. Fitting the boundary conditions to the variation of parameter solution gives

$$\begin{aligned} U(x, \lambda) = & \frac{1}{c\sqrt{\lambda} \sinh\left(\frac{l\sqrt{\lambda}}{c}\right)} \int_0^x ds \sinh\left(\frac{\sqrt{\lambda}}{c}s\right) \sinh\left(\frac{\sqrt{\lambda}}{c}(l-x)\right) f(s) \\ & + \frac{1}{c\sqrt{\lambda} \sinh\left(\frac{l\sqrt{\lambda}}{c}\right)} \int_x^l ds \sinh\left(\frac{\sqrt{\lambda}}{c}x\right) \sinh\left(\frac{\sqrt{\lambda}}{c}(l-s)\right) f(s). \end{aligned} \quad (735)$$

Finding the inverse transform of (735) is not easy, but we have a secret weapon. In the inversion formula for the Laplace transform (719) we can place the integration curve L anywhere as long as $\lambda_r > k$. Thus we may assume in (735) that $|\lambda|$ is as large as we want.

For such λ we have

$$\frac{1}{\sinh\left(\frac{\sqrt{\lambda}l}{c}\right)} = \frac{2}{e^{\frac{l\sqrt{\lambda}}{c}} - e^{-\frac{l\sqrt{\lambda}}{c}}} = \frac{2e^{-\frac{l\sqrt{\lambda}}{c}}}{1 - e^{-2\frac{l\sqrt{\lambda}}{c}}} = 2e^{-\frac{l\sqrt{\lambda}}{c}} \sum_{j=0}^{\infty} e^{-\frac{2jl\sqrt{\lambda}}{c}}. \quad (736)$$

The last equality sign in (736) is valid because for $|\lambda|$ large enough $|e^{-2l\sqrt{\lambda}/c}| < 1$ and therefore

$$\frac{1}{1-x} = \sum_{j=0}^{\infty} x^j. \quad (737)$$

The hyperbolic sine under the integral signs can also be expressed in terms of exponentials. Thus we end up with an infinite sum where each term is an exponential.

From a table of Laplace transforms we have

$$\int_0^\infty dt e^{-\lambda t} \frac{1}{\sqrt{\pi t}} e^{-\frac{\alpha^2}{4t}} = \frac{1}{\sqrt{\lambda}} e^{-\alpha\sqrt{\lambda}}, \quad \alpha > 0. \quad (738)$$

Inserting (738) into the infinite sum, formally interchanging summation and integration gives us finally

$$u(x, t) = \int_0^l ds \sum_{j=-\infty}^{\infty} \{G(x - s - 2jl, t) - G(x + s - 2jl, t)\} f(s), \quad (739)$$

where $G(y, t)$ is the fundamental solution to the heat equation. We have previously found the solution to (732) using a Fourier sine series

$$u(x, t) = \sqrt{\frac{2}{l}} \sum_{j=1}^{\infty} a_j e^{-(\frac{\pi j c}{l})^2 t} \sin\left(\frac{\pi j}{l} x\right). \quad (740)$$

Problem (732) has a unique solution so even if (739) and (740) *appear* to be very different they must in fact be identical.

The formulas can be proven to be identical using the *Poisson summation formula*[1].

The formulas are mathematically equivalent, but are useful for evaluating $u(x, t)$ for different ranges of t . Because of the exponential term in (740), evaluating $u(x, t)$ for large t using this formula requires only a few terms.

However for small t , j must be large before the exponential term kicks in. Thus we require many terms in (740) in order to evaluate $u(x, t)$ for small t .

On the other hand, for small t , the fundamental solution $G(y, t)$ is sharply peaked around $y = 0$. Thus for small t , only a few terms in (739) contribute to $u(x, t)$. Thus for small t (739) is fast to evaluate.

In general it is not easy to invert integral transforms. There are however powerful analytic approximation methods that can be brought to bear on this problem.

Observe that the formula for the Laplace transform

$$F(\lambda) = \int_0^\infty dt e^{-\lambda t} f(t) \quad (741)$$

clearly shows that $F(\lambda)$ for large $|\lambda|$ depends mainly on $f(t)$ for small t . This can be made more precise. From the body of results called *Abelian asymptotic theory* we have the following result:

If

$$f(t) \approx \sum_{n=0}^{\infty} \alpha_n t^n, \quad t \rightarrow 0, \quad (742)$$

then

$$F(\lambda) \approx \sum_{n=0}^{\infty} \alpha_n \frac{n!}{\lambda^{n+1}}, \quad |\lambda| \rightarrow \infty. \quad (743)$$

The converse is also true, (743) implies (742). Thus if we are only interested in $f(t)$ for small t we can find the expression for $f(t)$ by expanding $F(\lambda)$ in inverse powers of λ . This expansion gives us the coefficients α_n which we then insert into (742).

There are also results relating the behavior of $f(t)$ for large t to the behavior of $F(\lambda)$ for small λ . This body of results is called *Tauherian asymptotic theory*. Describing the simplest results from this theory requires however complex analysis so we will not do it here.

11.4 Group velocity 1

Let us consider the Klein-Gordon equation

$$u_{tt}(x, t) - \gamma^2 u_{xx}(x, t) + c^2 u(x, t) = 0, \quad t > 0, \quad -\infty < x < \infty. \quad (744)$$

Initial data is of the form

$$\begin{aligned} u(x, 0) &= f(x), \\ u_t(x, 0) &= g(x). \end{aligned} \quad (745)$$

Recall that the Klein-Gordon equation is of dispersive type. We have found this previously using normal mode analysis. Taking the Fourier transform of the Klein-Gordon equation we get

$$\partial_{tt}U(\lambda, t) + (\gamma^2 \lambda^2 + c^2)U(\lambda, t) = 0. \quad (746)$$

The general solution to this equation is

$$U(\lambda, t) = F_+(\lambda)e^{i\omega t} + F_-(\lambda)e^{-i\omega t}, \quad (747)$$

where $\omega = \omega(\lambda)$ is the dispersion relation for the Klein-Gordon equation

$$\omega = (\gamma^2 \lambda^2 + c^2)^{\frac{1}{2}}. \quad (748)$$

Using the inverse Fourier transform we get the following general solution to (744)

$$\begin{aligned} u(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_+(\lambda) e^{i(\omega(\lambda)t - \lambda x)} \\ &+ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_-(\lambda) e^{-i(\omega(\lambda)t + \lambda x)}. \end{aligned} \quad (749)$$

The functions $F_+(\lambda)$ and $F_-(\lambda)$ are determined from initial data.

From (745) the Fourier transform of the initial data is

$$U(\lambda, 0) = F(\lambda) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f(x) e^{i\lambda x}, \quad (750)$$

$$U_t(\lambda, 0) = G(\lambda) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx g(x) e^{i\lambda x}.$$

Using (747) we get

$$\begin{aligned} F_+(\lambda) + F_-(\lambda) &= F(\lambda), \\ i\omega F_+(\lambda) - i\omega F_-(\lambda) &= G(\lambda). \end{aligned} \quad (751)$$

Solving (751) gives us

$$\begin{aligned} F_+(\lambda) &= \frac{i\omega F(\lambda) + G(\lambda)}{2i\omega}, \\ F_-(\lambda) &= \frac{i\omega F(\lambda) - G(\lambda)}{2i\omega}. \end{aligned} \quad (752)$$

(752) together with (749) solves the initial value problem (745) for the Klein-Gordon equation (744).

The solution (749) is clearly a superposition of normal modes

$$F_{\pm}(\lambda) e^{i(\pm\omega(\lambda)t - \lambda x)}. \quad (753)$$

The phase-speed for the normal modes are

$$v_f(\lambda) = \frac{dx}{dt} = \pm \frac{\omega(\lambda)}{\lambda} = \pm \sqrt{\gamma^2 + \left(\frac{c}{\lambda}\right)^2}. \quad (754)$$

Recall that the phase speed of a normal mode is the speed we have to move at in order for the phase

$$\theta_{\pm}(x, t) = \pm\omega(\lambda)t - \lambda x, \quad (755)$$

to stay fixed. Thus we have

$$\theta_{\pm}(x(t), t) = \theta_{\pm}^0, \quad (756)$$

\Downarrow

$$\partial_x \theta_{\pm} \frac{dx}{dt} + \partial_t \theta_{\pm} = 0, \quad (757)$$

\Downarrow

$$-\lambda \frac{dx}{dt} \pm \omega(\lambda) = 0,$$

\Downarrow

$$\frac{dx}{dt} = \pm \frac{\omega(\lambda)}{\lambda}.$$

The Klein-Gordon equation has the same principle part as the wave equation. It thus have the same characteristic speed $\pm\gamma$.

As we have seen, discontinuities move at characteristic speed $\pm\gamma$ and $\pm\gamma$ represents the upper speed limit for the Klein-Gordon equation.

But from (754) we see that the phase speed is bigger than γ for all modes. The physical relevance of the phase speed is therefore in question. Furthermore, noting that the solution to the Klein-Gordon equation is a superposition of normal modes which all move at different phase velocities, it is not at all clear that the phase speed is a relevant quantity for this equation.

In order to investigate this problem we will consider special initial conditions of the form

$$\begin{aligned} u(x, 0) &= p(\epsilon x) e^{-ik_0 x}, \\ u_t(x, 0) &= q(\epsilon x) e^{-ik_0 x}, \end{aligned} \quad (758)$$

where k_0 is fixed. The function $e^{ik_0 x}$ is clearly periodic of period

$$L_0 = \frac{2\pi}{k_0}. \quad (759)$$

In (758) we assume that $0 < \epsilon \ll 1$ is a number so small that

$$\epsilon L_0 \ll 1. \quad (760)$$

Then the amplitudes $p(\epsilon x)$ and $q(\epsilon x)$ vary very little over a period L_0 . The initial conditions (758) are what we call *wave packets*. These are functions that locally look like plane waves but that has a (complex) amplitude that varies over scales much larger than L_0 .

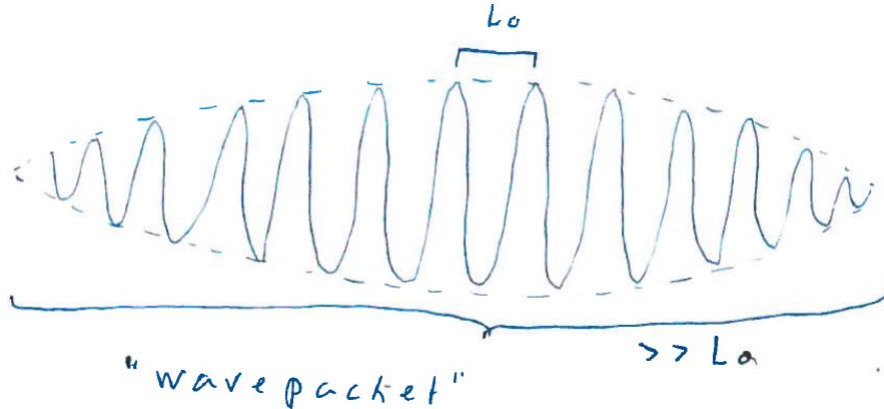


Figure 61

Using simple properties of the Fourier transform we find easily from (749) that

$$\begin{aligned} F(\lambda) &= \frac{1}{\epsilon} P\left(\frac{\lambda - k_0}{\epsilon}\right), \\ G(\lambda) &= \frac{1}{\epsilon} Q\left(\frac{\lambda - k_0}{\epsilon}\right), \end{aligned} \quad (761)$$

where P and Q are the Fourier transform of p and q . From Figure 62 it is clear

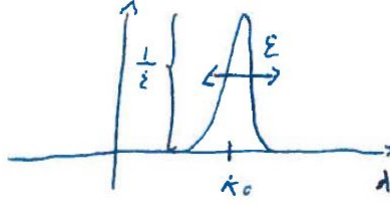


Figure 62

that in Fourier space, $F(\lambda)$ and $G(\lambda)$ are narrow "spikes" of width ϵ , height $\frac{1}{\epsilon}$ and centered at $\lambda = k_0$.

The functions $F_+(\lambda)$, $F_-(\lambda)$ will also be of this form

$$F_{\pm}(\lambda) = \frac{1}{\epsilon} A_{\pm}\left(\frac{\lambda - k_0}{\epsilon}\right), \quad (762)$$

for some functions A_{\pm} .

The solution of the Klein-Gordon equation consists of two terms of a very similar structure. The first term is

$$\begin{aligned} u_+(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_+(\lambda) e^{i(\omega(\lambda)t - \lambda x)} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{1}{\epsilon} A_+\left(\frac{\lambda - k_0}{\epsilon}\right) e^{i\omega(\lambda)t} e^{-i\lambda x}. \end{aligned} \quad (763)$$

Consider the following picture

From figure 63 it is clear that if the wavepacket is narrow enough, the dispersion relation $\omega(\lambda)$ is well approximated by its first order Taylor polynomial within the support of $F_+(\lambda)$. Thus we have

$$\begin{aligned} u_+(x, t) &\approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{1}{\epsilon} A_+\left(\frac{\lambda - k_0}{\epsilon}\right) e^{i(\omega(k_0) + \omega'(k_0)(\lambda - k_0))t} e^{-i\lambda x} \\ &= e^{i(\omega(k_0)t - k_0 x)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \frac{1}{\epsilon} A_+\left(\frac{\lambda - k_0}{\epsilon}\right) e^{-i((\lambda - k_0)x - \omega'(k_0)(\lambda - k_0)t)}. \end{aligned} \quad (764)$$

Changing integration variable in the integral

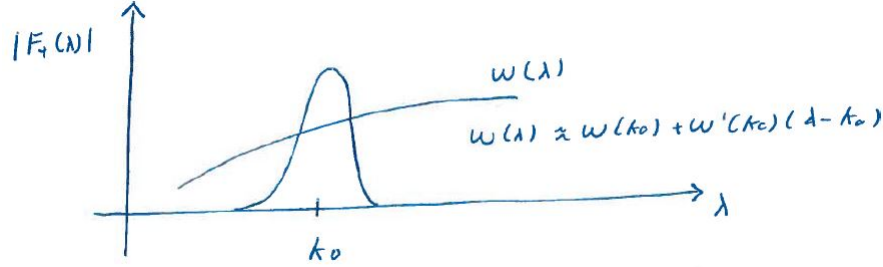


Figure 63

$$\eta = \frac{\lambda - k_0}{\epsilon} \Rightarrow d\eta = \frac{1}{\epsilon} d\lambda, \quad (765)$$

gives us

$$u_+(x, t) \approx e^{i(\omega(k_0)t - k_0 x)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\eta A_+(\eta) e^{-i\eta\epsilon(x - \omega'(k_0)t)}. \quad (766)$$

Defining

$$a_+(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\eta A_+(\eta) e^{-i\eta\epsilon}, \quad (767)$$

we get the final expression

$$u(x, t) = a_+(x - \omega'(k_0)t) e^{i(\omega(k_0)t - k_0 x)}. \quad (768)$$

We observe that the *envelope* of the wave packet, a_+ , translates without changing shape at a speed

$$v_g \equiv \frac{d\omega}{d\lambda}(k_0). \quad (769)$$

v_g is the *group velocity* for the wave packet. The group velocity, and not the phase velocity, is the relevant speed for a wave packet. This conclusion holds true for all equations of dispersive type.

For the Klein-Gordon equation

$$\begin{aligned} \omega &= (\gamma^2 \lambda^2 + c^2)^{\frac{1}{2}}, \\ &\Downarrow \\ v_g &= \frac{\gamma^2 \lambda}{(\gamma^2 \lambda^2 + c^2)^{\frac{1}{2}}} = \frac{\gamma}{\sqrt{1 + \left(\frac{c}{\gamma \lambda}\right)^2}} < \gamma. \end{aligned} \quad (770)$$

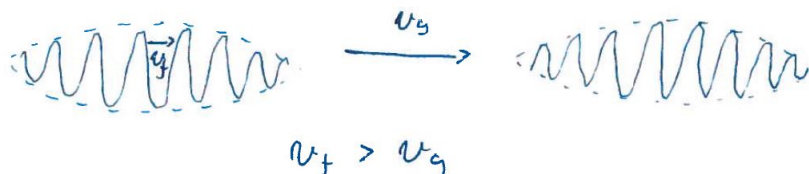


Figure 64

The group velocity is always smaller than the characteristic velocity γ for the Klein-Gordon equation

Note that the statement that the wave packet move at the group velocity is only approximately true. A closer investigation shows that eventually the wave packet spreads out and disperse.

We will shortly see that the group velocity plays a pivotal role in the description of general solutions to dispersive equations. However, before we do this we must describe a tool of outmost importance. This is the *method of stationary phase*.

11.5 Method of stationary phase

The method of stationary phase is part of a very large and important domain of applied mathematics called *asymptotic approximations*, or *perturbation methods*.

The method of stationary phase seeks to approximate integrals of complex integrands, where the *phase* of the integrand has a component that varies very quickly.

Let us consider the integral

$$I(k) = \int_{-\infty}^{\infty} dt f(t) e^{ik\varphi(t)}, \quad (771)$$

where $\varphi(t)$ is a real function and $f(t)$ can be a complex function. Both functions are assumed to be smooth.

We want to investigate the integral for large k . The formulas we derive will hold in the limit $k \rightarrow \infty$. In any actual application of the formulas, k is a fixed number, that is large, but not infinite. The errors we make in applying the asymptotic formulas using a given finite k , are not so easy to estimate. That is why we call it *method* of stationary phase rather than *theory* of stationary phase: Proofs are few and far between in this domain of applied mathematics. The method of stationary phase and asymptotic methods in general tend however to be surprisingly accurate even for values of k that are nowhere "close" to infinity.

Observe that when k is very large the phase term

$$e^{ik\varphi(t)}, \quad (772)$$

performs very fast oscillations. Since $f(t)$ is assumed to be smooth, the phase (772) will oscillate on a scale much faster than the scale at which the amplitude, $f(t)$, varies.

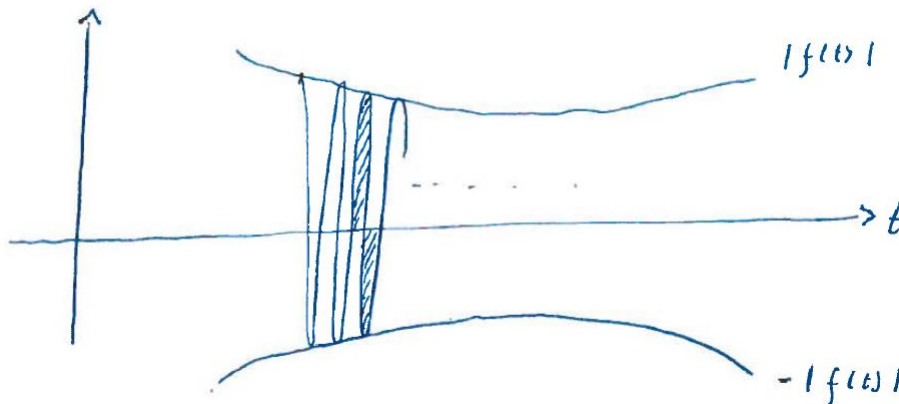


Figure 65

In this situation, as the figure illustrates, the positive and negative contributions in the integral tends to cancel each other so that we get

$$I(k) \rightarrow 0, \quad k \rightarrow \infty. \quad (773)$$

This argument can be made into precise mathematics using the *Riemann-Lebesques* lemma from complex analysis, but we will not pursue this any further here.

Let $t = t_0$. Close to t_0 we have

$$\varphi(t) = \varphi(t_0) + \varphi'(t_0)(t - t_0) + \frac{1}{2}\varphi''(t_0)(t - t_0)^2 + \dots \quad (774)$$

The largest contribution to $I(k)$ comes close to points where

$$\varphi'(t_0) = 0, \quad (775)$$

because around such points the phase varies more slowly than around points where $\varphi'(t_0) \neq 0$ for k large.

A point t_0 such that (775) hold is called a *stationary point*. If $t = t_0$ is a stationary point, the contribution to $I(k)$, which we call $I(k; t_0)$, is approximately given by

$$\begin{aligned} I(k; t_0) &\approx \int_{t_0-\epsilon}^{t_0+\epsilon} dt f(t) e^{i(\varphi(t_0) + \frac{1}{2}\varphi''(t_0)(t-t_0)^2)k} \\ &\approx f(t_0) e^{i\varphi(t_0)k} \int_{t_0-\epsilon}^{t_0+\epsilon} dt e^{i\frac{k}{2}\varphi''(t_0)(t-t_0)^2}, \end{aligned} \quad (776)$$

where we have used the fact that for a small enough interval $(t_0 - \epsilon, t_0 + \epsilon)$, $f(t)$ varies very little around the value $f(t_0)$. This is where we use the smoothness of $f(t)$. If $f(t)$ has a singularity at $t = t_0$ the approximation (776) is not valid. If k is large enough the integrand will oscillate very fast beyond $(t_0 - \epsilon, t_0 + \epsilon)$, so by the same argument as the one leading up to (773), we can conclude that the contribution from the integral

$$\int_{|t-t_0|>\epsilon} dt e^{i\frac{k}{2}\varphi''(t_0)(t-t_0)^2} \quad (777)$$

is negligible with respect to the contribution from the integral

$$\int_{|t-t_0|<\epsilon} dt e^{i\frac{k}{2}\varphi''(t_0)(t-t_0)^2}. \quad (778)$$

Thus we can without loss of generality write (776) in the form

$$I(k; t_0) \approx f(t_0) e^{i\varphi(t_0)k} \int_{-\infty}^{\infty} dt e^{i\frac{k}{2}\varphi''(t_0)(t-t_0)^2}. \quad (779)$$

The integral in (779) can be evaluated exactly using complex analysis

$$I(k; t_0) \approx f(t_0) e^{i\varphi(t_0)k} \sqrt{\frac{2\pi}{k|\varphi''(t_0)|}} e^{i\frac{\pi}{4} \frac{\varphi''(t_0)}{|\varphi''(t_0)|}}. \quad (780)$$

If you have several stationary points, each point gives a contribution of type (780) so that finally

$$I(k) \approx \sum_{j=1}^n I(k; t_j), \quad (781)$$

where $\varphi'(t_j) = 0$, $j = 1, 2, \dots, n$.

The method of stationary phase can be extended to integrals over finite domains and to multidimensional integrals.

The importance of the method of stationary phase can hardly be overstated. For example, it is by using this method that we understand how our macroscopic everyday world, ruled by classical physics, arise from a substratum ruled by the laws of quantum physics.

11.6 Group velocity 2

Let us now return to the Klein-Gordon equation (744), whose solution satisfying the general initial data (745) is

$$u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_+(\lambda) e^{i(\omega(\lambda)t - \lambda x)} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_-(\lambda) e^{-i(\omega(\lambda)t + \lambda x)}, \quad (782)$$

where F_+ and F_- are defined in terms of the initial data in (752). Both terms in (782) have a similar structure, so let us focus our attention on the first of them

$$I(x, t) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_+(\lambda) e^{i(\omega(\lambda)t - \lambda x)}, \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda F_+(\lambda) e^{it\varphi(\lambda)}. \quad (783)$$

Where we have defined the phase function

$$\varphi(\lambda) = \omega(\lambda) - \lambda \frac{x}{t}. \quad (784)$$

We want to investigate this expression for both x and t large. For such large values of x and t the phase varies very quickly as a function of λ and we will, according to the method of stationary phase, get the largest contributions to $I(x, t)$ close to values of λ where the phase is stationary. Thus we seek λ such that

$$\varphi'(\lambda) = 0, \\ \Downarrow \\ \omega'(\lambda) = \frac{x}{t}. \quad (785)$$

Thus for given x and t large we find the stationary points by solving the equation

For the Klein-Gordon equation we have from (748)

$$\omega(\lambda) = (\gamma^2 \lambda^2 + c^2)^{\frac{1}{2}}. \quad (786)$$

Thus the equation we must solve is

$$\frac{\gamma^2 \lambda}{(\gamma^2 \lambda^2 + c^2)^{\frac{1}{2}}} = \frac{x}{t}, \quad (787)$$

\Downarrow

$$\gamma^4 \lambda^2 = \left(\frac{x}{t}\right)^2 (\gamma^2 \lambda^2 + c^2),$$

\Updownarrow

$$\left(\gamma^4 - \gamma^2 \left(\frac{x}{t}\right)^2\right) \lambda^2 = c^2 \left(\frac{x}{t}\right)^2, \quad (788)$$

\Updownarrow

$$\gamma^2 \left(\gamma^2 - \left(\frac{x}{t}\right)^2\right) \lambda^2 = c^2 \left(\frac{x}{t}\right)^2. \quad (789)$$

Clearly this equation only has solutions if

$$\left|\frac{x}{t}\right| < \gamma. \quad (790)$$

But the light cone for the Klein-Gordon equation is $|x| = \gamma|t|$,

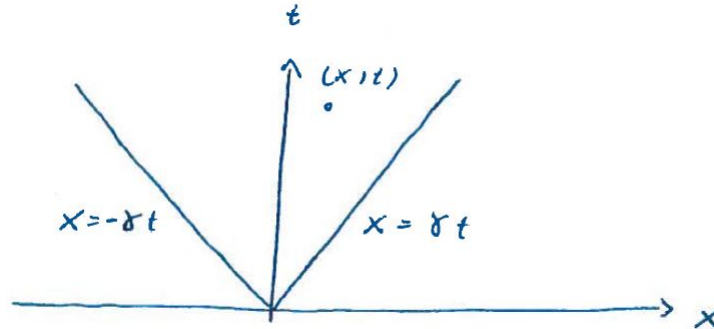


Figure 66

so stationary points only exists for (x, t) inside the light cone. Outside the light cone there are no stationary points and according to the Riemann Lebesques lemma $I(x, t) \approx 0$.

Assuming $\left|\frac{x}{t}\right| < \gamma$ we solve (789) and find

$$\lambda = \frac{c}{\gamma} \frac{x}{t} \left(\gamma^2 - \left(\frac{x}{t}\right)^2\right)^{-\frac{1}{2}}. \quad (791)$$

We do not get \pm in this formula because according to (787) λ and $\frac{x}{t}$ have the same sign.

Observe that

$$\omega''(\lambda) = \gamma^2 c^2 (\gamma^2 \lambda^2 + c^2)^{-\frac{3}{2}} > 0, \quad (792)$$

so that

$$\varphi''(\lambda) = \omega''(\lambda) > 0 \quad (\text{see 785}), \quad (793)$$

and therefore

$$\frac{\varphi''(\lambda)}{|\varphi''(\lambda)|} = 1. \quad (794)$$

Thus the stationary phase formula (780) gives us

$$I(x, t) \approx \frac{1}{\sqrt{t}} H(\lambda) e^{i(\omega(\lambda)t - \lambda x + \frac{\pi}{4})}, \quad (795)$$

where $H(\lambda) = F_+(\lambda)(\omega''(\lambda))^{-\frac{1}{2}}$ and where the phase in (783) has been written as

$$\omega(\lambda)t - \lambda x = t(\omega(\lambda) - \frac{x}{t}) = t\varphi, \quad (796)$$

so k in the stationary phase formula is identified with $t \gg 1$.

So what have we found? Observe that

$$\frac{x}{t} = \omega'(\lambda) = \frac{\gamma}{\left(1 + \left(\frac{c}{\gamma\lambda}\right)^2\right)^{\frac{1}{2}}} < \gamma. \quad (797)$$

We have the picture

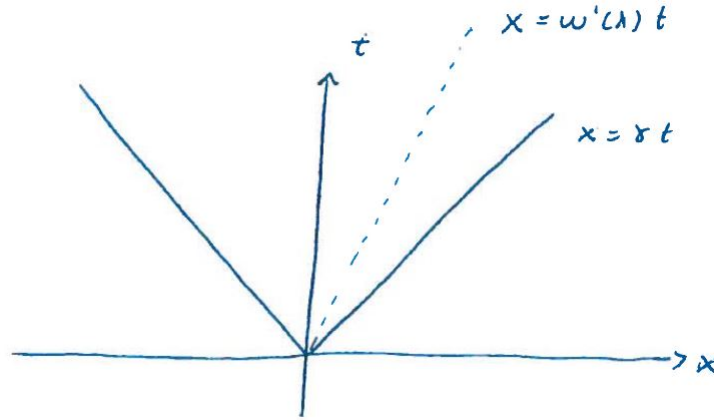


Figure 67

According to (795) and (779) for large x and t along the line $x = \omega'(\lambda)t$, the wave field take the form of a plane wave with frequency $\omega(\lambda)$, wavelength $\frac{2\pi}{\lambda}$ and amplitude

$$|I(x, t)| \sim \frac{1}{\sqrt{t}} |H(\lambda)|. \quad (798)$$

The amplitude of the plane wave will decay as $\frac{1}{\sqrt{t}}$. Or to put it another way; if we shift to a moving frame moving at speed $\omega'(\lambda)$ with respect to the lab frame we would see a plane wave whose amplitude decays as $\frac{1}{\sqrt{t}}$ and whose frequency $\omega(\lambda)$ and wavelength $\frac{2\pi}{\lambda}$ will depend on our speed. Observe that $\omega'(\lambda)$ is just the group velocity that we introduced in (769).

One can prove that for the Klein-Gordon equation, and also for many other dispersive equations, the energy in the wave field is transported at the group velocity.

The group velocity is the relevant velocity for dispersive wave equations. It is also a concept whose importance for the description of wave phenomena in general hardly can be overstated.

12 Projects

12.1 Project1

In this project the aim is to develop and implement on a computer a finite difference algorithm for the solution of the 1D wave equation. Deriving the finite difference algorithm and implementing it in a computer code is not really the challenging and time consuming step here. The real challenge is to make a convincing case for the correctness of the computer implementation. As I have stressed in my lectures we should never, ever, trust anything that comes out of a computer without extensive testing of the implemented code.

So how do we proceed to build the needed trust in our implementation? The best way to do this is to run the code in a situation where an analytic solution is known. Most of the time an analytic solution also needs to be implemented numerically, it typically include infinite sums of functions and integrals that can not be solved in terms of elementary functions. So what we are really doing is to compare the output of two different computer implementations for the same solution. If this is to make sense and build confidence, it is of paramount importance that the two codes implementing the finite difference algorithm and the analytical solution should be as independent as humanly possible. Under such circumstances closely similar output from the two codes will build confidence that our finite difference solution is correct.

So the challenge then is to find analytical solutions to the problem at hand, in our particular case this is the 1D wave equation subject to some boundary conditions that we will specify shortly. It is typically the case that we will not be able to find analytic solutions to the actual problem we are trying to solve but to a related problem or a special case of the problem where the domain, boundary conditions and/or equation is changed somewhat or fixed by some specific choices. After all, if we could find an analytic solution to the actual problem it would not be much point in spending a lot of effort in trying to

solve it using finite differences. It is however important that changing from the actual problem to the related problem does not involve major modifications to the core of the finite difference code. The reason for this is that such major modifications are their own source of error and as a consequence our test will not be very convincing as a test of the actual problem.

So what is the actual problem then? Here it is:

$$\begin{aligned}u_{tt} - u_{xx} &= 0 & -l < x < l \\u(-l, t) &= h(t) \\u(l, t) &= k(t) \\u(x, 0) &= \varphi(x) \\u_t(x, 0) &= \psi(x)\end{aligned}$$

a) Implement a finite difference method for this problem. The code should be flexible enough to handle arbitrary choices for the functions h, k, φ and ψ .

b) Show that

$$u(x, t) = \frac{1}{2}(\varphi(x+t) + \varphi(x-t))$$

is a solution to the problem

$$\begin{aligned}u_{tt} - u_{xx} &= 0 & -\infty < x < \infty \\u(x, 0) &= \varphi(x) \\u_t(x, 0) &= 0\end{aligned}$$

In your finite difference code choose l and φ in such a way that the interval is much larger than the support of φ . (By the support of φ I mean the set of points where the function is nonzero). A Gaussian function of the form

$$\varphi(x) = ae^{-b(x-x_0)^2}$$

is a good choice if the parameters a, b and x_0 are chosen in the right way. Also choose $h = k = 0$. As long as the support of the solution u is well away from the boundary of the interval $[0, l]$ our exact solution and the finite difference solution should match. This is our first test.

c) Our second test involves a rather general idea called an "artificial source". We consider the modified problem

$$\begin{aligned}u_{tt} - u_{xx} &= \rho(x, t) & -l < x < l \\u(-l, t) &= h(t) \\u(l, t) &= k(t) \\u(x, 0) &= \varphi(x) \\u_t(x, 0) &= \psi(x)\end{aligned}$$

This equation would appear as the evolution equation for a string clamped at both ends and where there is a prescribed force acting along the string. This

kind of prescribed influence is called a "source". In the artificial source test we turn the problem around and rather assume a solution of a particular type and then compute what the corresponding source must be.

We thus pick a function $u(x, t)$ and calculate the source $\rho(x, t)$, boundary functions $h(t), k(t)$ and initial data $\varphi(x)$ and $\psi(x)$ so that $u(x, t)$ is in fact a solution to our modified problem. Thus

$$\begin{aligned}\rho(x, t) &= u_{tt}(x, t) - u_{xx}(x, t) \\ \varphi(x) &= u(x, 0) \\ \psi(x) &= u_t(x, 0) \\ h(t) &= u(-l, t) \\ k(t) &= u(l, t)\end{aligned}$$

We now have a analytic solution to our modified problem and can compare this solution to the solution produced by our finite difference code. You are free to choose the function $u(x, t)$ for yourself but here is a possible family of choices

$$u(x, t) = ae^{-b(x-x_0 \cos(\omega t))^2}$$

This is a family of Gaussian functions whose center move back and forth in a periodic manner determined by the parameters x_0 and ω . Compare the numerical solution to the exact solution for several choices of parameters. Make sure that at least one of the test involve the boundary conditions in a significant way. By this I mean that $h(t)$ and $k(t)$ should be nonzero for some values of t .

d) Now use you code to solve the problem

$$\begin{aligned}u_{tt} - u_{xx} &= 0 & -l < x < l \\ u(-l, t) &= \sin(\omega t) \\ u(l, t) &= 0 \\ u(x, 0) &= 0 \\ u_t(x, 0) &= 0\end{aligned}$$

This model what happend when you periodically move the left end of the string, whose right end is clamped, up and down. What do you expect to happen? Do your code verify your intuition?

12.2 Project2

In this project the aim is to develop a stable numerical scheme for the initial value problem

$$\begin{aligned}au_t + bu_x &= \rho, & -L < x < L, \quad 0 < t < T \\ u(-L, t) &= 0 \\ u(L, t) &= 0 \\ u(x, 0) &= \varphi(x)\end{aligned}\tag{799}$$

where

$$\begin{aligned}a &= a(x, t) \\ b &= b(x, t, u) \\ \rho &= \rho(x, t)\end{aligned}$$

Remark: Since the equation is first order in time, one can actually not pose boundary conditions at both endpoints. Thus we should only use the condition at one end point, for example $u(-L, t) = 0$. However since we are using center difference we really also need a boundary condition at $x = L$. In this problem I ask you to use the condition $u(-L, t) = 0$. Using these two conditions you can construct a finite difference scheme using center difference in space. However, and this is of outmost importance, you must make sure that the solution does not become nonzero in the region close to the boundaries, if it does the scheme will go unstable no matter what we do. But this restriction is really ok since we really are trying to simulate the dynamics on an infinite domain, and thus the computations boundaries at $x = -L, L$ should be "invisible" at all time.

- a) Derive a numerical scheme for equation (799) using center differences for the time and space derivatives. Find the Von-Neuman condition for numerical stability of your scheme for the case when $a = a_0, b = b_0$ are constants and $\rho = 0$.
- b) Verify the numerical stability condition by running the your numerical scheme for some initial condition of your choosing. Run with grid parameters both in the Von Neumann stable and unstable regime Since our numerical scheme is meant to be a good approximation to solutions of equation (799) on an infinite interval, you must make sure that the initial condition is very small or zero close to the end points of the interval $[-L, L]$. A well localized Gaussian function will do the job.

- c) Solve the initial value problem

$$\begin{aligned}(3x^2 + 1)u_t + 2tu_x &= 0, \quad -\infty < x < \infty \\ u(x, 0) &= e^{-\gamma x^2}\end{aligned}$$

using the method of characteristics. Plot some of the base characteristic curves in the $x - t$ plane and also plot the solution $u(x, t)$ as a function of x for several $t > 0$.

- d) Solve the initial value problem from c) using your finite difference scheme. Compare with the solution exact solution from c) by plotting them on top of each other for several times.
- e) Let the following quasi linear initial value problem be given

$$\begin{aligned}u_t + uu_x &= 0, \quad -\infty < x < \infty \\ u(x, 0) &= \varphi(x)\end{aligned}$$

Find exact formulas for the shock time for solutions to the initial value problem from e) corresponding to the initial data

i)

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

ii)

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

iii)

$$\varphi(x) = \text{sech}(x)$$

Test the validity of your finite difference code, and exact formulas for the shock time, by running your code with the given initial conditions and verifying that the graph of the numerical solutions become approximately vertical for some x as the shock time is approached.

12.3 Project3

In this project we are going to solve a heat equation numerically using the finite difference method and the finite Fourier transform.

$$\begin{aligned} u_t - u_{xx} &= \rho(x, t), \quad 0 < x < l, \quad t > 0 \\ u(0, t) &= f(t) \\ u(l, t) &= g(t) \\ u(x, 0) &= \varphi(x) \end{aligned} \tag{800}$$

a) Implement a finite difference method for this problem. The code should be flexible enough to handle arbitrary choices for the functions f , g , ρ and φ .

b) The function

$$u_e(x, t) = \frac{1}{l} \left(x h(x, t) \left(\frac{g(t)}{h(l, t)} \right) + (l - x) h(x, t) \left(\frac{f(t)}{h(0, t)} \right) \right)$$

satisfy the boundary conditions. Calculate the corresponding artificial source ρ and initial condition φ . Now run your finite difference code with the calculated ρ and φ and compare the numerical solution to the exact solution $u_e(x, t)$. Do this with a couple of choices for the functions $f(t)$, $g(t)$ and $h(x, t)$ and for several values of the parameters in the problem. For each choice plot the numerical and exact solution in the same plot and thereby verify that they coincide.

- c) Solve (800) using the finite Fourier transform. Before you apply the finite Fourier transform you should convert (800) into a problem with homogeneous boundary conditions like in equation (434) in the lecture notes. Use the same choices for $f(t)$ and $g(t)$ that you used in problem b) and choose your own source. Compare the finite Fourier transform solution with the finite difference solution for some choices of initial conditions. Make sure that the initial conditions is consistent with the boundary conditions

$$\begin{aligned}\varphi(0) &= f(0) \\ \varphi(l) &= g(0)\end{aligned}$$

- d) Let the source in problem (800) be the nonlinear source from equation (445) in the lecture notes.

$$\rho(x, t) = \hat{\lambda}u(x, t)(1 - u^2(x, t))$$

Use the finite difference code to verify the linear and nonlinear stability results found in section 4.7 of our textbook. In particular you should shown that

- i) $u(x, t) = 0$ is stable for $\hat{\lambda} < 1$
- ii) $u(x, t)$ is unstable for $1 < \hat{\lambda} < 4$ and approach the solution 4.7.24 as time increase to infinity.

Run the finite difference code for values of $\hat{\lambda}$ higher than 4 and plot the resulting solutions. Any conclusions?

References

- [1] “Partial Differential Equations of Applied Mathematics”, Erich Sauderer, 2006, Wiley.
- [2] “A first Course in Partial Differential Equations”, H. F. Weinberger, 1965, John Wiley and Sons.
- [3] “Introduction to Partial Differential Equations”, Gerald B. Folland, 1976, Princeton University Press.
- [4] “Partial Differential Equations, An Introduction”, Walter A. Strauss, 2007, John Wiley and Sons.
- [5] “Methods of mathematical Physics, Volume II”, Courant and Hilbert, 1962, Interscience Publishers.
- [6] “Partial Differential Equations”, Fritz John, 1982, Springer.
- [7] “Methods of Applied Mathematics”, James P. Keener, 1988, Springer.
- [8] “Green’s functions and Boundary value problems”, Ivar Stakgold and Michael J. Holst, 2011, Wiley.