MA30044

Mathematical Methods I

Department of Mathematical Sciences, University of Bath

Dr. Philippe Trinh

Thursday 14th October, 2021, 08:29

Contents

Chapter 1

Motivation and Overview

The aim of this course is to teach several advanced mathematical methods, which have a wide range of applications in in science and engineering. Although we will prove key theorems, the focus will be on **applying** those methods and we will see how they work for several **examples**. In the following sections we will give an overview of the covered topics and their applications.

Sturm Liouville theory (Chapter 2)

In linear algebra you might have come across **eigenvalue problems** and methods for **solving linear systems**. Sturm Liouville theory can be seen as an extension of those concepts to infinite dimensional function spaces. The methods developed in this chapter can be used for example to describe the eigenmodes of a vibrating plate. If this plate is covered with sand which collects in the stationary regions, this gives rise to the famous Chladni Figures (Fig. 1.1, left). Harmonic analysis also underpins much of acoustics, and can be used to describe and optimise the fundamental vibrations of musical instruments (Fig. 1.1, right).

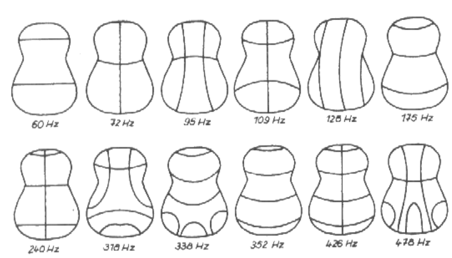


Figure 1.1: Chladni Figures on a circular disk (left[a]) and eigenmodes of a guitar (right[b]).

[a]Source: wikimedia commons, https://commons.wikimedia.org/wiki/File:Round_Chladni_plate_with_4_circular_nodes.JPG for license terms.

[b]Source: wikimedia commons, https://commons.wikimedia.org/wiki/File:Chladni_guitar.png for license terms.

The **self-adjoint** differential operators which arise in Sturm Liouville theory have special properties (for example, their eigenvalues are real and the eigenfunctions of different eigenvalues are orthogonal, as in the finite dimensional case, which is briefly reviewed in Appendix C). Those self-adjoint differential operators were crucial in the development of quantum mechanics in the beginning of the 20th century. For bound states those operators have discrete eigenvalues, which is related to the physical concepts of "quantisation". The key mathematical object encountered in quantum mechanics is the *wave function* (see for example Fig. 1.2 (left) which shows the wave function of a harmonic oscillator). This is the solution of the celebrated *Schrödinger equation*, which is another example of a Sturm-Liouville operator. The eigenmodes can then be interpreted as physical states, such as the different configurations of the hydrogen atom (Fig. 1.2, right). This insight was crucial for the development of modern chemistry and physics.

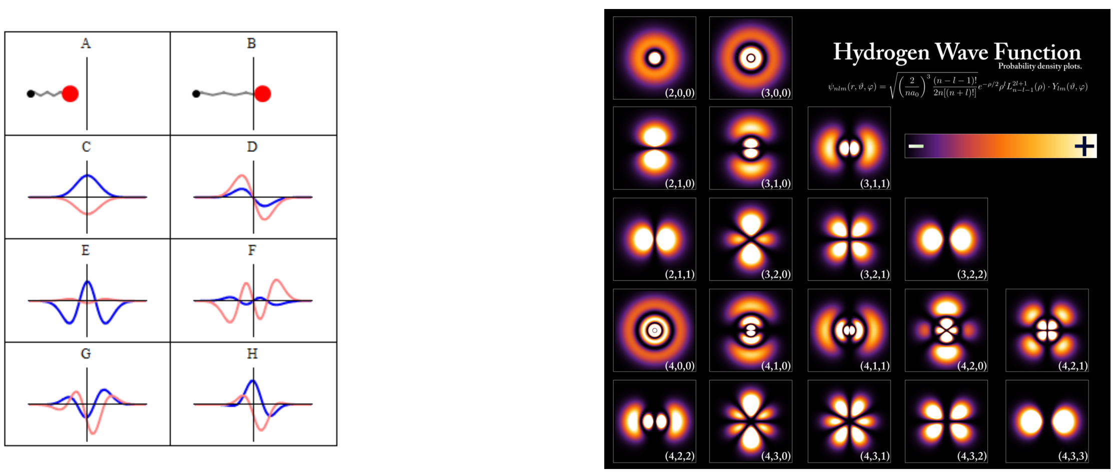


Figure 1.2: Quantum mechanics: harmonic oscillator (left[a]) and hydrogen atom (right[b])

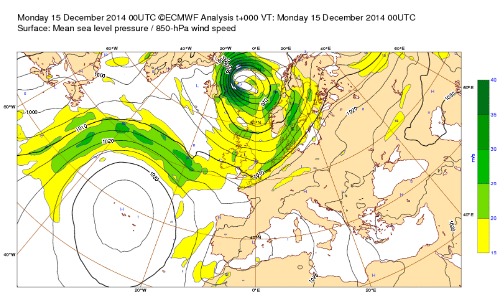
[a]Source: wikimedia commons, see https://commons.wikimedia.org/wiki/File:QuantumHarmonicOscillatorAnimation.gif for license terms.

[b]Source: wikimedia commons, see https://commons.wikimedia.org/wiki/File:Hydrogen_Density_Plots.png for license terms.

Another important application which we will encounter in this chapter is the **approximation of functions** in terms of simple basis functions (which are the solutions of Sturm-Liouville eigenvalue problems). This is often important for the fast evaluation of functions on a computer and for the practical solution of problems in physics, especially if those have to be solved in special geometries. We will apply this technique to approximate the electric field of an antenna.

The Fourier Transform (Chapter 3)

You might already have come across the Fourier series expansion and this chapter extends this to more general, non-periodic functions. We will study the Fourier transforms of several important functions and introduce techniques for the easy calculation of others. Fourier transforms can be used to solve differential equations. The key idea here is to **tranform the problem** into a different form, which can be solved with known techniques. It turns out that for some PDEs with constant coefficients, the problem reduces to solving an algebraic equation, i.e. finding the roots of a polynomial. This observation has important practical applications. For example, the European Centre for Medium Weather Forecasting (ECMWF) uses the Fourier-transform to solve certain long range correlations in the atmosphere; an example for a forecast and the related grid can be see in Fig 1.3. Computationally this is possible using the *Fast Fourier Transform* method by Cooley and Tuckey, which is one of the most celebrated mathematical algorithms of the 20th century.



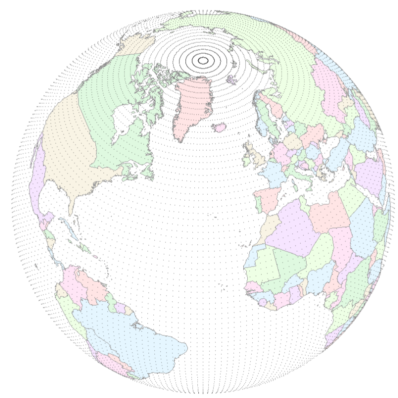


Figure 1.3: ECMWF weather forecast (left[a]) and spectral grid (right[b])

[a]Source: flickr https://www.flickr.com/photos/ecmwf/13384533105

[b]Source: wikimedia commons, see https://commons.wikimedia.org/wiki/File:NCEP_T62_gaussian_grid.png for license terms.

In this chapter we will also introduce the concept of **distributions**, which can be used to generalise functions. Although they might look strange (for example, because the value of a distribution can be infinite), they are very useful for mathematically describing physical concepts such as point charges.

Quasilinear first order PDEs (Chapter 4)

The PDEs studied in this chapter arise in many physical problems, for example in the formation of **shocks** in a fluid (see Fig. 1.4). We will show how they can be solved using contour lines and the **method of characteristics** and apply those techniques to Burgers' equation, which models the non-linear advection of a fluid. The key idea is to deduce the characteristics by solving a set of ODEs to obtain a family of curves which covers the entire domain of the PDE. The value at a particular point in space can then be deduced by following those curves back to a **data curve** on which the initial conditions are prescribed.

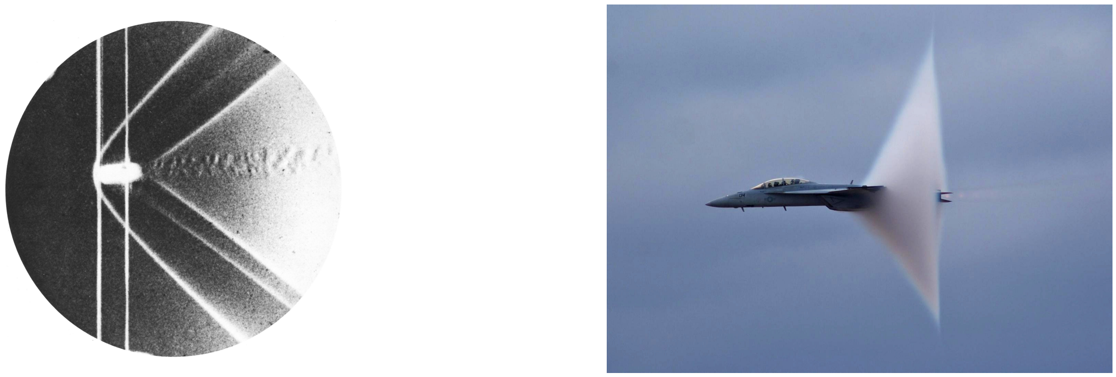


Figure 1.4: Examples of shock waves: supersonic bullet (left[a]) and jet (right[b])

[a]Source: wikimedia commons, see https://commons.wikimedia.org/wiki/File:Photography_of_bow_shock_waves_around_a_brass_bullet,_1888.jpg for license terms.

[b]Source: wikimedia commons, see https://commons.wikimedia.org/wiki/File:FA-18_going_transonic.JPG for license terms.

Second order hyperbolic PDEs (Chapter 5)

The brief final chapter will focus on an important subset of second order PDEs. Equations of this form arise for example when studying the propagation of water waves; Fig. 1.5 (left) shows a Tsunami model. There is a plethora of other physical phenomena which can be described by wave propagation. Recently there has been lots of interest in so-called inverse problems which require a thorough understanding of wave propagation in heterogeneous media. A classical example is a hunting bat (Fig. 1.5, right) which uses ultrasound to find its prey. By emitting a series of high-pitched cries and observing the reflected sound it can deduce the location of objects and flying insects. A similar technique is used to locate oil- and gas-fields under the sea bed. For this the propagation of pressure waves induced by small detonation charges through the underlying rock is measured by a set of observers. Understanding how wave propagation depends on geological features allow interpretation of this data.

Acknowledgements

These lecture notes have been inherited and modified from previous lecturers Chris Budd, Adrian Hill, Eike Mueller, and Kirill Cherednichenko.

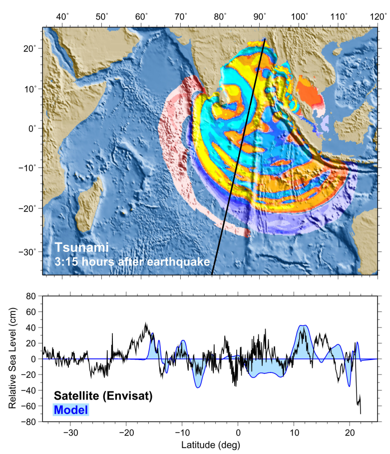


Figure 1.5: Wave propagation: Tsunami model (left[a]) and hunting bat (right[b])

[a]Source: NOAA news, http://www.noaanews.noaa.gov/stories2005/s2365.htm
[b]Source: wikipedia https://en.wikipedia.org/wiki/Bat#/media/File:Big-eared-townsend-fledermaus.jpg

Chapter 2

Sturm Liouville Theory

2021-22: This year, we will present a more applied treatment of Sturm-Liouville theory. See the accompanying video lectures. The basis of this treament will follow Chap. 5 from Richard Haberman's book, "Applied partial differential equations with Fourier series and boundary value problems".

Aims:

1. To introduce the basic theory of Sturm Liouville equations
2. Learn about orthogonal and orthonormal series
3. Solve Sturm Liouville eigenvalue problems with applications in electrostatics and quantum mechanics
4. Approximate functions with simpler basis functions
5. Introduce Legendre polynomials and generating functions
6. Solve forced ODEs and linear PDEs

Removed in 2021–22: definition of sq. integrable, definition of Wronskian, definition of positive definite, positive semi-definite, various lemmas regarding being positive definite and positive semi-definite.

2.1 Introduction

Previously in your study of partial differential equations, you encountered the heat equation, which models the heat distribution using a function, say $u(x,t)$. In a more general form, the heat equation along a line is

$$c\rho\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(K_0\frac{\partial u}{\partial x}\right) + Q. \tag{2.1}$$

You will likely remember the much simpler case where c, ρ, and K_0 are constant, and there are no additional heat sources, $Q \equiv 0$. In this case, you would obtain

$$\frac{\partial u}{\partial t} = \kappa\frac{\partial^2 u}{\partial x^2}, \tag{2.2}$$

with κ now constant. You would have treated the above equation using separation of variables, and then derived Fourier series solutions.

Let us consider what happens if we study the more general (2.1) in situations where the material properties may involve variable $c = c(x)$, $\rho = \rho(x)$, and $K_0 = K_0(x)$. We also assume that there is a source of heat $Q = \alpha u$. You do not have to worry about the physics, and focus purely on the analysis of this problem:

$$c(x)\rho(x)\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}\right) + \alpha u. \tag{2.3}$$

We now try to do separation of variables. Set

$$u(x,t) = h(t)\phi(x).$$

The PDE (2.3) now becomes after separating functions of t from functions of x:

$$\frac{h'}{h} = \frac{1}{c\rho\phi}\left(\frac{\mathrm{d}}{\mathrm{d}x}(K_0\phi') + \alpha\right) = -\lambda.$$

We have set the constant on the right hand-side to $-\lambda$ because of the anticipated solution.

If we first solve the differential equation for $h(t)$, we find that

$$\frac{h'}{h} = -\lambda \Longrightarrow h(t) = A\mathrm{e}^{-\lambda t},$$

where A is constant. Generally, based on physical principles, we expect that $\lambda > 0$ so that solutions do not grow exponentially in time. However, with a source of heat, $Q = \alpha u$, it may be possible for $\lambda < 0$. We must depend on boundary conditions in order to decide upon the possibilities for λ.

Our focus is rather on the equation for $\phi(x)$. We must seek to solve

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(K_0\frac{\mathrm{d}\phi}{\mathrm{d}x}\right) + \alpha\phi + \lambda c\rho\phi = 0. \tag{2.4}$$

This is not an easy equation to solve, now that the coefficients, K_0, α, c, and ρ are all potentially functions of x. In general, we would need to depend on numerical solutions of this problem.

2.2 Sturm-Liouville problems

The previous section provides one of many examples where it is important to study differential equations of the following form.

Definition 2.1 (Sturm-Liouville equation)**.**

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)+q(x)\phi+\lambda\sigma(x)\phi=0. \tag{2.5}$$

The equation (2.5) is known as a **Sturm-Liouville equation** or **SL equation**. We call λ the **eigenvalue**, ϕ the **eigenfunction**, and (λ,ϕ) an **eigenvalue-eigenfunction** pair and $\sigma(x)$ the **weight function**.

Note: the eigenfunctions are not unique, i.e. if $(\lambda,\phi(x))$ is and eigenvalue/eigenfunction pair and $C\neq 0$ is an arbitrary constant then $(\lambda, C\phi(x))$ is also an eigenvalue/eigenfunction pair.

A seemingly more general 2nd order ODE is given by

$$\alpha(x)\frac{d^2u}{dx^2}+\beta(x)\frac{du}{dx}+\gamma(x)u+\lambda\delta(x)u=0 \tag{2.6}$$

However, the following lemma shows how (2.6) can be re-written as a SL-equation under certain conditions.

Lemma 2.1 (Transformation to Sturm-Liouville form)**.** *If $\alpha(x)\neq 0$ then the 2nd order ODE in (**??**) can be written as a Sturm-Liouville equation in the form (2.5) with*

$$p(x)=\exp\left(\int_{x_0}^{x}\frac{\beta(t)}{\alpha(t)}dt\right),\qquad q(x)=p(x)\frac{\gamma(x)}{\alpha(x)},\qquad r(x)=p(x)\frac{\delta(x)}{\alpha(x)} \tag{2.7}$$

where x_0 is an arbitrary number.

Proof. Since $\alpha(x)\neq 0$ we can divide (2.6) by α to obtain

$$u''+\frac{\beta(x)}{\alpha(x)}u'+\frac{\gamma(x)}{\alpha(x)}u+\lambda\frac{\delta(x)}{\alpha(x)}u=0 \tag{2.8}$$

We apply the following trick. Notice that by the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[u'\mathrm{e}^{\int_{x_0}^{x}\frac{\beta(t)}{\alpha(t)}\,\mathrm{d}t}\right]=\mathrm{e}^{\int_{x_0}^{x}\frac{\beta(t)}{\alpha(t)}\,\mathrm{d}t}\left[u''+\frac{\beta(x)}{\alpha(x)}u'\right]. \tag{2.9}$$

So we simply multiply the entire equation (2.8) by the exponential factor above and re-write the first two terms to give

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)u'\right] + \underbrace{\left[p(x)\frac{\gamma(x)}{\alpha(x)}\right]}_{q(x)} u + \lambda \underbrace{\left[p(x)\frac{\delta(x)}{\alpha(x)}\right]}_{\sigma(x)} u = 0. \tag{2.10}$$

where

$$p(x) = \mathrm{e}^{\int_{x_0}^{x} \frac{\beta(t)}{\alpha(t)}\,\mathrm{d}t}.$$

We see that this now gives (2.5) with $u = \phi$. □

Here is an example of the above procedure.

Example 2.1. Let u satisfy

$$xu'' + \frac{1}{2}u' + x^2u + \lambda x^3 u = 0, \qquad \text{for } x \in (1, 2).$$

Divide by $x \neq 0$ to obtain

$$u'' + \frac{1}{2}u' + xu + \lambda x^2 u = 0.$$

As in the proof of the Lemma, we seek to write the first two terms as a single derivative. This is done by multiplying the equation by the integrating factor,

$$p(x) = \exp\left(\int_1^x \frac{dy}{2y}\right) = \exp\left(\frac{1}{2}\log(x)\right) = x^{1/2}.$$

Then the SL form of the equation is

$$\left(x^{1/2}u'\right)' + x^{3/2}u + \lambda x^{5/2}u = 0.$$

A great deal is known about the simplest case of Sturm-Liouville problems. We will thus define the following:

Definition 2.2 (Regular Sturm-Liouville Eigenvalue Problem)**.** A **regular** Sturm-Liouville eigenvalue problem consists of the Sturm-Liouville differential equation,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}\phi}{\mathrm{d}x}\right) + q(x)\phi + \lambda\sigma(x)\phi = 0, \tag{2.11}$$

subject to boundary conditions

$$\mathrm{BC}_1 \qquad \beta_1\phi(a) + \beta_2\phi'(a) = 0 \tag{2.12}$$

$$\mathrm{BC}_2 \qquad \beta_3\phi(b) + \beta_4\phi'(b) = 0. \tag{2.13}$$

All coefficients β_i are assumed to be real.

In addition, to be called **regular**, p, q, and σ are real and continuous everywhere (including the endpoints) and $p > 0$ and $\sigma > 0$ everywhere (including the endpoints).

Following Haberman (p.157), we cite all the theorems in one location.

Theorem 2.1 (Theorems on regular Sturm-Liouville eigenvalue problems)**.** *For the case of a regular SL problem:*

1. *All eigenvalues λ are real.*
2. *There exists an infinite number of eigenvalues:*

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

   a. There is a smallest eigenvalue, usually denoted λ_1.
   b. There is not a largest eigenvalue, and $\lambda_n \to \infty$ as $n \to \infty$.

3. *Corresponding to each eigenvalue, λ_n, there is an* ***eigenfunction****, denoted $\phi_n(x)$, which is unique to a multiplicative constant.*

   In addition ϕ_n has exactly $n-1$ zeros for $a < x < b$.

4. *The eigenfunctions form a* ***complete set****, meaning that any piecewise smooth function, $f(x)$, can be represented by an eigenfunction series of the form*

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

   The infinite series converges to $[f(x^+) + f(x^-)]/2$ if the coefficients a_n are properly chosen.

5. *Eigenfunctions belonging to different eigenvalues are* ***orthogonal relative to the weight function*** *$\sigma(x)$. In other words,*

$$\int_a^b \phi_n(x)\phi_m(x)\sigma(x)\,\mathrm{d}x = 0 \qquad \text{if } \lambda_n \neq \lambda_m.$$

6. *Any eigenvalue can be related to its eigenfunction by the* ***Rayleigh quotient:***

$$\lambda = \frac{-p\phi\phi'\Big|_a^b + \int_a^b [p((\phi')^2 - q\phi^2]\,\mathrm{d}x}{\int_a^b \phi^2\sigma\,\mathrm{d}x}, \tag{2.14}$$

   where the boundary conditions may simplify this expression.

Some, but not all, of these theorems will be proven and addressed in the sections to follow.

2.2.1 Example of simple harmonic motion

Example 2.2 (Demonstration of properties of SL theorem)**.** Let us do an example to demonstrate all of the items of the above theorem. Consider the following regular SL problem which arises in the simple study of the heat equation (2.2) with fixed end conditions:

$$\phi'' + \lambda\phi = 0, \tag{2.15a}$$

$$\phi(0) = 0, \tag{2.15b}$$

$$\phi(L) = 0. \tag{2.15c}$$

For $\lambda > 0$ equation (2.15) has the solution

$$\phi(x) = A\cos\left(\lambda^{1/2}x\right) + B\sin\left(\lambda^{1/2}x\right), \qquad \text{with arbitrary } A, B \in \mathbb{R}. \tag{2.16}$$

The left boundary condition $\phi(0) = 0$ implies $A = 0$.

Then $\phi(\pi) = B\sin\left(\lambda^{1/2}\pi\right)$, and the right boundary condition implies either $B = 0$ (which corresponds to the trivial case of $\phi = 0$) or $\sin\left(\lambda^{1/2}L\right) = 0$. The zeros of the sin-function can be found at $n\pi$ with $n \in \mathbb{Z}$. Therefore $\lambda^{1/2}\pi = n\pi/L$ and

$$\lambda = \left(\frac{n\pi}{L}\right)^2 \qquad \text{for } n = 1, 2, \ldots \tag{2.17}$$

(note that the case $n = 0$ would correspond to a zero solution again). Therefore, there is an infinite, countable number of eigenvalue, eigenfunction pairs (see Fig. 2.1).

$$(\lambda_n, \phi_n) = \left(\left[\frac{n\pi}{L}\right]^2, \sin\left[\frac{n^2\pi^2}{L^2}x\right]\right), \qquad \text{with } n \in \mathbb{N}. \tag{2.18}$$

Watch the video (Applied Sturm-Liouville Problems Part 2) to see a demonstration of each of the properties of the theorem.

Student exercise. Show that there is no solution of (2.15) if $\lambda \leq 0$.

Note: The example (2.15) represents the equation for the eigenfunctions for phenomena beyond heat flow (essentially many systems connected with the Laplacian, $\nabla\phi$). For example, the analysis of one-dimensional waves on a string, $u_{tt} = c^2 u_{xx}$ also produces the same SL equation.

The example equally represents the quantum motion of a particle in a one-dimensional, infinitely deep box of length L. The steady-state Schrödinger equation in this case is

$$-\frac{\hbar^2}{2m}\phi'' = E\phi,$$

where m is the mass of the particle and E is its energy with $\lambda = \frac{2mE}{\hbar^2}$. The number $\hbar = 6.63 \cdot 10^{-34}\,\mathrm{m^2\,kg\,s^{-1}}$ is known as Planck's constant. The wave function ϕ has to be zero at the boundaries $x = 0$ and $x = L$ (if the wave function extended beyond this interval, the particle would have an infinite energy since we assumed that the walls are infinitely high). For simplicity, consider $L = \pi$. Then the calculation of eigenvalues shows that the energy of a particle confined to a box is quantized, with $E_n = \frac{\hbar^2}{2m}n^2$.

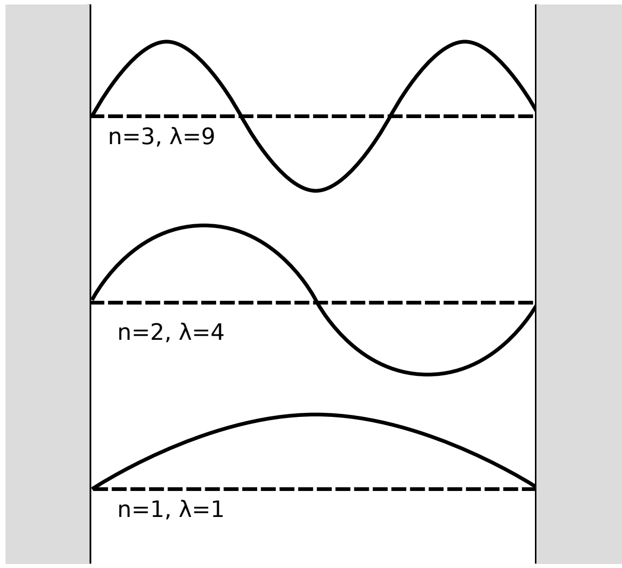


Figure 2.1: Eigenfunctions and eigenvalues for the SL system (**??**).

2.3 Self-adjoint operators, Green's identity, and orthogonality

Let us examine the regular SL problem:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}\phi}{\mathrm{d}x}\right) + q(x)\phi + \lambda\sigma(x)\phi = 0, \tag{2.19}$$

subject to boundary conditions

$$\begin{aligned} &\mathrm{BC}_1 \quad \beta_1\phi(a) + \beta_2\phi'(a) = 0, \\ &\mathrm{BC}_2 \quad \beta_3\phi(b) + \beta_4\phi'(b) = 0. \end{aligned} \tag{2.20}$$

It is time to set up some useful notation and establish certain identities.

Linear operators

First, it is convenient for us to define a linear differential operator for the SL equation so that we do not have to write it out in full.

Definition 2.3 (Linear operator)**.** Define the operator L so that

$$L\phi = \frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}\phi}{\mathrm{d}x}\right) + q(x)\phi.$$

Consequently the SL equation becomes

$$\boxed{L\phi + \lambda\sigma(x)\phi = 0.}$$

Many of the proofs of Theorem 2.1 are immediate consequences of a formula known as **Lagrange's identity**.

Lemma 2.2 (Lagrange's identity)**.** *Let u and v be any two functions (not necessarily eigenfunctions). Then*

$$\boxed{uL(v) - vL(u) = \frac{\mathrm{d}}{\mathrm{d}x}\left[p\big(uv' - vu'\big)\right].} \tag{2.21}$$

Proof. Using the definition of L, we have for the LHS,

$$\begin{aligned} uL(v) - vL(u) &= u\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}v}{\mathrm{d}x}\right) + q(x)v\right] - v\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right) + q(x)u\right] \\ &= u\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)v'\right)\right] - v\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)u'\right)\right] \end{aligned}$$

We can expand the above derivatives and ensure they agree with the expanded derivatives of the right hand-side of (2.21). ☐

If we integrate Lagrange's identity on $\int_a^b$, we obtain Green's identity.

Lemma 2.3 (Green's identity)**.** *Let u and v be any two functions (not necessarily eigenfunctions). Then*

$$\boxed{\int_a^b \left[uL(v) - vL(u)\right]\mathrm{d}x = \left[p\big(uv' - vu'\big)\right]_a^b.} \tag{2.22}$$

It goes without saying that both Green's identity and Lagrange's identity requires that u and v be sufficiently well behaved so that the quantities are defined (e.g. they are continuous and differentiable).

Self-adjointness

Green's identity works for general functions u and v where the integrals are defined. If, in addition, u and v satisfy boundary conditions at $x = a$ and $x = b$ such that

$$\big[p\big(uv' - vu'\big)\big]_a^b = 0, \tag{2.23}$$

then notice that Green's identity gives $\int_a^b [uL(v) - vL(u)]\,\mathrm{d}x = 0$.

In fact, for the case of regular SL problems, this is exactly what occurs if u and v are chosen as functions that satisfy BC_1 and BC_2 in (2.20).

Lemma 2.4 (Regular SL problem is self-adjoint)**.** *If u and v are any two functions satisfying the same set of homogeneous boundary conditions (of the regular Sturm-Liouville type), then*

$$\int_a^b [uL(v) - vL(u)]\mathrm{d}x = 0. \tag{2.24}$$

When (2.24) is valid, we say that the operator L, with corresponding boundary conditions, is ***self-adjoint***[1]*.*

Proof. From (2.22) we have for the right hand-side,

$$\text{RHS} = p(b)\big(u(b)v'(b) - v(b)u'(b)\big) - p(a)\big(u(a)v'(a) - v(a)u'(a)\big)$$

It will be sufficient to examine the first grouping of terms. We know from BC_2 that both u and v satisfy

$$\beta_3 u(b) + \beta_4 u'(b) = 0$$
$$\beta_3 v(b) + \beta_4 v'(b) = 0,$$

which can instead be written in the matrix form

$$\begin{pmatrix} u(b) & u'(b) \\ v(b) & v'(b) \end{pmatrix} \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix} = 0.$$

By assumption, not both β_3 and β_4 are zero, and so the determinant of the matrix must be zero. Hence

$$W = u(b)v'(b) - v(b)u'(b) = 0.$$

In fact, there is a special name given to the quantity above: it is known as the **Wronskian**, W. You may encounter it in various courses, including this one.

In any case, by an analogous argument, we see that $u(a)v'(a) - v(a)u'(a) = 0$. Hence the boundary terms are equal to zero, confirming (2.23).

□

Note: the boundary terms (2.23) can also vanish in circumstances other than of the regular SL type BC_1 and BC_2. Therefore, the simplicity of (2.24) may also be true for more general boundary conditions and problems beyond the regular SL type.

2.3.1 Orthogonality

We have introduced in previous courses the notion of orthogonality, which is defined in Theorem 2.1 for two functions ϕ_n and ϕ_m on the domain $[a, b]$ as

$$\int_a^b \sigma \phi_n \phi_m \,\mathrm{d}x = 0.$$

There is a useful notation that we can set up for this.

Definition 2.4 (Inner product)**.** For two complex-valued functions u and v, the inner product is

$$\boxed{(u, v) = \int_a^b u\overline{v}\ \mathrm{d}x.} \tag{2.25}$$

Note that for real-valued functions, this is simply

$$\boxed{(u, v) = \int_a^b uv\ \mathrm{d}x.} \tag{2.26}$$

Similarly,

Definition 2.5 (Weighted Inner product)**.** For two complex-valued functions u and v, the **σ-weighted inner product** is

$$\boxed{\langle u, v\rangle_\sigma = \int_a^b \sigma u\overline{v}\ \mathrm{d}x,} \tag{2.27}$$

and is defined for $\sigma > 0$.

Therefore Result 5 from Theorem 2.1 says that $\langle \phi_n, \phi_m\rangle_\sigma = 0$. This is simply a more convenient notation.

2.4 Proofs of Result 5 and 1 from Theorem 2.1

The previous section sets up the notation and main results that we need in order to prove a number of results from Theorem 2.1. The proofs of these results range in difficulty and here we shall prove only two.

2.4.1 Real eigenvalues (Result 1)

Let us prove that the eigenvalues must be real (Result 1).

Proof. First, let us assume that (λ, ϕ) is an eigenvalue-eigenfunction pair and that λ and ϕ are allowed to be complex-valued. We shall demonstrate that $(\overline{\lambda}, \overline{\phi})$ are also eigenvalue-eigenfunction pairs. Here, the overline corresponds to the complex conjugate. Thus for instance, $\overline{z} = \overline{x + \mathrm{i}y} = x - \mathrm{i}y$.

We take the complex conjugate of both sides of the differential equation:

$$\overline{L(\phi) + \lambda\sigma\phi} = \overline{L(\phi)} + \overline{\lambda\sigma\phi} = 0.$$

However, $\overline{L(\phi)} = L(\overline{\phi})$ since the coefficients of the differential equation are all real. Also, $\overline{\sigma} = \sigma$ since σ is real-valued. Hence we are left with:

$$L(\overline{\phi}) + \overline{\lambda}\sigma\overline{\phi} = 0.$$

We can also take the complex conjugate of BC_1 and BC_2, and since all the coefficients β_i are real, we also conclude that $\overline{\phi}$ correctly satisfies the necessary boundary conditions.

Therefore, we have concluded that if (λ, ϕ) is an eigenvalue-eigenfunction pair, then so too must be $(\overline{\lambda}, \overline{\phi})$. However, we can now apply the orthogonality relationship (2.30) to conclude that

$$(\lambda - \overline{\lambda}) \int_a^b \phi\overline{\phi}\sigma \, \mathrm{d}x = 0.$$

But the quantity $\phi\overline{\phi} = |\phi|^2 \geq 0$ and $\sigma > 0$. So the expression above is ≥ 0. Then it must be the case that $\lambda = \overline{\lambda}$ and this is equivalent to the fact that λ is real. □

2.4.2 Orthogonal eigenfunctions (Result 5)

Let (λ_n, ϕ_n) and (λ_m, ϕ_m) be two eigenvalue-eigenfunction pairs. We wish to prove Result 5 from Theorem 2.1. That is, eigenfunctions corresponding to different eigenvalues are orthogonal with weight $\sigma(x)$.

Proof. By Green's identity (2.22), we have that

$$\int_a^b [\phi_m L(\phi_n) - \phi_n L(\phi_m)] \, \mathrm{d}x = p(x) \Big[\phi_m \phi_n' - \phi_n \phi_m'\Big]_a^b. \tag{2.28}$$

Since ϕ_n and ϕ_m satisfy the SL equation, $L\phi = -\lambda\sigma\phi$, we can then replace the left hand-side with

$$(\lambda_m - \lambda_n) \int_a^b \phi_n \phi_m \sigma \,\mathrm{d}x = p(x) \Big[\phi_m \phi_n' - \phi_n \phi_m' \Big]_a^b. \tag{2.29}$$

However, since ϕ_n and ϕ_m satisfy BC_1 and BC_2, then we have established via (2.24) that the boundary terms are zero. Thus

$$(\lambda_m - \lambda_n) \int_a^b \phi_n \phi_m \sigma \,\mathrm{d}x = 0. \tag{2.30}$$

If $\lambda_m \neq \lambda_n$ then we are left with the orthogonality relationship,

$$\boxed{\langle \phi_n, \phi_m \rangle_\sigma = 0.}$$

Note that we have already established that the eigenvalues are real (Result 1) and therefore there is no need to worry about complex conjugation in the definition of $\langle \cdot, \cdot \rangle$.

Note: as we have mentioned previously, this orthogonality relationship (2.30) may be satisfied for problems with more general boundary conditions than those of regular SL problems—as long as the boundary terms contribute to zero. □

2.4.3 Rayleigh quotient (Result 6)

The Rayleigh quotient in (2.14) is proven in PS1.

2.5 Example of the damped harmonic oscillator

Let us go through another example of a regular Sturm-Liouville problem.

Example 2.3 (The damped harmonic oscillator)**.** Let

$$u'' + u' + \lambda u = 0, \qquad \text{with } u(0) = u(1) = 0. \tag{2.31}$$

This equation is not in SL form. We first use Lemma 2.1 to obtain

$$p(x) = \exp\left(\int_0^x \frac{1}{1} \, dy \right) = e^x,$$

and rewrite Eq. (2.31) as

$$(e^x u')' + \lambda e^x u = 0. \tag{2.32}$$

Notice that this is a **regular** SL problem and in particular, we have $p = e^x$ and $\sigma = e^x$ both positive in the domain $[0, 1]$. In this example, we will solve for the (λ_n, u_n) pairs, and then we will check certain results from Theorem 2.1.

2.5.1 Solve for eigenfunctions and eigenvalues

To solve the equation, we go back to (2.31) and make the ansatz $u(x) = Ce^{\mu x}$. Inserting this, we obtain

$$\left(\mu^2 + \mu + \lambda\right) Ce^{\mu x} = 0 \tag{2.33}$$

and hence

$$\mu^2 + \mu + \lambda = 0, \qquad \Rightarrow \mu_\pm = -\frac{1}{2} \pm \frac{\sqrt{1-4\lambda}}{2}. \tag{2.34}$$

If $\lambda < \frac{1}{4}$ both roots $\mu_\pm$ are different and real, and the general solution is

$$u(x) = C_+ e^{\mu_+ x} + C_- e^{\mu_- x}. \tag{2.35}$$

To satisfy the boundary conditions would require $C_+ + C_- = 0$ and $C_+ e^{\mu_+} + C_- e^{\mu_-} = 0$. It is easy to see that this is only possible for $C_+ = C_- = 0$ and hence there is no non-zero solution in this case. If $\lambda = \frac{1}{4}$ there is only one exponentially decaying solution $C_1 e^{-\frac{x}{2}}$ and one solution of the form $C_2 x e^{-\frac{x}{2}}$. Again, no linear combination $(C_1 + C_2 x)e^{-\frac{x}{2}}$ of those solutions solutions can fulfill the boundary conditions unless $C_1 = C_2 = 0$. We now assume $\lambda > \frac{1}{4}$ so that $\mu_\pm = -\frac{1}{2} \pm \frac{\sqrt{4\lambda-1}}{2} i = -\frac{1}{2} \pm \omega i$ where we defined the angular frequency

$$\omega \equiv \frac{\sqrt{4\lambda - 1}}{2}. \tag{2.36}$$

Hence the general solution is of the form

$$u(x) = e^{-\frac{x}{2}} \left(A\cos(\omega x) + B\sin(\omega x)\right) \tag{2.37}$$

The boundary conditions imply

$$u(0) = 0 \Rightarrow A = 0, \qquad u(1) = 0 \Rightarrow \sin(\omega) = 0 \Rightarrow \omega = n\pi, n \in \mathbb{Z}, n \neq 0. \tag{2.38}$$

Using the definition of ω in (2.36) and solving for λ we obtain

$$\frac{\sqrt{4\lambda - 1}}{2} = n\pi \Rightarrow \lambda = \frac{1}{4} + \pi^2 n^2. \tag{2.39}$$

The eigenvalue/eigenfunction pairs of the SL equation in (2.32) are therefore

$$(\lambda_n, \phi_n) = \left(\frac{1}{4} + \pi^2 n^2, e^{-\frac{x}{2}} \sin(n\pi x)\right) \qquad \text{for } n \in \mathbb{N}. \tag{2.40}$$

The lowest eigenfunctions are plotted in Fig. 2.2.

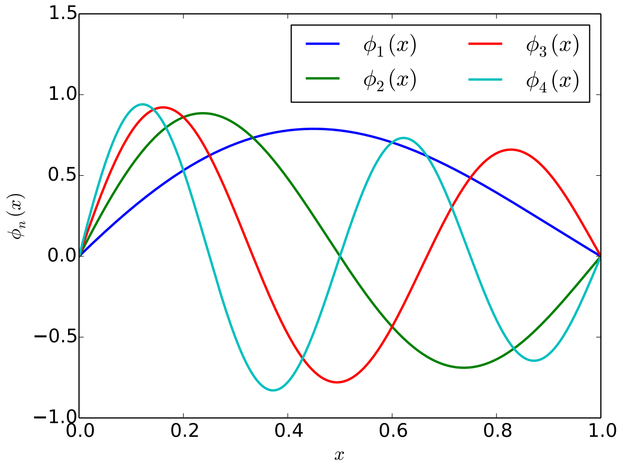


Figure 2.2: Eigenfunctions in (2.40).

2.5.2 Verification of properties of Theorem 2.1

(Result 1, 2) Notice that the eigenvalues are all real and form the increasing sequence as indicated. In this case, the eigenvalues are all positive as well.

(Result 5) We can also explicitly check those pairs for orthogonality

$$\begin{aligned} \langle \phi_n, \phi_m \rangle_\sigma &= \int_0^1 e^{-\frac{x}{2}} \sin(n\pi x) e^{-\frac{x}{2}} \sin(m\pi x) e^x \, dx \\ &= \int_0^1 \sin(n\pi x) \sin(m\pi x) \, dx = 0 \qquad \text{if } n \neq m \end{aligned} \tag{2.41}$$

as predicted by Theorem 2.1.

2.6 Approximation of functions

Eigenfunctions of SL operators are very useful as they give us the building block to build much more complex functions. This allows us to generalise Fourier series and solve very challenging problems. Before we discuss this concept, let us make a few definitions of orthogonal and orthonormal **sets**.

Definition 2.6 (Orthogonality)**.** Let $r(x) > 0$ for all $x \in (a, b)$. A set $\{\psi_n(x)\}$, $n \in \mathbb{N}$ of functions with $\psi_n \neq 0$ is **orthogonal** (with respect to the weight $r(x)$) if

$$\langle \psi_n, \psi_m \rangle_r = 0 \qquad \text{for } n \neq m. \tag{2.42}$$

There is a special kind of orthogonal set known as an orthonormal set.

Definition 2.7 (Orthonormality)**.** An orthogonal set $\{\phi_n\}$, $n \in \mathbb{N}$ of functions is **orthonormal** if in addition

$$\langle \phi_n, \phi_n \rangle_r = 1. \tag{2.43}$$

And finally, we can always construct orthonormal sets by properly normalising them.

Lemma 2.5 (Orthonormalisation)**.** *If the set $\{\psi_n\}$ is orthogonal and*

$$\phi_n(x) = \frac{\psi_n(x)}{||\psi_n(x)||_r} \qquad \textit{where } ||\psi_n(x)||_r = \langle \psi_n, \psi_n \rangle_r^{1/2} \tag{2.44}$$

then $\{\phi_n\}$ is orthonormal. In other words, every orthogonal set can be used to construct an orthonormal set by suitable re-normalisation of the functions.

Proof. First, we can check that $\{\phi_n\}$ is orthogonal due to the orthogonality of $\{\psi_n\}$. Remember that the division factor in the normalisation is a scalar. Then using properties of the inner product,

$$\langle \phi_n, \phi_m \rangle_r = \left\langle \frac{\psi_n}{||\psi_n||_r}, \frac{\psi_m}{||\psi_m||_r} \right\rangle = \frac{1}{||\psi_n||_r \cdot ||\psi_m||_r} \langle \psi_n, \psi_m \rangle_r = 0.$$

Orthormality follows from checking

$$\langle \phi_n, \phi_n \rangle_r = \left\langle \frac{\psi_n}{||\psi_n||_r}, \frac{\psi_n}{||\psi_n||_r} \right\rangle = \frac{\langle \psi_n, \psi_n \rangle_r}{\langle \psi_n, \psi_n \rangle_r^{1/2} \cdot \langle \psi_n, \psi_n \rangle_r^{1/2}} = 1.$$

□

Example 2.4. Consider the set $\{\psi_n(x)\} = \{\sin(nx)\}$, $n \in \mathbb{N}$ with $x \in [0, \pi]$ and weight function $r(x) = 1$. We have already shown that this set is orthogonal (Sec. 2.2.1). Since

$$\langle \psi_n, \psi_n \rangle_r = \int_0^\pi \sin^2(nx)\, dx = \frac{\pi}{2} \tag{2.45}$$

the set $\{\phi_n(x)\} = \{\sqrt{\frac{2}{\pi}} \sin(nx)\}$ is orthonormal.

A major application of orthogonal functions is their use to approximate much more complex functions. This is very important in the efficient computation of special functions on a computer. In this section we assume that all functions are real-valued. We choose to approximate f by the finite sum

$$f(x) \approx f_N(x) \equiv \sum_{n=1}^{N} \alpha_n \phi_n(x) \tag{2.46}$$

where $\alpha_n \in \mathbb{R}$ and $\{\phi_n\}$, $n = 1, \ldots, N$ is an orthonormal set. Obviously, this expansion is only helpful if we have some way of quantifying how "close" f_N is to f. To this end, we define the *approximation error* by

$$E_N = ||f - f_N||_r^2. \tag{2.47}$$

The two obvious questions are now:

Q1: Which choice of the coefficients α_n minimises the approximation error?

Q2: What is the error in this case?

In essence, these two questions are given by the following theorem, which indicates that the best approximation is one where the coefficients, α_n, are chosen in a specific way. Moreover the error can be quantified.

Theorem 2.2 (Optimal approximation)**.** *The optimal expansion coefficients in (2.46) can be chosen as follows:*

- *E_N is minimised if $\alpha_n = \langle f, \phi_n \rangle_r$*
- *In this case the error is*

$$E_N = ||f||_r^2 - \sum_{n=1}^{N} \alpha_n^2 = ||f||_r^2 - \sum_{n=1}^{N} \langle f, \phi_n \rangle_r^2 \tag{2.48}$$

Proof. Using the definition of the approximation error in (2.47) we have

$$E_N = \langle f - f_N, f - f_N\rangle_r = \langle f, f\rangle_r - 2\langle f, f_N\rangle_r + \langle f_N, f_N\rangle_r. \tag{2.49}$$

Now consider the last term and the expansion in (2.46) and split the double sum over n, m into two sums with $n = m$ and $n \neq m$

$$\begin{aligned}
\langle f_N, f_N\rangle_r &= \left\langle \sum_{n=1}^{N} \alpha_n\phi_n, \sum_{m=1}^{N} \alpha_m\phi_m \right\rangle_r \\
&= \sum_{n=1}^{N} \alpha_n^2 \cdot \underbrace{\langle \phi_n, \phi_n\rangle_r}_{=1 \text{ (by orthonormality)}} + \sum_{n\neq m} \alpha_n\alpha_m \cdot \underbrace{\langle \phi_n, \phi_m\rangle_r}_{=0 \text{ (by orthogonality)}} \\
&= \sum_{n=1}^{N} \alpha_n^2.
\end{aligned} \tag{2.50}$$

We also have

$$\langle f, f_N\rangle_r = \left\langle f, \sum_{n=1}^{N} \alpha_n\phi_n \right\rangle_r = \sum_{n=1}^{N} \alpha_n \langle f, \phi_n\rangle_r. \tag{2.51}$$

Insert this all into (2.49) to obtain

$$E_N = ||f||_r^2 - 2\sum_{n=1}^{N} \alpha_n\langle f, \phi_n\rangle_r + \sum_{n=1}^{N} \alpha_n^2. \tag{2.52}$$

We wish to choose the coefficients, α_n, so that they minimise the above error. Hence we calculate each partial derivative and set it equal to zero:

$$0 = \frac{\partial E_N}{\partial \alpha_n} = -2\langle f, \phi_n\rangle_r + 2\alpha_n \qquad \text{for all } n \in \{1, \ldots, N\} \tag{2.53}$$

or

$$\alpha_n = \langle f, \phi_n\rangle_r. \tag{2.54}$$

By inserting $\alpha_n = \langle f, \phi_n\rangle_r$ into (2.52) we can write down the error in this case as

$$E_N = ||f||_r^2 - 2\sum_{n=1}^{N}\langle f, \phi_n\rangle_r^2 + \sum_{n=1}^{N}\langle f, \phi_n\rangle_r^2 = ||f||_r^2 - \sum_{n=1}^{N}\langle f, \phi_n\rangle_r^2. \tag{2.55}$$

Essentially, it remains to check that if α_n is chosen as above, then the error is indeed at a minimum; this check is akin to checking the second derivative test when examining minima of functions. In fact, this result can be strengthened to indicate that the best approximation achievable occurs when the coefficients are chosen to be the generalised Fourier coefficients in (2.54). We do not give the proofs of these statements, but they can be found on p. 213 of the textbook by Haberman. □

Corollary 2.1 (Bessel's inequality). *As $E_N \geq 0$ we have*

$$0 \leq \sum_{n=1}^{N} \langle f, \phi_n \rangle_r^2 \leq ||f||_r^2. \tag{2.56}$$

Definition 2.8. An orthonormal set $\{\phi_n\}$ is *complete* if $E_N \to 0$ as $N \to \infty$.

Theorem 2.3. *The eigenfunctions of a Sturm-Liouville operator form a complete set.*

This is a very deep theorem, and the proof is beyond the scope of this course. However, from Theorem 2.3 we deduce the following result:

Theorem 2.4 (Parseval). *If the orthonormal set $\{\phi_n\}$ is complete then*

$$\sum_{n=1}^{\infty} \langle f, \phi_n \rangle_r^2 = \sum_{n=1}^{\infty} \alpha_n^2 = ||f||_r^2 \tag{2.57}$$

Proof. Let $N \to \infty$ in (2.48) and use Definition 2.8. $\square$

Example 2.5. Let $u'' + \lambda u = 0$, $u(0) = u(\pi) = 0$, $r(x) = 1$. The eigenfunctions $\{\phi_n\}$ with $\phi_n = \sqrt{\frac{2}{\pi}} \sin(nx)$ form an orthormal set with eigenvalues $\lambda_n = n^2$ (see Example 2.4). Let f be a hat-function (see Fig. 2.3)

$$f(x) = \begin{cases} x & \text{for } 0 \leq x < \frac{\pi}{2} \\ \pi - x & \text{for } \frac{\pi}{2} \leq x \leq \pi \end{cases} \tag{2.58}$$

α_n is readily evaluated as the following integral over x

$$\begin{aligned} \alpha_n = \sqrt{\frac{2}{\pi}} \int_0^{\pi} f(x) \sin(nx)\, dx &= \sqrt{\frac{2}{\pi}} \left(\int_0^{\frac{\pi}{2}} x \sin(nx)\, dx + \int_{\frac{\pi}{2}}^{\pi} (\pi - x) \sin(nx)\, dx \right) \\ &= \begin{cases} 2\sqrt{\frac{2}{\pi}} \int_0^{\frac{\pi}{2}} x \sin(nx)\, dx = 2\sqrt{\frac{2}{\pi}} \frac{(-1)^{\frac{n-1}{2}}}{n^2} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}. \end{aligned} \tag{2.59}$$

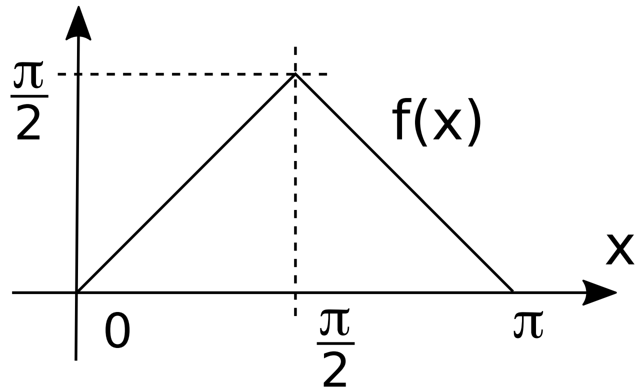


Figure 2.3: Hat function defined in (2.58)

Therefore the hat function f in (2.58) can be approximated as

$$f(x) \approx f_{2N-1}(x) = \sum_{n=1}^{N} \alpha_{2n-1}\phi_{2n-1}(x) = -\frac{4}{\pi}\sum_{n=1}^{N}\frac{(-1)^n}{(2n-1)^2}\sin\left((2n-1)x\right). \tag{2.60}$$

The first few functions f_N and the associated error $f(x) - f_N(x)$ are plotted in Fig. 2.4. Also

$$||f||_r^2 = \int_0^\pi f^2\,dx = \int_0^{\frac{\pi}{2}} x^2\,dx + \int_{\frac{\pi}{2}}^{\pi}(\pi - x)^2\,dx = 2\int_0^{\frac{\pi}{2}} x^2\,dx = \frac{\pi^3}{12}. \tag{2.61}$$

(2.57) gives the interesting result

$$\begin{aligned}
\frac{\pi^3}{12} &= ||f||_r^2 = \sum_{n=1}^{\infty}\alpha_n^2 = \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^4} \qquad \text{and therefore} \\
\frac{\pi^4}{96} &= \sum_{n=1}^{\infty}\frac{1}{(2n-1)^4} = 1 + \frac{1}{3^4} + \frac{1}{5^4} + \frac{1}{7^4} + \frac{1}{9^4} + \dots
\end{aligned} \tag{2.62}$$

For large $N \gg 1$ we can also estimate the error E_N as follows:

$$\begin{aligned}
E_{2N} &= ||f||_r^2 - \sum_{n=1}^{2N}\alpha_n^2 = \sum_{n=2N+1}^{\infty}\alpha_n^2 = \frac{8}{\pi}\sum_{n=N+1}^{\infty}\frac{1}{(2n-1)^4} \\
&= \frac{8}{\pi}\left(\sum_{k=2N+1}^{\infty}\frac{1}{k^4} - \sum_{k=N+1}^{\infty}\frac{1}{(2k)^4}\right) \approx \frac{8}{\pi}\left(\int_{2N+1}^{\infty}\frac{dx}{x^4} - \frac{1}{16}\int_{N+1}^{\infty}\frac{dx}{x^4}\right) \\
&= \frac{8}{3\pi}\left(\frac{1}{(2N+1)^3} - \frac{1}{16}\cdot\frac{1}{(N+1)^3}\right) \approx \frac{1}{6\pi N^3}.
\end{aligned} \tag{2.63}$$

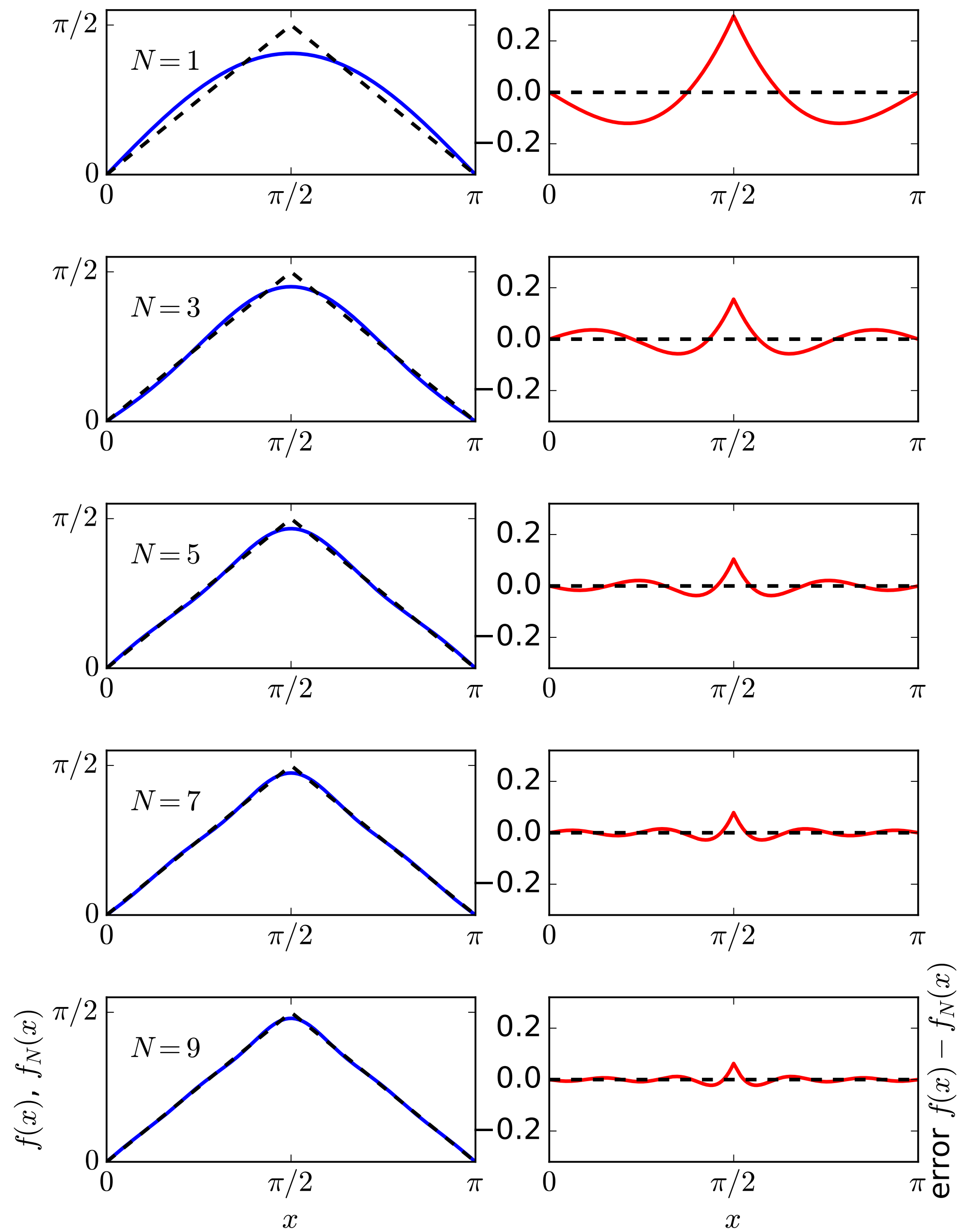


Figure 2.4: Expansion of the hat function $f(x)$ (2.58) in terms of basis functions. The error $f(x) - f_N(x)$ is shown on the right.

Example 2.6. The previous example might appear a bit arbitrary since the function itself is piecewise linear and very easy to evaluate directly. We therefore also consider the more complicated function

$$f(x) = \exp\left[-x^2 + \sin(x)\right] \qquad \text{for } x \in [-1, 1]. \tag{2.64}$$

and expand it in terms of the eigenfunctions $T_n(x)$ of the Chebyshev equation

$$\left(\sqrt{1-x^2}u'\right)' + \lambda \frac{u}{\sqrt{1-x^2}} = 0 \qquad \text{for } x \in [-1, 1]. \tag{2.65}$$

It can be shown that $T_n(x)$ is a polynomial of degree n. Hence expanding $f(x)$ in this basis amounts to representing it as a polynomial. Fig. 2.5 shows this expansion together with the error $f(x) - f_N(x)$. The dependence of the total approximation error E_N on the order N of the approximation is shown in Fig. 2.6. The plot demonstrates that the function can be approximated with an error $E_N < 10^{-8}$ if polynomial of degree nine is used.

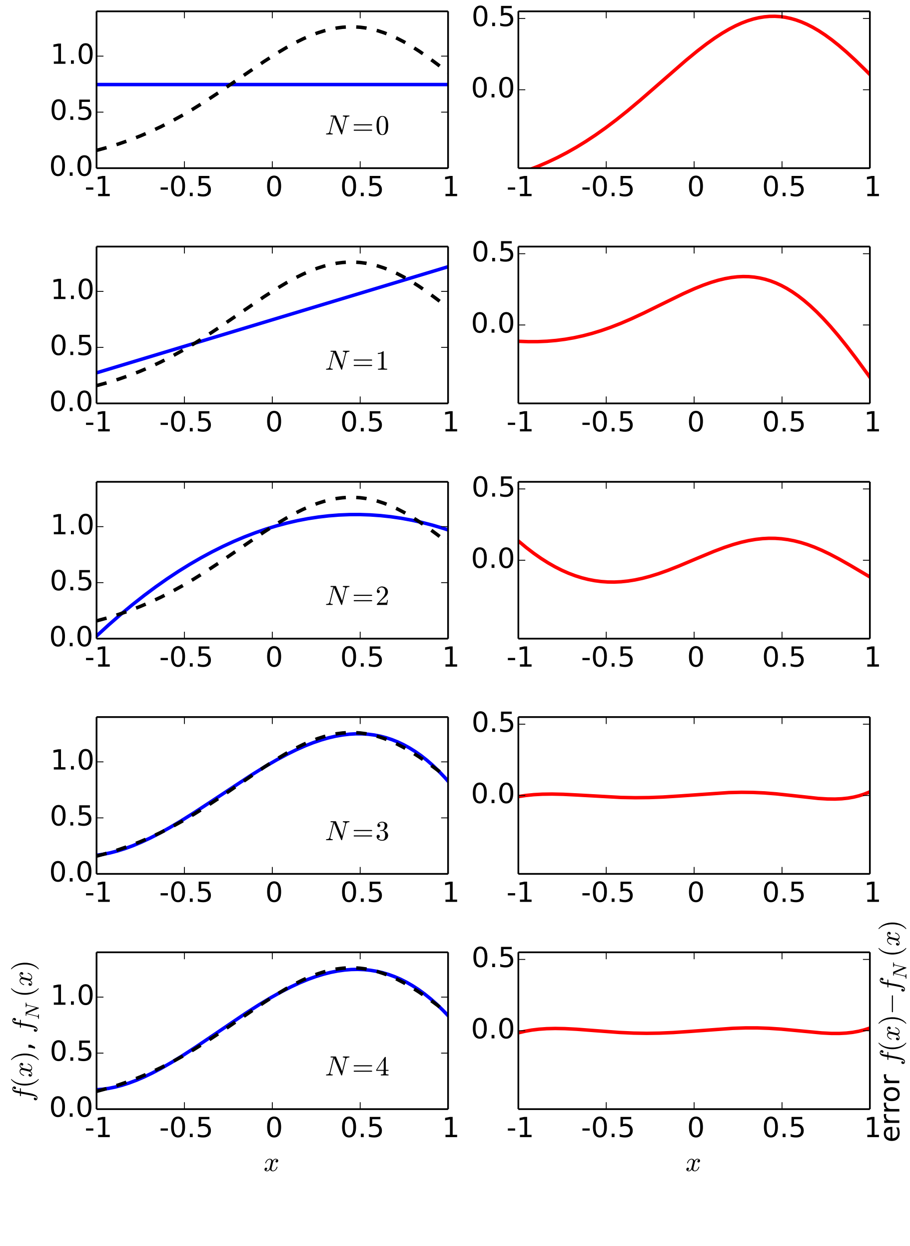


Figure 2.5: Expansion of the function $f(x) = \exp\left[-x^2 + \sin(x)\right]$ in terms of Chebyshev polynomials. The error $f(x) - f_N(x)$ is shown on the right.

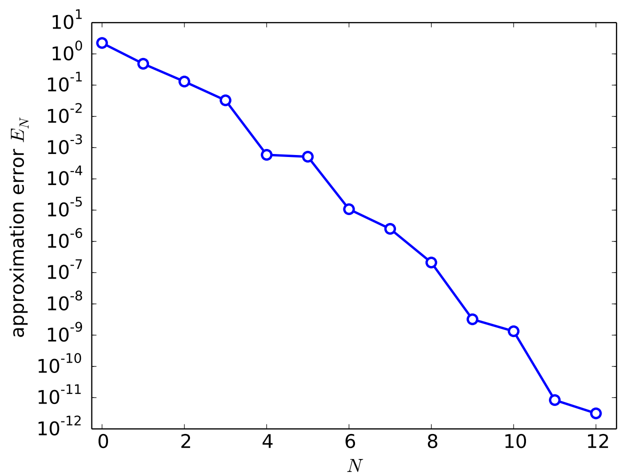


Figure 2.6: Approximation error E_N for the approximation of the function $f(x) = \exp\left[-x^2 + \sin(x)\right]$ in terms of Chebyshev polynomials.

2.7 Legendre polynomials

2021-22: The note below is correct and can serve as the lecture notes. It can be supplemented by the clear presentation in Chap. 7.10 of Haberman's *Applied Partial Differential Equations* textbook. In particular, the case below corresponds to the $m = 0$ Legendre polynomials in this reference.

We now consider the solution of a particular SL system which arises, for example, in the solution of PDEs in spherical coordinates. The solution are the Legendre polynomials. We will show how these eigenfunctions can be expressed very compactly as the expansion coefficients of a so-called generating function and how this can be used to derive recurrence relations between different Legendre polynomials.

The derivation of the below Legendre equation (can) begin from studying the three-dimensional wave equation that would *e.g.* describe the vibrations in the Earth:

$$\frac{\partial^2 \psi}{\partial t^2} = c^2 \nabla^2 \psi,$$

where ψ is the displacement. Separation of variables is applied to the above problem, of the form, $\psi = w(r, \theta, \phi)h(t)$, for spherical coordinates (r, θ, ϕ). After a series of transformations, one emerges with the Legendre equation.

Lemma 2.6. *The n-th eigenvalue/eigenfunction pair of the Legendre equation*

$$((1 - x^2)u')' + \lambda u = 0 \qquad \text{on } I = [-1, 1] \tag{2.66}$$

is given by

$$(\lambda_n, \phi_n) = (n(n+1), P_n) \qquad \text{with } P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[(x^2 - 1)^n\right] \quad \text{for } n = 0, 1, 2, \ldots \tag{2.67}$$

and P_n is a polynomial of degree n in x with $P_n(1) = 1$.

Note: Equation (2.66) is a **singular** Sturm-Liouville equation because of the conditions of $p(\pm 1) = 0$. It can be verified that, as long as we require that u is bounded on the two endpoints, it is not necessary to specify boundary conditions, and we may fix the normalisation of the function by defining its value at $x = 1$.

We claim that the usual properties of eigenvalues and eigenfunctions are valid in this case. In particular, there is an infinite set of eigenfunctions corresponding to different eigenvalues, and these eigenfunctions will be an orthogonal set with weight 1.

Proof. Use the identity

$$(1 - x^2)\frac{d}{dx}\left[(x^2 - 1)^n\right] + 2nx(x^2 - 1)^n = 0 \tag{2.68}$$

and take the $n+1$-st derivative of equation (2.68) as follows:

$$\frac{d}{dx}\frac{d^n}{dx^n}\left[(1-x)^2\frac{d}{dx}\left[(x^2-1)^n\right]\right]+2n\frac{d^{n+1}}{dx^{n+1}}\left[x(x^2-1)^n\right]=0 \tag{2.69}$$

Now evaluate the n-th derivative in the first term and the $n+1$-st derivative in the 2nd term using the generalised product rule (see appendix B)

$$\frac{d^n(f\cdot g)}{dx^n}=\sum_{k=0}^{n}\binom{n}{k}\frac{d^k f}{dx^k}\cdot\frac{d^{n-k}g}{dx^{n-k}} \tag{2.70}$$

with $f(x)=(1-x^2)$, $g(x)=\frac{d}{dx}(x^2-1)$ and $f(x)=x$, $g(x)=(x^2-1)^n$ to obtain[2]

$$\frac{d}{dx}\left\{(1-x^2)\frac{d}{dx}\frac{d^n}{dx^n}\left[(x^2-1)^n\right]-2nx\frac{d^n}{dx^n}\left[(x^2-1)^n\right]-2\frac{n(n-1)}{2}\frac{d^{n-1}}{dx^{n-1}}\left[(x^2-1)^n\right]\right\}$$
$$+2nx\frac{d^{n+1}}{dx^{n+1}}\left[(x^2-1)^n\right]+2n^2\frac{d^n}{dx^n}\left[(x^2-1)^n\right]=0 \tag{2.71}$$

and thus

$$\frac{d}{dx}\left\{(1-x^2)\frac{d}{dx}\frac{d^n}{dx^n}\left[(x^2-1)^n\right]\right\}+n(n+1)\frac{d^n}{dx^n}\left[(x^2-1)^n\right]=0$$

Hence the function $\frac{d^n}{dx^n}\left[(x^2-1)^n\right]$ and therefore $P_n(x)$ satisfies the Legendre equation with $\lambda=n(n+1)$.

Note also that since $p(-1)=p(1)=0$, it can be verified that, via a derivation of the analogous Rayleigh Quotient, $\lambda\geq 0$. Note also that $P_n(x)$ is a polynomial of degree n since we can write $(x^2-1)^n=x^{2n}+\mathcal{O}(x^{2n-1})$ and $\frac{d^n(x^{2n})}{dx^n}=\frac{(2n)!}{n!}x^n$.

To see that $P_n(1)=1$ write $x=1-\epsilon$, $d/dx=-d/d\epsilon$. Note that for $\epsilon\ll 1$ we can expand $((1-\epsilon)^2-1)^n=(-2\epsilon)^n+\mathcal{O}(\epsilon^{n-1})$ and hence

$$P_n(1)=\lim_{\epsilon\to 0}\frac{1}{2^n n!}(-1)^n\frac{d^n}{d\epsilon^n}\left[((1-\epsilon)^2-1)^n\right]=\lim_{\epsilon\to 0}\frac{1}{2^n n!}\frac{d^n}{d\epsilon^n}(2\epsilon)^n=1. \tag{2.72}$$

□

The first six Legendre polynomials are shown in Fig. 2.7.

Lemma 2.7 (Generating function)**.** *The Legendre polynomials can be obtained as coefficients of the following power series of the generating function $g(x;t)$*

$$g(x;t)=\frac{1}{\sqrt{1-2xt+t^2}}=\sum_{n=0}^{\infty}P_n(x)t^n \qquad \text{for } x\in[-1,1] \text{ and } |t|<1. \tag{2.73}$$

[2] note that only the first three terms with $k=0,1,2$ contribute in the sum over k in the first case and only the first two terms with $k=0,1$ contribute in the second case

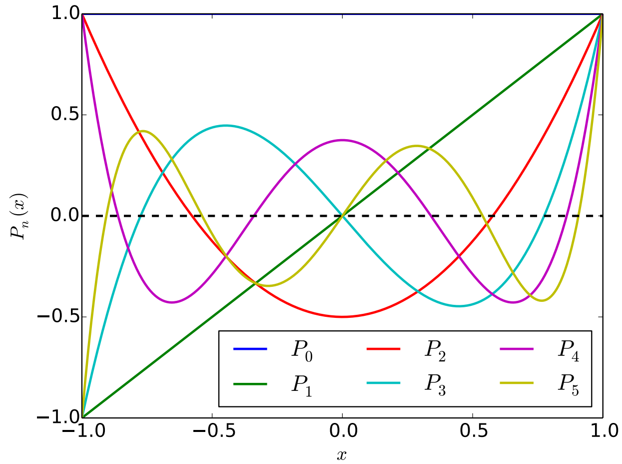


Figure 2.7: First six Legendre polynomials $P_n(x)$

Proof. A straightforward but tedious calculation confirms that $g(x;t) = 1/\sqrt{1-2xt+t^2}$ defined for $x \in [-1,1]$ and $t \in (-1,1)$ fulfills the following differential equation:

$$\frac{\partial}{\partial x}\left[(1-x^2)\frac{\partial g}{\partial x}\right] + t\frac{\partial^2(tg)}{\partial t^2} = 0. \tag{2.74}$$

Assume that $g(x;t)$ has the following Taylor expansion in t:

$$g(x;t) = \sum_{n=0}^{\infty} a_n(x)t^n. \tag{2.75}$$

Evaluating the first term of (2.74) gives

$$\frac{\partial}{\partial x}\left[(1-x^2)\frac{\partial g}{\partial x}\right] = \sum_{n=0}^{\infty}\left(\frac{d}{dx}\left[(1-x^2)\frac{da_n(x)}{dx}\right]\right)t^n \tag{2.76}$$

Evaluating the second term of (2.74) gives

$$\sum_{n=0}^{\infty} a_n(x)t\frac{d^2}{dt^2}t^{n+1} = \sum_{n=0}^{\infty} n(n+1)a_n(x)t^n \tag{2.77}$$

This has to hold for all t, and therefore at every order in t^n. Comparing the coefficient of the t^n term gives

$$\frac{d}{dx}\left[(1-x^2)\frac{da_n(x)}{dx}\right] + n(n+1)a_n(x) = 0 \tag{2.78}$$

and hence a_n fulfills the Legendre equation and has to be proportional to P_n. Now consider $x = 1$. Then

$$g(1;t) = \frac{1}{1-t} = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} a_n(1)t^n \tag{2.79}$$

and hence $a_n(1) = 1$. This implies that indeed $a_n = P_n$. $\square$

Lemma 2.8. *The Legendre polynomials defined in Lemma 2.6 fulfill the recursion relation*

$$(n+1)P_{n+1} = (2n+1)xP_n(x) - nP_{n-1}(x) \qquad \text{for } n = 0, 1, 2, \dots \tag{2.80}$$

if we set $P_{-1} = 0$.

Proof. Take the derivative of the generating function in (2.73) with respect to t to obtain

$$\frac{x-t}{(1-2xt+t^2)^{3/2}} = \sum_{n=0}^{\infty} nP_n(x)t^{n-1}. \tag{2.81}$$

Multiplying both sides by $1 - 2xt + t^2$ we find that

$$\frac{x-t}{\sqrt{1-2xt+t^2}} = (1-2xt+t^2)\sum_{n=0}^{\infty} nP_n(x)t^{n-1}. \tag{2.82}$$

Replacing $1/\sqrt{1-2xt+t^2}$ by the sum in (2.73) we have

$$(x-t)\sum_{n=0}^{\infty} P_n(x)t^n = (1-2xt+t^2)\sum_{n=0}^{\infty} nP_n(x)t^{n-1}. \tag{2.83}$$

For this to hold for every t, the equation has to hold at each order of t^n. From this the recursion relation (2.80) can be read off by comparing coefficients. $\square$

Generating functions exist for other Sturm-Liouville operators. For example, the Chebyshev equation

$$(\sqrt{1-x^2}u')' + \lambda\frac{u}{\sqrt{1-x^2}} = 0, \tag{2.84}$$

which we already encountered in Example 2.6, has eigenvalue/eigenfunctions $\{n^2, T_n(x)\}$ where $T_n(x)$ is a polynomial which can be read off from

$$\frac{1-xt}{1-2xt+t^2} = \sum_{n=0}^{\infty} T_n(x)t^n \tag{2.85}$$

Again, this allows deducing a recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{2.86}$$

We now show how the Legendre polynomials naturally arise in the expansion of the solution to a physical problem.

2.8 An application to electrostatics

Consider a point particle with charge Q at position $\boldsymbol{r}_0 = (x_0, y_0, z_0)$. The electric potential at point $\boldsymbol{r} = (x, y, z) \neq \boldsymbol{r}_0$ is then given by

$$V_{\text{point}}(\boldsymbol{r}; \boldsymbol{r}_0) = \frac{Q}{4\pi} \frac{1}{||\boldsymbol{r} - \boldsymbol{r}_0||} = \frac{Q}{4\pi} \frac{1}{\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}}. \tag{2.87}$$

where here $|| \cdot ||$ denotes the normal Euclidean norm of a vector in $\mathbb{R}^3$. It can be shown that this potential satisfies the Poisson equation

$$\nabla^2 V_{\text{point}}(\boldsymbol{r}; \boldsymbol{r}_0) = 0 \quad \text{for } \boldsymbol{r} \neq \boldsymbol{r}_0,$$

where recall the Laplacian operator, ∇^2, is defined via

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2},$$

and is also sometimes written as $\Delta = \nabla^2$.

By the superposition principle, the potential of a distribution of charges is the sum of the potentials of individual point charges. Thus the potential of a charge distribution with density $\rho(\boldsymbol{r})$ is given by (see Fig. 2.9):

$$V(\boldsymbol{r}) = \iiint_{\mathbb{R}^3} \rho(\boldsymbol{r'}) V_{\text{point}}(\boldsymbol{r}; \boldsymbol{r'})\, dV = \frac{1}{4\pi} \iiint_{\mathbb{R}^3} \frac{\rho(\boldsymbol{r'})}{||\boldsymbol{r} - \boldsymbol{r'}||}\, dV. \tag{2.88}$$

With the above definition of V, it can be shown that

$$\nabla^2 V(\boldsymbol{r}) = -\rho(\boldsymbol{r}).$$

Let us take a particular problem with a particular geometry. Consider a vertical rod with a charge of $\rho(z)$ per unit length with $\rho(z) = 0$ for $|z| > R$. We would like to apply the above formula (2.88) and simplify the integral using the specification of ρ.

It is easiest to solve this problem in spherical coordinates (r, θ, ϕ) (see Fig. 2.8). Here r is the distance from the origin, θ is the angle relative to the vertical axis, and ϕ is the angle between the positive x-axis and the projection of $\boldsymbol{r}$ onto the xy plane.

In spherical coordinates, then,

$$\boldsymbol{r} = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta)$$

while the positions along the rod are given by

$$\boldsymbol{r'} = (0, 0, z).$$

A calculation then shows that

$$||\boldsymbol{r} - \boldsymbol{r'}||^2 = \boldsymbol{r} \cdot \boldsymbol{r} + \boldsymbol{r'} \cdot \boldsymbol{r'} - 2\boldsymbol{r'} \cdot \boldsymbol{r} = z^2 + r^2 - 2rz \cos\theta.$$

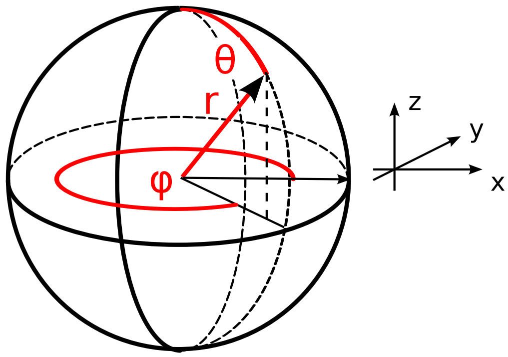


Figure 2.8: Spherical coordinates

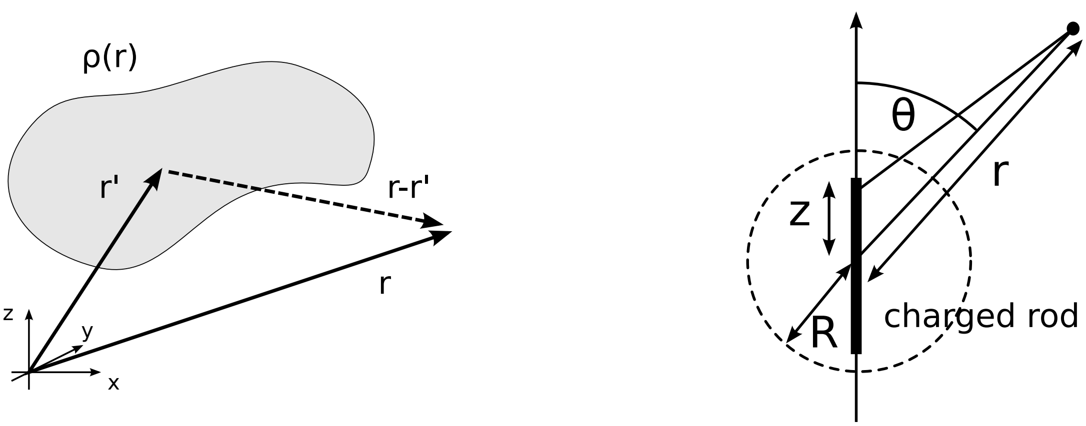


Figure 2.9: Illustration of the superposition principle (2.88) for a charge distribution $\rho(r)$ (left) and potential of a charged rod (right).

This is the cosine theorem from trigonometry.

The potential at a point (r, θ) which has a distance r from the origin and an angle of θ relative to the z-axis is therefore given by

$$V_{\text{rod}}(r, \theta) = \frac{1}{4\pi} \int_{-R}^{+R} \frac{\rho(z)\, dz}{\sqrt{z^2 + r^2 - 2rz\cos\theta}}. \tag{2.89}$$

Suppose we want to calculate the potential far away from the rod, i.e. $r > R$. Then we can express the potential in terms of the generating function (2.73) and therefore as a sum of

Legendre polynomials as

$$\begin{aligned}
V_{\text{rod}}(r,\theta) &= \frac{1}{4\pi}\int_{-R}^{+R}\frac{\rho(z)\,dz}{r\sqrt{1-2(z/r)\cos\theta+(z/r)^2}} \\
&= \frac{1}{4\pi r}\int_{-R}^{+R}\rho(z)g(\cos\theta,z/r)\,dz \\
&= \frac{1}{4\pi r}\int_{-R}^{+R}\rho(z)\sum_{n=0}^{\infty}P_n(\cos\theta)\left(\frac{z}{r}\right)^n\,dz \\
&= \frac{1}{4\pi}\sum_{n=0}^{\infty}Q_n r^{-(n+1)}P_n(\cos\theta) \qquad \text{with } Q_n=\int_{-R}^{+R}\rho(z)z^n\,dz.
\end{aligned} \tag{2.90}$$

The moments Q_n are called *multipole moments*. The zero-th moment is simply the total charge $Q_0=\int_{-R}^{+R}\rho(z)\,dz$ and the first moment $Q_1=\int_{-R}^{+R}\rho(z)z\,dz$ is known as the *dipole moment*. Note that since (for dimensional reasons) $Q_n \propto R^n$ the moments are suppressed by powers of $(R/r)^n$, and hence for $r \gg R$ only the lowest order moments contribute.

2.9 Inhomogeneous Sturm Liouville equations

So far we have only considered homogeneous SL equations $Lu+\lambda r(x)u=0$ and the associated eigenvalue problem. An *inhomogeneous* SL equation satisfies

$$Lu+\lambda r(x)u=f(x) \qquad + \text{ boundary conditions} \tag{2.91}$$

for a given $\lambda\in\mathbb{R}$ and a function f. For example $u''+\lambda u=e^x$, $u(0)=u(1)=0$. The following result tells us when we can solve (2.91) uniquely or not.

Theorem 2.5 (Fredholm Alternative)**.** *Let $Lu+\lambda r(x)u=f(x)$ and let L have the eigenvalues $\{\lambda_n\}$ and corresponding orthonormal eigenfunctions $\{\phi_n\}$ for $n\in\mathbb{N}$.*

1. *If $\lambda\neq\lambda_n$ for all $n\in\mathbb{N}$ then (2.91) has the unique solution*

$$u(x)=\int_a^b G(x,y)f(y)\,dy=\sum_{n=1}^{\infty}\frac{\phi_n(x)}{\lambda-\lambda_n}\int_a^b\phi_n(y)f(y)\,dy \tag{2.92}$$

where

$$G(x,y)=\sum_{n=1}^{\infty}\frac{\phi_n(x)\phi_n(y)}{\lambda-\lambda_n}. \tag{2.93}$$

2. *If $\lambda=\lambda_n$ for some n then the number of solutions depends on the integral*

$$\beta_n\equiv\int_a^b f(x)\phi_n(x)\,dx. \tag{2.94}$$

(a) If $\beta_n \neq 0$ the inhomogeneous SL equation (2.91) has no solution.

(b) If, on the other hand, $\beta_n = 0$ it has an infinite number of solutions.

Proof. Let's assume that a solution of (2.91) exists. As $\{\phi_n\}$ is complete (see Theorem 2.3) and orthonormal, we can expand u and the ratio f/r in this basis as

$$u(x) = \sum_{n=1}^{\infty} \alpha_n \phi_n(x), \qquad \frac{f(x)}{r(x)} = \sum_{n=1}^{\infty} \beta_n \phi_n(x). \tag{2.95}$$

The expansion coefficients β_n in Eq. (2.95) are indeed the same as in Eq. (2.94) since orthonormality implies

$$\int_a^b f(x)\phi_n(x)\,dx = \int_a^b \frac{f(x)}{r(x)}\phi_n(x)r(x)\,dx = \langle \frac{f}{r}, \phi_n\rangle_r = \langle \sum_{m=1}^{\infty} \beta_m \phi_m, \phi_n\rangle_r = \beta_n. \tag{2.96}$$

Then since $L\phi_n = -\lambda_n r\phi_n$ and L is linear:

$$Lu + \lambda r u = \sum_{n=1}^{\infty} \alpha_n \left(L\phi_n + \lambda r \phi_n\right) = \sum_{n=1}^{\infty} \alpha_n(\lambda - \lambda_n) r\phi_n \tag{2.97}$$

But, since $Lu + \lambda r u = f(x) = \sum_{n=1}^{\infty} \beta_n \phi_n(x) r(x)$ by (2.95)

$$\sum_{n=1}^{\infty} \alpha_n(\lambda - \lambda_n) r\phi_n = \sum_{n=1}^{\infty} \beta_n \phi_n r. \tag{2.98}$$

Due to the orthonormality of the basis functions ϕ_n this equation has to be fulfilled separately for each term in the sum[3],

$$\alpha_n(\lambda - \lambda_n) = \beta_n \qquad \text{for all } n \in \mathbb{N}. \tag{2.99}$$

If we have $\lambda \neq \lambda_n$ then $\alpha_n = \beta_n/(\lambda - \lambda_n)$ and we can write[4]

$$u(x) = \sum_{n=1}^{\infty} \frac{\beta_n \phi_n(x)}{\lambda - \lambda_n}. \tag{2.100}$$

Inserting the expression in Eq. (2.96) for β_n into (2.100) we obtain the desired result

$$\begin{aligned} u(x) &= \sum_{n=1}^{\infty} \left(\int_a^b f(y)\phi_n(y)\,dy\right) \frac{\phi_n(x)}{\lambda - \lambda_n} \\ &= \int_a^b G(x,y) f(y)\,dy \qquad \text{with } G(x,y) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(y)}{\lambda - \lambda_n}. \end{aligned} \tag{2.101}$$

[3] To see this, take the inner product with ϕ_m, i.e. multiply by ϕ_m and integrate over the interval $[a,b]$.

[4] This argument might appear circular since we assumed that a solution exists. However, we could equally well have started by construction u as in (2.100) and then show that this fulfills the SL equation (2.91).

Now consider case 2., i.e. $\lambda = \lambda_n$ for some n. If $\beta_n \neq 0$ then in (2.99) we have

$$\alpha_n \underbrace{(\lambda - \lambda_n)}_{=0} = \underbrace{\beta_n}_{\neq 0} \tag{2.102}$$

so α_n is undefined and there is no solution $u(x)$ for the inhomogeneous SL equation (2.91). If, on the other hand $\beta_n = 0$, we have

$$\alpha_n \underbrace{(\lambda - \lambda_n)}_{=0} = 0 \tag{2.103}$$

and hence α_n is arbitrary. There is an infinite number of solutions to the inhomogeneous SL equation (2.91), parameterised by the free parameter α_n:

$$\begin{aligned} u(x) = u_0(x) + \alpha_n \phi_n(x) \qquad &\text{with } u_0(x) = \int_a^b \hat{G}(x,y) f(y)\, dy \text{ and } \alpha_n \in \mathbb{R} \\ \text{where} \quad \hat{G}(x,y) = &\sum_{m=1, m\neq n}^{\infty} \frac{\phi_m(x)\phi_m(y)}{\lambda - \lambda_m}. \end{aligned} \tag{2.104}$$

□

Example 2.7. Consider the inhomogeneous wave equation with Dirichlet boundary conditions on the interval $[0, \pi]$:

$$u'' + \lambda u = f(x), \qquad \text{with } u(0) = u(\pi) = 0 \tag{2.105}$$

Let f be a hat-function (see Fig. 2.3)

$$f(x) = \begin{cases} x & \text{for } 0 \leq x < \frac{\pi}{2} \\ \pi - x & \text{for } \frac{\pi}{2} \leq x \leq \pi \end{cases} \tag{2.106}$$

From Theorem 2.1 and Example 2.4 we know that the SL problem $u'' + \lambda u = 0$ has orthonormal eigenfunctions $\phi_n(x) = \sqrt{\frac{2}{\pi}} \sin(nx)$ with corresponding eigenvalues $\lambda_n = n^2$. So, if $\lambda \neq n^2$ for any n we have

$$\begin{aligned} G(x,y) &= \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(nx)\sin(ny)}{\lambda - n^2} \\ \Rightarrow u(x) &= \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(nx)}{\lambda - n^2} \int_0^{\pi} \sin(ny) f(y)\, dy \end{aligned} \tag{2.107}$$

The integral over y can be evaluated as in (2.59) to obtain

$$\int_0^{\pi} \sin(ny) f(y)\, dy = \begin{cases} \frac{2(-1)^{\frac{n-1}{2}}}{n^2} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}. \tag{2.108}$$

Therefore

$$u(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2(\lambda - (2n+1)^2)} \sin((2n+1)x) \tag{2.109}$$

Example 2.8. Now consider the same equation but with Neumann boundary conditions at both ends:

$$u'' + \lambda u = f(x), \qquad \text{with } u'(0) = u'(\pi) = 0 \tag{2.110}$$

It can be shown that this equation has the orthonormal eigenfunctions $\{\phi_n\}$ with $\phi_1(x) = \sqrt{\pi}$ and $\phi_n(x) = \sqrt{\frac{2}{\pi}}\cos((n-1)x)$ for $n = 2, 3, \ldots$ with corresponding eigenvalues $\lambda_n = (n-1)^2$ for $n = 1, 2, 3, \ldots$. Note that the first eigenfunction ϕ_1 is constant and has eigenvalue $\lambda_1 = 0$. Now if $\lambda = 0$ we are in case 2. of Theorem 2.5 and hence (2.110) only has a solution if

$$\int_0^\pi f(x)\phi_1(x)\, dx = 0, \qquad \Leftrightarrow \qquad \int_0^\pi f(x)\, dx = 0. \tag{2.111}$$

Furthermore, in this case the solution is not unique and only fixed up to a constant shift, i.e. $u(x) = u_0(x) + \tilde{C}\phi_1(x) = u_0(x) + C$. This can be confirmed by explicitly solving the equation $u'' = f$ by integration:

$$u'(x) = \int_0^x f(y)\, dy + C' \tag{2.112}$$

but, due to the assumption (2.111) we have $u'(\pi) = \int_0^\pi f(y)\, dy = 0$ and hence the integration constant $C' = 0$. Integrating again we find that

$$u(x) = \int_0^x \left(\int_0^z f(y)\, dy \right) dz + C \tag{2.113}$$

i.e. as expected the solution is only unique up to a constant shift.

Finally note that if $\int_0^\pi f(y)\, dy \neq 0$ we would have from (2.112) that either $u'(0) \neq 0$ or $u'(\pi) \neq 0$, i.e. one of the boundary conditions is violated and we don't have a solution. Again this is consistent with the prediction from Theorem 2.5.

2.10 Solution of PDEs using Sturm-Liouville theory

Consider the three dimensional Laplace equation in spherical coordinates (r, θ, φ). If we are only interested in solutions which are rotationally invariant around the z-axis, i.e. don't depend on the angle φ, this equation is

$$\nabla^2 V(r, \cos\theta) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r} \right) + \frac{1}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial V}{\partial \theta} \right) = 0 \tag{2.114}$$

Suppose that we only look for solutions in the domain with $r > R$. We multiply this equation by r^2 and change variable to $x = \cos\theta$, noting that $\frac{\partial}{\partial\theta} = -\sin\theta \frac{\partial}{\partial\cos\theta} = -\sin\theta\frac{\partial}{\partial x}$

$$r^2\nabla^2 V(r, x) = \frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial x}\left(\sin^2\theta \frac{\partial V}{\partial x} \right) = \frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial x}\left((1 - x^2) \frac{\partial V}{\partial x} \right) = 0. \tag{2.115}$$

Now make the ansatz $V(r,x) = \zeta(r)u(x)$ and subsequently divide the equation by $\zeta(r)u(x)$ to obtain

$$r^2 \frac{\nabla^2 V(r,x)}{V(r,x)} = \frac{1}{\zeta(r)} \frac{d}{dr} \left(r^2 \frac{d\zeta(r)}{dr} \right) + \frac{1}{u(x)} \frac{d}{dx} \left((1-x^2) \frac{du(x)}{dx} \right) = 0. \tag{2.116}$$

Since the first term only depends on r and the second term only depends on x, this equation can only be fulfilled for all r, x if there is a constant λ such that

$$\frac{1}{\zeta(r)} \frac{d}{dr} \left(r^2 \frac{d\zeta(r)}{dr} \right) = \lambda = -\frac{1}{u(x)} \frac{d}{dx} \left((1-x^2) \frac{du(x)}{dx} \right). \tag{2.117}$$

In other words, u has to fulfill a Sturm-Liouville system

$$((1-x^2)u')' + \lambda u = 0. \tag{2.118}$$

However, from Lemma 2.6 we know that the eigenfunctions of this system are the Legendre polynomials $P_n(x)$ with corresponding eigenvalues $\lambda_n = n(n+1)$. For every n there is a corresponding radial function ζ_n which satisfies

$$(r^2 \zeta_n')' = n(n+1)\zeta_n \tag{2.119}$$

Making the ansatz $\zeta = Cr^\kappa$, it is easy to verify that $\kappa = n$ or $\kappa = -(n+1)$ and hence the solution is

$$\zeta_n(r) = \tilde{Q}_n r^n + Q_n r^{-(n+1)}. \tag{2.120}$$

Only the second term does not diverge as $r \to \infty$. In summary, the solution is

$$V(r, x = \cos\theta) = \sum_{n=0}^{\infty} Q_n r^{-(n+1)} P_n(\cos\theta). \tag{2.121}$$

Up to a factor of 4π this is exactly the same expression as (2.90). This is no coincidence since

$$\nabla^2 V_{\text{rod}}(r, \cos\theta) = 0 \qquad \text{for } r > R. \tag{2.122}$$

Chapter 3

The Fourier Transform

Aims:

1. To introduce you to a very useful mathematical transform
2. Discuss the basic properties and techniques for calculating the Fourier transform
3. To introduce distributions and the Dirac δ-function
4. Apply the Fourier transform to solve linear systems and PDEs

3.1 Introduction

Let $u(x)$ be a periodic, real-valued function with period L so that $u(x+L) = u(x)$ for all $x \in \mathbb{R}$ (see Fig. 3.1 for an example). Note that u does not necessarily have to be continuous. To address any jumps in the function we have the following

Definition 3.1. Let u be a not necessarily continuous function. Then we define

$$\tilde{u}(x) = \lim_{\epsilon \to 0} \frac{1}{2} \left(u(x+\epsilon) + u(x-\epsilon) \right) \qquad \text{with } \epsilon > 0. \tag{3.1}$$

For simplicity, we will only consider functions with $u(x) = \tilde{u}(x)$ in this chapter, but the Fourier transform can also be defined for discontinuous functions for which this is not the case.

Given a function u with period L, the Fourier coefficients a_n, b_n with $n \in \mathbb{N}^0$ are

$$a_n = \frac{2}{L} \int_{-L/2}^{+L/2} \cos\left(\frac{2n\pi x}{L}\right) u(x)\, dx, \qquad b_n = \frac{2}{L} \int_{-L/2}^{+L/2} \sin\left(\frac{2n\pi x}{L}\right) u(x)\, dx. \tag{3.2}$$

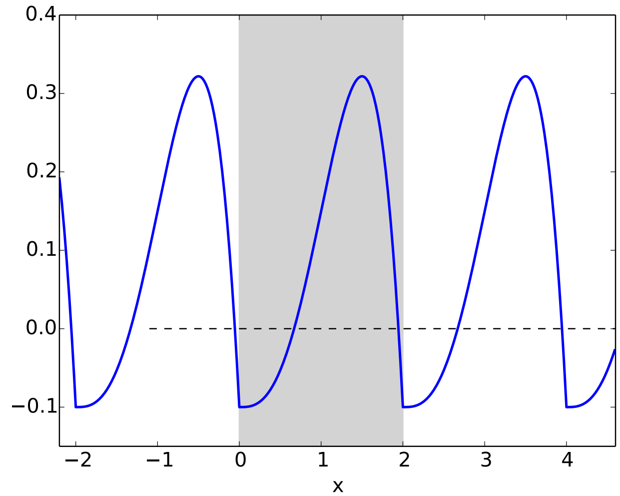


Figure 3.1: Example of a periodic function with $L = 2$.

Theorem 3.1 (Fourier series)**.** *A periodic function u with $u(x + L) = u(x)$ and $\tilde{u}(x) = u(x)$ for all x can be written as the infinite series*

$$u(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2n\pi x}{L}\right) + b_n \sin\left(\frac{2n\pi x}{L}\right)\right). \tag{3.3}$$

where a_n and b_n are given in (3.2).

We will not prove this theorem in this course, but we will use it to motivate the Fourier transform, which is an extension to more general, non-periodic function and central for this chapter. Note that if we drop the assumption that $u(x) = \tilde{u}(x)$ for all x, then the expansion in (3.3) is still valid if the left hand side is replaced by $\tilde{u}(x)$.

Example 3.1. If the hat function in Example 2.6 in Chapter 2 is periodically extended as $f(x + \pi) = f(x)$, the Fourier coefficients are given implicitly in (2.60).

We can also express the Fourier series (3.3) in complex form by using the well known identity $e^{i\alpha} = \cos(\alpha) + i \sin(\alpha)$ to obtain

$$\begin{aligned} u(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(\frac{a_n}{2} \left(e^{2\pi inx/L} + e^{-2\pi inx/L}\right) + \frac{b_n}{2i} \left(e^{2\pi inx/L} - e^{-2\pi inx/L}\right)\right) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(\frac{a_n - ib_n}{2} e^{2\pi inx/L} + \frac{a_n + ib_n}{2} e^{-2\pi inx/L}\right). \end{aligned} \tag{3.4}$$

If we define $c_n = \frac{1}{2}(a_n - ib_n)$, $c_{-n} = \frac{1}{2}(a_n + ib_n)$ for $n \in \mathbb{N}$ and $c_0 = \frac{a_0}{2}$ then we have the complex Fourier series

$$c_n = \frac{1}{L}\int_{-L/2}^{+L/2} e^{-2\pi i n x/L} u(x)\, dx, \qquad u(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/L}. \tag{3.5}$$

So far we still have the restriction that this only applies to functions which are periodic with period L. What happens if we lift this restriction by letting $L \to \infty$?

To see this, introduce the angular frequency $\omega_n \equiv 2n\pi/L$ and $\Delta\omega \equiv \omega_{n+1} - \omega_n = 2\pi/L$ and observe that $\Delta\omega \to 0$ as $L \to \infty$. Further define

$$F_L(\omega_n) \equiv \int_{-L/2}^{L/2} e^{-i\omega_n x} u(x)\, dx, \quad \text{then } c_n = \frac{1}{L} F_L(\omega_n) \tag{3.6}$$

Note that so far the function F_L is only defined at the discrete points ω_n. Now let $L \to \infty$ and set

$$F(\omega) \equiv F_\infty(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} u(x)\, dx \tag{3.7}$$

which is defined for all $\omega \in \mathbb{R}$. Recall also that from (3.5) with the definition of $\Delta\omega$ and $F_L(\omega_n)$ above we have for any function with period L

$$u(x) = \sum_{n=-\infty}^{\infty} F_L(\omega_n) e^{i\omega_n x} \frac{\Delta\omega}{2\pi}. \tag{3.8}$$

As $L \to \infty$, $\Delta\omega \to 0$ this converges to the integral.

$$u(x) \to \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega) e^{i\omega x}\, d\omega \tag{3.9}$$

This is summarised in the following

Definition 3.2. Let u be a complex valued function with $\tilde{u}(x) = u(x)$ for all $x \in \mathbb{R}$. Then we have

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} e^{-i\omega x} u(x)\, dx \\ u(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{+i\omega x} F(\omega)\, d\omega \end{aligned} \tag{3.10}$$

$F(\omega)$ is called the *Fourier transform* of u. The expression for u given in the second line of (3.10) is the *inverse Fourier transform* of F. Note that up to a factor [a] of $1/(2\pi)$ and a change of sign in the exponential the transforms are the same.
The following notation is used

$$F(\omega) = \mathcal{F}[u](\omega) = \hat{u}(\omega), \qquad u(x) = \mathcal{F}^{-1}[\hat{u}](x). \tag{3.11}$$

[a] In the literature you might come across other normalisation factors in front of the Fourier-transform and its inverse. The product of those factors is always $1/(2\pi)$.

As before, it is possible to relax the condition that $u(x) = \tilde{u}(x)$ for all x; in this case $\mathcal{F}^{-1}\left[\mathcal{F}[u]\right](x) = \tilde{u}(x)$. In other words, by applying the Fourier transform and its inverse we recover $\tilde{u}$ instead of the original function u.

Definition 3.2 is a bit ad-hoc, it can be shown rigorously that this all works giving the following

Theorem 3.2 (Fourier inversion). *Let $\tilde{u}(x) = u(x)$ for all $x \in \mathbb{R}$. Then*

$$u(x) = \mathcal{F}^{-1}[\mathcal{F}[u]](x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\omega(x-y)}u(y)\;dy\;d\omega \tag{3.12}$$

Note: The theorem is usually proved assuming that $u \in L_2(\mathbb{R})$ is square-integrable, but it can be extended to more general functions, in particular distributions (see section 3.3).

Example 3.2 (The decaying pulse). Set

$$u(x) = p_\alpha(x) = \begin{cases} e^{-\alpha x} & \text{for } x > 0 \text{ where } \alpha > 0 \\ \frac{1}{2} & \text{for } x = 0 \\ 0 & \text{for } x < 0 \end{cases}. \tag{3.13}$$

This example is very important and we will use this function and its Fourier transform in serveral places below; it is plotted in Fig. 3.2. The Fourier-transform of this function is

$$F(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x}u(x)\;dx = \int_{0}^{\infty} e^{-i\omega x}e^{-\alpha x}\;dx = \int_{0}^{\infty} e^{-(\alpha+i\omega)x}\;dx = \frac{1}{\alpha+i\omega}. \tag{3.14}$$

We also find that

$$u(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\omega x}}{\alpha+i\omega}\;d\omega. \tag{3.15}$$

as can be verified explicitly with the Residue Theorem from Complex Analysis.

Student exercise: Show that if $u(x) = e^{-\alpha|x|}$ the Fourier transform is

$$F(\omega) = \frac{2\alpha}{\alpha^2+\omega^2}. \tag{3.16}$$

Example 3.3 (Gaussian). Consider the following function, which is a Gaussian with width σ centred at $x = 0$:

$$u(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{x^2}{2\sigma^2}\right] = \mathcal{N}(0,\sigma) \tag{3.17}$$

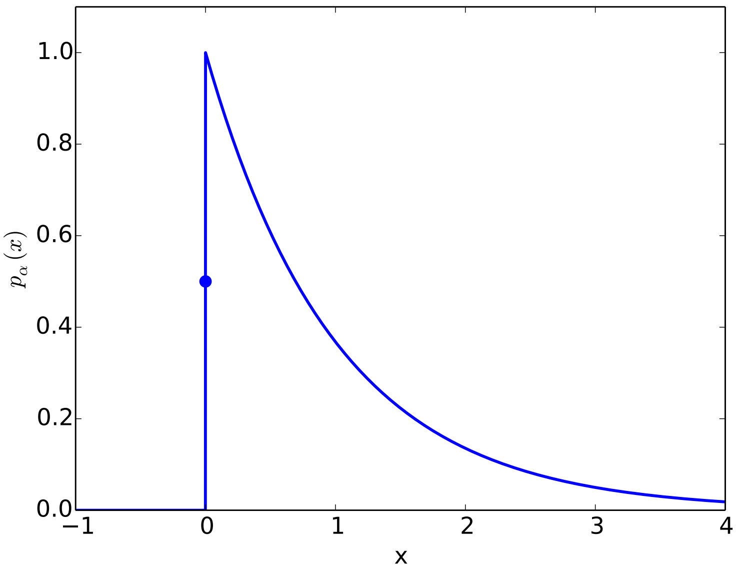


Figure 3.2: Decaying pulse $p_\alpha(x)$ defined in (3.13) in Example 3.2.

The Fourier transform is

$$F(\omega) = \mathcal{F}[u](\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{x^2}{2\sigma^2}\right] e^{-i\omega x}\, dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{x^2}{2\sigma^2} - i\omega x\right] dx \tag{3.18}$$

The trick is to complete the square in the exponential

$$\begin{aligned} -\frac{x^2}{2\sigma^2} - i\omega x &= -\frac{1}{2\sigma^2}\left(x^2 + 2i\omega\sigma^2 x\right) \\ &= -\frac{1}{2\sigma^2}\left(x^2 + 2i\omega\sigma^2 x - \sigma^4\omega^2\right) - \frac{1}{2}\sigma^2\omega^2 \\ &= -\frac{1}{2}\left(\frac{x + i\omega\sigma^2}{\sigma}\right)^2 - \frac{1}{2}\sigma^2\omega^2 \end{aligned} \tag{3.19}$$

Then the Fourier transform can be calculated with with the change of variable $z = \frac{x+i\omega\sigma^2}{\sigma}$, $dx = \sigma\, dz$:

$$\begin{aligned} F(\omega) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\sigma^2\omega^2\right] \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{x + i\omega\sigma^2}{\sigma}\right)^2\right] dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\sigma^2\omega^2\right] \sigma \underbrace{\int_{-\infty}^{\infty} e^{-z^2/2}\, dz}_{=\sqrt{2\pi}} \\ &= \exp\left[-\frac{1}{2}\sigma^2\omega^2\right] = \sqrt{\frac{2\pi}{\sigma^2}} \mathcal{N}(0, \sigma^{-1}) \end{aligned} \tag{3.20}$$

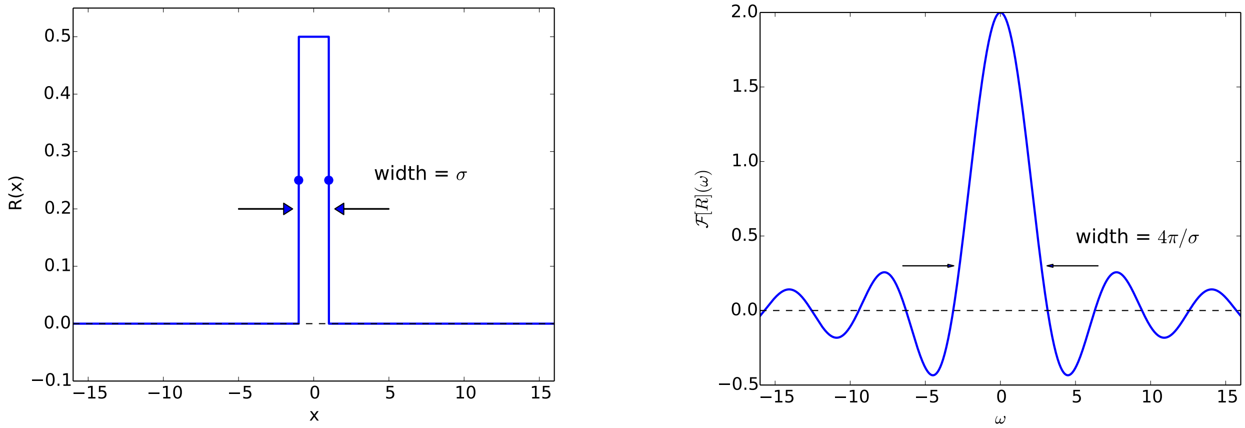


Figure 3.3: Rectangular pulse $R_\sigma(x)$ defined in (3.21) (left) and its Fourier-transform $\mathcal{F}[R_\sigma](\omega) = 2\frac{\sin\left(\frac{\omega\sigma}{2}\right)}{\omega\sigma}$ (right).

In other words, the Fourier transform of a Gaussian with width σ is (up to a constant) another Gaussian with width $1/\sigma$. It can be shown that in general the Fourier transform of an arbitrary function with width σ is a function with width $\propto 1/\sigma$, i.e. the width if the function itself and the width of the Fourier transform are inversely related. This is the mathematical reason for the famous uncertainty principle in Quantum Physics, which states that (for example) the position x and momentum p of a particle can not both be measured to arbitrary precision, $\Delta x \cdot \Delta p \geq \hbar/2$ where $\hbar$ is Planck's constant.

Example 3.4 (Rectangular pulse)**.** Let

$$u(x) = R_\sigma(x) = \begin{cases} 0 & \text{for } |x| > \frac{\sigma}{2} \\ \frac{1}{\sigma} & \text{for } |x| < \frac{\sigma}{2} \\ \frac{1}{2\sigma} & \text{for } |x| = \frac{\sigma}{2} \end{cases} \tag{3.21}$$

The function $R_\sigma(x)$ is plotted in Fig. 3.3 (left). The Fourier transform is

$$F(\omega) = \mathcal{F}[R_\sigma](\omega) = \int_{-\frac{\sigma}{2}}^{+\frac{\sigma}{2}} \frac{e^{-i\omega x}}{\sigma}\, dx = -\frac{1}{i\omega\sigma} e^{-i\omega x}\Big|_{-\frac{\sigma}{2}}^{+\frac{\sigma}{2}} = \frac{e^{i\frac{\omega\sigma}{2}} - e^{-i\frac{\omega\sigma}{2}}}{i\omega\sigma} = 2\frac{\sin\left(\frac{\omega\sigma}{2}\right)}{\omega\sigma} = \operatorname{sinc}\left(\frac{\omega\sigma}{2}\right) \tag{3.22}$$

and this is shown in Fig. 3.3 (right). The width of the Fourier transform can be defined as the first zero of the sinc() function which occurs at $\frac{\omega\sigma}{2} = \pi$. Hence this width is $\frac{4\pi}{\sigma}$ which is inversely proportional to the width of $R_\sigma(x)$. Note also that from the definition of the inverse Fourier transform

$$R_\sigma(x) = \mathcal{F}^{-1}[F](x) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{i\omega x}\frac{\sin\left(\frac{\omega\sigma}{2}\right)}{\omega\sigma}\, d\omega \tag{3.23}$$

In particular for $\sigma = 2$ and $x = 0$ this gives the nice result

$$\pi R(0) = \pi = \int_{-\infty}^{\infty} \frac{\sin(\omega)}{\omega}\, d\omega. \tag{3.24}$$

3.2 Elementary properties of the Fourier transform

Three obvious properties of the Fourier transform are:

- **Linearity**:
$$\mathcal{F}[\lambda u + \mu v] = \lambda \mathcal{F}[u] + \mu \mathcal{F}[v] \tag{3.25}$$
- **Symmetry**: If $v(x) \equiv u(-x)$ then
$$\mathcal{F}[u](-\omega) = \mathcal{F}[v](\omega) \tag{3.26}$$
- **Complex conjugation**:
$$\mathcal{F}[u^*](\omega) = \mathcal{F}[u]^*(-\omega) \tag{3.27}$$

Theorem 3.3 (First shift theorem)**.**
$$\mathcal{F}\left[e^{i\varpi x}u\right](\omega) = \mathcal{F}[u](\omega - \varpi) \qquad \textit{for all real shifts } \varpi \in \mathbb{R}. \tag{3.28}$$

Proof.
$$\mathcal{F}\left[e^{i\varpi x}u\right](\omega) = \int_{-\infty}^{\infty} e^{-i\omega x}e^{i\varpi x}u(x)\,dx = \int_{-\infty}^{\infty} e^{-i(\omega-\varpi)x}u(x)\,dx = \mathcal{F}[u](\omega-\varpi) \tag{3.29}$$

□

Example 3.5. Consider
$$v(x) = \begin{cases} e^{-\alpha x}\cos(\beta x) & \text{for } x > 0 \\ \frac{1}{2} & \text{for } x = 0 \\ 0 & \text{for } x < 0 \end{cases}, \qquad \Rightarrow \qquad v(x) = \frac{1}{2}e^{i\beta x}p_\alpha(x) + \frac{1}{2}e^{-i\beta x}p_\alpha(x). \tag{3.30}$$

with $p_\alpha(x)$ as defined in (3.13). Since $\mathcal{F}[p_\alpha](\omega) = 1/(\alpha + i\omega)$ for Example 3.2 we have
$$\mathcal{F}[v] = \frac{1}{2}\left(\frac{1}{\alpha + i(\omega - \beta)} + \frac{1}{\alpha + i(\omega+\beta)}\right). \tag{3.31}$$

Theorem 3.4 (Second shift theorem)**.** *Let* $u_{\overline{x}}(x) = u(x - \overline{x})$ *for any real shift* $\overline{x} \in \mathbb{R}$*. Then*

$$\mathcal{F}[u_{\overline{x}}](\omega) = e^{-i\omega\overline{x}}\mathcal{F}[u](\omega) \tag{3.32}$$

Proof.

$$\begin{aligned} \mathcal{F}[u_{\overline{x}}](\omega) &= \int_{-\infty}^{\infty} e^{-i\omega x} u(x - \overline{x})\, dx, \qquad \text{let } y = x - \overline{x}, \\ &= \int_{-\infty}^{\infty} e^{-i\omega(y+\overline{x})} u(y)\, dy = e^{-i\omega\overline{x}} \int_{-\infty}^{\infty} e^{-i\omega y} u(y)\, dy \\ &= e^{-i\omega\overline{x}}\mathcal{F}[u](\omega) \end{aligned} \tag{3.33}$$

□

We now establish a deep result which is essential in the application of the Fourier transform to differential equations.

Theorem 3.5 (Derivative theorem)**.** *Let* $|u(x)| \to 0$ *as* $|x| \to \infty$ *then*

$$\mathcal{F}[u'](\omega) = i\omega\mathcal{F}[u](\omega) \tag{3.34}$$

To paraphrase: "Differentiation in x-space ⇔ Multiplication in ω-space". Note that there is a similar result for the Laplace transform.

Proof. Integrate by parts to obtain

$$\begin{aligned} \mathcal{F}[u'](\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} u'(x)\, dx &= \underbrace{\left[e^{-i\omega x} u(x)\right]_{-\infty}^{\infty}}_{=0 \text{ as } \lim_{x\to\pm\infty} |u(x)| = 0} + i\omega \int_{-\infty}^{\infty} e^{-i\omega x} u(x)\, dx \\ = i\omega \int_{-\infty}^{\infty} e^{-i\omega x} u(x)\, dx &= i\omega\mathcal{F}[u](\omega). \end{aligned} \tag{3.35}$$

□

Corollary 3.1. *If both* $|u(x)| \to 0$ *and* $|u'(x)| \to 0$ *as* $|x| \to \infty$ *then*

$$\mathcal{F}[u''](\omega) = -\omega^2\mathcal{F}[u](\omega) \tag{3.36}$$

Proof. Apply Theorem 3.5 twice. □

Theorem 3.6 (Reverse derivative theorem)**.** *Let $v(\omega)$ be a function such that $|v(\omega)| \to 0$ as $|\omega| \to \infty$. Then we have*

$$\mathcal{F}^{-1}[v'](x) = -ix\mathcal{F}^{-1}[v](x) \tag{3.37}$$

Proof. This is proven in exactly the same way as Theorem 3.5. □

Example 3.6. Let

$$v(\omega) = \frac{1}{(\alpha + i\omega)^2}. \tag{3.38}$$

We want to find $u(x)$ such that $\mathcal{F}[u](\omega) = v(\omega)$, i.e. $u = \mathcal{F}^{-1}[v]$. For this, note that

$$\begin{aligned} & u(x) = \mathcal{F}^{-1}\left[\frac{1}{(\alpha + i\omega)^2}\right](x) = i\mathcal{F}^{-1}\left[\frac{\partial}{\partial\omega}\frac{1}{\alpha + i\omega}\right](x) = x\mathcal{F}^{-1}\left[\frac{1}{\alpha + i\omega}\right](x) \\ \Rightarrow \qquad & u(x) = xp_\alpha(x) = \begin{cases} xe^{-\alpha x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0. \end{cases} \end{aligned} \tag{3.39}$$

3.3 Introduction to Distributions

So far we have considered the Fourier transform for "well defined" functions, in particular we have only considered square-integrable functions. It turns out that the Fourier transform can be defined for more general objects if we introduce a generalisation of what we mean by a function.

To motivate this, obvious question to ask is: what is the Fourier transform of the constant function $u(x) = 1$? Naively we get

$$\mathcal{F}[u](\omega) = \int_{-\infty}^{\infty} e^{-i\omega x}\, dx = \left[-\frac{1}{i\omega}e^{-i\omega x}\right]_{-\infty}^{\infty} \tag{3.40}$$

and this is not well-defined since $e^{-i\omega x}$ has no meaningful value for $x \to \pm\infty$. Note also that $u(x)$ is not square-integrable on the entire real axis.

It turns out that there is a way of defining the Fourier transform of a constant function, but the result a **distribution**, not an ordinary function. To see this, consider

$$u_\epsilon(x) = \exp\left[-\frac{1}{2}\epsilon^2 x^2\right] = \sqrt{2\pi\epsilon^{-2}}\mathcal{N}(0, 1/\epsilon). \tag{3.41}$$

It is easy to see that for fixed x we have $\lim_{\epsilon\to 0} u_\epsilon(x) = 1$, i.e. in the limit $\epsilon \to 0$ this function approaches the constant function. We also know from Example 3.3 that

$$\mathcal{F}[u_\epsilon](\omega) = 2\pi \frac{1}{\sqrt{2\pi\epsilon^2}} \exp\left[-\frac{\omega^2}{2\epsilon^2}\right] = 2\pi \mathcal{N}(0, \epsilon). \tag{3.42}$$

Now define $\delta_\epsilon(\omega) = \mathcal{N}(0, \epsilon)$ which is a Gaussian of width ϵ. Note that

- $\delta_\epsilon(0) = \frac{1}{\sqrt{2\pi}}\epsilon^{-1} \to \infty$ as $\epsilon \to 0$
- $\delta_\epsilon(\omega) \propto \epsilon^{-1} \exp\left[-\frac{\omega^2}{2\epsilon^2}\right] \to 0$ as $\epsilon \to 0$ for $\omega \neq 0$. To see this take $\log\left(\epsilon^{-1} \exp\left[-\frac{\omega^2}{2\sigma^2}\right]\right) = \log \epsilon^{-1} - \frac{\omega^2}{2\epsilon^2} \to -\infty$ as $\epsilon \to 0$ since $\log x$ grows slower than x^2 for $x \to \infty$.
- $\int_{-\infty}^{\infty} \delta_\epsilon(\omega)\, d\omega = 1$ for all ϵ, since the integral over $\mathcal{N}$ is normalised to 1.

We now formally define the Dirac δ-function

$$\delta(\omega) = \lim_{\epsilon\to 0} \delta_\epsilon(\omega). \tag{3.43}$$

Although it is called the δ-“function“ it is a **distribution** with the following properties:

- $\delta(0) = \infty$
- $\delta(\omega) = 0$ for all $\omega \neq 0$
- $\int_{-\infty}^{\infty} \delta(\omega)\, d\omega = 1$

Even though this looks a bit pathological since it is infinite at 0, we usually do not need to worry about this since usually the δ function only appears under the integral.

The above allows us to define the Fourier transformation of a constant function.

Theorem 3.7 (Fourier transform of the constant function)**.** *If $u(x) = 1$ then*

$$\mathcal{F}[u](\omega) = 2\pi\delta(\omega) \tag{3.44}$$

with

$$\delta(\omega) = \lim_{\epsilon\to 0} \delta_\epsilon(\omega) \qquad \textit{where } \delta_\epsilon(\omega) = \mathcal{N}(0, \epsilon) = \frac{1}{\sqrt{2\pi\epsilon^2}} \exp\left[-\frac{\omega^2}{2\epsilon^2}\right]. \tag{3.45}$$

Note that although we defined the δ-function as the limit of a Gaussian, there are other equivalent definitions which use different limiting functions. The Dirac δ-function has many physical meanings:

- Density of a point mass

- Charge density of a point particle (electron, proton)
- Impulsive force

In particular the δ-function has the following property

Theorem 3.8. *If $f(x)$ is a differentiable and bounded function then*

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx = f(a). \tag{3.46}$$

Proof. Write $f(x) = f(a) + \tilde{f}(x-a)$ with $\tilde{f}(0) = 0$. Then

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx &= f(a)\int_{-\infty}^{\infty} \delta(x-a)\,dx + \int_{-\infty}^{\infty} \tilde{f}(x-a)\delta(x-a)\,dx \\ &= f(a)\underbrace{\int_{-\infty}^{\infty} \delta(y)\,dy}_{=1} + \int_{-\infty}^{\infty} \tilde{f}(y)\delta(y)\,dy\end{aligned} \tag{3.47}$$

where we changed the integration variable to $y = x - a$. We hence have to show that the second integral vanishes. For this consider the limit $\epsilon \to 0$ of the integral

$$\int_{-\infty}^{\infty} \tilde{f}(y)\delta_\epsilon(y)\,dy. \tag{3.48}$$

For this, split the integral into three parts:

$$\int_{-\infty}^{\infty} \tilde{f}(y)\delta_\epsilon(y)\,dy = \underbrace{\int_{-\infty}^{-\sqrt{\epsilon}} \tilde{f}(y)\delta_\epsilon(y)\,dy}_{I_-} + \underbrace{\int_{-\sqrt{\epsilon}}^{\sqrt{\epsilon}} \tilde{f}(y)\delta_\epsilon(y)\,dy}_{I_0} + \underbrace{\int_{\sqrt{\epsilon}}^{\infty} \tilde{f}(y)\delta_\epsilon(y)\,dy}_{I_+} \tag{3.49}$$

Since f and therefore $\tilde{f}$ is differentiable, it is Lipschitz-continuous on compact intervals and there exists a positive constant L such that $|\tilde{f}(y)| \leq L|y|$ for all $y \in [-\epsilon_0, \epsilon_0]$ for some $\epsilon_0 > 0$. Now consider $\epsilon < \epsilon_0$. We can bound the central integral I_0 as follows:

$$|I_0| \leq \int_{-\sqrt{\epsilon}}^{\sqrt{\epsilon}} |\tilde{f}(y)|\delta_\epsilon(y)\,dy \leq L\int_{-\sqrt{\epsilon}}^{\sqrt{\epsilon}} |y|\delta_\epsilon(y)\,dy = 2L\int_{0}^{\sqrt{\epsilon}} y\delta_\epsilon(y)\,dy \tag{3.50}$$

But then with the substitution $y = z\epsilon$

$$\begin{aligned}\int_0^{\sqrt{\epsilon}} y\delta_\epsilon(y)\,dy = \int_0^{\sqrt{\epsilon}} y\frac{1}{\sqrt{2\pi\epsilon^2}}\exp\left[-\frac{y^2}{2\epsilon^2}\right]\,dy &= \frac{\epsilon}{\sqrt{2\pi}}\int_0^{\epsilon^{-1/2}} z\exp\left[-\frac{z^2}{2}\right]\,dz \\ &= -\frac{\epsilon}{\sqrt{2\pi}}\exp\left[-\frac{z^2}{2}\right]\Big|_0^{\epsilon^{-1/2}} \\ &= \frac{\epsilon}{\sqrt{2\pi}}\left(1-\exp\left[-\frac{1}{2\epsilon^2}\right]\right) \to 0 \quad \text{as } \epsilon \to 0.\end{aligned} \tag{3.51}$$

Hence we find that $\lim_{\epsilon\to 0} I_0 = 0$. For the rightmost integral I_+ we have since f and therefore $\tilde{f}$ is bounded, $|\tilde{f}| \leq C$ for some positive C:

$$|I_+| \leq \int_{\sqrt{\epsilon}}^{\infty} |f(y)|\delta_\epsilon(y)\, dy \leq C \int_{\sqrt{\epsilon}}^{\infty} \delta_\epsilon(y)\, dy \tag{3.52}$$

But then:

$$\int_{\sqrt{\epsilon}}^{\infty} \frac{1}{\sqrt{2\pi\epsilon^2}} \exp\left[-\frac{y^2}{2\epsilon^2}\right]\, dy = \frac{1}{\sqrt{2\pi}} \int_{\epsilon^{-1/2}}^{\infty} \exp\left[-\frac{z^2}{2}\right]\, dz \to 0 \quad \text{as } \epsilon \to 0 \tag{3.53}$$

and therefore $\lim_{\epsilon\to 0} I_+ = 0$. Similarly we can show that $\lim_{\epsilon\to 0} I_- = 0$. Taking everything together we find that

$$\int_{-\infty}^{\infty} \tilde{f}(y)\delta(y)\, dy = \lim_{\epsilon\to 0} \int_{-\infty}^{\infty} \tilde{f}(y)\delta_\epsilon(y)\, dy = 0. \tag{3.54}$$

□

Example 3.7. We can use this result to define the Fourier-transform of the δ-function

$$\mathcal{F}[\delta](x) = \int_{-\infty}^{\infty} e^{-i\omega x}\delta(\omega)\, d\omega = e^{0\cdot i} = 1. \tag{3.55}$$

Using this, we can extend our “zoo” of Fourier transforms.

Corollary 3.2. *We have the following table:*

$u(x)$	$\mathcal{F}[u](\omega)$
$\delta(x)$	1
$e^{i\alpha x}$	$2\pi\delta(x-a)$
$\cos(\alpha x)$	$\pi(\delta(\omega-\alpha)+\delta(\omega+\alpha))$
$\sin(\alpha x)$	$-i\pi(\delta(\omega-\alpha)-\delta(\omega+\alpha))$

Proof. This follows immediately from straightforward applications of Theorem 3.3 and Theorem 3.8. □

Theorem 3.9 (δ-function chain rule)**.** *Let g be a real-valued function with a finite number of zeros x_i, $i = 1, \ldots, N$, such that $g(x_i) = 0$, $g'(x_i) \neq 0$ and g is continuously differentiable at x_i. Then we have:*

$$\delta(g(x)) = \sum_{i=1}^{N} \frac{\delta(x-x_i)}{|g'(x_i)|} \tag{3.56}$$

Proof. We show that for any bounded and differentiable function f

$$\int_{-\infty}^{\infty} f(x)\delta(g(x))\,dx = \sum_{i=1}^{N} \frac{f(x_i))}{|g'(x_i)|} = \sum_{i=1}^{N} \int_{-\infty}^{\infty} \frac{f(x)\delta(x-x_i)}{|g'(x_i)|}. \tag{3.57}$$

Since g is continuously differentiable at x_i and $g'(x_i) \neq 0$, it follows from the inverse function theorem that for each i there is an ϵ_i such that g is invertible on the interval $\Omega_i = [x_i - \epsilon_i, x_i + \epsilon_i]$. Let $\overline{\Omega} = [-\infty, \infty] \setminus \cup_{i=1}^{N} \Omega_i$, then we also have that $g(x) \neq 0$ for all $x \in \overline{\Omega}$. Split the integral in (3.57) as follows:

$$\int_{-\infty}^{\infty} f(x)\delta(g(x))\,dx = \sum_{i=1}^{N} \int_{x_i-\epsilon_i}^{x_i+\epsilon_i} f(x)\delta(g(x))\,dx + \int_{\overline{\Omega}} f(x)\delta(g(x))\,dx \tag{3.58}$$

The rightmost integral vanishes since $g(x) \neq 0$ and hence $\delta(g(x)) = 0$ for all $x \in \overline{\Omega}$. Now consider the integrals in the sum. First assume that $g'(x_i) > 0$, and hence $g(x_i + \epsilon_i) > g(x_i - \epsilon_i)$. We can use the chain rule with $y = g(x)$ and $dy = g'(x)dx$ to obtain by using $y = 0 \Leftrightarrow g^{-1}(y) = x = x_i$ on Ω_i:

$$\int_{x_i-\epsilon_i}^{x_i+\epsilon_i} f(x)\delta(g(x))\,dx = \int_{g(x_i-\epsilon_i)}^{g(x_i+\epsilon_i)} \frac{f(g^{-1}(y))}{g'(g^{-1}(y))}\delta(y)\,dy = \frac{f(x_i)}{g'(x_i)} \tag{3.59}$$

If $g'(x_i) < 0$ and hence $g(x_i + \epsilon_i) < g(x_i - \epsilon_i)$ then there is an additional sign since the boundaries of the integral over y need to be swapped. In summary

$$\int_{x_i-\epsilon_i}^{x_i+\epsilon_i} f(x)\delta(g(x))\,dx = \frac{f(x_i)}{|g'(x_i)|} = \int_{-\infty}^{\infty} \frac{f(x)\delta(x-x_i)}{|g'(x_i)|}. \tag{3.60}$$

□

Corollary 3.3. *$\delta(\alpha x) = \frac{\delta(x)}{|\alpha|}$ and in particular $\delta(-x) = \delta(x)$.*

Proof. This follows directly from the previous theorem with $g(x) = \alpha x$ and $\alpha = -1$. □

Example 3.8. Consider the integral

$$I = \int_{-\infty}^{\infty} x^3 \delta(x^3 - 3x^2 - 4x)\,dx. \tag{3.61}$$

In this case we have

$$g(x) = x^3 - 3x^2 - 4x = x(x+1)(x-4) \qquad \text{with } g'(x) = 3x^2 - 6x - 4. \tag{3.62}$$

The polynomial $g(x)$ has zeros at $x = 0, -1, 4$. We also have that

$$g'(0) = -4, \qquad g'(-1) = 5, \qquad g'(4) = 20. \tag{3.63}$$

Using Theorem 3.9 we find that the integral I in Eq. (3.61) is

$$\begin{aligned} I &= \int_{-\infty}^{\infty} x^3 \frac{\delta(x)}{|g'(0)|}\, dx + \int_{-\infty}^{\infty} x^3 \frac{\delta(x+1)}{|g'(-1)|}\, dx + \int_{-\infty}^{\infty} x^3 \frac{\delta(x-4)}{|g'(4)|}\, dx \\ &= \frac{1}{4}\int_{-\infty}^{\infty} x^3\delta(x)\, dx + \frac{1}{5}\int_{-\infty}^{\infty} x^3\delta(x+1)\, dx + \frac{1}{20}\int_{-\infty}^{\infty} x^3\delta(x-4)\, dx \\ &= \frac{1}{4}\cdot 0 + \frac{1}{5}\cdot(-1)^3 + \frac{1}{20}\cdot 4^3 = 3. \end{aligned} \tag{3.64}$$

3.4 The convolution theorem and its applications

Convolution is a way for combining functions. It arises also as correlation in statistics, in the solution of ODEs and in image analysis.

Definition 3.3. Let f and g be integrable functions. The *convolution* $h = f \star g$ of f and g is defined by

$$h(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\, dy \equiv (f \star g)(x) \tag{3.65}$$

Exercise: show that convolution is a symmetric operation, i.e. $f \star g = g \star f$.

Example 3.9.

Let $f(x) = p_\alpha(x)$ be the decaying pulse as in (3.13) and $g(x) = x$. Then with the change of variable $z = \alpha y$

$$\begin{aligned} (f \star g)(x) &= \int_0^{\infty} e^{-\alpha y}(x-y)\, dy = x\int_0^{\infty} e^{-\alpha y}\, dy - \int_0^{\infty} ye^{-\alpha y}\, dy \\ &= \frac{x}{\alpha}\int_0^{\infty} e^{-z}\, dz - \frac{1}{\alpha^2}\int_0^{\infty} ze^{-z}\, dz = \frac{\alpha x - 1}{\alpha^2} \end{aligned} \tag{3.66}$$

Lemma 3.1. *$f \star \delta = f = \delta \star f$, i.e. the Dirac δ-function plays the role of a "unit" in multiplication.*

Proof. This follows immediately from Theorem 3.8 and a simple change of variables:

$$(f \star \delta)(x) = \int_{-\infty}^{\infty} f(y)\delta(x-y)\, dy = \int_{-\infty}^{\infty} f(x-y)\delta(y)\, dy = f(x) \tag{3.67}$$

□

Convolution is a central concept in the study of linear systems, optics, electronics, ODEs and PDEs. The key reason for this is the following, which is the "theorem of the course":

Theorem 3.10 (Convolution theorem)**.** *The Fourier transform of the convolution of two functions is the product of the Fourier transforms. If $f(x)$ and $g(x)$ are two functions then*

$$\mathcal{F}[f \star g] = \mathcal{F}[f] \cdot \mathcal{F}[g]. \tag{3.68}$$

Proof. Write down the integrals explicitly and use the transformation of variables $z = x - y$ with $dx\, dy = dz\, dy$

$$\begin{aligned}
\mathcal{F}[f \star g](\omega) &= \int_{-\infty}^{\infty} e^{-i\omega x}(f \star g)(x)\, dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\omega x} f(y) g(x-y)\, dx\, dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\omega y} e^{-i\omega(x-y)} f(y) g(x-y)\, dx\, dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\omega y} f(y) e^{-i\omega z} g(z)\, dz\, dy \\
&= \int_{-\infty}^{\infty} e^{-i\omega y} f(y)\, dy \int_{-\infty}^{\infty} e^{-i\omega z} g(z)\, dz \\
&= \mathcal{F}[f](\omega) \cdot \mathcal{F}[g](\omega)
\end{aligned} \tag{3.69}$$

□

Note that this allows us to calculate the convolution of two functions as follows:

$$f \star g = \mathcal{F}^{-1}\left[\mathcal{F}[f] \cdot \mathcal{F}[g]\right]. \tag{3.70}$$

It is easy to see that if we want to calculate the convolution of two functions on a computer, then naively this would take $\mathcal{O}(n^2)$ operations if the functions are sampled at n points. This can be reduced to $\mathcal{O}(n \log n)$ with the representation in (3.70). The reason for this is that the Fourier transform (and its inverse) can be evaluated in $\mathcal{O}(n \log n)$ operations. The multiplication in Fourier space only requires an additional $\mathcal{O}(n)$ operations.

Example 3.10. Consider $f(x) = p_\alpha(x)$, $g(x) = p_\beta(-x)$ as defined in (3.13). To calculate the convolution of those two functions we first write down the Fourier transforms of f and g as

$$\mathcal{F}[f](\omega) = \frac{1}{\alpha + i\omega}, \qquad \mathcal{F}[g](\omega) = \frac{1}{\beta - i\omega} \tag{3.71}$$

Expand the product using partial fractions

$$\begin{aligned}\mathcal{F}[f \star g](\omega) = \mathcal{F}[f] \cdot \mathcal{F}[g] &= \frac{1}{\alpha + i\omega} \cdot \frac{1}{\beta - i\omega} = \frac{1}{\alpha + \beta}\left(\frac{1}{\alpha + i\omega} + \frac{1}{\beta - i\omega}\right) \\ &= \frac{\mathcal{F}[f](\omega) + \mathcal{F}[g](\omega)}{\alpha + \beta}\end{aligned} \tag{3.72}$$

and we can read off

$$(f \star g)(x) = \frac{f(x) + g(x)}{\alpha + \beta} = \frac{1}{\alpha + \beta} \cdot \begin{cases} e^{\beta x} & \text{for } x < 0 \\ e^{-\alpha x} & \text{for } x \geq 0 \end{cases} \tag{3.73}$$

Can you see how this is related to the Fourier transform in (3.16) in the case $\alpha = \beta$?

Application: deblurring of an image. Let f be the picture (signal) and g a blurring function (for example a Gaussian with width σ, $g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2}\right]$). Then the blurred image is $h = f \star g$. If we only have the blurred image, how can we reconstruct the original, unblurred image? This is covered by the following Corrollary.

Corollary 3.4. *If $h = f \star g$, then*

$$f = \mathcal{F}^{-1}\left[\frac{\mathcal{F}[h]}{\mathcal{F}[g]}\right] = \mathcal{F}^{-1}\left[\frac{\mathcal{F}[f \star g]}{\mathcal{F}[g]}\right]. \tag{3.74}$$

Proof. This follows immediately from Theorem 3.10 by applying the Fourier transform to both sides of (3.74) and multiplying by $\mathcal{F}[g]$. □

Note that this only works if $\mathcal{F}[g](\omega) > 0$ for all ω.

3.5 Using the Fourier transform to solve ODEs

Many mechanical and electrical systems can be expressed as solutions $u(x)$ of linear ODEs with right hand side $h(x)$. Although the methods discussed here can in principle be applied to systems of arbitrary order, here we only consider first- and second-order ODEs:

$$\begin{aligned} a\frac{du}{dx} + bu &= h(x) \qquad \text{(first order)} \\ a\frac{d^2u}{dx^2} + b\frac{du}{dx} + cu &= h(x) \qquad \text{(second order).}\end{aligned} \tag{3.75}$$

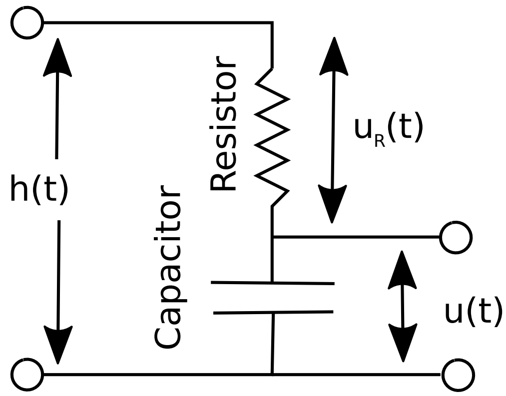


Figure 3.4: Electrical circuit which can be represented by (3.76)

In some cases the independent variable will be time t instead of x, but of course the method is the same. For example, consider the circuit shown in Fig. 3.4 and study its evolution over time. $h(t)$ is the time-dependent total electric potential across the resistor and capacitor. If u_R denotes the potential across the resistor and u the potential measured at the capacitor, then $h = u + u_R$. We also know that $u_R = RI$ (Ohm's law) and that the current is $I = C\frac{du}{dt}$, therefore $u_R = RC\frac{du}{dt}$ and we get the time dependent first order ODE

$$RC\frac{du}{dt} + u = h(t). \tag{3.76}$$

In the following we will always assume that the solutions of (3.75) have Fourier transforms; this excludes exponentially growing solutions.

We now look at a generalisation of the first order ODE in (3.76).

3.5.1 First order equations

Consider

$$a\frac{du}{dx} + bu = h(x) \qquad \text{with constant } a, b \in \mathbb{R},\ a, b > 0. \tag{3.77}$$

To solve this ODE for u for a given right hand side $h(x)$, take the Fourier transform of both sides and apply Theorem 3.5 to give

$$\begin{aligned} a\mathcal{F}\left[\frac{du}{dx}\right](\omega) + b\mathcal{F}[u](\omega) &= \mathcal{F}[h](\omega) \\ \Rightarrow \qquad ai\omega\mathcal{F}[u](\omega) + b\mathcal{F}[u](\omega) &= \mathcal{F}[h](\omega) \\ \Rightarrow \mathcal{F}[u](\omega) &= \frac{\mathcal{F}[h](\omega)}{b + a\omega i} \end{aligned} \tag{3.78}$$

Now, suppose there is a function $g(x)$ with $\mathcal{F}[g] = G(\omega) = \frac{1}{b+a\omega i}$. From (3.78) and the convolution theorem it then follows that

$$u = h \star g. \tag{3.79}$$

In this context we call

- $g(x)$: transfer function/impulse response
- $G(\omega)$: frequency response.

In other words, once we know the transfer function, we can calculate the solution for any $h(x)$ with the convolution in (3.79).

The name *impulse response* comes from the following

Lemma 3.2. *If $h(x) = \delta(x)$ then the solution of (3.77) is $g(x)$ with $\mathcal{F}[g](\omega) = \frac{1}{b+a\omega i}$.*

Proof. This follows immediately from the fact that $\mathcal{F}[h] = \mathcal{F}[\delta] = 1$ and applying the inverse Fourier transform to the last line of (3.78). □

For the ODE in (3.77) we can calculate the transfer function exactly:

Lemma 3.3. *The transfer function of the first order ODE*

$$a\frac{du}{dx} + bu = h(x) \tag{3.80}$$

with $a, b > 0$ is

$$g(x) = \frac{1}{a}p_{b/a}(x) = \begin{cases} \frac{1}{a}\exp\left[-\frac{bx}{a}\right] & \text{for } x > 0 \\ \frac{1}{2a} & \text{for } x = 0 \\ 0 & \text{for } x < 0. \end{cases} \tag{3.81}$$

Proof. This follows immediately from the fact that the frequency response is

$$G(\omega) = \frac{1}{b + a\omega i} = \frac{1}{a} \cdot \frac{1}{\frac{b}{a} + i\omega} \tag{3.82}$$

and Example 3.2. □

Note that $g(x) > 0$ only for $x \geq 0$. This is called a *causal* impulse since in this case

$$u(x) = (g \star h)(x) = \int_0^\infty g(y)h(x-y)\,dy \qquad \text{(note the lower bound of 0 instead of } \infty) \tag{3.83}$$

and hence only values $h(z)$ of the forcing with $z = x - y < x$ contribute to the solution at time x.

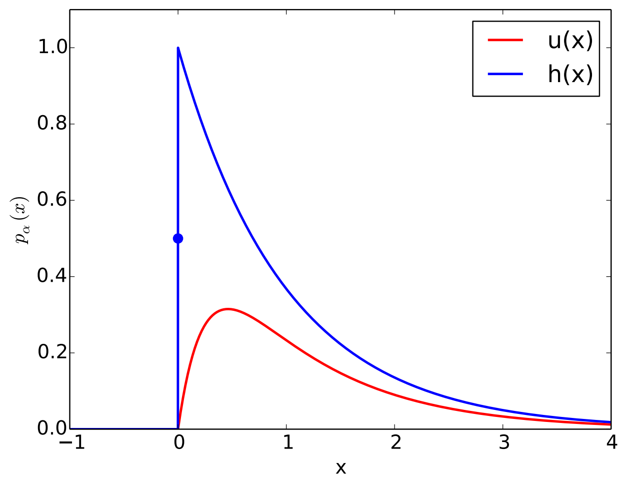


Figure 3.5: Solution of the first order ODE in (3.77).

Example 3.11. Let

$$h(x) = p_\gamma(x) = \begin{cases} e^{-\gamma x} & \text{for } x > 0 \\ \frac{1}{2} & \text{for } x = 0 \\ 0 & \text{for } x < 0. \end{cases} \tag{3.84}$$

Then $\mathcal{F}[h](\omega) = \frac{1}{\gamma + \omega i}$ and

$$\begin{aligned} \mathcal{F}[u](\omega) &= \frac{1}{\gamma + \omega i} \cdot \frac{1}{b + a\omega i} = \frac{1}{b - a\gamma}\left(\frac{1}{\gamma + \omega i} - \frac{1}{\frac{b}{a} + \omega i}\right) \\ &= \frac{1}{b - a\gamma}\left(\mathcal{F}[p_\gamma] - \mathcal{F}[p_{b/a}]\right) \\ \Rightarrow \qquad u(x) &= \begin{cases} \frac{1}{b - a\gamma}\left(\exp\left[-\gamma x\right] - \exp\left[-\frac{bx}{a}\right]\right) & \text{for } x > 0 \\ 0 & \text{for } x \leq 0. \end{cases} \end{aligned} \tag{3.85}$$

The solution $u(x)$ and the forcing function $h(x)$ are plotted in Fig. 3.5. Note that this is only the solution if $b \neq a\gamma$. For $\gamma = b/a$ we have a **resonance** and the solution has a different form. In this case

$$\mathcal{F}[u](\omega) = \frac{1}{a}\frac{1}{(\gamma + \omega i)^2} \tag{3.86}$$

and the expansion in (3.85) will not work. However, with Example 3.6 we obtain

$$u(x) = \frac{x}{a} p_\gamma(x) = \begin{cases} \frac{x}{a} e^{-\gamma x} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0. \end{cases} \tag{3.87}$$

3.5.2 Second order equations

We can study equations of the form

$$a\frac{d^2u}{dx^2} + b\frac{du}{dx} + cu = h(x) \qquad \text{with constant } a, b, c \in \mathbb{R},\ a, b, c > 0. \tag{3.88}$$

in exactly the same way. This equation describes the motion of a damped oscillator, where b is proportional to the damping force. Fourier-transforming both sides of (3.88) using Theorem 3.5 and Corollary 3.1 we get

$$\begin{aligned} -a\omega^2\mathcal{F}[u] + i\omega b\mathcal{F}[u] + c\mathcal{F}[u] &= \mathcal{F}[h] \\ \Rightarrow \qquad \mathcal{F}[u](\omega) &= \frac{\mathcal{F}[h](\omega)}{-a\omega^2 + b\omega i + c} \end{aligned} \tag{3.89}$$

So, as before,

$$u = h \star g, \qquad \qquad \mathcal{F}[g](\omega) = \frac{1}{-a\omega^2 + b\omega i + c}. \tag{3.90}$$

We now expand the Fourier-space transfer function as

$$\mathcal{F}[g](\omega)\frac{1}{-a\omega^2 + b\omega i + c} = \frac{1}{a(\omega_- - \omega_+)}\left(\frac{1}{\omega - \omega_+} - \frac{1}{\omega - \omega_-}\right) \tag{3.91}$$

where

$$\omega_\pm = \frac{ib}{2a}\left(1 \pm \sqrt{1 - \frac{4ac}{b^2}}\right) = \frac{ib}{2a}\left(1 \pm \sqrt{D}\right) \qquad \text{with } D = 1 - \frac{4ac}{b^2} \tag{3.92}$$

are the solutions to the quadratic equation $-a\omega^2 + b\omega i + c = 0$. Note that since b is related to the damping, very strong damping ($b \gg 1$) will result in $D \to 1$ and very weak damping ($b \to 0$) in $D \to -\infty$. There is also a special case $b = 2\sqrt{ac}$ for which $D = 0$. It turns out that in this case the solution converges to $u = 0$ most rapidly. In total, there are three cases, depending on the sign of D:

Case I (overdamped) $b^2 > 4ac$, $D > 0$.
There are two different purely imaginary solutions,

$$\omega_\pm = \rho_\pm i, \quad \rho_\pm = \frac{b}{2a}\left(1 \pm \sqrt{1 - \frac{4ac}{b^2}}\right), \quad \omega_- - \omega_+ = -\frac{ib}{a}\sqrt{1 - \frac{4ac}{b^2}} = -i\frac{b}{a}\sqrt{D}. \tag{3.93}$$

Then

$$G(\omega) = \frac{i}{b\sqrt{D}}\left(\frac{1}{\omega - i\rho_+} - \frac{1}{\omega - i\rho_-}\right) = \frac{1}{b\sqrt{D}}\left(\frac{1}{\rho_- + i\omega} - \frac{1}{\rho_+ + i\omega}\right) = \frac{\mathcal{F}[p_{\rho_-}](\omega) - \mathcal{F}[p_{\rho_+}](\omega)}{b\sqrt{D}} \tag{3.94}$$

Therefore we have that $g(x) = 0$ for $x \leq 0$ and

$$\begin{aligned} g(x) &= \frac{1}{b\sqrt{D}}\left(\exp\left[-\rho_- x\right] - \exp\left[-\rho_+ x\right]\right) \\ &= \frac{1}{b\sqrt{D}}\exp\left[-\frac{b}{2a}x\right]\underbrace{\left(\exp\left[\frac{b}{2a}\sqrt{D}x\right] - \exp\left[-\frac{b}{2a}\sqrt{D}\right]\right)}_{2\sinh\left(\frac{b}{2a}\sqrt{D}x\right)} \\ &= \frac{2}{b\sqrt{1-\frac{4ac}{b^2}}}\exp\left[-\frac{b}{2a}x\right]\sinh\left(\frac{b}{2a}\sqrt{1-\frac{4ac}{b^2}}x\right) \qquad \text{for } x > 0. \end{aligned} \tag{3.95}$$

Case II (optimal damping) $b^2 = 4ac$, $D = 0$.
There is only one solution, which is purely imaginary

$$\omega_\pm = \frac{b}{2a}i \tag{3.96}$$

and hence

$$G(\omega) = -\frac{1}{a(\omega - \frac{b}{2a}i)^2} = \frac{1}{a}\cdot\frac{1}{(\frac{b}{2a}+\omega i)^2}. \tag{3.97}$$

We can use Example 3.6 to obtain:

$$g(x) = \frac{x}{a}p_{\frac{b}{2a}}(x) = \begin{cases} \frac{x}{a}\exp\left[-\frac{b}{2a}x\right] & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases} \tag{3.98}$$

Case III (underdamped) $b^2 < 4ac$, $D < 0$.
There are two distinct solutions which have both a real and an imaginary part,

$$\omega_\pm = \frac{b}{2a}i \pm \varpi, \qquad \omega_- - \omega_+ = -2\varpi, \qquad \varpi \equiv \frac{b}{2a}\sqrt{\frac{4ac}{b^2}-1} = \frac{b}{2a}\sqrt{-D}. \tag{3.99}$$

$$\begin{aligned} G(\omega) &= \frac{1}{2a\varpi}\left(\frac{1}{\omega+\varpi-\frac{b}{2a}i} - \frac{1}{\omega-\varpi-\frac{b}{2a}i}\right) = \frac{i}{2a\varpi}\left(\frac{1}{\frac{b}{2a}+(\omega+\varpi)i} - \frac{1}{\frac{b}{2a}+(\omega-\varpi)i}\right) \\ &= \frac{i}{2a\varpi}\left(\mathcal{F}[p_{\frac{b}{2a}}](\omega+\varpi) - \mathcal{F}[p_{\frac{b}{2a}}](\omega-\varpi)\right) \end{aligned} \tag{3.100}$$

Therefore we find with Theorem 3.3

$$\begin{aligned} g(x) &= \frac{i}{2a\varpi}\left(e^{-i\varpi x} - e^{i\varpi x}\right)p_{\frac{b}{2a}}(x) = \frac{\sin(\varpi x)}{a\varpi}p_{\frac{b}{2a}}(x) \\ &= \begin{cases} \frac{2}{b\sqrt{\frac{4ac}{b^2}-1}}\sin\left(\frac{b}{2a}\sqrt{\frac{4ac}{b^2}-1x}\right)\exp\left[-\frac{b}{2a}x\right] & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases} \end{aligned} \tag{3.101}$$

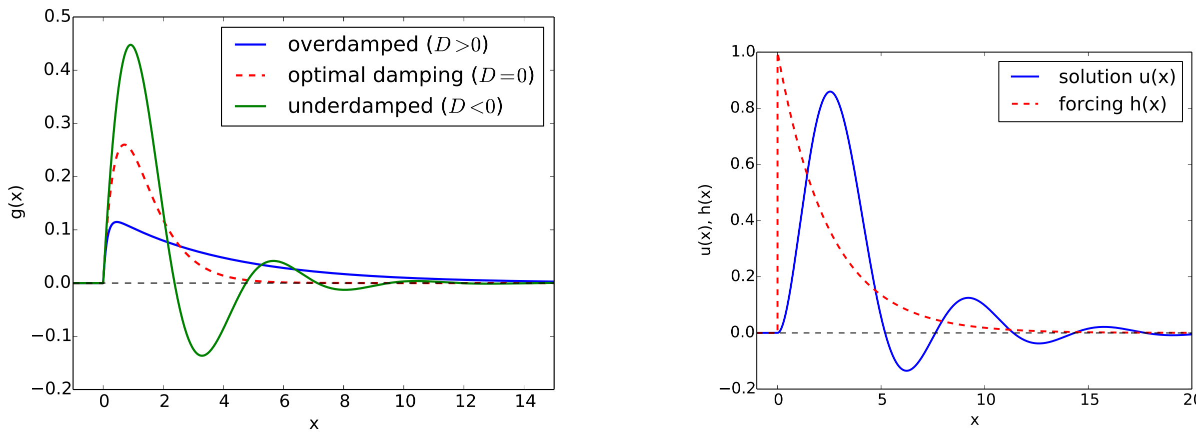


Figure 3.6: Transfer functions $g(x)$ for the damped oscillator (left) and solution $u(x)$ for the forcing $h(x)$ in the underdamped case in Example 3.12 (right).

In contrast to the other two cases, the solution oscillates between positive and negative values. The transfer functions for all three cases are shown in Fig. 3.6. Now that we have worked out the general form of the transfer function we can solve the second order ODE (3.88) for a particular right hand side $h(x)$.

Example 3.12. Consider the *underdamped* oscillator with forcing

$$h(x) = p_\gamma(x) = \begin{cases} e^{-\gamma x} & \text{for } x > 0 \\ \frac{1}{2} & \text{for } x = 0 \\ 0 & \text{for } x < 0. \end{cases} \tag{3.102}$$

The solution of the second order ODE in (3.88) can then be obtained from the transfer function in (3.100), again using partial fraction expansion. The calculation is tedious but straightforward:

$$\begin{aligned} \mathcal{F}[u](\omega) = G(\omega)\mathcal{F}[h](\omega) &= \frac{i}{2a\overline{\omega}}\left(\frac{1}{\frac{b}{2a}+(\omega+\overline{\omega})i} - \frac{1}{\frac{b}{2a}+(\omega-\overline{\omega})i}\right)\frac{1}{\gamma+\omega i} \\ &= \frac{i}{2a\overline{\omega}}\left[\frac{1}{\gamma-\frac{b}{2a}-\overline{\omega}i}\left(\frac{1}{\frac{b}{2a}+(\omega+\overline{\omega})i} - \frac{1}{\gamma+\omega i}\right)\right. \\ &\quad - \left.\frac{1}{\gamma-\frac{b}{2a}+\overline{\omega}i}\left(\frac{1}{\frac{b}{2a}+(\omega-\overline{\omega})i} - \frac{1}{\gamma+\omega i}\right)\right] \end{aligned} \tag{3.103}$$

If we introduce the notation $1/(\gamma - \frac{b}{2a} - \overline{\omega} i) \equiv Ce^{i\theta}$ for simplicity, this implies

$$\begin{aligned}
\mathcal{F}[u](\omega) &= \frac{Ci}{2a\overline{\omega}} \left[e^{i\theta} \mathcal{F}[p_{\frac{b}{2a}}](\omega + \overline{\omega}) - e^{-i\theta} \mathcal{F}[p_{\frac{b}{2a}}](\omega - \overline{\omega}) + \left(e^{-i\theta} - e^{i\theta} \right) \mathcal{F}[p_\gamma](\omega) \right] \\
\Rightarrow \qquad u(x) &= \frac{Ci}{2a\overline{\omega}} \left[\left(e^{-i(\overline{\omega}x - \theta)} - e^{i(\overline{\omega}x-\theta)} \right) p_{\frac{b}{2a}}(x) + \left(e^{-i\theta} - e^{i\theta} \right) p_\gamma(x) \right] \\
&= \frac{C}{a\overline{\omega}} \left[\sin(\overline{\omega}x - \theta) p_{\frac{b}{2a}}(x) + \sin(\theta) p_\gamma(x) \right] \\
&= \begin{cases} \frac{C}{a\overline{\omega}} \left[\sin(\overline{\omega}x - \theta) \exp\left[-\frac{b}{2a} x \right] + \sin(\theta) \exp\left[-\gamma x \right] \right] & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases}
\end{aligned} \tag{3.104}$$

This solution is shown in Fig. 3.6 (right).

3.6 Plancheral's Theorem

This theorem gives a connection between the size of a function and its Fourier-transform. For this define the L_2 norm

$$||u||_2^2 = \int_{-\infty}^{\infty} |u(x)|^2 \, dx, \qquad\qquad ||v||_2^2 = \int_{-\infty}^{\infty} |v(\omega)|^2 \, d\omega. \tag{3.105}$$

Theorem 3.11 (Plancherel). *Iff $||u||_2^2 < \infty$ then $||\mathcal{F}[u]||_2^2 < \infty$ and*

$$||u||_2^2 = \frac{1}{2\pi} ||\mathcal{F}[u]||_2^2. \tag{3.106}$$

If we interpret functions as elements of an (infinite dimensional) vector space, then the Fourier-transform preserves the length of those vectors up to a constant factor $1/(2\pi)$. For finite dimensional vectors $\in \mathbb{R}^n$ an example of a length-preserving linear operator is a rotation. For complex-valued functions, a length-preserving operation is called "unitary".

Proof. Define $v(x) \equiv u^*(-x)$.

$$||u||_2^2 = \int_{-\infty}^{\infty} u(x) u^*(x) \, dx = \int_{-\infty}^{\infty} u(x) v(-x) \, dx = (u \star v)(0) \tag{3.107}$$

But from Theorem 3.10 and with the definition of the (inverse) Fourier transform in (3.2)

$$\begin{aligned}
u \star v &= \mathcal{F}^{-1} \left[\mathcal{F}[u] \cdot \mathcal{F}[v] \right] \\
\Rightarrow \qquad (u \star v)(0) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega \cdot 0} \mathcal{F}[u](\omega) \mathcal{F}[v](\omega) \, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}[u](\omega) \mathcal{F}[v](\omega) \, d\omega
\end{aligned} \tag{3.108}$$

The Fourier transform of v is (transforming variables $x' = -x$):

$$\mathcal{F}[v](\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} u^*(-x)\, dx = \int_{-\infty}^{\infty} e^{i\omega x'} u^*(x')\, dx' = \mathcal{F}[u]^*(\omega) \tag{3.109}$$

This implies that

$$(u \star v)(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}[u](\omega)\mathcal{F}[u]^*(\omega)\, d\omega = \frac{1}{2\pi} ||\mathcal{F}[u]||_2^2. \tag{3.110}$$

□

Example 3.13. Imagine we want to calulate the integral

$$\int_{-\infty}^{\infty} \operatorname{sinc}^2(\omega)\, d\omega = \int_{-\infty}^{\infty} \frac{\sin^2(\omega)}{\omega^2}\, d\omega. \tag{3.111}$$

Recall from Example 3.4 that the sinc function is the Fourier-transform of the rectangular pulse R_2 with width 2, i.e. $\mathcal{F}[R_2](\omega) = \operatorname{sinc}(\omega)$. The definition of R_2 can be found in Eq. (3.21) and is $R_2(x) = \frac{1}{2}$ for $|x| < 1$ and $R_2(x) = 0$ for $|x| > 1$. With this we find for the integral in Eq. (3.111), using Plancherel's Theorem:

$$\int_{-\infty}^{\infty} \frac{\sin^2(\omega)}{\omega^2}\, d\omega = ||\mathcal{F}[R_2]||_2^2 = 2\pi ||R_2||_2^2 = 2\pi \int_{-\infty}^{\infty} R_2^2(x)\, dx = 2\pi \int_{-1}^{1} \frac{1}{4}\, dx = \pi. \tag{3.112}$$

3.7 Solving PDEs with the Fourier transform

We will be looking at PDEs for the rest of this course. The Fourier transform gives a powerful way to solve *linear* PDEs with *constant coefficients* which is widely used in analysis and applications, for example

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad\qquad \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + a\frac{\partial u}{\partial x} \tag{3.113}$$

with suitable initial-/boundary-conditions. The idea to solve those equations is as follows:

1. Start with a PDE in (x, t) (or (x, y)) with solution $u(x, t)$, usually defined over $-\infty < x < \infty$.

2. Take the Fourier transform with respect to x to obtain $F(\omega, t)$.

3. Solve the corresponding ODE for $F(\omega, t)$ in t for all ω and apply initial conditions.

4. Fourier transform back to obtain the solution $u(x, t)$

We will look at two important examples:

Example 3.14 (Heat equation)**.** Consider

$$\begin{aligned} \frac{\partial u}{\partial t} &= k\frac{\partial^2 u}{\partial x^2}, \qquad \text{for } -\infty < x < \infty \text{ and } t > 0 \\ u(x,t=0) &= h(x) \qquad \text{where } k \in \mathbb{R},\ k > 0. \end{aligned} \tag{3.114}$$

$u(x,t)$ describes the temperature or density of the diffusing material for a given initial distribution $u(x,0) = h(x)$. Take the Fourier transform of (3.114) and use Theorem 3.5 to obtain an ODE for $F(\omega,t) = \mathcal{F}[u](\omega)$

$$\begin{aligned} \frac{\partial F}{\partial t} &= -\omega^2 kF, \qquad \text{with } -\infty < \omega < \infty. \\ F(\omega,0) &= \mathcal{F}[h](\omega) \end{aligned} \tag{3.115}$$

This is an ODE in t which is easily solved to give

$$F(\omega,t) = \mathcal{F}[h](\omega)\exp\left[-k\omega^2 t\right] \equiv \mathcal{F}[h](\omega)\mathcal{F}[g](\omega) \tag{3.116}$$

Using Example 3.3 with $\sigma^2 = 2kt$ it follows that

$$g(x,t) = \frac{1}{\sqrt{4\pi kt}}\exp\left[-\frac{x^2}{4kt}\right] \tag{3.117}$$

and hence

$$\begin{aligned} u(x,t) = (h\star g)(x,t) &= \frac{1}{\sqrt{4\pi kt}}\int_{-\infty}^{\infty}\exp\left[-\frac{(z-x)^2}{4kt}\right]h(z)\,dz \\ &\equiv \int_{-\infty}^{\infty} K(x-z,t)h(z)\,dz \end{aligned} \tag{3.118}$$

The function $K(x,t) = g(x,t)$ in (3.118) is called the **heat kernel**. As a function of x it is a Gaussian with width $\sigma = \sqrt{2kt}$, i.e. the width grows with the square root of time, which is typical for diffusion processes. The function $K(x,t)$ solves (3.114) for $h(x) = \delta(x)$. It is used for example in the Black-Scholes formula in economics. A plot of $K(x,t)$ for different times is shown in Fig. 3.7.

Example 3.15 (Laplace equation on the half-plane)**.** The electric field above a charged plate $u(x,y)$ is given by

$$\begin{aligned} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= 0, \qquad \text{for } -\infty < x < \infty,\ y > 0 \\ u(x,0) &= h(x) \end{aligned} \tag{3.119}$$

Here $h(x)$ is the potential on the plate $y = 0$; we also assume that u is bounded as $|x|, y \to \infty$. Taking the Fourier transform in x we obtain

$$-\omega^2 F + \frac{\partial^2 F}{\partial y^2} = 0, \qquad\qquad F(\omega,0) = \mathcal{F}[h](\omega) \tag{3.120}$$

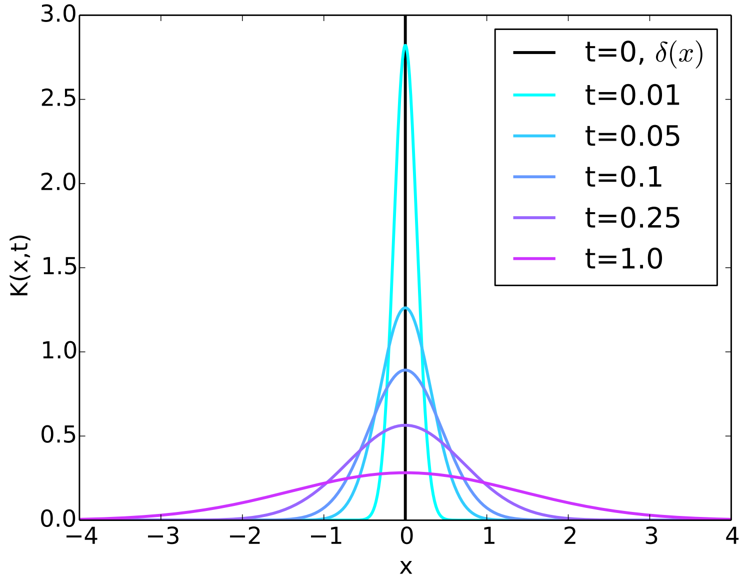


Figure 3.7: Heat kernel $K(x,t)$ defined by $g(x,t)$ in (3.117).

The most general solution of the first equation is

$$F(\omega, y) = Ae^{-\omega y} + Be^{\omega y} \tag{3.121}$$

However, depending on the sign of ω, only one of those two contributions does not diverge and therefore

$$\begin{aligned} F(\omega, y) &= \begin{cases} \mathcal{F}[h](\omega)e^{-\omega y} & \text{for } \omega \geq 0 \\ \mathcal{F}[h](\omega)e^{\omega y} & \text{for } \omega < 0 \end{cases} \\ &= \mathcal{F}[h](\omega)e^{-|\omega| y}. \end{aligned} \tag{3.122}$$

We need to find $g(x,y)$ such that $\mathcal{F}[g](\omega, y) = e^{-|\omega|y}$. For this, use the inverse Fourier transform to obtain

$$\begin{aligned} g(x,y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega x}e^{-|\omega|y}\, d\omega = \frac{1}{2\pi}\left(\int_{-\infty}^{0} e^{(ix+y)\omega}\, d\omega + \int_{0}^{\infty} e^{(ix-y)\omega}\, d\omega\right) \\ &= \frac{1}{2\pi}\left(\frac{1}{y+ix} + \frac{1}{y-ix}\right) = \frac{1}{\pi}\cdot\frac{y}{x^2+y^2} \end{aligned} \tag{3.123}$$

It is easy to show that $g(x,y)$ is constant on circles of radius R with centre at $(x = 0, y = R)$ and that $g(x,y) = \frac{1}{2\pi R}$ on those circles. Based on this the solution for a given $h(x)$ can be obtained via convolution

$$u(x,y) = \int_{-\infty}^{\infty} g(z-x,y)h(z)\, dz = \frac{y}{\pi}\int_{-\infty}^{\infty} \frac{h(z)}{(z-x)^2+y^2}\, dz \tag{3.124}$$

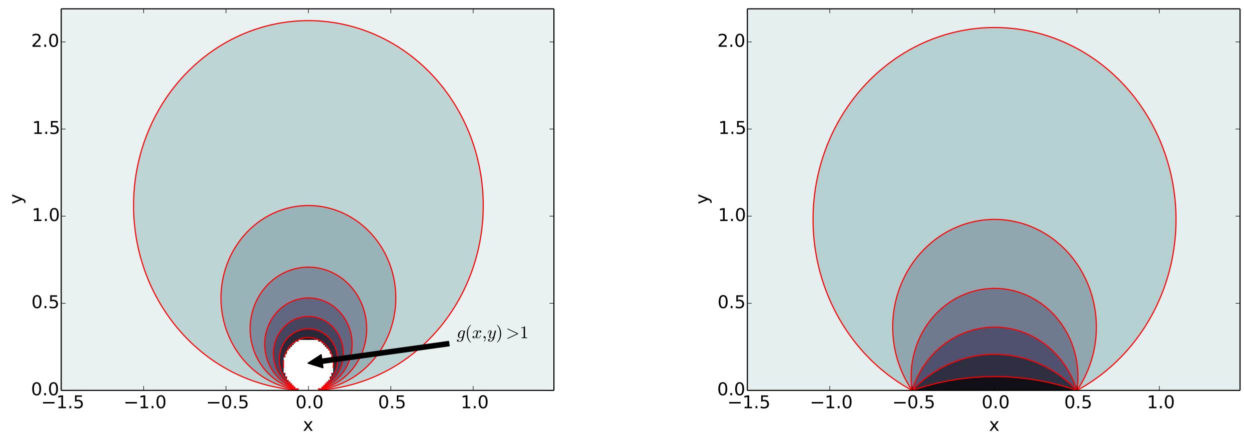


Figure 3.8: Contour plot of the fundamental solution $g(x, y)$ of the Laplace equation in a half-plane (left) and solution for $h(x)$ as defined in (3.125) for $a = 0.5$.

Example 3.16. Suppose that the potential on the plate is fixed such that

$$u(x, 0) = h(x) = \begin{cases} 1 & \text{for } |x| < a \\ 0 & \text{otherwise.} \end{cases} \tag{3.125}$$

Then we obtain the solution by convolution $u = h \star g$ under the change of variable $z' = x - z$

$$\begin{aligned} u(x, y) &= \frac{y}{\pi} \int_{-a}^{a} \frac{1}{(z - x)^2 + y^2}\, dz = \frac{y}{\pi} \int_{x-a}^{x+a} \frac{1}{z'^2 + y^2}\, dz' \\ &= \frac{1}{\pi} \left[\arctan \left(\frac{z}{y} \right) \right]_{x-a}^{x+a} = \frac{1}{\pi} \left(\arctan \left(\frac{x + a}{y} \right) - \arctan \left(\frac{x - a}{y} \right) \right). \\ &= \frac{1}{\pi} \arctan \left(\frac{2ay}{y^2 + (x^2 - a^2)} \right) \end{aligned} \tag{3.126}$$

Again it can be shown that the contour lines of $u(x, y)$ are circles.

Note: You will meet these ideas again in the Methods II course when you look at Greens functions which satisfy higher-dimensional PDEs with $\delta(\boldsymbol{x})$ functions as source terms.

Chapter 4

Quasilinear first order PDEs

Aims:

- To introduce first order quasilinear PDEs
- To introduce the concepts of contour-lines and data curves
- To solve these using the method of characteristics
- To understand the effects of initial data and domains of dependence

Since the PDEs we study in this chapter will depend on several variables, we will use the notation

$$u_x = \frac{\partial u}{\partial x}, \qquad u_y = \frac{\partial u}{\partial y}, \qquad u_t = \frac{\partial u}{\partial t} \qquad \text{etc.} \tag{4.1}$$

to denote partial derivatives. Sometimes we will also use

$$\dot{u} = \frac{\partial u}{\partial t} \tag{4.2}$$

for derivatives with respect to the variable t. The functions in this chapter will be real-valued.

4.1 Introduction

Definition 4.1. A *first order* PDE for a real-valued function $u(x, y)$ with independent variables $(x, y) \in \mathbb{R}^2$ takes the form

$$F(x, y, u, u_x, u_y) = 0 \tag{4.3}$$

where F is a real-valued functional (while the argument of a function is a variable such as x or y, a real-valued functional maps a function to a real number). The PDE in (4.3) is called

1. *linear* if the functional F is linear in u, u_x and u_y:
$$F(x, y, u, u_x, u_y) = a(x, y)u_x + b(x, y)u_y + c(x, y)u - d(x, y) = 0 \tag{4.4}$$
2. *quasi-linear* if F is a linear functional of u_x and u_y (but F is not necessarily linear in u):
$$F(x, y, u, u_x, u_y) = P(x, y, u)u_x + Q(x, y, u)u_y - R(x, y, u) = 0 \tag{4.5}$$
3. *fully non-linear* otherwise.

A special case of a linear PDE is if the coefficients are constant,
$$au_x + bu_y + cu = d(x, y). \tag{4.6}$$
In this case the PDE can be solved using the Fourier transform method from Chapter 3.

Example 4.1. The following two PDEs are fully nonlinear:
$$u_x u_y = u^2, \qquad u_x^2 + u_y^2 = 1. \tag{4.7}$$
The second equation is the Eikonal equation of optics. Nonlinear PDEs are usually very hard to solve and will not be considered in this course.

The following wave equation is linear, but of second order, and will therefore also not be considered here:
$$u_{xx} - c^2 u_{yy} = \frac{\partial^2 u}{\partial x^2} - c^2 \frac{\partial^2 u}{\partial y^2} = 0 \tag{4.8}$$
We will discuss this equation in Chapter 5.

Example 4.2. A famous quasi-linear equation is Burgers' equation
$$u_y + uu_x = c(x, y) \tag{4.9}$$
where u is the fluid velocity and y is time. This equation can be used to describe shocks in a fluid.

Example 4.3. Another example of a quasi-linear equation is
$$xu_y - yu_x = u^2. \tag{4.10}$$

In this chapter we will study *quasi-linear PDEs*. These have many applications for example in

- Fluid mechanics
- Meteorology

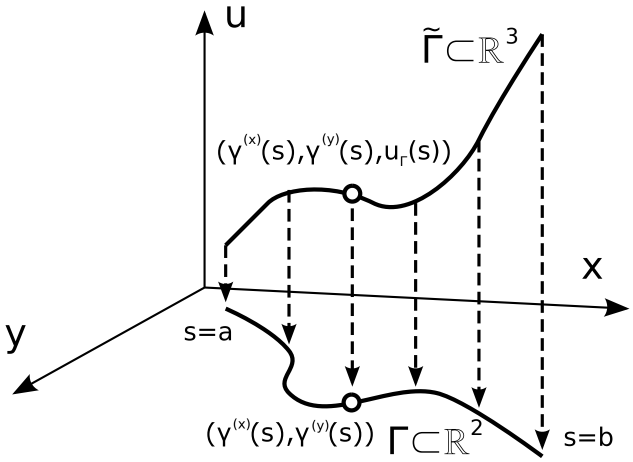


Figure 4.1: Data curve Γ and $\tilde{\Gamma}$

- Gas dynamics (detonations and explosions)

To uniquely solve a quasi-linear PDE we must provide initial data on a **data curve** $\Gamma \subset \mathbb{R}^2$

$$\Gamma = \{(x, y) : x = \gamma^{(x)}(s), y = \gamma^{(y)}(s), a \leq s \leq b\} \subset \mathbb{R}^2 \tag{4.11}$$

i.e. we need to specify the value of $u(x, y)$ for all $(x, y) \in \Gamma$. In other words, we need to define a curve $\tilde{\Gamma} \subset \mathbb{R}^3$ such that

$$\tilde{\Gamma} = \{(x, y, u) : x = \gamma^{(x)}(s), y = \gamma^{(y)}(s), u = u_\Gamma(s), a \leq s \leq b\} \subset \mathbb{R}^3 \tag{4.12}$$

Note that the projection of $\tilde{\Gamma}$ onto $\mathbb{R}^2$ is Γ (see Fig. 4.1) and in the following we will refer to both Γ and $\tilde{\Gamma}$ as "data curves". We will assume that Γ is smooth and it does not self-intersect.

Example 4.4. Consider the homogeneous Burgers' equation in (4.9)

$$u_y + uu_x = 0. \tag{4.13}$$

We can choose $\Gamma = \{(x, y) : x = s, y = 0, -\infty < s < \infty\}$ and $u_\Gamma(s) = \tanh(s)$. Since y can be interpreted as a time in this example, this is equivalent to prescribing an initial fluid velocity.

4.2 Contour lines

Definition 4.2. A **contour line** for a function $u(x, y)$ is a curve $\mathcal{C}_\kappa$ along which $u(x, y)$ is constant:

$$\mathcal{C}_\kappa = \{(x, y) : u(x, y) = \kappa\} \subset \mathbb{R}^2 \tag{4.14}$$

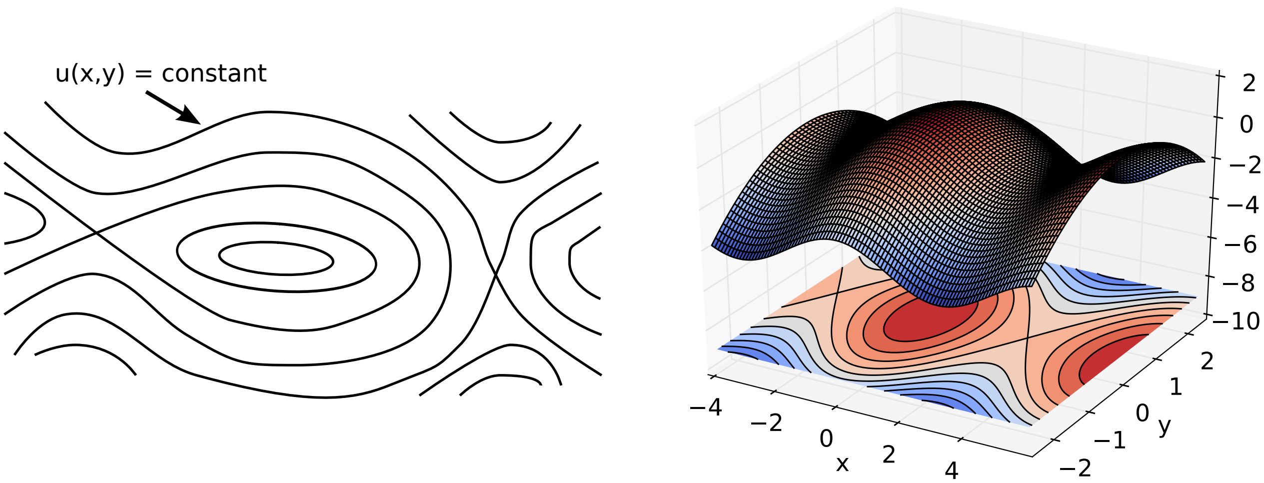


Figure 4.2: Schematic sketch of a family of contour lines (left) and contour lines of the function $u(x,y) = \cos(x) - \frac{1}{2}y^2$ (right).

Note that a contour line can self-intersect and can degenerate into a point. Usually we have a whole family of contour lines, $\mathcal{C} = \{\mathcal{C}_\kappa, \kappa \in S \subset \mathbb{R}\}$ where S is the set of all values that the function u can assume. It turns out that the contour lines can be found by solving a set of ordinary differential equations, as the following theorem shows.

Theorem 4.1. *Let the function $u(x,y)$ have the property that it is constant on any curves $(x(t),y(t))$ which satisfy the differential equations*

$$\frac{dx}{dt} = P(x,y,u), \qquad \frac{dy}{dt} = Q(x,y,u). \tag{4.15}$$

Then u satisfies the quasi-linear PDE

$$P(x,y,u)u_x + Q(x,y,u)u_y = 0. \tag{4.16}$$

Proof. By assumption, u is constant on each curve defined by (4.15), i.e the total derivative of u vanishes along a contour line:

$$\frac{du(x(t),y(t))}{dt} = 0. \tag{4.17}$$

The total derivative with respect to t can be evaluated with the chain rule

$$0 = \frac{du(x(t),y(t))}{dt} = \frac{dx}{dt}u_x + \frac{dy}{dt}u_y = P(x,y,u)u_x + Q(x,y,u)u_y \tag{4.18}$$

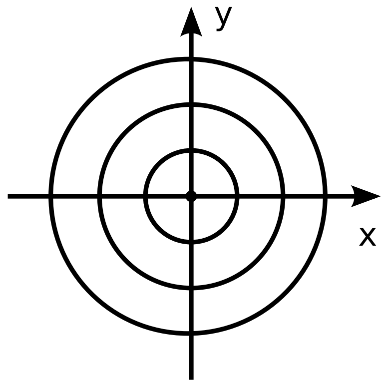


Figure 4.3: Concentric contour lines in Example 4.5.

where the last equality follows from (4.15). □

Example 4.5. Suppose that we consider the equation

$$yu_x - xu_y = 0 \tag{4.19}$$

i.e. $P(x, y, u) = y$ and $Q(x, y, u) = -x$. The contour lines satisfy the equations

$$\dot{x} = \frac{dx}{dt} = y, \qquad \dot{y} = \frac{dy}{dt} = -x. \tag{4.20}$$

To solve this system of two first-order coupled equations for $x(t)$ and $y(t)$ take the derivative of the first equation and insert $\dot{y}$ from the second equation to obtain

$$\frac{d^2x}{dt^2} + x = 0 \qquad \Rightarrow \qquad x = A\cos(t) + B\sin(t),\ y = -A\sin(t) + B\cos(t). \tag{4.21}$$

Since $x^2 + y^2 = A^2 + B^2 = const$ the contour lines are concentric circles around the point $(0, 0)$, see Fig. 4.3.

Given a data curve $\tilde{\Gamma}$ we can find the solution $u(x, y)$ of the quasi-linear PDE

$$P(x, y, u)u_x + Q(x, y, u)u_y = 0. \tag{4.22}$$

at a particular point (x, y) as follows:

1. Calculate the set of contour lines by solving (4.15).
2. Follow the contour line which contains (x, y) to the point $(\gamma^{(x)}(s), \gamma^{(y)}(s))$ where it intersects the data curve Γ.
3. Evaluate the function u_Γ for this s to obtain $u(x, y) = u(\gamma^{(x)}(s), \gamma^{(y)}(s)) = u_\Gamma(s)$.

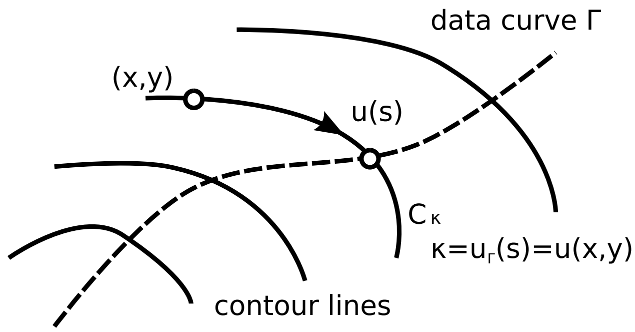


Figure 4.4: Calculating $u(x, y)$ by following contour lines $(x(t), y(t))$ and evaluating on the data curve.

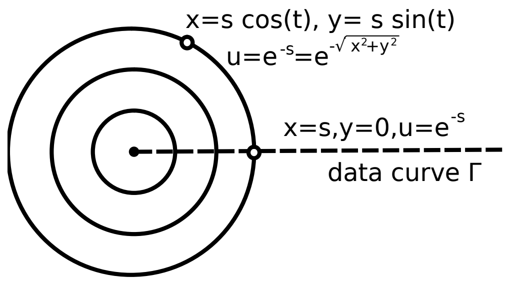


Figure 4.5: Calculation of $u(x, y)$ in Example 4.6

Note that this only works if the data curve Γ intersects every contour at least once. If this is not the case, it might only be possible to compute the solution in the part of the domain which is covered by contour lines which intersect the data curve.

Example 4.6. Consider again the equation in (4.19) and choose the data curve

$$\tilde{\Gamma} = \{(x, y, u) : x = s, y = 0, u = u_\Gamma(s) = e^{-s}, 0 < s < \infty\}. \tag{4.23}$$

Note that Γ is the positive x-axis and it intesects each contour line (circles around the origin) exactly once. Every point (x, y) can be written in polar coordinates (s, t) as $(x, y) = (s\cos(t), -s\sin(t))$ where $s = \sqrt{x^2 + y^2}$. Note that each s corresponds to a particular contour line with radius s and that following a contour to the data curve simply amounts to setting $t = 0$. Since we have $u_\Gamma(s) = e^{-s}$, we find that

$$u(x, y) = u_\Gamma(s) = u_\Gamma(\sqrt{x^2 + y^2}) = \exp\left[-\sqrt{x^2 + y^2}\right]. \tag{4.24}$$

It is easy to manually verify that this indeed satisfies (4.19):

$$u_x = -\frac{x}{\sqrt{x^2 + y^2}} \exp\left[-\sqrt{x^2 + y^2}\right], \qquad u_y = -\frac{y}{\sqrt{x^2 + y^2}} \exp\left[-\sqrt{x^2 + y^2}\right] \tag{4.25}$$

and therefore

$$yu_x - xu_y = \frac{yx}{\sqrt{x^2+y^2}} \exp\left[-\sqrt{x^2+y^2}\right] - \frac{xy}{\sqrt{x^2+y^2}} \exp\left[-\sqrt{x^2+y^2}\right] = 0. \tag{4.26}$$

Furthermore, on the data curve where $x = s$, $y = 0$ with $s > 0$: $u(s, 0) = \exp\left[-\sqrt{s^2}\right] = e^{-s}$.

Let's iterate the general method for solving $P(x, y, u)u_x + Q(x, y, u)u_y = 0$ in more detail:

Step 1: Set u to a constant value κ and solve the system of two coupled ODEs

$$\frac{dx}{dt} = P(x, y, \kappa), \qquad \frac{dy}{dt} = Q(x, y, \kappa) \tag{4.27}$$

to obtain the curve $(x_\kappa(t), y_\kappa(t))$. The solution will involve two arbitrary constants A and B.

Step 2: Set $t = 0$ and find A, B such that $(x_\kappa(t = 0), y_\kappa(t = 0), \kappa) = (\gamma^{(x)}(s), \gamma^{(y)}(s), u_\Gamma(s)) \in \tilde{\Gamma}$ for some $a \leq s \leq b$. For each s this gives a unique contour line $(x(t, s), y(t, s)) = (x_\kappa(t), y_\kappa(t))$.

Step 3: For a general point (x, y) invert the expressions

$$x = x(t, s), \qquad y = y(t, s) \tag{4.28}$$

to find $t = t(x, y)$ and $s = s(x, y)$. Then $u(x, y) = u_\Gamma(s(x, y))$ is the solution.

Example 4.7. Consider the equation and data curve

$$-x^2u_x + uu_y = 0, \qquad \tilde{\Gamma} = \{(x, y, u) : x = y = u = s, s > 0\} \tag{4.29}$$

Step 1: We have $P(x, y, u) = -x^2$ and $Q(x, y, u) = u$. Solve

$$\frac{dx}{dt} = -x^2, \qquad \frac{dy}{dt} = \kappa \tag{4.30}$$

to obtain the curves

$$A + \frac{1}{x_\kappa(t)} = t, \qquad y_\kappa(t) = B + \kappa t \tag{4.31}$$

with arbitrary constants A and B.

Step 2: Setting $t = 0$ and requiring that $(x_\kappa(0), y_\kappa(0), \kappa) \in \tilde{\Gamma}$, i.e. $x_\kappa(0) = y_\kappa(0) = \kappa = s$ results in $A = -1/s$, $B = s$. Therefore the contour lines satisfy

$$\frac{1}{x(t, s)} - \frac{1}{s} = t, \qquad y(t, s) = s + st \tag{4.32}$$

and are given explicitly as

$$x(t, s) = \frac{s}{1 + st}, \qquad y(t, s) = s(1 + t). \tag{4.33}$$

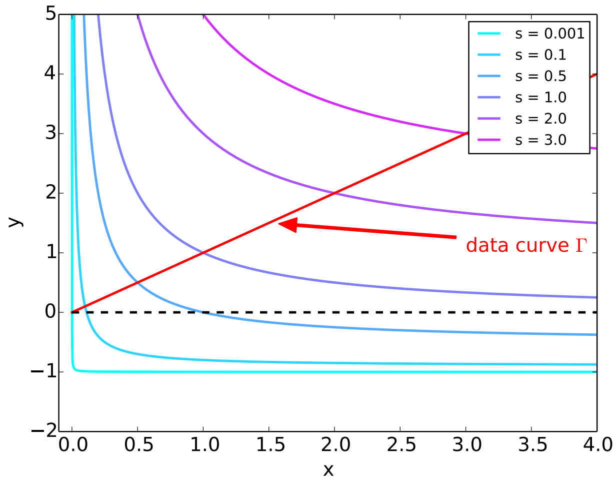


Figure 4.6: Contour lines and data curve Γ for Example 4.7.

Step 3: To find the solution for a general (x, y) we must invert (4.33). A straightforward calculation gives

$$s(x,y) = \frac{y+1}{x+1}x, \qquad t(x,y) = \frac{y}{x}\cdot\frac{x+1}{y+1} - 1. \tag{4.34}$$

Hence, for $x > 0$ and $y > -1$ the solution is

$$u(x,y) = u_\Gamma(s(x,y)) = s(x,y) = \frac{y+1}{x+1}x. \tag{4.35}$$

Exercise: Check that the solution in (4.35) really solves the PDE (4.29).

4.3 The method of characteristics

So far we have only considered equations with vanishing right hand side, $R(x, y, u) = 0$. The following approach generalises the previous method to general quasilinear equations of the form

$$P(x,y,u)u_x + Q(x,y,u)u_y = R(x,y,u). \tag{4.36}$$

Definition 4.3. A **characteristic curve** $(x(t), y(t), v(t))$ solves the ordinary differential equations

$$\frac{dx}{dt} = P(x,y,v), \qquad \frac{dy}{dt} = Q(x,y,v), \qquad \frac{dv}{dt} = R(x,y,v) \tag{4.37}$$

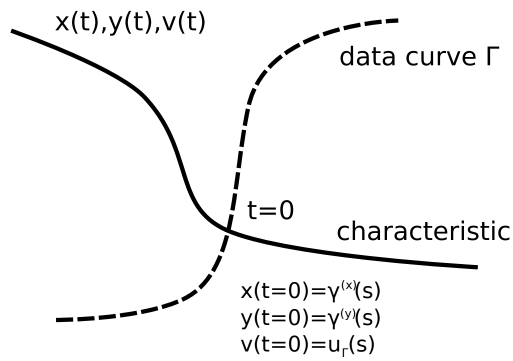


Figure 4.7: Solving equations with characteristics

Note that if $R(x, y, u) = 0$ the projection of characteristic curves onto the x, y-plane are contour lines.

Lemma 4.1. *Let $(x(t), y(t), v(t))$ be a characteristic which satisfies (4.37). Then the function $u(x, y)$ defined by $u(x(t), y(t)) \equiv v(t)$ solves (4.36) for all t on a characteristic.*

Proof. By the chain rule we have

$$\frac{dv}{dt} = \frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt}. \tag{4.38}$$

Replacing the time derivatives dv/dt, dx/dt and dy/dt according to (4.37) gives the desired result. □

We say that *"information propagates along a characteristic"*. To solve (4.36) with characteristics we can proceed as follows:

Step 1: Solve (4.37) to obtain $(x(t), y(t), v(t))$ with arbitrary integration constants A, B and C.

Step 2: Find the point s at which the characteristic intersects the data curve. Assume that $t = 0$ at this point and fix A, B and C from the initial data by requiring that

$$x(t = 0) = \gamma^{(x)}(s), \qquad y(t = 0) = \gamma^{(y)}(s), \qquad v(t = 0) = u_\Gamma(s). \tag{4.39}$$

Since this can be done for any s, the characteristic curves can be parametrised by s as $(x(t, s), y(t, s), v(t, s))$.

Step 3: Invert the expressions

$$x = x(t, s), \qquad y = y(t, s) \tag{4.40}$$

to obtain

$$t = t(x, y), \qquad s = s(x, y) \tag{4.41}$$

Then the solution for arbitrary x, y is given by

$$u(x, y) = v(t(x, y), s(x, y)). \tag{4.42}$$

Example 4.8. Consider the equation

$$xu_x + yu_y = u, \qquad \tilde{\Gamma} = \{(x, y, u) : x = s, y = s^2, u = s^3, s > 0\}. \tag{4.43}$$

Step 1: Solving

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = y, \qquad \frac{dv}{dt} = v \tag{4.44}$$

we obtain

$$x = Ae^t, \qquad y = Be^t, \qquad v = Ce^t. \tag{4.45}$$

Step 2: Set $t = 0$ and apply the initial data to obtain

$$x(0) = A = s, \qquad y(0) = B = s^2, \qquad v(0) = C = s^3 \tag{4.46}$$

and the characteristics can be written as

$$x(t, s) = se^t, \qquad y(t, s) = s^2e^t = sx(t, s), \qquad v(t, s) = s^3e^t = s^2x(t, s). \tag{4.47}$$

Step 3: Solving (4.47) for s and t gives

$$s(x, y) = \frac{y}{x}, \qquad t(x, y) = \log\left(\frac{x^2}{y}\right). \tag{4.48}$$

Therefore the solution is

$$u(x, y) = v(t(x, y), s(x, y)) = s(x, y)^3e^{t(x,y)} = \left(\frac{y}{x}\right)^3 \cdot \frac{x^2}{y} = \frac{y^2}{x} \qquad \text{for } x > 0. \tag{4.49}$$

Again it is easy to verify manually that this indeed satisfies (4.43) and the initial data.

Example 4.9 (harder)**.** Consider

$$(x + u)u_x + yu_y = u + y^2, \quad \tilde{\Gamma} = \{(x, y, u) : x = 0, y = s, u = s^2\frac{\log(s)}{1 + \log(s)}, e^{-1} < s < \infty\}. \tag{4.50}$$

Steps 1 and 2: The equations we have to solve to obtain the characteristics are

$$\frac{dx}{dt} = x + v, \qquad \frac{dy}{dt} = y, \qquad \frac{dv}{dt} = v + y^2. \tag{4.51}$$

Solving the second equation and using the initial data gives $y(t,s) = se^t$. Inserting this into the third equation gives

$$\begin{aligned} \frac{dv}{dt} &= v + s^2e^{2t} \\ \frac{dv}{dt} - v &= s^2e^{2t} \qquad \text{multiply by } e^{-t} \\ e^{-t}\frac{dv}{dt} - e^{-t}v = \frac{d}{dt}\left(e^{-t}v\right) &= s^2e^t \\ \Rightarrow \qquad v &= s^2e^{2t} + Ce^t. \end{aligned} \tag{4.52}$$

To obtain C set $t = 0$ and use the initial data

$$v(0) = s^2 + C = s^2\frac{\log(s)}{1+\log(s)}, \qquad \Rightarrow \qquad C = -\frac{s^2}{1+\log(s)} \tag{4.53}$$

and therefore

$$v(t,s) = s^2e^{2t} - \frac{s^2e^t}{1+\log(s)} \tag{4.54}$$

With this, we can finally solve the equation for x

$$\begin{aligned} \frac{dx}{dt} &= x + s^2e^{2t} - \frac{s^2e^t}{1+\log(s)} \qquad \text{again, multiply by } e^{-t} \\ e^{-t}\frac{dx}{dt} - e^{-t}x = \frac{d}{dt}\left(e^{-t}x\right) &= s^2e^t - \frac{s^2}{1+\log(s)} \\ \Rightarrow \qquad x &= s^2e^{2t} - \frac{s^2te^t}{1+\log(s)} + Ae^t \end{aligned} \tag{4.55}$$

Setting $t = 0$ and using the initial data we find that $A = -s^2$ and hence

$$x(t,s) = s^2e^{2t} - \frac{s^2te^t}{1+\log(s)} - s^2e^t \tag{4.56}$$

Putting everything together, the characteristics are given by

$$\begin{aligned} x(t,s) &= s^2e^{2t} - \frac{s^2te^t}{1+\log(s)} - s^2e^t, \\ y(t,s) &= se^t, \\ v(t,s) &= s^2e^{2t} - \frac{s^2e^t}{1+\log(s)}. \end{aligned} \tag{4.57}$$

Step 3: Now invert the relations in (4.57) to obtain v as a function of x and y for each t. We have

$$x = s^2e^{2t} - s^2e^t\left(1 + \frac{t}{1+\log(s)}\right) \tag{4.58}$$

Using $y = se^t$ and $\log(y) = t + \log(s)$ we can rewrite this as

$$x = y^2 - s^2e^t\frac{1+\log(y)}{1+\log(s)}. \tag{4.59}$$

Similarly we find that

$$v = y^2 - s^2e^t\frac{1}{1+\log(s)} \tag{4.60}$$

$$\begin{aligned} v - y^2 &= -\frac{s^2e^t}{1+\log(s)} = \frac{x-y^2}{1+\log(y)} \\ \Rightarrow \qquad v(t,s) &= u(x(t,s),y(t,s)) = \frac{x(t,s)+y(t,s)^2\log(y(t,s))}{1+\log(y(t,s))} \qquad (4.61) \\ \Rightarrow \qquad u(x,y) &= \frac{x+y^2\log(y)}{1+\log(y)} \qquad \text{for } y > e^{-1} \end{aligned}$$

Exercise: check directly that this satisfies (4.50) and the initial data.

4.4 Envelopes and domains of influence

The initial data might only be sufficient to determine the solution in part of the domain. In this section we will develop some tools to work out where a solution exists.

Definition 4.4. The domain of influence Ω is the set of all points at which we can uniquely determine the solution $u(x,y)$, given the initial data.

To determine Ω, consider Step 3 of our algorithm: *"Find (s,t) given $x = x(t,s)$, $y = y(t,s)$."*

Theorem 4.2 (Inverse function)**.** *The relation $x(t,s)$, $y(t,s)$ can be inverted (locally uniquely) if the Jacobian*

$$J = \frac{\partial(x,y)}{\partial(s,t)} = \begin{pmatrix} x_s & x_t \\ y_s & y_t \end{pmatrix} \tag{4.62}$$

is invertible, i.e.

$$E = \det(J) = x_sy_t - y_sx_t \neq 0. \tag{4.63}$$

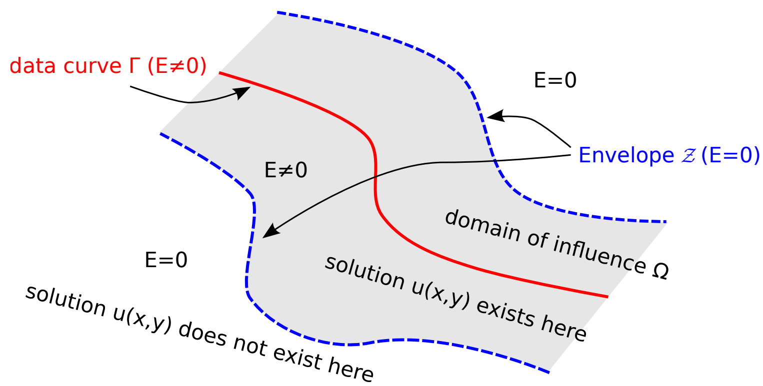


Figure 4.8: Domain of influence Ω and envelope $\mathcal{Z}$.

If condition (4.63) fails so that $y_s x_t = y_t x_s$ then we cannot invert the system uniquely and thus cannot determine $u(x, y)$ uniquely.

Definition 4.5. The curve

$$\mathcal{Z} = \{(x, y) : y_s x_t = y_t x_s\} \tag{4.64}$$

is called the *envelope of the characteristics.*

This leads to an important condition on the data curve Γ. For this, recall that by construction the characteristics $x(t, s), y(t, s)$ intersect Γ at $t = 0$, i.e. $x(0, s) = \gamma^{(x)}(s)$ and $y(0, s) = \gamma^{(y)}(s)$.

Definition 4.6. The data curve Γ has *well posed initial data* if

$$E = x_s y_t - y_s x_t \neq 0 \tag{4.65}$$

at any point $(x(t = 0, s), y(t = 0, s)) = (\gamma^{(x)}(s), \gamma^{(y)}(s))$ on Γ.

Note: The tangent to Γ is $(x_s, y_s)^T$. The characteristics through Γ have tangent (x_t, y_t). Condition (4.63) says that *"The data curve* Γ *is never parallel to a characteristic at the point where they touch."* (see Fig. 4.9). Usually the envelope $\mathcal{Z}$ separates $\mathbb{R}^2$ into two regions, one of which contains the data curve Γ. The domain of influence Ω is a subset of the region which contains Γ and the envelope $\mathcal{Z}$ bounds Ω.

Example 4.10. Suppose the characteristics of a quasilinear PDE are given by

$$x = s + t, \qquad\qquad y = st \tag{4.66}$$

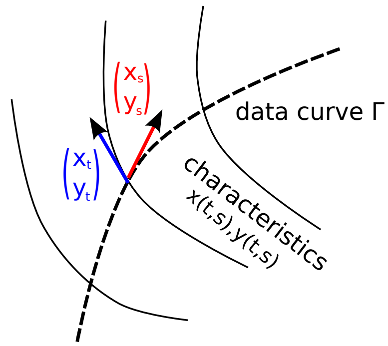


Figure 4.9: Tangents on characteristics and data curve.

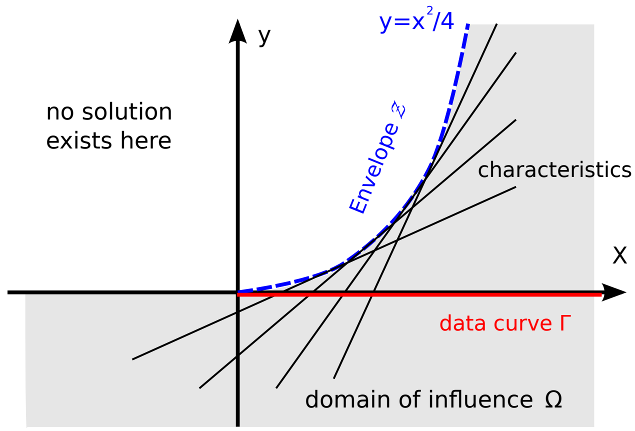


Figure 4.10: Envelope, characteristics lines and domain of influence for example 4.69.

for $s > 0$ with data given on the curve $\Gamma = \{(x, y) : x = s, y = 0, s > 0\}$. What is the envelope of the characteristics and the domain of influence? For this calculate

$$
\begin{aligned}
J &= \begin{pmatrix} 1 & 1 \\ t & s \end{pmatrix} \\
E &= \det(J) = s - t
\end{aligned}
\tag{4.67}
$$

Hence the determinant vanishes if $E = 0 \Leftrightarrow s = t$. Then (4.66) implies that $x = 2s$, $y = s^2$ and hence the envelope is given by $y = x^2/4$ (see Fig. 4.10). If we solve (4.66) for s, a simple calculation shows that

$$
s = \frac{x}{2} \pm \frac{1}{2}\sqrt{x^2 - 4y}. \tag{4.68}
$$

This confirms that we can find s if $y < x^2/4$, as predicted. In particular, if data is given for

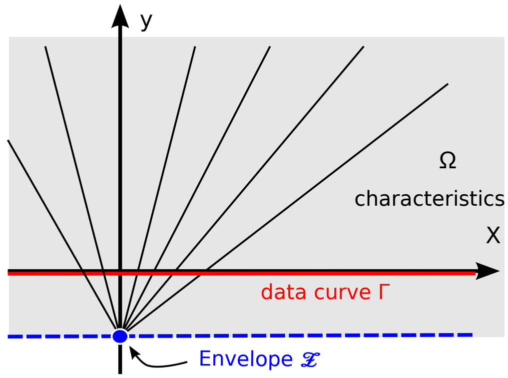


Figure 4.11: Characteristics for the Burgers' equation in (4.71).

$x = s > 0$, then for all $x > 0$ we can take

$$s = \frac{x}{2} + \frac{1}{2}\sqrt{x^2 - 4y}, \qquad t = \frac{x}{2} - \frac{1}{2}\sqrt{x^2 - 4y}. \tag{4.69}$$

For $x \le 0$ on the other hand there is only a positive solution for s if $y < 0$. Hence the domain of influence is given by (see Fig. 4.10)

$$\Omega = \{(x, y) : (x > 0 \text{ and } y < x^2/4) \text{ or } (x \le 0 \text{ and } y < 0)\} \subset \mathbb{R}^2. \tag{4.70}$$

Example 4.11 (Burgers' equation)**.** Consider

$$uu_x + u_y = 0, \qquad \Gamma = \{(x, y) : y = 0, x = s, -\infty < s < \infty\} \tag{4.71}$$

where y can be interpreted as time. We will look at two cases:
Case I: $u = s$ on Γ.
Calculate the characteristics by solving

$$\frac{dx}{dt} = v, \qquad \frac{dy}{dt} = 1, \qquad \frac{dv}{dt} = 0 \tag{4.72}$$

to obtain

$$x(t) = A + Ct, \qquad y(t) = B + t, \qquad v(t) = C. \tag{4.73}$$

Using the initial data we fix the constants A, B and C and find the characteristics

$$x(t, s) = s + st, \qquad y(t, s) = t, \qquad u(t, s) = s. \tag{4.74}$$

The characteristics are straight lines which all go through the point $(0, -1)$; they are shown in Fig. 4.11. The Jacobian and its determinant are

$$J = \begin{pmatrix} 1 + t & s \\ 0 & 1 \end{pmatrix}, \qquad E = \det(J) = 1 + t. \tag{4.75}$$

Hence $E = 0$ if $t = -1$, i.e. the envelope is given by the point $(x, y) = (0, -1)$ at which all characteristics intersect. For $y > -1$ the solution of (4.71) is

$$u(x, y) = s = \frac{x}{1 + y} \tag{4.76}$$

For $y \leq -1$ the data curve does not determine a solution. The reason for this is that since the right hand side is zero, every characteristic is also a contour line. If we were to trace characteristics from the region $y < -1$ back to the data curve we would always have to go through the point $(0, -1)$, independent of the value of s. However, $u_\Gamma(s)$ is not constant, this would imply that the solution takes multiple values at the point $(0, -1)$.

Case II: $u = f(s)$ on Γ.
Generalising the previous calculation we find that the characteristics are again straight lines and given by

$$x(t, s) = s + f(s)t, \qquad y(t, s) = t, \qquad u(t, s) = f(s) \tag{4.77}$$

and the Jacobian and its determinant are

$$J = \begin{pmatrix} 1 + f'(s)t & f(s) \\ 0 & 1 \end{pmatrix}, \qquad E = \det(J) = 1 + f'(s)t = 1 + f'(s)y. \tag{4.78}$$

Hence $E = 0$ for $y = -1/f'(s)$ and we can not extend the solution for all x beyond

$$y_{\max} = \min_s \left\{ -\frac{1}{f'(s)} \right\} \tag{4.79}$$

For example, let $f(s) = 1 + \tanh(-s)$. Physically this describes a "bore" advancing into a fluid at rest. We find that $f'(s) = -\operatorname{sech}^2(-s)$, $-1/f'(s) = \cosh^2(-s) \geq 1$ and the minimum is attained for $s = 0$. Hence the solution exists for $y < 1$ but not for $y \geq 1$. Physically we get a "shock" at $y = 1$. The solution as a function of x is shown for different values of y in Fig. 4.12 and the characteristics and envelope are plotted in Fig. 4.13. Note that as $y \to 1$ the solution function as a function of x approaches a step function, i.e. the shock develops. The slope u_x at $x = 1$ become infinite and the Burgers' equation can not be used beyond this point. In a more realistic model this is prevented by including additional second order diffusion terms. This, however, is well beyond the scope of this course.

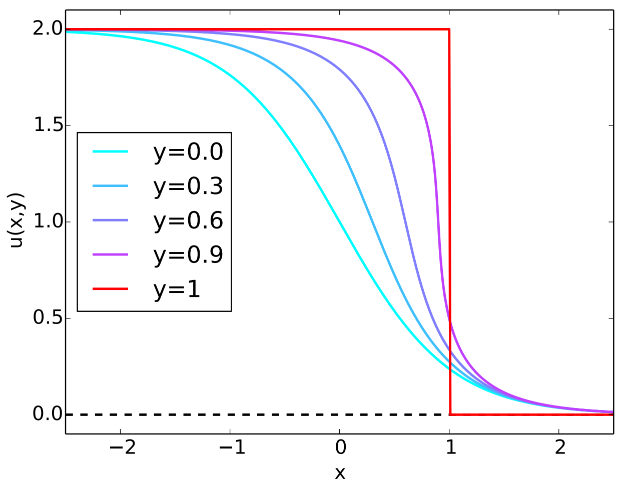


Figure 4.12: Solution of Burgers' equation for different values of y

.

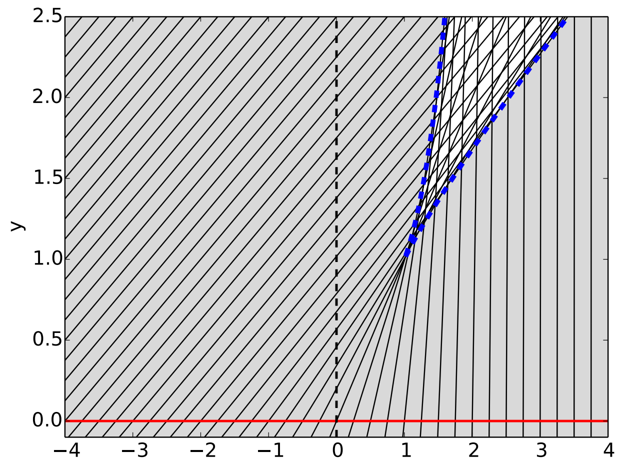


Figure 4.13: Characteristics and domain of influence of the Burgers' equation

Chapter 5

Second order hyperbolic PDEs

Aims:

- To introduce the basic theory of linear second order hyperbolic PDEs
- To study the wave equation using d'Alembert's solution and the Fourier transform

5.1 Introduction

A general *semi-linear* PDE of *second order* is given by

$$a(x,y)u_{xx} + b(x,y)u_{xy} + c(x,y)u_{yy} = d(x,y,u,u_x,u_y) \tag{5.1}$$

where we use the obvious notation

$$u_{xx} \equiv \frac{\partial^2 u}{\partial x^2}, \qquad u_{xy} \equiv \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}, \qquad u_{yy} \equiv \frac{\partial^2 u}{\partial y^2} \tag{5.2}$$

for the second order derivatives. This defines a linear differential operator $\mathscr{L}$

$$\mathscr{L}u \equiv au_{xx} + bu_{xy} + cu_{yy} = d(x,y,u,u_x,u_y) \tag{5.3}$$

where we have dropped the x, y dependency of a, b and c (but we still assume that they are functions of x and y). This equation has the **discriminant**

$$D \equiv b^2 - 4ac. \tag{5.4}$$

The behaviour of the solution depends crucially on the sign of D.

Definition 5.1. If $D > 0$ then $\mathscr{L}$ is *hyperbolic*.

In this chapter we will study hyperbolic operators, which are associated with a finite speed of propagation and describe wavelike phenomena such as: electromagnetism (Maxwell's equations), compressible fluid flow (Euler equation), relativity and acoustics.

Definition 5.2. If $D = 0$ then $\mathscr{L}$ is *parabolic.*

Parabolic equations have an infinite speed of propagation. They are associated with diffusion processes, the heat equation, the Black Scholes formula of mathematical finance and Brownian motion.

Definition 5.3. If $D < 0$ then $\mathscr{L}$ is *elliptic.*

Elliptic PDEs arise in electrostatics and potential flow/theory.

Example 5.1. The wave equation

$$u_{xx} - \lambda^2 u_{yy} = 0, \qquad \lambda \in \mathbb{R} \tag{5.5}$$

is hyperbolic since $a = 1, b = 0, c = -\lambda^2 \Rightarrow D = 4\lambda^2 > 0$.

Example 5.2. The heat equation

$$u_{xx} - u_y = 0 \tag{5.6}$$

is parabolic since $a = 1, b = 0, c = 0 \Rightarrow D = 0$.

Example 5.3. The Laplace equation

$$\Delta u = u_{xx} + u_{yy} = 0 \tag{5.7}$$

is elliptic since $a = 1, b = 0, c = 1 \Rightarrow D = -4 < 0$.

Example 5.4. The canonical wave equation

$$u_{xy} - u_x - u_y = u \tag{5.8}$$

is hyperbolic since $a = 0, b = 1, c = 0 \Rightarrow D = 1 > 0$.

5.2 Characteristics and changes of variable

We will mostly consider the homogeneous linear PDE with constant coefficients, $a, b, c \in \mathbb{R}$ and $a > 0$

$$\mathscr{L}u = au_{xx} + bu_{xy} + cu_{yy} = 0. \tag{5.9}$$

To solve this equation, make the following ansatz for the solution:

$$u(x, y) = f(y + \lambda x) \tag{5.10}$$

where f is a twice differentiable function. It follows that

$$u_{xx} = \lambda^2 f'', \qquad u_{xy} = \lambda f'', \qquad u_{yy} = f''. \tag{5.11}$$

Inserting this into (5.9) leads to an algebraic equation for λ

$$\mathscr{L}u = au_{xx} + bu_{xy} + cu_{yy} = \left(a\lambda^2 + b\lambda + c\right) f'' \tag{5.12}$$

Hence if

$$a\lambda^2 + b\lambda + c = 0 \tag{5.13}$$

then $\mathscr{L}u = 0$ for *any function* f. The solution of (5.13) is

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-b \pm \sqrt{D}}{2a}. \tag{5.14}$$

If $\mathscr{L}$ is hyperbolic ($D > 0$) then $\sqrt{D}$ is real and positive. We therefore have two distinct, nonzero real roots λ_+ and λ_- with

$$\lambda_+ = \frac{-b + \sqrt{D}}{2a} \neq 0, \qquad \lambda_- = \frac{-b - \sqrt{D}}{2a} \neq 0 \tag{5.15}$$

Hence, if f_+ and f_- are two twice differentiable functions, the most general solution of the PDE (5.9) is expected to be $u(x, y) = f_+(y + \lambda_+ x) + f_-(y + \lambda_- x)$. We now prove this.

Theorem 5.1. *If $\mathscr{L}u = 0$ and $\mathscr{L}$ defined by (5.9) is hyperbolic with $a > 0$, then*

$$u(x, y) = f_+(y + \lambda_+ x) + f_-(y + \lambda_- x) \tag{5.16}$$

for two arbitrary, twice differentiable functions f_+ and f_- is the most general solution; the constants λ_+ and λ_- are given by (5.15).

Proof. Introduce new independent variables as

$$s_+ \equiv y + \lambda_+ x, \qquad s_- \equiv y + \lambda_- x \tag{5.17}$$

Note that since $\lambda_+ \neq \lambda_-$ and $\lambda_+ - \lambda_- = \sqrt{D}/a \neq 0$, this relationship can be inverted to obtain

$$x = x(s_+, s_-) = \frac{s_+ - s_-}{\lambda_+ - \lambda_-}, \qquad y = y(s_+, s_-) = \frac{\lambda_+ s_- - \lambda_- s_+}{\lambda_+ - \lambda_-} \tag{5.18}$$

with $\lambda_\pm$ defined by (5.15). Let $u(x,y)$ be the solution of $\mathscr{L}u = 0$ and introduce the function v defined by

$$v(s_+, s_-) \equiv u(x(s_+, s_-), y(s_+, s_-)). \tag{5.19}$$

Then by the chain rule

$$\begin{aligned}
\frac{\partial v}{\partial s_+} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s_+} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s_+} = \frac{1}{\lambda_+ - \lambda_-}\left(\frac{\partial}{\partial x} - \lambda_- \frac{\partial}{\partial y}\right) u \\
\frac{\partial v}{\partial s_-} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s_-} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s_-} = \frac{1}{\lambda_+ - \lambda_-}\left(-\frac{\partial}{\partial x} + \lambda_+ \frac{\partial}{\partial y}\right) u \\
\frac{\partial^2 v}{\partial s_+ \partial s_-} &= \frac{1}{(\lambda_+ - \lambda_-)^2}\left(\frac{\partial}{\partial x} - \lambda_- \frac{\partial}{\partial y}\right)\left(-\frac{\partial}{\partial x} + \lambda_+ \frac{\partial}{\partial y}\right) u \\
&= \frac{1}{(\lambda_+ - \lambda_-)^2}\left(-\frac{\partial^2 u}{\partial x^2} + (\lambda_+ + \lambda_-)\frac{\partial^2 u}{\partial x \partial y} - \lambda_+ \lambda_- \frac{\partial^2 u}{\partial y^2}\right)
\end{aligned} \tag{5.20}$$

But since λ_+ and λ_- are the two solutions of $a\lambda^2 + b\lambda + c = 0$ we have that $a\lambda^2 + b\lambda + c = a(\lambda - \lambda_+)(\lambda - \lambda_-)$ and $\lambda_+ + \lambda_- = -b/a$ and $\lambda_+ \lambda_- = c/a$. Therefore

$$\frac{\partial^2 v}{\partial s_+ \partial s_-} = \frac{1}{(\lambda_+ - \lambda_-)^2}\left(-u_{xx} - \frac{b}{a}u_{xy} - \frac{c}{a}u_{yy}\right) = -\frac{a}{D}\left(a u_{xx} + b u_{xy} + c u_{yy}\right) = 0 \tag{5.21}$$

i.e. the function v has to satisfy

$$\frac{\partial^2 v}{\partial s_+ \partial s_-} = 0. \tag{5.22}$$

Defining $v_+(s_+, s_-) \equiv \partial v / \partial s_+$ and noting that $\partial v_+ / \partial s_- = 0$, it follows that the derivative $v_+(s_+, s_-)$ can only by a function of s_+, i.e.

$$v_+(s_+, s_-) = \frac{\partial v}{\partial s_+} = f'_+(s_+) \tag{5.23}$$

Integrating this, we obtain

$$u(x,y) = v(s_+, s_-) = \int v_+(s_+, s_-)\, ds_+ = f_+(s_+) + f_-(s_-) = f_+(y + \lambda_+ x) + f_-(y + \lambda_- x). \tag{5.24}$$

Note, however, that the functions f_+ and f_- are only defined up to a constant, i.e. if $\tilde{f}_+(s_+) = f_+(s_+) + C$, $\tilde{f}_-(s_-) = f_-(s_-) - C$, then $u(x,y)$ can also be written as

$$u(x,y) = f_+(s_+) + f_-(s_-) = \tilde{f}_+(s_+) + \tilde{f}_-(s_-) = \tilde{f}_+(y + \lambda_+ x) + \tilde{f}_-(y + \lambda_- x). \tag{5.25}$$

□

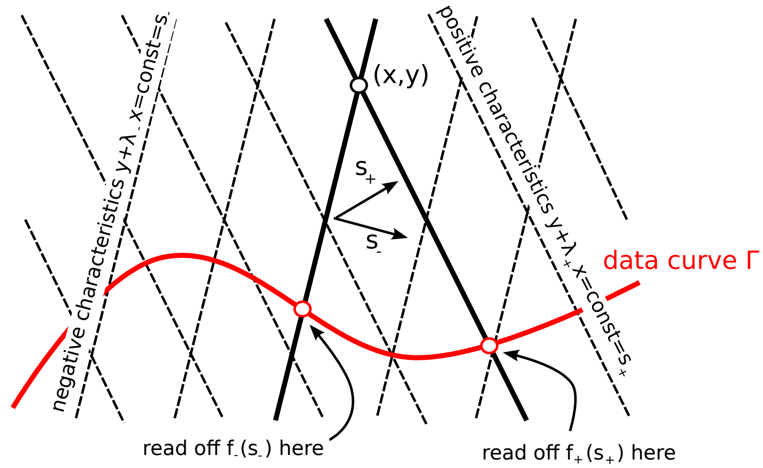


Figure 5.1: Characteristics and data curves for hyperbolic equations.

Definition 5.4. The form

$$\frac{\partial^2 v}{\partial s_+ \partial s_-} = 0 \tag{5.26}$$

is called the *canonical form* for the second order hyperbolic PDE in Eq. (5.9) and the lines

$$s_+ = y + \lambda_+ x = \text{const.}, \qquad s_- = y + \lambda_- x = \text{const.} \tag{5.27}$$

are called the *characteristics*. Note that there are two sets of characteristics, which we will refer to as "positive" and "negative".

In general the data for the PDE is specified on a data curve Γ which is intersected by the characteristics. However, in contrast to a first order PDE, it is now necessary to specify both the value of the field and its derivative in the direction which is not parallel to the data curve. Characteristics convey information from Γ to the point (x, y). More specifically, the recipe for calculating the solution to the PDE is as follows:

Step 1: Use the value of the field and its derivative to calculate the functions $f_+(s_+)$ and $f_-(s_-)$ on the data curve.

Step 2: At a given point (x, y) trace the positive characteristic back to the data curve and evaluate f_+ at this point (which we call s_+) to obtain $f_+(s_+)$. In the same way, trace the negative characteristic back to the data curve and read off the value $f_-(s_-)$.

Step 3: The solution at the point (x, y) is then given by $u(x, y) = f_+(s_+) + f_-(s_-)$.

Example 5.5. Let's consider the special case where the data curve is the x-axis, i.e.

$$\Gamma = \{(x, y) : x = s, y = 0, -\infty < s < \infty\} \tag{5.28}$$

On Γ we specify

$$u(x, 0) = \psi(x), \qquad\qquad u_y(x, 0) = \phi(x) \tag{5.29}$$

for some functions ψ and ϕ. We can use this information to deduce the functions f_+ and f_- since we know that

$$u(x, 0) = f_+(\lambda_+ x) + f_-(\lambda_- x) = \psi(x), \quad u_y(x, 0) = f'_+(\lambda_+ x) + f'_-(\lambda_- x) = \phi(x). \tag{5.30}$$

Integrating the second equation by x from a reference point x_0 we find that

$$\int_{x_0}^{x} \phi(z)\, dz = \frac{1}{\lambda_+} f_+(\lambda_+ x) + \frac{1}{\lambda_-} f_-(\lambda_- x) + C \tag{5.31}$$

Rearranging and solving for f_+ and f_- we find that

$$\begin{aligned}
\psi(x) - \lambda_- \left(\int_{x_0}^{x} \phi(z)\, dz - C \right) &= \left(1 - \frac{\lambda_-}{\lambda_+} \right) f_+(\lambda_+ x) \\
\psi(x) - \lambda_+ \left(\int_{x_0}^{x} \phi(z)\, dz - C \right) &= \left(1 - \frac{\lambda_+}{\lambda_-} \right) f_-(\lambda_- x)
\end{aligned} \tag{5.32}$$

Now use the fact that on the data curve $s_+ = \lambda_+ x$ and $s_- = \lambda_- x$ and hence

$$\begin{aligned}
f_+(s_+) &= \frac{\lambda_+}{\lambda_+ - \lambda_-} \left(\psi(s_+/\lambda_+) - \lambda_- \int_{x_0}^{s_+/\lambda_+} \phi(z)\, dz + \lambda_- C \right) \\
f_-(s_-) &= \frac{\lambda_-}{\lambda_- - \lambda_+} \left(\psi(s_-/\lambda_-) - \lambda_+ \int_{x_0}^{s_-/\lambda_-} \phi(z)\, dz + \lambda_+ C \right) \\
u(x, y) &= f_+(y + \lambda_+ x) + f_-(y + \lambda_- x) \\
&= \frac{1}{\lambda_+ - \lambda_-} \left(\lambda_+ \psi(x + y/\lambda_+) - \lambda_- \psi(x + y/\lambda_-) - \lambda_+ \lambda_- \int_{x+y/\lambda_-}^{x+y/\lambda_+} \phi(z)\, dz \right).
\end{aligned} \tag{5.33}$$

Note that for $u(x, y) = f_+(s_+) + f_-(s_-)$ the constant C and the choice of x_0 are irrelevant and cancel out.

Example 5.6. Consider the second order PDE

$$u_{xx} - 4u_{xy} + 3u_{yy} = 0, \qquad \text{with } u(x, x) = x^3 \text{ and } u_x(x, x) = 0. \tag{5.34}$$

In this case we specified the solution and its derivative in the x-direction on the line $y = x$. The characteristic equation is

$$(\lambda - 1)(\lambda - 3) = 0 \tag{5.35}$$

which implies $\lambda_+ = 1$ and $\lambda_- = 3$. Hence the general solution is

$$u(x,y) = f_+(s_+) + f_-(s_-) = f_+(y+\lambda_+ x) + f_-(y+\lambda_- x) = f_+(y+x) + f_-(y+3x). \tag{5.36}$$

To find the functions f_+ and f_-, use the initial conditions on the line $y = x$:

$$u(x,x) = f_+(2x) + f_-(4x) = x^3, \qquad u_x(x,x) = f_+'(2x) + 3f_-'(4x) = 0. \tag{5.37}$$

Take the total derivative of the first equation[1] in Eq. (5.37) to obtain

$$\frac{d}{dx}u(x,x) = 2f_+'(2x) + 4f_-'(4x) = 3x^2. \tag{5.38}$$

Subtracting twice the second equation in Eq. (5.37) from Eq. (5.38) we find

$$\frac{d}{dx}u(x,x) - 2u_x(x,x) = -2f_-'(4x) = 3x^2. \tag{5.39}$$

Changing variables and introducing $z := 4x$ we conclude that

$$\begin{aligned} f_-'(z) \quad &= -\tfrac{3}{32}z^2 \\ \Rightarrow f_-(z) \quad &= -\tfrac{1}{32}z^3 + C. \end{aligned} \tag{5.40}$$

The function f_+ can be found by inserting this into the first equation of Eq. (5.37):

$$\begin{aligned} f_+(2x) &= x^3 - f_-(4x) = 3x^3 - C \qquad \text{and hence with } \tilde{z} := 2x \\ \Rightarrow f_+(\tilde{z}) &= \frac{3}{8}\tilde{z}^3 - C \end{aligned} \tag{5.41}$$

Inserting the expressions for f_+ and f_- from Eqs (5.41) and (5.40) into the solution in Eq. (5.36), we obtain

$$u(x,y) = \frac{3}{8}(x+y)^3 - \frac{1}{32}(3x+y)^3. \tag{5.42}$$

Let's double check that this indeed the correct solution, and that it satisfies the initial conditions on the line $y = x$. For this, calculate the derivates

$$\begin{aligned} u_x(x,y) &= \frac{9}{8}(x+y)^2 - \frac{9}{32}(3x+y)^2 \\ u_y(x,y) &= \frac{9}{8}(x+y)^2 - \frac{3}{32}(3x+y)^2 \\ u_{xx}(x,y) &= \frac{9}{4}(x+y) - \frac{27}{16}(3x+y) = \frac{9}{16}(-5x+y) \\ u_{xy}(x,y) &= \frac{9}{4}(x+y) - \frac{9}{16}(3x+y) = \frac{9}{16}(x+3y) \\ u_{yy}(x,y) &= \frac{9}{4}(x+y) - \frac{3}{16}(3x+y) = \frac{9}{16}\left(3x + \frac{11}{3}y\right) \end{aligned} \tag{5.43}$$

[1] Make sure you understand the difference between the total derivative $\frac{d}{dx}u(x,x)$ and the partial derivative evaluated at (x,x), $u_x(x,x) = \frac{\partial}{\partial x}u(x,y)|_{(x,x)}$.

Inserting into Eq. (5.34) gives

$$u_{xx} - 4u_{xy} + 3u_{yy} = \frac{9}{16}\left(-5x + y - 4(x+3y) + 3\left(3x + \frac{11}{3}y\right)\right) = 0. \tag{5.44}$$

We also confirm that on the line $y = x$

$$\begin{aligned} u(x,x) &= \frac{3}{8}(2x)^3 - \frac{1}{32}(4x)^3 = x^3 \\ u_x(x,x) &= \frac{9}{8}(2x)^2 - \frac{9}{32}(4x)^2 = 0. \end{aligned} \tag{5.45}$$

5.3 d'Alembert's solution of the wave equation

By far the most important hyperbolic PDE is the linear wave equation. In this case $x \in [-\infty, \infty]$ is a *spatial* variable and instead of y we use $t > 0$ as a *temporal (time)* variable. The wave equation takes the usual form

$$u_{tt} - c^2 u_{xx} = 0, \qquad \text{for } -\infty < x < \infty,\ t > 0 \tag{5.46}$$

with initial conditions

$$u(x, t=0) = \psi(x), \qquad u_t(x, t=0) = \phi(x) \qquad \text{for } -\infty < x < \infty. \tag{5.47}$$

Here $c > 0$ is the wave speed. The solutions to the characteristics equation $\lambda^2 c^2 - 1 = 0$ are $\lambda_+ = 1/c$ and $\lambda_- = -1/c$ and hence the characteristics are described by

$$s_+ = t + \frac{x}{c} = \text{const}, \qquad s_- = t - \frac{x}{c} = \text{const}. \tag{5.48}$$

Hence the general solution is

$$u(x,t) = f_+\left(\frac{x}{c} + t\right) + f_-\left(\frac{x}{c} - t\right) \tag{5.49}$$

This expresses u as the sum of a *leftward* travelling wave f_+ and a rightward travelling wave f_-. Using (5.33) this can be expressed in terms of the initial conditions as

$$u(x,t) = \frac{1}{2}\left(\psi(x+ct) + \psi(x-ct)\right) + \frac{1}{2c}\int_{x-ct}^{x+ct} \phi(z)\, dz. \tag{5.50}$$

If $\phi(x) = 0$, half of the initial distribution travels to the left and the other half to the right. This is shown in Fig. 5.2 for the initial condition

$$u(x,0) = \psi(x) = \begin{cases} 1 + x & \text{for } -1 \le x < 0 \\ 1 - x & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}, \qquad u_t(x,0) = \phi(x) = 0 \tag{5.51}$$

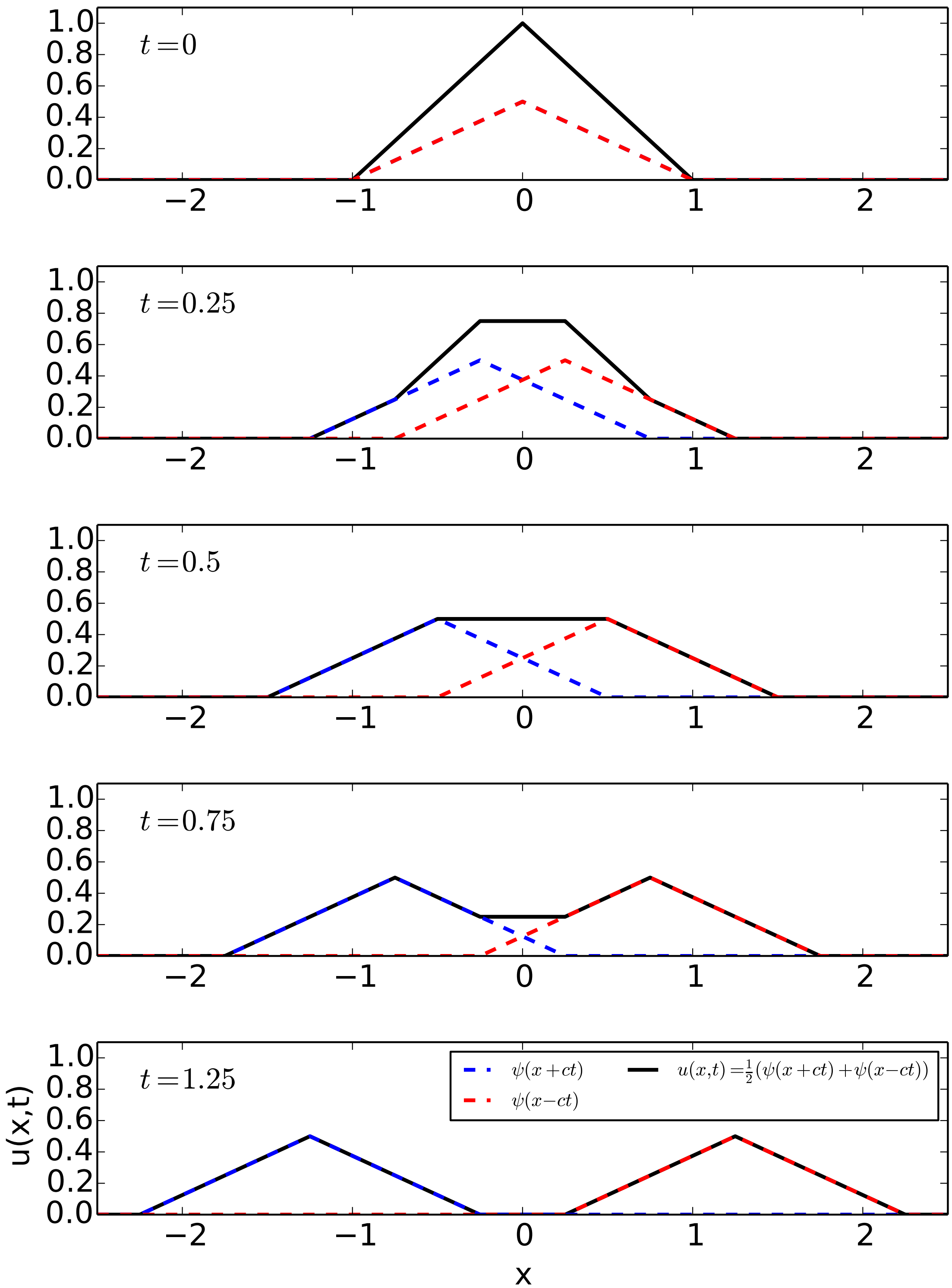


Figure 5.2: Solution of the wave equation $u(x,t)$ given in (5.50) for different times t.

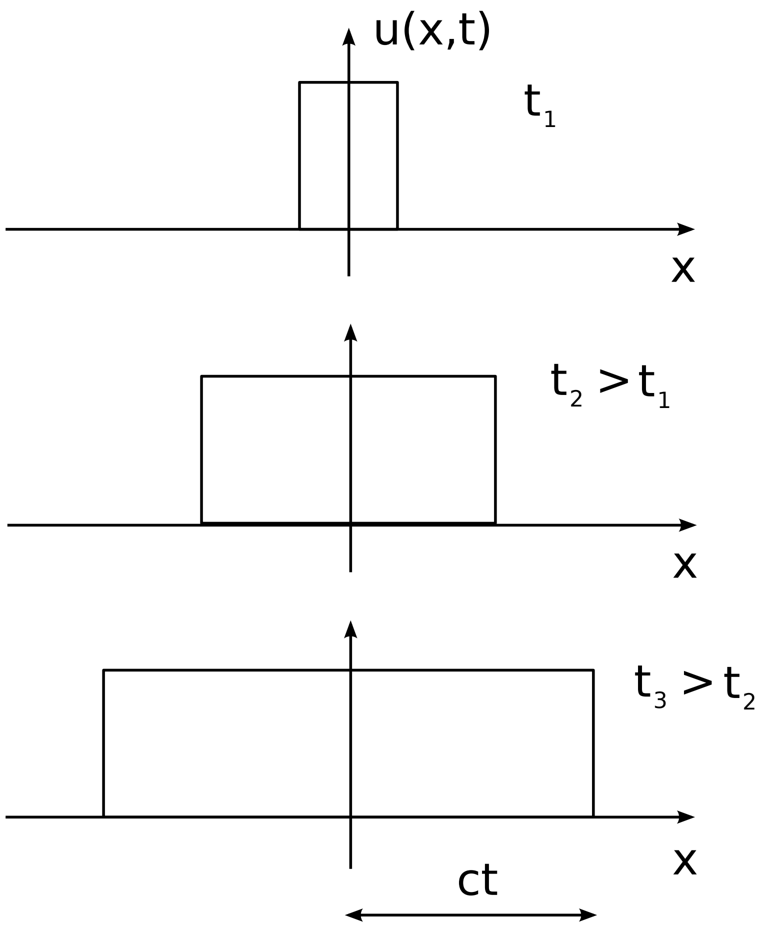


Figure 5.3: Solution of the wave equation $u(x,t)$ for the impulsively struck string in Example 5.7. The solution is shown for different times t.

Example 5.7 (Impulsively struck string)**.** If, on the other hand $\psi(x) = 0$ and $\phi(x) = \delta(x)$, this describes an infinite string which is struck at the point $x = 0$ at time $t = 0$. In this case the solution is (see Fig. 5.3)

$$u(x,t) = \frac{1}{2c}\int_{x-ct}^{x+ct} \delta(z)\, dz = \frac{1}{2c} \quad \text{if } |x| < ct \text{ and } u(x,t) = 0 \text{ otherwise.} \tag{5.52}$$

This follows immediately from the fact that the integral is nonzero only if $[x - ct, x + ct]$ contains 0.

Appendix A

Notation

- The natural logarithm is written as $y = \log(x)$, i.e. $x = e^y$.
- $\mathbb{R}$ denotes the set of all real numbers.
- $\mathbb{N} = 1, 2, 3, \ldots$ denotes the set of positive integer numbers.
- For a complex number $z = x + iy \in \mathbb{C}$ with $x, y \in \mathbb{R}$ the complex conjugate is defined as $z^* = x - iy$.
- Partial derivatives are sometimes abbreviated as

$$u_x = \frac{\partial u}{\partial x}, \qquad u_y = \frac{\partial u}{\partial y}, \qquad u_{xx} = \frac{\partial^2 u}{\partial x^2}, \qquad u_{xy} = \frac{\partial^2 u}{\partial x \partial y}, \qquad \dot{u} = u_t = \frac{\partial u}{\partial t} \tag{A.1}$$

Appendix B

Generalised product rule

Theorem B.1. *If f and g are two n-times differentiable functions, the n-th derivative of their product $f \cdot g$ is given by*

$$\frac{d^n(f \cdot g)}{dx^n} = \sum_{k=0}^{n} \binom{n}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} \tag{B.1}$$

Proof. This can be proven by induction. Start with the base case $n = 1$, which is just the normal product rule

$$\frac{d(f \cdot g)}{dx} = \frac{df}{dx} \cdot g + f \cdot \frac{dg}{dx} = \sum_{k=0}^{1} \frac{d^k f}{dx^k} \cdot \frac{d^{1-k} g}{dx^{1-k}} \tag{B.2}$$

Assuming (B.1) is true for $n - 1$, we can obtain the case n by differentiating the equation (B.1) for $n - 1$:

$$\begin{aligned}
\frac{d^n(f \cdot g)}{dx^n} &= \frac{d}{dx}\left(\frac{d^{n-1}(f \cdot g)}{dx^{n-1}}\right) = \frac{d}{dx}\left(\sum_{k=0}^{n-1} \binom{n-1}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-1-k} g}{dx^{n-1-k}}\right) \\
&= \sum_{k=0}^{n-1} \binom{n-1}{k} \left(\frac{d^{k+1} f}{dx^{k+1}} \cdot \frac{d^{n-1-k} g}{dx^{n-1-k}} + \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}}\right) \\
&= \sum_{k=1}^{n} \binom{n-1}{k-1} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} + \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} \\
&= \frac{d^n f}{dx^n} \cdot g + \sum_{k=1}^{n-1} \left(\binom{n-1}{k-1} + \binom{n-1}{k}\right) \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} + f \cdot \frac{d^n g}{dx^n} \\
&= \sum_{k=0}^{n} \binom{n}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}}
\end{aligned} \tag{B.3}$$

since $\binom{n-1}{0} = \binom{n-1}{n-1} = \binom{n}{0} = \binom{n}{n} = 1$ and Pascal's rule

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \tag{B.4}$$

□

Appendix C

Self-adjoint matrices

The following results on finite dimensional self-adjoint matrices are included for completeness.

Definition C.1. Let $u, v \in \mathbb{C}^d$ be complex vectors. Then the scalar product $\langle \cdot, \rangle \cdot$ is defined as

$$\langle u, v\rangle \equiv \sum_{i=1}^{d} u_i^* v_i. \tag{C.1}$$

Two vectors u, v are said to be **orthogonal** if $\langle u, v\rangle = 0$.

It is easy to see that

$$\langle u, v\rangle^* = \langle u^*, v^*\rangle = \langle v, u\rangle \tag{C.2}$$

and that

$$\langle u, u\rangle \geq 0 \qquad \text{for all } u \in \mathbb{C}^d \tag{C.3}$$

with

$$\langle u, u\rangle > 0 \qquad \text{for all } u \in \mathbb{C}^d \text{ with } u \neq 0. \tag{C.4}$$

Definition C.2. A $d \times d$ matrix H is called **self-adjoint** or **Hermitian** if $\langle u, Hv\rangle = \langle Hu, v\rangle$ for all $u, v \in \mathbb{C}^d$.

Lemma C.1. *If H is self-ajoint, then $H_{ij} = H_{ji}^*$.*

Proof. Let $e^{(i)}$ be the unit vector in direction i, i.e. $e^{(i)}_k = \delta_{ik}$. Choose $u = e^{(i)}$ and $v = e^{(j)}$ and use the fact that H is self-adjoint to obtain

$$H_{ij} = \langle e^{(i)}, He^{(j)}\rangle = \langle He^{(i)}, e^{(j)}\rangle = H^*_{ji}. \tag{C.5}$$

□

Definition C.3. A $d \times d$ matrix A is called positive semi-definite if for all $u \in \mathbb{C}^d$

$$\langle u, Au\rangle \geq 0. \tag{C.6}$$

We can make several interesting statements about Hermitian matrices:

Theorem C.1. *Let H be a Hermitian matrix and $\{(\lambda_n, u_n)\}$ its eigenvalue/eigenvector pairs. Then the following holds:*

1. *All eigenvalues are real, $\lambda_n \in \mathbb{R}$.*
2. *If H is positive semi-positive then all eigenvalues are non-negative, $\lambda_n \geq 0$.*
3. *If $\lambda_n \neq \lambda_m$ are two distinct eigenvalues then the corresponding eigenvectors are orthogonal, $\langle u_n, u_m\rangle = 0$.*

This is the finite dimensional version of Theorem **??**.

Proof. **Part 1.**

If λ_n is an eigenvalue with $u_n \neq 0$ then $Hu_n = \lambda_n u_n$. Multiplying by u_n from the left we find that

$$\langle u_n, Hu_n\rangle = \lambda_n\langle u_n, u_n\rangle. \tag{C.7}$$

Taking the complex conjugate of the left hand side of this equation, using (C.2) and the fact that H is Hermitian gives

$$\langle u_n, Hu_n\rangle^* = \langle Hu_n, u_n\rangle = \langle u_n, Hu_n\rangle = \lambda_n\langle u_n, u_n\rangle. \tag{C.8}$$

On the other hand, taking the complex conjugate of the right hand side of (C.7) gives

$$\lambda^*_n\langle u_n, u_n\rangle^* = \lambda^*_n\langle u_n, u_n\rangle \tag{C.9}$$

and it follows that $\lambda_n\langle u_n, u_n\rangle = \lambda^*_n\langle u_n, u_n\rangle$. Since $u_n \neq 0$ and hence $\langle u_n, u_n\rangle \neq 0$ this implies that $\lambda_n = \lambda^*_n$.

Part 2.

This follows immediately with (C.7). As H is positive semi-definite $\langle u_n, Hu_n\rangle \geq 0$, and therefore $\lambda_n\langle u_n, u_n\rangle \geq 0$. Since $u_n \neq 0$ (C.4) we have that $\langle u_n, u_n\rangle > 0$ and hence $\lambda_n \geq 0$.

Part 3.

Consider two eigenvalue/eigenvectors pairs (λ_n, u_n), (λ_m, u_m) with $\lambda_n \neq \lambda_m$. Then we have:

$$\langle u_m, Hu_n\rangle = \lambda_n\langle u_m, u_n\rangle, \qquad \langle u_n, Hu_m\rangle = \lambda_m\langle u_n, u_m\rangle \tag{C.10}$$

Take the complex conjugate of the second equation and use (C.2), the fact that H is Hermitian and $\lambda_m^* = \lambda_m$ to obtain

$$\begin{aligned} & \langle u_n, Hu_m\rangle^* = \lambda_m\langle u_n, u_m\rangle^* \\ \Leftrightarrow \qquad & \langle Hu_m, u_n\rangle = \langle u_m, Hu_n\rangle = \lambda_m\langle u_m, u_n\rangle \end{aligned} \tag{C.11}$$

Together with (C.10) this implies that $\lambda_n\langle u_m, u_n\rangle = \lambda_m\langle u_m, u_n\rangle$. Since, by assumption, $\lambda_n \neq \lambda_m$, this can only be fufilled if $\langle u_m, u_n\rangle = 0$, i.e. the eigenvectors u_n and u_m are orthogonal. $\square$