
MA30044
Mathematical Methods I

Department of Mathematical Sciences, University of Bath

Dr. Philippe Trinh

Monday 13th September, 2021, 18:42h
version:

Contents

1 Motivation and Overview	3
2 Sturm Liouville Theory	8
2.1 Introduction	8
2.2 Eigenfunctions and orthogonality	14
2.3 Orthogonal and orthonormal expansions	20
2.4 Approximation of functions	22
2.4.1 Legendre polynomials	28
2.5 Inhomogeneous Sturm Liouville equations	33
2.6 Solution of PDEs using Sturm-Liouville theory	37
3 The Fourier Transform	39
3.1 Introduction	39
3.2 Elementary properties of the Fourier transform	45
3.3 Introduction to Distributions	47
3.4 The convolution theorem and its applications	52
3.5 Using the Fourier transform to solve ODEs	54
3.5.1 First order equations	55
3.5.2 Second order equations	58
3.6 Plancheral's Theorem	61
3.7 Solving PDEs with the Fourier transform	62
4 Quasilinear first order PDEs	66
4.1 Introduction	66
4.2 Contour lines	68
4.3 The method of characteristics	73
4.4 Envelopes and domains of influence	77
5 Second order hyperbolic PDEs	83
5.1 Introduction	83
5.2 Characteristics and changes of variable	85
5.3 d'Alembert's solution of the wave equation	90

CONTENTS

A Notation	93
B Generalised product rule	94
C Self-adjoint matrices	96

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 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).

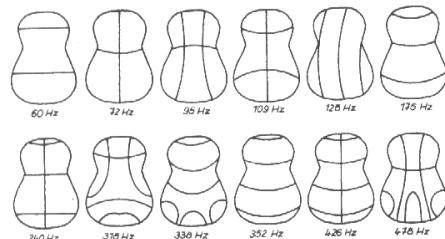


Figure 1.1: Chladni Figures on a circular disk (left^a) and eigenmodes of a guitar (right^b).

^aSource: wikimedia commons, https://commons.wikimedia.org/wiki/File:Round_Chladni_plate_with_4_circular_nodes.JPG for license terms.

^bSource: 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.

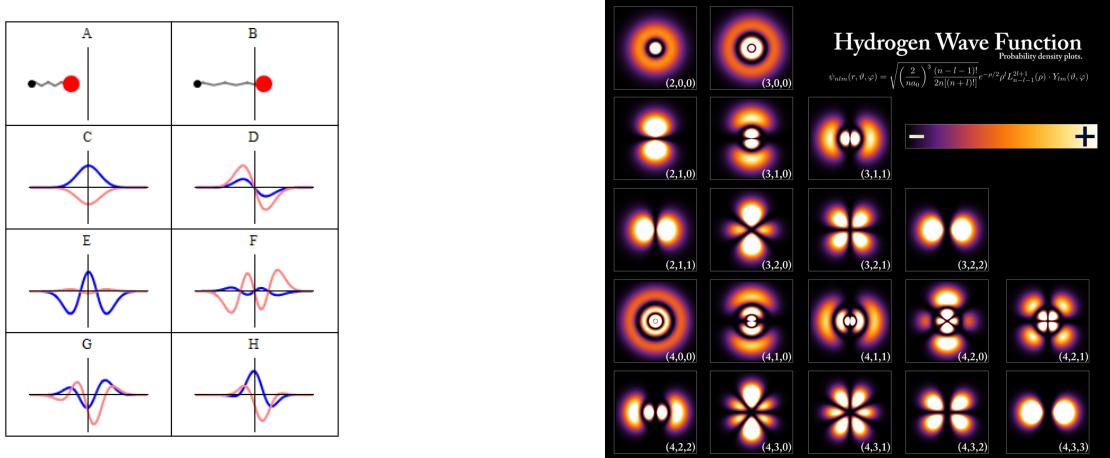


Figure 1.2: Quantum mechanics: harmonic oscillator (left^a) and hydrogen atom (right^b)

^aSource: wikipedia commons, see <https://commons.wikimedia.org/wiki/File:QuantumHarmonicOscillatorAnimation.gif> for license terms.

^bSource: wikipedia 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 **transform 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.

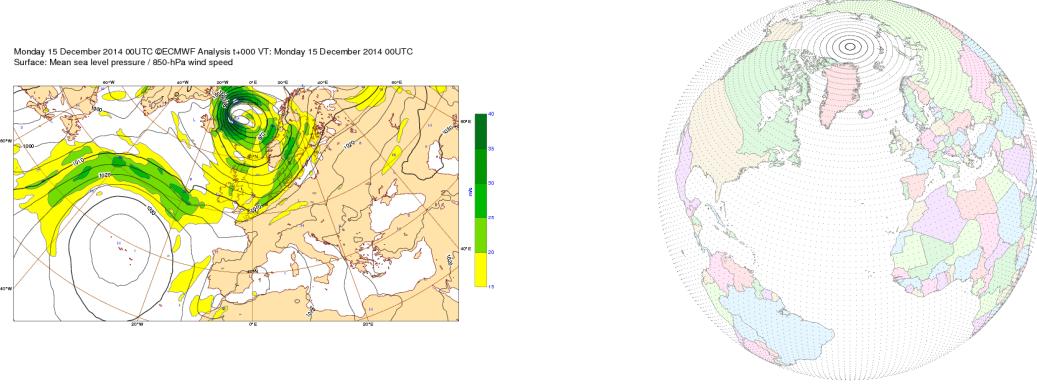


Figure 1.3: ECMWF weather forecast (left^a) and spectral grid (right^b)

^aSource: flickr <https://www.flickr.com/photos/ecmwf/13384533105>

^bSource: 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.

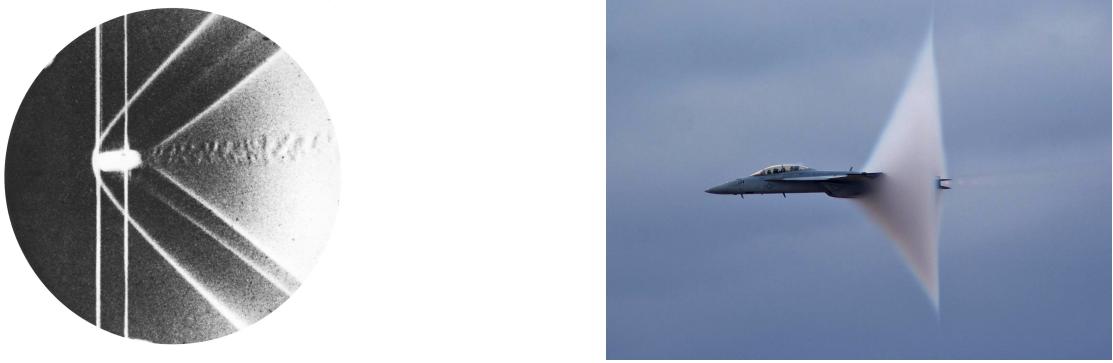


Figure 1.4: Examples of shock waves: supersonic bullet (left^a) and jet (right^b)

^aSource: wikimedia commons, see https://commons.wikimedia.org/wiki/File:Photography_of_bow_shock_waves_around_a_brass_bullet,_1888.jpg for license terms.

^bSource: 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.

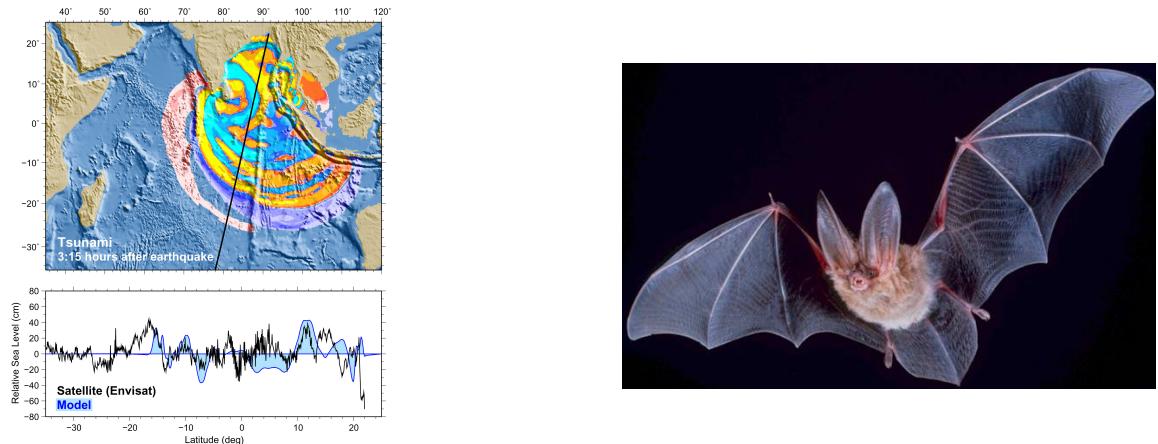


Figure 1.5: Wave propagation: Tsunami model (left^a) and hunting bat (right^b)

^aSource: NOAA news, <http://www.noaanews.noaa.gov/stories2005/s2365.htm>

^bSource: wikipedia <https://en.wikipedia.org/wiki/Bat#/media/File:Big-eared-townsend-fledermaus.jpg>

Chapter 2

Sturm Liouville Theory

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

2.1 Introduction

You will probably already have met linear second order ODEs

$$\alpha \frac{d^2u}{dx^2} + \beta \frac{du}{dx} + \gamma u = 0 \quad (2.1)$$

where $u(x)$ is an unknown function defined on an interval $[a, b]$ and α, β and γ are constants. To solve this equation uniquely, boundary conditions need to be specified. For example, we could fix the solution at the endpoints,

$$u(a) = c, \quad u(b) = d, \quad \text{with } c, d \in \mathbb{R}. \quad (2.2)$$

A **Sturm-Liouville (SL) equation** is a generalisation of this and has the form

$$\frac{d}{dx} \left(p(x) \frac{du}{dx} \right) - q(x)u + \lambda r(x)u = 0 \quad (\text{plus boundary conditions}). \quad (2.3)$$

The function p has to be at least once differentiable. We seek a complex-valued solution $u(x) \in \mathbb{C}$ with $x \in I$ on an interval $I = [a, b]$ with boundary conditions (see Definition 2.1 for details) at a and b .

- We usually take $p(x), r(x) > 0$. This is called a **regular** SL-equation
- Otherwise the SL-equation is called **singular**
- $r(x)$ is called the **weight function**, the number λ is called the **eigenvalue**

A seemingly more general 2nd order ODE is given by

$$\alpha(x) \frac{d^2u}{dx^2} + \beta(x) \frac{du}{dx} + \gamma(x)u + \lambda\delta(x)u = 0 \quad (2.4)$$

However, the following lemma shows how (2.4) can be re-written as a SL-equation under certain conditions:

Lemma 2.1 (Transformation to Sturm-Liouville form). *If $\alpha(x) > 0$ then the 2nd order ODE in (2.4) can be written as a Sturm-Liouville equation in the form (2.3) with*

$$p(x) = \exp \left(\int_{x_0}^x \frac{\beta(y)}{\alpha(y)} dy \right), \quad q(x) = -p(x) \frac{\gamma(x)}{\alpha(x)}, \quad r(x) = p(x) \frac{\delta(x)}{\alpha(x)} \quad (2.5)$$

where $x_0 \in [a, b]$ is an arbitrary number. Note that $p(x) > 0$ and the resulting SL-equation is regular if $\delta(x) > 0$. In this context $p(x)$ is sometimes called an **integrating factor**.

Usually $x_0 = a$, but choosing a different x_0 (say, \tilde{x}_0) will only result in multiplication of $p(x)$, $q(x)$ and $r(x)$ by a constant positive factor $C = \exp \left(\int_{\tilde{x}_0}^{x_0} \frac{\beta(y)}{\alpha(y)} dy \right)$ which cancels out. In some cases we also can not choose $x_0 = a$, for example if this leads to an ill-defined integral. In this case we need to use our intuition to find a good value for x_0 .

Proof. Since $\alpha(x) > 0$ we can divide (2.4) by α to obtain

$$u'' + \frac{\beta(x)}{\alpha(x)}u' + \frac{\gamma(x)}{\alpha(x)}u + \lambda \frac{\delta(x)}{\alpha(x)}u = 0 \quad (2.6)$$

(for simplicity, we use the notation $u' = \frac{du}{dx}$ to denote derivatives). Now let $p(x)$ as defined in (2.5), then by the chain rule

$$p'(x) = \frac{\beta(x)}{\alpha(x)} \exp \left(\int_{x_0}^x \frac{\beta(y)}{\alpha(y)} dy \right) = \frac{\beta(x)}{\alpha(x)} p(x) \quad (2.7)$$

and hence $\beta/\alpha = p'/p$. Further, from (2.5) we also have $\gamma/\alpha = -q/p$ and $\delta/\alpha = r/p$. Inserting those ratios into (2.6) results in

$$\begin{aligned} u'' + \frac{p'(x)}{p(x)}u' - \frac{q(x)}{p(x)}u + \lambda \frac{r(x)}{p(x)}u &= 0 \\ \Leftrightarrow p(x)u'' + p'(x)u' - q(x)u + \lambda r(x)u &= 0 \\ \Leftrightarrow (p(x)u')' - q(x)u + \lambda r(x)u &= 0 \end{aligned} \tag{2.8}$$

as required. \square

Example 2.1. Let u satisfy

$$xu'' + \frac{1}{2}u' + x^2u + \lambda x^3u = 0, \quad \text{for } x \in (1, 2) \tag{2.9}$$

Then

$$\alpha(x) = x, \quad \beta(x) = \frac{1}{2}, \quad \gamma(x) = x^2, \quad \delta(x) = x^3 \tag{2.10}$$

This implies that

$$p(x) = \exp\left(\int_1^x \frac{dy}{2y}\right) = \exp\left(\frac{1}{2}\log(x)\right) = x^{1/2}, \quad q(x) = -x^{3/2}, \quad r(x) = x^{5/2}. \tag{2.11}$$

Therefore the SL form of the equation is

$$(x^{1/2}u')' + x^{3/2}u + \lambda x^{5/2}u = 0. \tag{2.12}$$

Definition 2.1 (Sturm Liouville operator). Let p, q and r be real valued functions with $p(x), r(x) > 0$. If u is any twice differentiable complex-valued function on the interval $I = [a, b]$ and

$$Lu \equiv (p(x)u')' - q(x)u \tag{2.13}$$

with boundary conditions

$$\alpha u(a) + \beta u'(a) = 0, \quad \gamma u(b) + \delta u'(b) = 0 \tag{2.14}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, $\alpha^2 + \beta^2 > 0$, $\gamma^2 + \delta^2 > 0$ then L is called a **Sturm-Liouville (SL) operator** with associated **weight function** $r(x)$.

Note that since we have assumed that $p(x), r(x) > 0$, strictly speaking this defines a *regular* Sturm Liouville operator. L is a linear operator, i.e. for any real numbers c_1, c_2 and function u, v we have

$$L(c_1u + c_2v) = c_1Lu + c_2Lv. \tag{2.15}$$

Exercise: verify this.

Example 2.2. The following operators arise naturally in the study of PDEs:

- *Harmonic* $Lu = u''$, $r(x) = 1$
(vibrating string, quantum mechanical particle in a box, see Example 2.3)
- *Legendre* $Lu = ((1 - x^2)u')'$, $r(x) = 1$, $[a, b] = [-1, 1]$
(electrostatics, see Section 2.4.1 and problem sheet 2)
- *Chebyshev* $Lu = (\sqrt{1 - x^2}u')'$, $r(x) = \frac{1}{\sqrt{1 - x^2}}$, $[a, b] = [-1, 1]$
(function approximation, see Example 2.7 and problem sheet 2)
- *Hermite* $Lu = \left(e^{-x^2}u'\right)', r(x) = \exp(-\frac{x^2}{2})$, $x \in \mathbb{R}$
(quantum mechanical harmonic oscillator, see problem sheet 2).

Hermite polynomials are defined for all $x \in \mathbb{R}$, i.e. the interval $[a, b]$ is infinite in this case.

You might have come across self-adjoint linear operators on finite dimensional vector spaces (see Appendix C). SL operators can be seen as an equivalent of this on the infinite dimensional vector space of functions on an interval.

Definition 2.2 (Square integrability). A complex valued function u defined on the interval $[a, b]$ is called **square integrable** if

$$\int_a^b |u(x)|^2 < \infty \quad (2.16)$$

We denote the space of square integrable functions on $[a, b]$ by $L_{[a,b]}^2$.

Definition 2.3 (Self-adjointness). Let $u, v \in L_{[a,b]}^2$ be two complex valued square integrable functions on the interval $[a, b]$. We can define the following scalar product (compare this to the finite-dimensional version in Definition C.1):

$$(v, u) \equiv \int_a^b v^*(x)u(x) dx \quad (2.17)$$

Let $D_H \subset L_{[a,b]}^2$ be the set of all functions for which the operator H can be defined (i.e. the “domain” of the operator). Assume also that H maps to the space of square-integrable functions. Then $H : D_H \rightarrow L_{[a,b]}^2$ is called **self-adjoint** if

$$(v, Hu) = (Hv, u) \quad \text{for all } u, v \in D_H. \quad (2.18)$$

Before we can show that SL operators are indeed self-adjoint, we prove the following Lemma:

Lemma 2.2 (Greens identity). *Let L be a Sturm-Liouville operator and $u, v \in D_L \subset L^2_{[a,b]}$ arbitrary complex-valued square integrable functions on $[a, b]$. Then*

$$(v, Lu) - (Lv, u) = \int_a^b (v^* Lu - u Lv^*) dx = [p(v^* u' - uv'^*)]_a^b. \quad (2.19)$$

Proof. From the definition of the SL operator and using that $(Lv)^* = Lv^*$ (since p and q are real valued) we have that

$$(v, Lu) - (Lv, u) = \int_a^b (v^* Lu - u Lv^*) = \int_a^b (v^*(pu')' - u(pv'^*)) dx. \quad (2.20)$$

Integration by parts gives the desired result

$$= [v^* pu' - upv'^*]_a^b - \int_a^b \underbrace{(v'^* pu' - u'pv'^*)}_{=0} dx = [p(v^* u' - uv'^*)]_a^b. \quad (2.21)$$

□

Definition 2.4 (Wronskian). For convenience, we introduce the **Wronskian** $W(u, v)$ of two functions u, v defined by

$$W(u, v) \equiv uv' - vu' = \det \begin{pmatrix} u & u' \\ v & v' \end{pmatrix}. \quad (2.22)$$

Lemma 2.3. *If*

$$\alpha u(a) + \beta u'(a) = \alpha v(a) + \beta v'(a) = 0 \quad \text{and } \alpha^2 + \beta^2 > 0 \text{ with } \alpha, \beta \in \mathbb{R} \quad (2.23)$$

then $W(u, v^)(a) = 0$. Similarly, if*

$$\gamma u(b) + \delta u'(b) = \gamma v(b) + \delta v'(b) = 0 \quad \text{and } \gamma^2 + \delta^2 > 0 \text{ with } \gamma, \delta \in \mathbb{R} \quad (2.24)$$

then $W(u, v^)(b) = 0$.*

Proof. After taking the complex conjugate of the right hand side of equation (2.23) it is equivalent to the linear system

$$\begin{pmatrix} u(a) & u'(a) \\ v^*(a) & v'^*(a) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \quad (2.25)$$

This system only has a non-zero solution if the determinant of the 2×2 matrix vanishes, i.e.

$$0 = \det \begin{pmatrix} u(a) & u'(a) \\ v^*(a) & v'^*(a) \end{pmatrix} = W(u, v^*)(a) \quad (2.26)$$

The corresponding equation $W(u, v^*)(b) = 0$ at the right boundary b is proven in the same way. \square

Theorem 2.1. *Let L be a Sturm Liouville operator and $u, v \in D_L$ satisfy the boundary conditions in (2.14). Then*

$$(v, Lu) - (Lv, u) = \int_a^b (v^* Lu - u Lv^*) dx = 0 \quad (2.27)$$

i.e. the Sturm Liouville operator L is self-adjoint.

Proof. By the Green's identity in (2.19) we have

$$\begin{aligned} \int_a^b (v^* Lu - u Lv^*) dx &= [p(v^* u' - uv'^*)]_a^b \\ &= p(a)W(u, v^*)(a) - p(b)W(u, v^*)(b) \\ &= 0 \quad \text{by Lemma 2.3.} \end{aligned} \quad (2.28)$$

\square

Definition 2.5. An operator H is called **negative semi-definite** if

$$(u, Hu) \leq 0 \quad \text{for all functions } u \in D_H. \quad (2.29)$$

Similarly we can define:

- **H negative definite** $\Leftrightarrow (u, Hu) < 0$ for all $u \in D_H, u \neq 0$
- **H positive definite** $\Leftrightarrow (u, Hu) > 0$ for all $u \in D_H, u \neq 0$
- **H positive semi-definite** $\Leftrightarrow (u, Hu) \geq 0$ for all $u \in D_H$

This should be compared to the finite-dimensional case in Definition C.3.

Lemma 2.4. A (regular) Sturm-Liouville operator L is negative semi-definite if $q(x) \geq 0$ and the boundary conditions are chosen such that either the function itself or its derivative vanish at both boundaries. If the regularity condition is relaxed in the sense that we only assume $p(x) \geq 0$, the operator can also be negative semi-definite if $p(x)$ is zero at both boundaries.

Proof. The assumptions in the theorem imply that $p(a)u^*(a)u'(a) = p(b)u^*(b)u'(b) = 0$ since one of the terms in each product will be zero. Integrating by parts gives

$$\begin{aligned} (u, Lu) &= \int_a^b (u^*(p(x)u')' - q(x)u^*u) dx = - \int_a^b (p(x)u'^*u' + q(x)u^*u) dx + \underbrace{[pu^*u']_a^b}_{\substack{=p(b)u^*(b)u'(b) \\ -p(a)u^*(a)u'(a)=0}} \\ &= - \int_a^b (p(x)|u'|^2 + q(x)|u|^2) dx \leq 0 \quad \text{since } p(x) \geq 0 \text{ and } q(x) \geq 0. \end{aligned} \tag{2.30}$$

□

2.2 Eigenfunctions and orthogonality

Definition 2.6 (Eigenvalue, eigenfunction pairs). An **eigenvalue, eigenfunction pair** $(\lambda, \phi(x))$ of an SL operator L with weight function $r(x) > 0$ is a non-trivial solution (i.e. $\phi \neq 0$) of the ordinary differential equation

$$L\phi + \lambda r\phi = 0 \tag{2.31}$$

with boundary conditions

$$\alpha\phi(a) + \beta\phi'(a) = 0, \quad \gamma\phi(b) + \delta\phi'(b) = 0 \quad \text{with } \alpha, \beta, \gamma, \delta \in \mathbb{R}, \alpha^2 + \beta^2 > 0, \gamma^2 + \delta^2 > 0. \tag{2.32}$$

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

Example 2.3. Let $Lu \equiv u''$, $r(x) \equiv 1$, $a = 0$, $b = \pi$ and the corresponding ODE

$$\phi'' + \lambda\phi = 0, \quad \phi(0) = \phi(\pi) = 0. \tag{2.33}$$

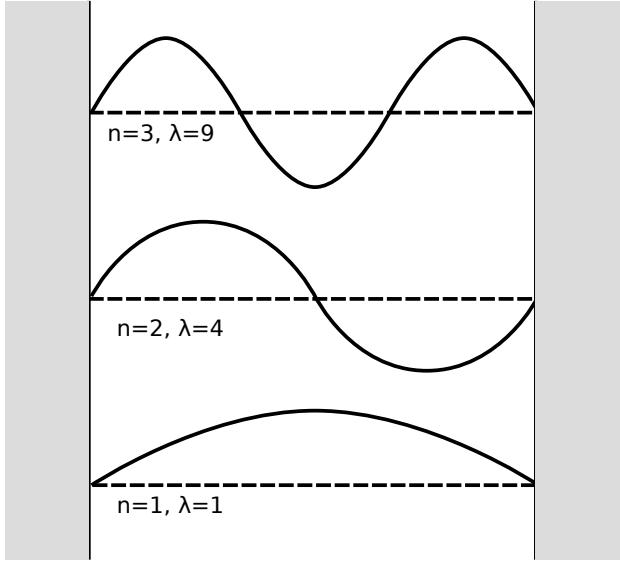


Figure 2.1: Eigenfunctions and eigenvalues for the SL system (2.33).

This equation describes the eigenmodes of a string which is attached at both ends or the quantum mechanical motion of a particle in a one-dimensional, infinitely deep box of length $L = \pi$. The 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 $\hbar = 6.63 \cdot 10^{-34} \text{ m}^2 \text{ kg s}^{-1}$ is known as Planck's constant. The wave function ϕ has to be zero at the boundaries $x = 0$ and $x = \pi$ (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 $\lambda > 0$ equation (2.33) has the solution

$$\phi(x) = A \cos(\lambda^{1/2}x) + B \sin(\lambda^{1/2}x), \quad \text{with arbitrary } A, B \in \mathbb{R}. \quad (2.34)$$

Since $\phi(0) = A$ the left boundary condition $\phi(0) = 0$ implies $A = 0$. Since then $\phi(\pi) = B \sin(\lambda^{1/2}\pi)$ the right boundary condition implies either $B = 0$ (which corresponds to $\phi = 0$, which we don't want since we are looking for non-zero solutions) or $\sin(\lambda^{1/2}\pi) = 0$. The zeros of the sin-function can be found at $n\pi$ with $n \in \mathbb{Z}$. Therefore $\lambda^{1/2}\pi = n\pi$ and

$$\lambda = n^2 \quad \text{for } n = 1, 2, \dots \quad (2.35)$$

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

$$(\lambda_n, \phi_n) = (n^2, \sin(nx)), \quad \text{with } n \in \mathbb{N}. \quad (2.36)$$

We call the n -th eigenfunction $\phi_n(x) = \sin(nx)$ and the corresponding eigenvalue λ_n . Coming back to the quantum mechanical interpretation, the calculation shows that the energy of a

particle confined to a box is quantized, with $E_n = \frac{\hbar^2}{2m}n^2$.

Student exercise: Show that there is no solution of (2.33) if $\lambda \leq 0$.

Lemma 2.5. *The functions ϕ_n defined in (2.36) are orthogonal in the sense that*

$$\int_0^\pi \phi_n^*(x)\phi_m(x) dx = 0 \quad \text{if } n \neq m \quad (2.37)$$

and we also have that

$$\int_0^\pi \phi_n^*(x)\phi_n(x) dx = \int_0^\pi |\phi_n(x)|^2 dx = \frac{\pi}{2}. \quad (2.38)$$

Proof. First assume that $n \neq m$. Then

$$\begin{aligned} \int_0^\pi \phi_n^*(x)\phi_m(x) dx &= \int_0^\pi \sin(nx)\sin(mx) dx \\ &= \frac{1}{2} \int_0^\pi (\cos((n-m)x) - \cos((n+m)x)) dx \\ &= \frac{1}{2} \left[\frac{\sin((n-m)x)}{n-m} - \frac{\sin((n+m)x)}{n+m} \right]_0^\pi = 0. \end{aligned} \quad (2.39)$$

Since $\sin(k\pi) = 0$ for all $k \in \mathbb{N}_0$. For $n = m$ we have

$$\begin{aligned} \int_0^\pi |\phi_n(x)|^2 dx &= \int_0^\pi \sin^2(nx) dx \\ &= \frac{1}{2} \int_0^\pi (1 - \cos(2nx)) dx \\ &= \frac{\pi}{2} - \frac{1}{2} \int_0^\pi \cos(2nx) dx = \frac{\pi}{2} - \frac{1}{4n} [\sin(2nx)]_0^\pi = \frac{\pi}{2}. \end{aligned} \quad (2.40)$$

□

This result is no coincidence. It is in fact an example of a much deeper and more general result which we will consider in the following. We first define a more general concept of orthogonality.

Definition 2.7 (Weighted inner product). Let $r(x) > 0$ on $[a, b]$ and define a weighted inner product for all functions u, v by

$$\langle u, v \rangle_r \equiv \int_a^b u^*(x)v(x)r(x) dx. \quad (2.41)$$

if $\langle u, v \rangle_r = 0$ then u and v are called **orthogonal** with respect to the weight function r .

The weighted inner product naturally induces a weighted norm.

Definition 2.8 (Weighted norm). Let u be a complex valued function on the interval $[a, b]$. We also assume that $r(x) > 0$ for all $x \in (a, b)$. The norm (weighted L_2 -norm) of u is

$$\|u\|_r \equiv \sqrt{\langle u, u \rangle_r} = \left(\int_a^b |u(x)|^2 r(x) dx \right)^{\frac{1}{2}}. \quad (2.42)$$

Note that since $r(x) > 0$ we have that $\|u\|_r = 0 \Leftrightarrow u = 0$, i.e. the norm of a function is zero if and only if the function is zero.

We are now ready to prove an important theorem for Sturm Liouville operators.

Theorem 2.2. *Let L be a self-adjoint Sturm-Liouville operator with associated weight function $r(x) > 0$ on (a, b) and $\{(\lambda_n, \phi_n)\}$ a set of eigenvalue/eigenfunction pairs such that $L\phi_n + \lambda_n r(x)\phi_n = 0$. Then we have*

1. $\lambda_n \in \mathbb{R}$
2. If L is negative semi-definite, all eigenvalues are strictly non-negative.
3. If $\lambda_n \neq \lambda_m$ then the corresponding eigenfunctions are orthogonal

$$\langle \phi_n, \phi_m \rangle_r = 0 \quad (2.43)$$

This major theorem is a direct consequence of the fact that L is a self-adjoint operator. It is instructive to compare this theorem and its proof to the finite-dimensional version which is given as Theorem C.1 in Appendix C.

Proof. **Part 1.**

Let $L\phi_n + \lambda_n r(x)\phi_n = 0$. Taking the complex conjugate of this equation yields the following equation for ϕ_n^* : $L\phi_n^* + \lambda_n^* r(x)\phi_n^* = 0$. Furthermore, if ϕ_n satisfies the boundary conditions in (2.14), so does the complex conjugate ϕ_n^* .

It follows that

$$\begin{aligned} & \phi_n^* L\phi_n - \phi_n L\phi_n^* + (\lambda_n - \lambda_n^*) r(x) \phi_n^* \phi_n = 0 \\ \Rightarrow & \int_a^b (\phi_n^* L\phi_n - \phi_n L\phi_n^*) dx + (\lambda_n - \lambda_n^*) \int_a^b r(x) |\phi_n|^2 dx = 0. \end{aligned} \quad (2.44)$$

But the first term is $(\phi_n, L\phi_n) - (L\phi_n, \phi_n) = 0$ since L is self-adjoint. Hence the second term has to vanish

$$(\lambda_n - \lambda_n^*) \int_a^b r(x) |\phi_n|^2 dx = (\lambda_n - \lambda_n^*) \cdot \underbrace{\|\phi_n\|_r^2}_{>0 \text{ as } \phi_n \neq 0} = 0. \quad (2.45)$$

Therefore $\lambda_n - \lambda_n^* = 0$ and $\lambda_n \in \mathbb{R}$.

Part 2.

Let (ϕ_n, λ_n) be an eigenfunction/eigenvalue pair. Multiply $L\phi_n + \lambda_n r(x)\phi_n = 0$ by ϕ_n^* and integrate over $[a, b]$ to obtain

$$(\phi_n, L\phi_n) + \lambda_n \langle \phi_n, \phi_n \rangle_r = 0 \Leftrightarrow \lambda_n \|\phi_n\|_r^2 = -(\phi_n, L\phi_n) \geq 0 \quad (2.46)$$

if L is negative semi-definite. Since $\|\phi_n\|_r > 0$ it follows that $\lambda_n \geq 0$.

Part 3.

Let

$$L\phi_n + \lambda_n r\phi_n = 0, \quad L\phi_m + \lambda_m r\phi_m = 0 \quad (2.47)$$

Multiply the complex conjugate of the first equation with ϕ_m and the second equation with ϕ_n^* and subtract to obtain (recall that $\lambda_n^* = \lambda_n$ from Part 1)

$$\phi_m L\phi_n^* - \phi_n^* L\phi_m + (\lambda_n - \lambda_m) r\phi_n^* \phi_m = 0 \quad (2.48)$$

Integrate over the interval $[a, b]$ to obtain, using the self-adjointness,

$$\begin{aligned} & \underbrace{\int_a^b (\phi_m L\phi_n^* - \phi_n^* L\phi_m) dx}_{=(L\phi_n,\phi_m)-(L\phi_m,\phi_n)=0} + (\lambda_n - \lambda_m) \int_a^b r\phi_n^* \phi_m dx = 0 \\ \Rightarrow & (\lambda_n - \lambda_m) \int_a^b r\phi_n^* \phi_m dx = 0 \end{aligned} \quad (2.49)$$

If $\lambda_n - \lambda_m \neq 0$ this equation can only be satisfied if

$$0 = \int_a^b r\phi_n^*\phi_m dx = \langle \phi_n, \phi_m \rangle_r. \quad (2.50)$$

This concludes the proof of Theorem 2.2. \square

Example 2.4 (The damped harmonic oscillator). Let

$$u'' + u' + \lambda u = 0, \quad \text{with } u(0) = u(1) = 0. \quad (2.51)$$

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

$$p(x) = \exp\left(\int_0^x \frac{1}{1} dy\right) = e^x, \quad q(x) = 0, \quad r(x) = e^x \quad (2.52)$$

and rewrite Eq. (2.51) as

$$(e^x u')' + \lambda e^x u = 0. \quad (2.53)$$

We will use the SL form later to explicitly check the orthogonality of the eigenfunctions. To solve the equation, we go back to (2.51) and make the ansatz $u(x) = Ce^{\mu x}$. Inserting this, we obtain

$$(\mu^2 + \mu + \lambda) Ce^{\mu x} = 0 \quad (2.54)$$

and hence

$$\mu^2 + \mu + \lambda = 0, \quad \Rightarrow \mu_{\pm} = -\frac{1}{2} \pm \frac{\sqrt{1-4\lambda}}{2}. \quad (2.55)$$

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

$$u(x) = C_+ e^{\mu_+ x} + C_- e^{\mu_- x}. \quad (2.56)$$

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 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

$$\omega \equiv \frac{\sqrt{4\lambda-1}}{2}. \quad (2.57)$$

Hence the general solution is of the form

$$u(x) = e^{-\frac{x}{2}} (A \cos(\omega x) + B \sin(\omega x)) \quad (2.58)$$

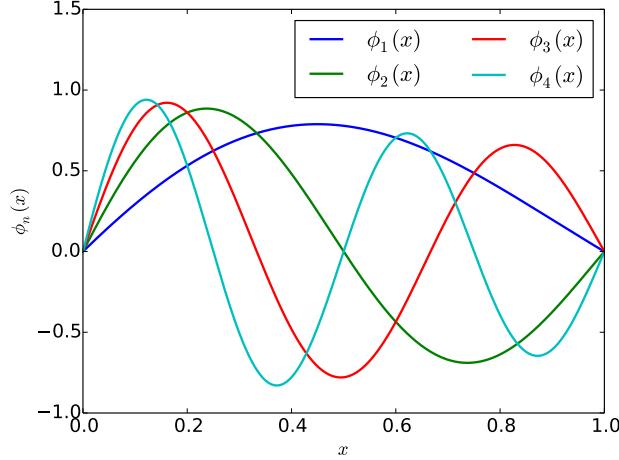


Figure 2.2: Eigenfunctions in (2.61).

The boundary conditions imply

$$u(0) = 0 \Rightarrow A = 0, \quad u(1) = 0 \Rightarrow \sin(\omega) = 0 \Rightarrow \omega = n\pi, n \in \mathbb{Z}, n \neq 0. \quad (2.59)$$

Using the definition of ω in (2.57) and solving for λ we obtain

$$\frac{\sqrt{4\lambda - 1}}{2} = n\pi \Rightarrow \lambda = \frac{1}{4} + \pi^2 n^2. \quad (2.60)$$

The eigenvalue/eigenfunction pairs of the SL equation in (2.53) are therefore

$$(\lambda_n, \phi_n) = \left(\frac{1}{4} + \pi^2 n^2, e^{-\frac{x}{2}} \sin(n\pi x) \right) \quad \text{for } n \in \mathbb{N}. \quad (2.61)$$

The lowest eigenfunctions are plotted in Fig. 2.2. In particular, with this choice of boundary conditions the SL operator is negative definite and hence all eigenvalues are strictly positive. We can also explicitly check those pairs for orthogonality

$$\begin{aligned} \langle \phi_n, \phi_m \rangle_r &= \int_0^1 e^{-\frac{x}{2}} \sin(n\pi x) e^{-\frac{x}{2}} \sin(m\pi x) e^x dx \\ &= \int_0^1 \sin(n\pi x) \sin(m\pi x) dx = 0 \quad \text{if } n \neq m \end{aligned} \quad (2.62)$$

as predicted by Theorem 2.2.

2.3 Orthogonal and orthonormal expansions

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.

Definition 2.9 (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

$$\langle \psi_n, \psi_m \rangle_r = 0 \quad \text{for } n \neq m. \quad (2.63)$$

Definition 2.10 (Orthonormality). An orthogonal set $\{\phi_n\}$, $n \in \mathbb{N}$ of functions is *orthonormal* if in addition

$$\langle \phi_n, \phi_n \rangle_r = 1. \quad (2.64)$$

Lemma 2.6 (Orthonormalisation). If the set $\{\psi_n\}$ is orthogonal and

$$\phi_n(x) = (\langle \psi_n, \psi_n \rangle_r)^{-1/2} \psi_n(x) \quad (2.65)$$

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. It is obvious that $\{\phi_n\}$ is orthogonal due to the bi-linearity of the inner product $\langle \cdot, \cdot \rangle_r$:

$$\langle \phi_n, \phi_m \rangle_r = (\langle \psi_n, \psi_n \rangle_r)^{-1/2} (\langle \psi_m, \psi_m \rangle_r)^{-1/2} \langle \psi_n, \psi_m \rangle_r = 0. \quad (2.66)$$

Orthonormality follows from

$$\langle \phi_n, \phi_n \rangle_r = ((\langle \psi_n, \psi_n \rangle_r)^{-1/2})^2 \langle \psi_n, \psi_n \rangle_r = 1. \quad (2.67)$$

□

Example 2.5. 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 (see Lemma 2.5). Since

$$\langle \psi_n, \psi_n \rangle_r = \int_0^\pi \sin^2(nx) dx = \frac{\pi}{2} \quad (2.68)$$

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

2.4 Approximation of functions

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

$$f(x) \approx f_N(x) \equiv \sum_{n=1}^N \alpha_n \phi_n(x) \quad (2.69)$$

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

$$E_N = \|f - f_N\|_r^2. \quad (2.70)$$

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?

Theorem 2.3 (Optimal approximation). *The optimal expansion coefficients in (2.69) can be chosen as follows:*

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

$$E_N = \|f\|_r^2 - \sum_{n=1}^N \alpha_n^2 = \|f\|_r^2 - \sum_{n=1}^N \langle f, \phi_n \rangle_r^2 \quad (2.71)$$

Proof. Using the definition of the approximation error in (2.70) we have

$$E_N = \langle f - f_N, f - f_N \rangle_r = \langle f, f \rangle_r - 2\langle f, f_N \rangle_r + \langle f_N, f_N \rangle_r. \quad (2.72)$$

Now consider the last term and the expansion in (2.69) and split the double sum over n, m into two sums with $n = m$ and $n \neq m$

$$\begin{aligned} \langle f_N, f_N \rangle_r &= \left\langle \sum_{n=1}^N \alpha_n \phi_n, \sum_{m=1}^N \alpha_m \phi_m \right\rangle_r \\ &= \sum_{n=1}^N \alpha_n^2 \cdot \underbrace{\langle \phi_n, \phi_n \rangle_r}_{=1 \text{ (by orthonormality)}} + \sum_{n \neq m} \alpha_n \alpha_m \cdot \underbrace{\langle \phi_n, \phi_m \rangle_r}_{=0 \text{ (by orthogonality)}} = \sum_{n=1}^N \alpha_n^2. \end{aligned} \quad (2.73)$$

We also have

$$\langle f, f_N \rangle_r = \langle f, \sum_{n=1}^N \alpha_n \phi_n \rangle_r = \sum_{n=1}^N \alpha_n \langle f, \phi_n \rangle_r. \quad (2.74)$$

Insert this all into (2.72) to obtain

$$E_N = \|f\|_r^2 - 2 \sum_{n=1}^N \alpha_n \langle f, \phi_n \rangle_r + \sum_{n=1}^N \alpha_n^2. \quad (2.75)$$

The coefficients α_n which minimise the error have to satisfy

$$0 = \frac{\partial E_N}{\partial \alpha_n} = -2 \langle f, \phi_n \rangle_r + 2\alpha_n \quad \text{for all } n \in \{1, \dots, N\} \quad (2.76)$$

or $\alpha_n = \langle f, \phi_n \rangle_r$. Furthermore, the Hessian matrix H with

$$H_{nm} = \frac{\partial^2 E_N}{\partial \alpha_n \partial \alpha_m} = \begin{cases} 2 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases} \quad (2.77)$$

is proportional to the N -dimensional unit-matrix and therefore positive definite. Hence the α_n above indeed minimise the approximation error.

By inserting $\alpha_n = \langle f, \phi_n \rangle_r$ into (2.75) we can write down the error in this case as

$$E_N = \|f\|_r^2 - 2 \sum_{n=1}^N \langle f, \phi_n \rangle_r^2 + \sum_{n=1}^N \langle f, \phi_n \rangle_r^2 = \|f\|_r^2 - \sum_{n=1}^N \langle f, \phi_n \rangle_r^2. \quad (2.78)$$

□

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

$$0 \leq \sum_{n=1}^N \langle f, \phi_n \rangle_r^2 \leq \|f\|_r^2. \quad (2.79)$$

Definition 2.11. An orthonormal set $\{\phi_n\}$ is *complete* if $E_N \rightarrow 0$ as $N \rightarrow \infty$.

Theorem 2.4. *The eigenfunctions of a Sturm-Liouville operator form a complete set.*

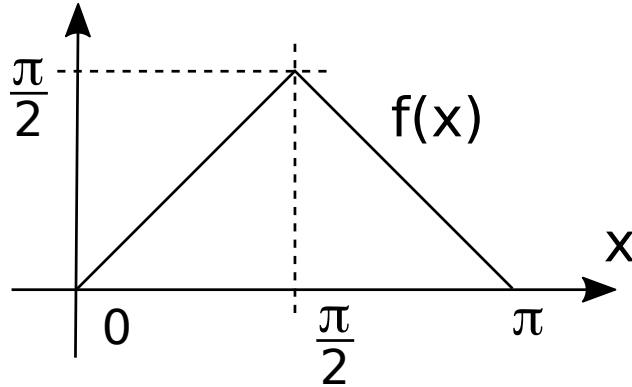


Figure 2.3: Hat function defined in (2.81)

This is a very deep theorem, and the proof is beyond the scope of this course. However, from Theorem 2.4 we deduce the following result:

Theorem 2.5 (Parseval). *If the orthonormal set $\{\phi_n\}$ is complete then*

$$\sum_{n=1}^{\infty} \langle f, \phi_n \rangle_r^2 = \sum_{n=1}^{\infty} \alpha_n^2 = \|f\|_r^2 \quad (2.80)$$

Proof. Let $N \rightarrow \infty$ in (2.71) and use Definition 2.11. □

Example 2.6. 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 orthonormal set with eigenvalues $\lambda_n = n^2$ (see Example 2.5). Let f be a hat-function (see Fig. 2.3)

$$f(x) = \begin{cases} x & \text{for } 0 \leq x < \frac{\pi}{2} \\ \pi - x & \text{for } \frac{\pi}{2} \leq x \leq \pi \end{cases} \quad (2.81)$$

α_n is readily evaluated as the following integral over x

$$\begin{aligned} \alpha_n &= \sqrt{\frac{2}{\pi}} \int_0^\pi f(x) \sin(nx) dx = \sqrt{\frac{2}{\pi}} \left(\int_0^{\frac{\pi}{2}} x \sin(nx) dx + \int_{\frac{\pi}{2}}^\pi (\pi - x) \sin(nx) dx \right) \\ &= \begin{cases} 2\sqrt{\frac{2}{\pi}} \int_0^{\frac{\pi}{2}} x \sin(nx) dx = 2\sqrt{\frac{2}{\pi}} \frac{(-1)^{\frac{n-1}{2}}}{n^2} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}. \end{aligned} \quad (2.82)$$

Therefore the hat function f in (2.81) can be approximated as

$$f(x) \approx f_{2N-1}(x) = \sum_{n=1}^N \alpha_{2n-1} \phi_{2n-1}(x) = -\frac{4}{\pi} \sum_{n=1}^N \frac{(-1)^n}{(2n-1)^2} \sin((2n-1)x). \quad (2.83)$$

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

$$\|f\|_r^2 = \int_0^\pi f^2 dx = \int_0^{\frac{\pi}{2}} x^2 dx + \int_{\frac{\pi}{2}}^\pi (\pi - x)^2 dx = 2 \int_0^{\frac{\pi}{2}} x^2 dx = \frac{\pi^3}{12}. \quad (2.84)$$

(2.80) gives the interesting result

$$\begin{aligned} \frac{\pi^3}{12} &= \|f\|_r^2 = \sum_{n=1}^{\infty} \alpha_n^2 = \frac{8}{\pi} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^4} \quad \text{and therefore} \\ \frac{\pi^4}{96} &= \sum_{n=1}^{\infty} \frac{1}{(2n-1)^4} = 1 + \frac{1}{3^4} + \frac{1}{5^4} + \frac{1}{7^4} + \frac{1}{9^4} + \dots \end{aligned} \quad (2.85)$$

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

$$\begin{aligned} E_{2N} &= \|f\|_r^2 - \sum_{n=1}^{2N} \alpha_n^2 = \sum_{n=2N+1}^{\infty} \alpha_n^2 = \frac{8}{\pi} \sum_{n=N+1}^{\infty} \frac{1}{(2n-1)^4} \\ &= \frac{8}{\pi} \left(\sum_{k=2N+1}^{\infty} \frac{1}{k^4} - \sum_{k=N+1}^{\infty} \frac{1}{(2k)^4} \right) \approx \frac{8}{\pi} \left(\int_{2N+1}^{\infty} \frac{dx}{x^4} - \frac{1}{16} \int_{N+1}^{\infty} \frac{dx}{x^4} \right) \\ &= \frac{8}{3\pi} \left(\frac{1}{(2N+1)^3} - \frac{1}{16} \cdot \frac{1}{(N+1)^3} \right) \approx \frac{1}{6\pi N^3}. \end{aligned} \quad (2.86)$$

Example 2.7. 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

$$f(x) = \exp[-x^2 + \sin(x)] \quad \text{for } x \in [-1, 1]. \quad (2.87)$$

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

$$\left(\sqrt{1-x^2} u' \right)' + \lambda \frac{u}{\sqrt{1-x^2}} = 0 \quad \text{for } x \in [-1, 1]. \quad (2.88)$$

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.

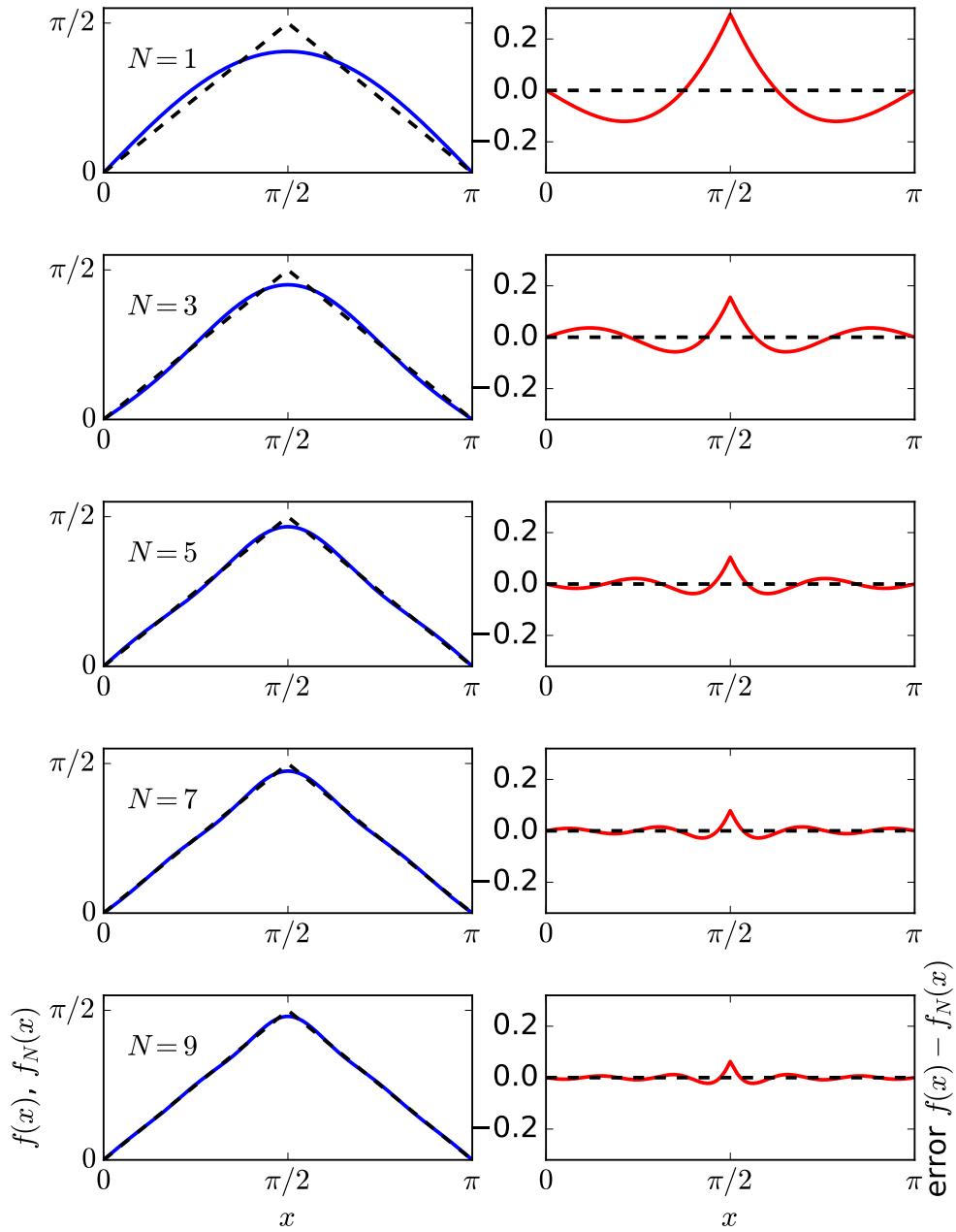


Figure 2.4: Expansion of the hat function $f(x)$ (2.81) in terms of basis functions. The error $f(x) - f_N(x)$ is shown on the right.

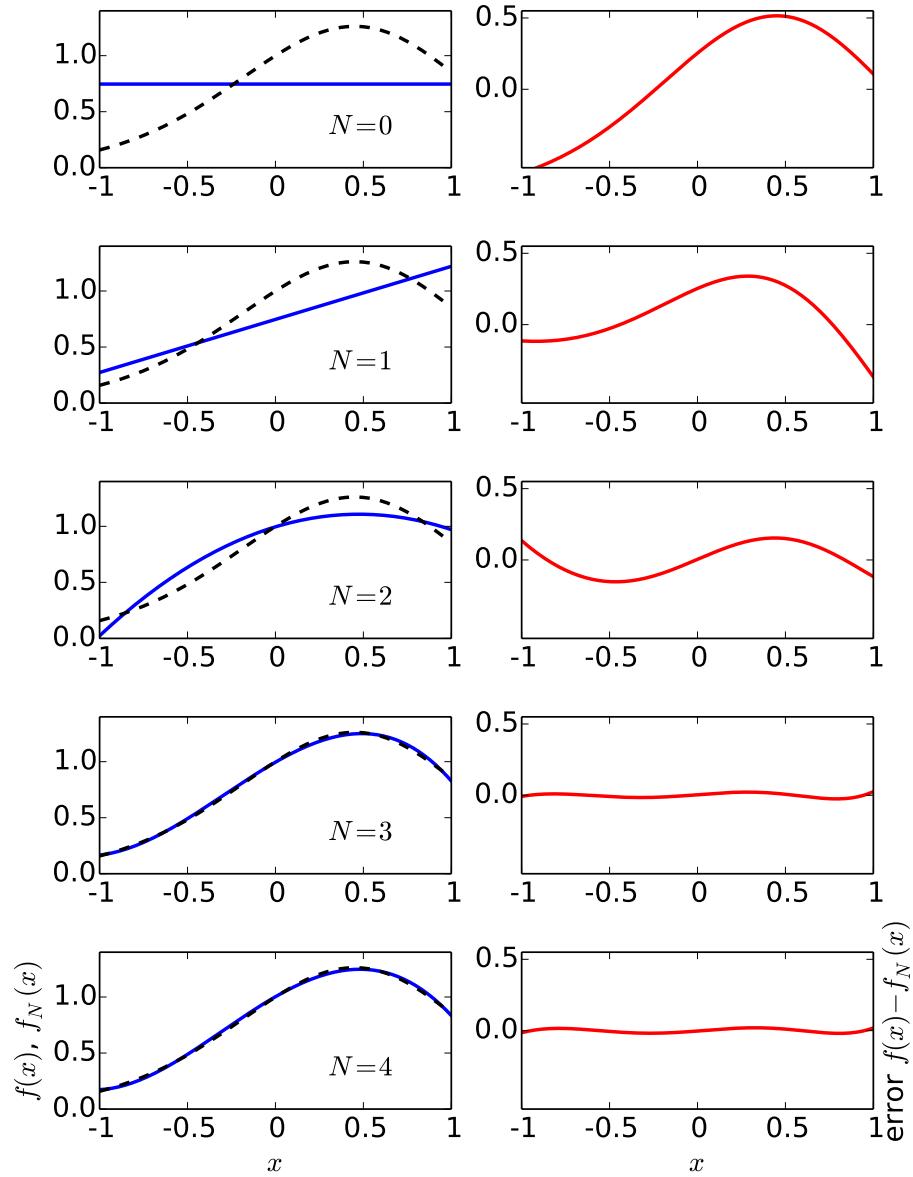


Figure 2.5: Expansion of the function $f(x) = \exp[-x^2 + \sin(x)]$ in terms of Chebyshev polynomials. The error $f(x) - f_N(x)$ is shown on the right.

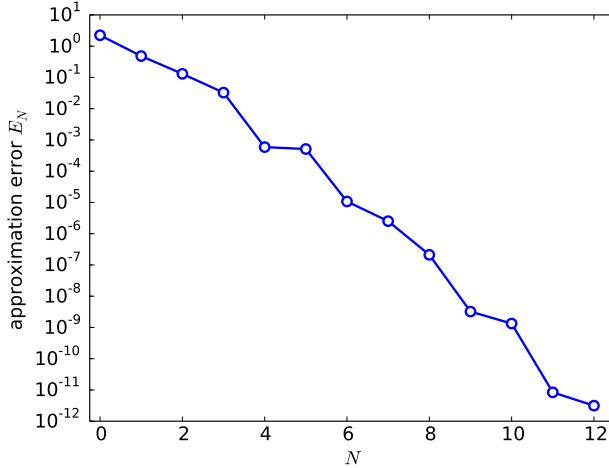


Figure 2.6: Approximation error E_N for the approximation of the function $f(x) = \exp[-x^2 + \sin(x)]$ in terms of Chebyshev polynomials.

2.4.1 Legendre polynomials

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.

Lemma 2.7. *The n -th eigenvalue/eigenfunction pair of the Legendre equation*

$$((1 - x^2)u')' + \lambda u = 0 \quad \text{on } I = [-1, 1] \quad (2.89)$$

is given by

$$(\lambda_n, \phi_n) = (n(n+1), P_n) \quad \text{with } P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n] \quad \text{for } n = 0, 1, 2, \dots \quad (2.90)$$

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

Note that $p(\pm 1) = 0$ and hence it is not necessary to specify boundary conditions. We only have to fix the normalisation of the function by defining its value at $x = 1$.

Proof. Use the identity

$$(1 - x^2) \frac{d}{dx} [(x^2 - 1)^n] + 2nx(x^2 - 1)^n = 0 \quad (2.91)$$

and take the $n + 1$ -st derivative of equation (2.91) as follows:

$$\frac{d}{dx} \frac{d^n}{dx^n} \left[(1-x)^2 \frac{d}{dx} [(x^2-1)^n] \right] + 2n \frac{d^{n+1}}{dx^{n+1}} [x(x^2-1)^n] = 0 \quad (2.92)$$

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)

$$\frac{d^n(f \cdot g)}{dx^n} = \sum_{k=0}^n \binom{n}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k}g}{dx^{n-k}} \quad (2.93)$$

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 obtain¹

$$\begin{aligned} & \frac{d}{dx} \left\{ (1-x^2) \frac{d}{dx} \frac{d^n}{dx^n} [(x^2-1)^n] - 2nx \frac{d^n}{dx^n} [(x^2-1)^n] - 2 \frac{n(n-1)}{2} \frac{d^{n-1}}{dx^{n-1}} [(x^2-1)^n] \right\} \\ & \qquad \qquad \qquad + 2nx \frac{d^{n+1}}{dx^{n+1}} [(x^2-1)^n] + 2n^2 \frac{d^n}{dx^n} [(x^2-1)^n] = 0 \\ \Leftrightarrow & \frac{d}{dx} \left\{ (1-x^2) \frac{d}{dx} \frac{d^n}{dx^n} [(x^2-1)^n] \right\} + n(n+1) \frac{d^n}{dx^n} [(x^2-1)^n] = 0 \end{aligned} \quad (2.94)$$

Hence the function $\frac{d^n}{dx^n} [(x^2-1)^n]$ and therefore $P_n(x)$ satisfies the Legendre equation with $\lambda = n(n+1)$. Note also that since $p(-1) = p(1) = 0$ it follows from Lemma 2.4 that the SL operator is negative semi-definite and therefore $\lambda \geq 0$. $P_n(x)$ is clearly 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

$$P_n(1) = \lim_{\epsilon \rightarrow 0} \frac{1}{2^n n!} (-1)^n \frac{d^n}{d\epsilon^n} [((1-\epsilon)^2-1)^n] = \lim_{\epsilon \rightarrow 0} \frac{1}{2^n n!} \frac{d^n}{d\epsilon^n} (2\epsilon)^n = 1. \quad (2.95)$$

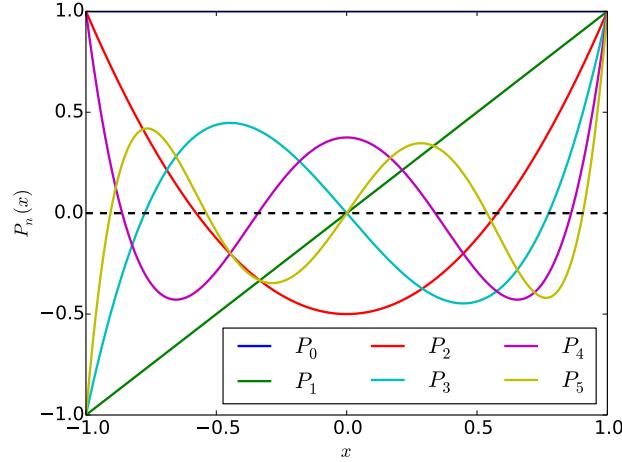
□

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

Lemma 2.8 (Generating function). *The Legendre polynomials can be obtained as coefficients of the following power series of the generating function $g(x; t)$*

$$g(x; t) = \frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x) t^n \quad \text{for } x \in [-1, 1] \text{ and } |t| < 1. \quad (2.96)$$

¹note 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:

$$\frac{\partial}{\partial x} \left[(1 - x^2) \frac{\partial g}{\partial x} \right] + t \frac{\partial^2 (tg)}{\partial t^2} = 0. \quad (2.97)$$

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

$$g(x; t) = \sum_{n=0}^{\infty} a_n(x) t^n. \quad (2.98)$$

Evaluating the first term of (2.97) gives

$$\frac{\partial}{\partial x} \left[(1 - x^2) \frac{\partial g}{\partial x} \right] = \sum_{n=0}^{\infty} \left(\frac{d}{dx} \left[(1 - x^2) \frac{da_n(x)}{dx} \right] \right) t^n \quad (2.99)$$

Evaluating the second term of (2.97) gives

$$\sum_{n=0}^{\infty} a_n(x) t \frac{d^2}{dt^2} t^{n+1} = \sum_{n=0}^{\infty} n(n+1) a_n(x) t^n \quad (2.100)$$

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

$$\frac{d}{dx} \left[(1 - x^2) \frac{da_n(x)}{dx} \right] + n(n+1)a_n(x) = 0 \quad (2.101)$$

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

$$g(1; t) = \frac{1}{1-t} = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} a_n(1) t^n \quad (2.102)$$

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

Lemma 2.9. *The Legendre polynomials defined in Lemma 2.7 fulfill the recursion relation*

$$(n+1)P_{n+1} = (2n+1)xP_n(x) - nP_{n-1}(x) \quad \text{for } n = 0, 1, 2, \dots \quad (2.103)$$

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

Proof. Take the derivative of the generating function in (2.96) with respect to t to obtain

$$\frac{x-t}{(1-2xt+t^2)^{3/2}} = \sum_{n=0}^{\infty} nP_n(x)t^{n-1}. \quad (2.104)$$

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

$$\frac{x-t}{\sqrt{1-2xt+t^2}} = (1-2xt+t^2) \sum_{n=0}^{\infty} nP_n(x)t^{n-1}. \quad (2.105)$$

Replacing $1/\sqrt{1-2xt+t^2}$ by the sum in (2.96) we have

$$(x-t) \sum_{n=0}^{\infty} P_n(x)t^n = (1-2xt+t^2) \sum_{n=0}^{\infty} nP_n(x)t^{n-1}. \quad (2.106)$$

For this to hold for every t , the equation has to hold at each order of t^n . From this the recursion relation (2.103) can be read off by comparing coefficients. \square

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

$$(\sqrt{1-x^2}u')' + \lambda \frac{u}{\sqrt{1-x^2}} = 0, \quad (2.107)$$

which we already encountered in Example 2.7, has eigenvalue/eigenfunctions $\{n^2, T_n(x)\}$ where $T_n(x)$ is a polynomial which can be read off from

$$\frac{1-xt}{1-2xt+t^2} = \sum_{n=0}^{\infty} T_n(x)t^n \quad (2.108)$$

Again, this allows deducing a recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \quad (2.109)$$

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

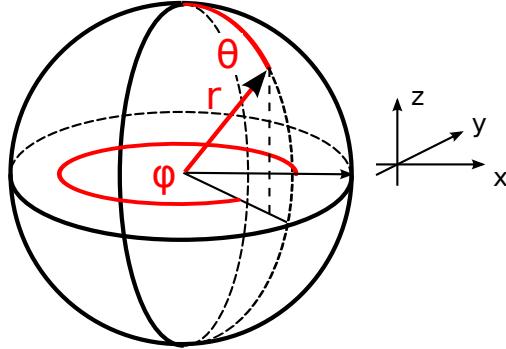


Figure 2.8: Spherical coordinates

Example 2.8. Consider a point particle with charge Q at position $\mathbf{r}_0 = (x_0, y_0, z_0)$. The electric potential at point $\mathbf{r} = (x, y, z) \neq \mathbf{r}_0$ is then given by

$$V_{\text{point}}(\mathbf{r}; \mathbf{r}_0) = \frac{Q}{4\pi} \frac{1}{\|\mathbf{r} - \mathbf{r}_0\|} = \frac{Q}{4\pi} \frac{1}{\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}}. \quad (2.110)$$

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 $\Delta V_{\text{point}}(\mathbf{r}; \mathbf{r}_0) = 0$ for $\mathbf{r} \neq \mathbf{r}_0$. By the superposition principle, the potential of a charge distribution with density $\rho(\mathbf{r})$ is given by (see Fig. 2.9):

$$V(\mathbf{r}) = \int_{\mathbb{R}^3} \rho(\mathbf{r}') V_{\text{point}}(\mathbf{r}; \mathbf{r}') dV = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} dV \quad (2.111)$$

and it can be shown that $\Delta V(\mathbf{r}) = -\rho(\mathbf{r})$. Consider a vertical rod with a charge of $\rho(z)$ per unit length with $\rho(z) = 0$ for $|z| > R$. 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 and θ is the angle relative to the vertical axis. ϕ is the angle between the positive x -axis and the projection of \mathbf{r} onto the xy plane. We can write $\mathbf{r} = (r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta)$ and $\mathbf{r}' = (0, 0, z)$. An easy calculation shows that $\|\mathbf{r} - \mathbf{r}'\|^2 = \mathbf{r} \cdot \mathbf{r} + \mathbf{r}' \cdot \mathbf{r}' - 2\mathbf{r}' \cdot \mathbf{r} = z^2 + r^2 - 2rz \cos \theta$; 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

$$V_{\text{rod}}(r, \theta) = \frac{1}{4\pi} \int_{-R}^{+R} \frac{\rho(z) dz}{\sqrt{z^2 + r^2 - 2rz \cos \theta}}. \quad (2.112)$$

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 (2.96) and therefore as a sum of

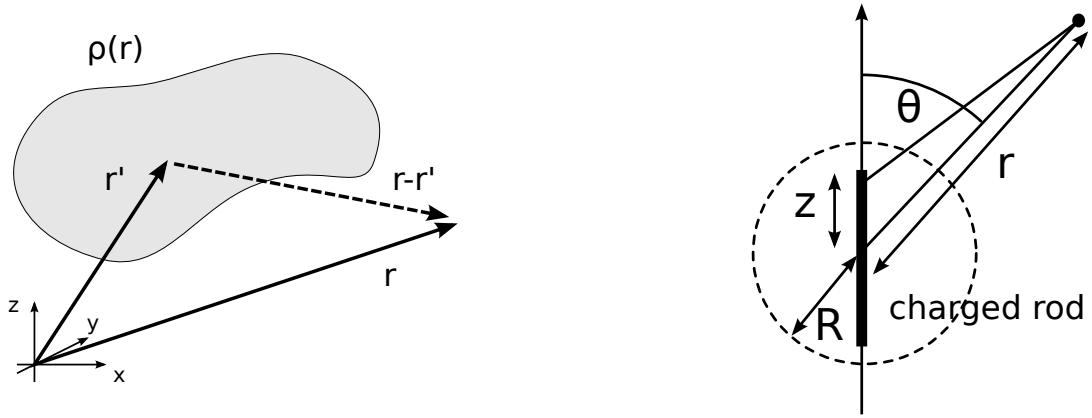


Figure 2.9: Illustration of the superposition principle (2.111) for a charge distribution $\rho(r)$ (left) and potential of a charged rod (right).

Legendre polynomials as

$$\begin{aligned}
 V_{\text{rod}}(r, \theta) &= \frac{1}{4\pi} \int_{-R}^{+R} \frac{\rho(z) dz}{r \sqrt{1 - 2(z/r) \cos \theta + (z/r)^2}} \\
 &= \frac{1}{4\pi r} \int_{-R}^{+R} \rho(z) g(\cos \theta, z/r) dz \\
 &= \frac{1}{4\pi r} \int_{-R}^{+R} \rho(z) \sum_{n=0}^{\infty} P_n(\cos \theta) \left(\frac{z}{r}\right)^n dz \\
 &= \frac{1}{4\pi} \sum_{n=0}^{\infty} Q_n r^{-(n+1)} P_n(\cos \theta) \quad \text{with } Q_n = \int_{-R}^{+R} \rho(z) z^n dz.
 \end{aligned} \tag{2.113}$$

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.

2.5 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

$$Lu + \lambda r(x)u = f(x) \quad + \text{boundary conditions} \tag{2.114}$$

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 (2.114) uniquely or not.

Theorem 2.6 (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 (2.114) has the unique solution*

$$u(x) = \int_a^b G(x, y) f(y) dy = \sum_{n=1}^{\infty} \frac{\phi_n(x)}{\lambda - \lambda_n} \int_a^b \phi_n(y) f(y) dy \quad (2.115)$$

where

$$G(x, y) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(y)}{\lambda - \lambda_n}. \quad (2.116)$$

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

$$\beta_n \equiv \int_a^b f(x)\phi_n(x) dx. \quad (2.117)$$

(a) *If $\beta_n \neq 0$ the inhomogeneous SL equation (2.114) 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 (2.114) exists. As $\{\phi_n\}$ is complete (see Theorem 2.4) and orthonormal, we can expand u and the ratio f/r in this basis as

$$u(x) = \sum_{n=1}^{\infty} \alpha_n \phi_n(x), \quad \frac{f(x)}{r(x)} = \sum_{n=1}^{\infty} \beta_n \phi_n(x). \quad (2.118)$$

The expansion coefficients β_n in Eq. (2.118) are indeed the same as in Eq. (2.117) since orthonormality implies

$$\int_a^b f(x)\phi_n(x) dx = \int_a^b \frac{f(x)}{r(x)} \phi_n(x)r(x) dx = \left\langle \frac{f}{r}, \phi_n \right\rangle_r = \left\langle \sum_{m=1}^{\infty} \beta_m \phi_m, \phi_n \right\rangle_r = \beta_n. \quad (2.119)$$

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

$$Lu + \lambda ru = \sum_{n=1}^{\infty} \alpha_n (L\phi_n + \lambda r\phi_n) = \sum_{n=1}^{\infty} \alpha_n (\lambda - \lambda_n) r\phi_n \quad (2.120)$$

But, since $Lu + \lambda ru = f(x) = \sum_{n=1}^{\infty} \beta_n \phi_n(x)r(x)$ by (2.118)

$$\sum_{n=1}^{\infty} \alpha_n (\lambda - \lambda_n) r\phi_n = \sum_{n=1}^{\infty} \beta_n \phi_n r. \quad (2.121)$$

Due to the orthonormality of the basis functions ϕ_n this equation has to be fulfilled separately for each term in the sum²,

$$\alpha_n(\lambda - \lambda_n) = \beta_n \quad \text{for all } n \in \mathbb{N}. \quad (2.122)$$

If we have $\lambda \neq \lambda_n$ then $\alpha_n = \beta_n/(\lambda - \lambda_n)$ and we can write³

$$u(x) = \sum_{n=1}^{\infty} \frac{\beta_n \phi_n(x)}{\lambda - \lambda_n}. \quad (2.123)$$

Inserting the expression in Eq. (2.119) for β_n into (2.123) we obtain the desired result

$$\begin{aligned} u(x) &= \sum_{n=1}^{\infty} \left(\int_a^b f(y) \phi_n(y) dy \right) \frac{\phi_n(x)}{\lambda - \lambda_n} \\ &= \int_a^b G(x, y) f(y) dy \quad \text{with } G(x, y) = \sum_{n=1}^{\infty} \frac{\phi_n(x) \phi_n(y)}{\lambda - \lambda_n}. \end{aligned} \quad (2.124)$$

Now consider case 2., i.e. $\lambda = \lambda_n$ for some n . If $\beta_n \neq 0$ then in (2.122) we have

$$\alpha_n \underbrace{(\lambda - \lambda_n)}_{=0} = \underbrace{\beta_n}_{\neq 0} \quad (2.125)$$

so α_n is undefined and there is no solution $u(x)$ for the inhomogeneous SL equation (2.114). If, on the other hand $\beta_n = 0$, we have

$$\alpha_n \underbrace{(\lambda - \lambda_n)}_{=0} = 0 \quad (2.126)$$

and hence α_n is arbitrary. There is an infinite number of solutions to the inhomogeneous SL equation (2.114), parameterised by the free parameter α_n :

$$\begin{aligned} u(x) &= u_0(x) + \alpha_n \phi_n(x) \quad \text{with } u_0(x) = \int_a^b \hat{G}(x, y) f(y) dy \text{ and } \alpha_n \in \mathbb{R} \\ \text{where } \hat{G}(x, y) &= \sum_{m=1, m \neq n}^{\infty} \frac{\phi_m(x) \phi_m(y)}{\lambda - \lambda_m}. \end{aligned} \quad (2.127)$$

□

Example 2.9. Consider the inhomogeneous wave equation with Dirichlet boundary conditions on the interval $[0, \pi]$:

$$u'' + \lambda u = f(x), \quad \text{with } u(0) = u(\pi) = 0 \quad (2.128)$$

²To see this, take the inner product with ϕ_m , i.e. multiply by ϕ_m and integrate over the interval $[a, b]$.

³This argument might appear circular since we assumed that a solution exists. However, we could equally well have started by construction u as in (2.123) and then show that this fulfills the SL equation (2.114).

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

$$f(x) = \begin{cases} x & \text{for } 0 \leq x < \frac{\pi}{2} \\ \pi - x & \text{for } \frac{\pi}{2} \leq x \leq \pi \end{cases} \quad (2.129)$$

From Lemma 2.5 and Example 2.5 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

$$\begin{aligned} G(x, y) &= \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(nx) \sin(ny)}{\lambda - n^2} \\ \Rightarrow u(x) &= \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(nx)}{\lambda - n^2} \int_0^{\pi} \sin(ny) f(y) dy \end{aligned} \quad (2.130)$$

The integral over y can be evaluated as in (2.82) to obtain

$$\int_0^{\pi} \sin(ny) f(y) dy = \begin{cases} \frac{2(-1)^{\frac{n-1}{2}}}{n^2} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}. \quad (2.131)$$

Therefore

$$u(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2(\lambda - (2n+1)^2)} \sin((2n+1)x) \quad (2.132)$$

Example 2.10. Now consider the same equation but with Neumann boundary conditions at both ends:

$$u'' + \lambda u = f(x), \quad \text{with } u'(0) = u'(\pi) = 0 \quad (2.133)$$

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 $n = 2, 3, \dots$ with corresponding eigenvalues $\lambda_n = (n-1)^2$ for $n = 1, 2, 3, \dots$. 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 2.6 and hence (2.133) only has a solution if

$$\int_0^{\pi} f(x) \phi_1(x) dx = 0, \quad \Leftrightarrow \quad \int_0^{\pi} f(x) dx = 0. \quad (2.134)$$

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:

$$u'(x) = \int_0^x f(y) dy + C' \quad (2.135)$$

but, due to the assumption (2.134) we have $u'(\pi) = \int_0^{\pi} f(y) dy = 0$ and hence the integration constant $C' = 0$. Integrating again we find that

$$u(x) = \int_0^x \left(\int_0^z f(y) dy \right) dz + C \quad (2.136)$$

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 (2.135) 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 2.6.

2.6 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

$$\Delta V(r, \cos \theta) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial V}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial V}{\partial \theta} \right) = 0 \quad (2.137)$$

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 $\frac{\partial}{\partial \theta} = -\sin \theta \frac{\partial}{\partial \cos \theta} = -\sin \theta \frac{\partial}{\partial x}$

$$r^2 \Delta V(r, x) = \frac{\partial}{\partial r} \left(r^2 \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial x} \left(\sin^2 \theta \frac{\partial V}{\partial x} \right) = \frac{\partial}{\partial r} \left(r^2 \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial x} \left((1 - x^2) \frac{\partial V}{\partial x} \right) = 0. \quad (2.138)$$

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

$$r^2 \frac{\Delta V(r, x)}{V(r, x)} = \frac{1}{\zeta(r)} \frac{d}{dr} \left(r^2 \frac{d\zeta(r)}{dr} \right) + \frac{1}{u(x)} \frac{d}{dx} \left((1 - x^2) \frac{du(x)}{dx} \right) = 0. \quad (2.139)$$

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

$$\frac{1}{\zeta(r)} \frac{d}{dr} \left(r^2 \frac{d\zeta(r)}{dr} \right) = \lambda = -\frac{1}{u(x)} \frac{d}{dx} \left((1 - x^2) \frac{du(x)}{dx} \right). \quad (2.140)$$

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

$$((1 - x^2)u')' + \lambda u = 0. \quad (2.141)$$

However, from Lemma 2.7 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

$$(r^2 \zeta_n')' = n(n + 1)\zeta_n \quad (2.142)$$

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

$$\zeta_n(r) = \tilde{Q}_n r^n + Q_n r^{-(n+1)}. \quad (2.143)$$

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

$$V(r, x = \cos \theta) = \sum_{n=0}^{\infty} Q_n r^{-(n+1)} P_n(\cos \theta). \quad (2.144)$$

Up to a factor of 4π this is exactly the same expression as (2.113). This is no coincidence since

$$\Delta V_{\text{rod}}(r, \cos \theta) = 0 \quad \text{for } r > R. \quad (2.145)$$

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

$$\tilde{u}(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{2} (u(x + \epsilon) + u(x - \epsilon)) \quad \text{with } \epsilon > 0. \quad (3.1)$$

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

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

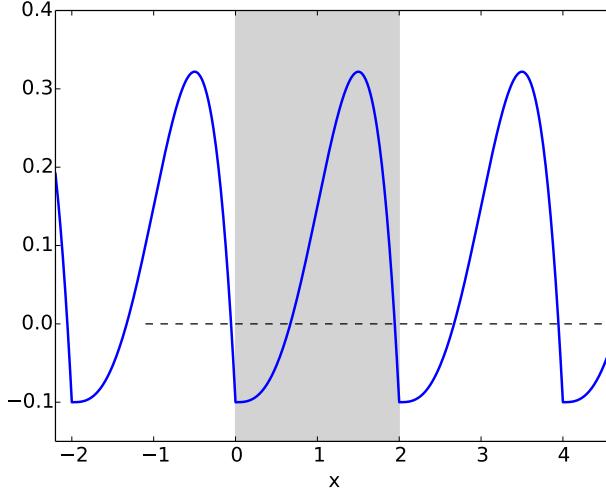


Figure 3.1: Example of a periodic function with $L = 2$.

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*

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

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 2.4 in Chapter 2 is periodically extended as $f(x + \pi) = f(x)$, the Fourier coefficients are given implicitly in (2.83).

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

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

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

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

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 $L \rightarrow \infty$?

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 $\Delta\omega \rightarrow 0$ as $L \rightarrow \infty$. Further define

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

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

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

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

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

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

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

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

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

$F(\omega)$ 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

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

^aIn 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 $\mathcal{F}^{-1}[\mathcal{F}[u]](x) = \tilde{u}(x)$. 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*

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

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

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

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

$$F(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} u(x) dx = \int_0^{\infty} e^{-i\omega x} e^{-\alpha x} dx = \int_0^{\infty} e^{-(\alpha+i\omega)x} dx = \frac{1}{\alpha + i\omega}. \quad (3.14)$$

We also find that

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

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

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

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

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

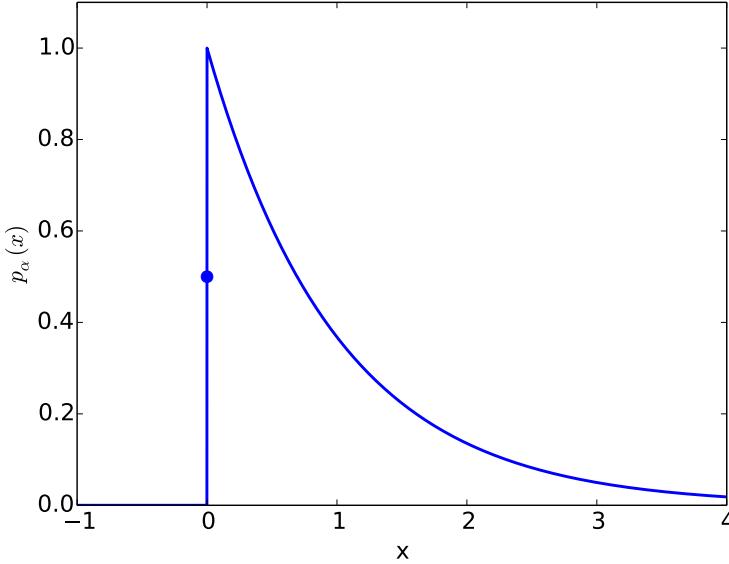


Figure 3.2: Decaying pulse $p_\alpha(x)$ defined in (3.13) in Example 3.2.

The Fourier transform is

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

The trick is to complete the square in the exponential

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

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

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

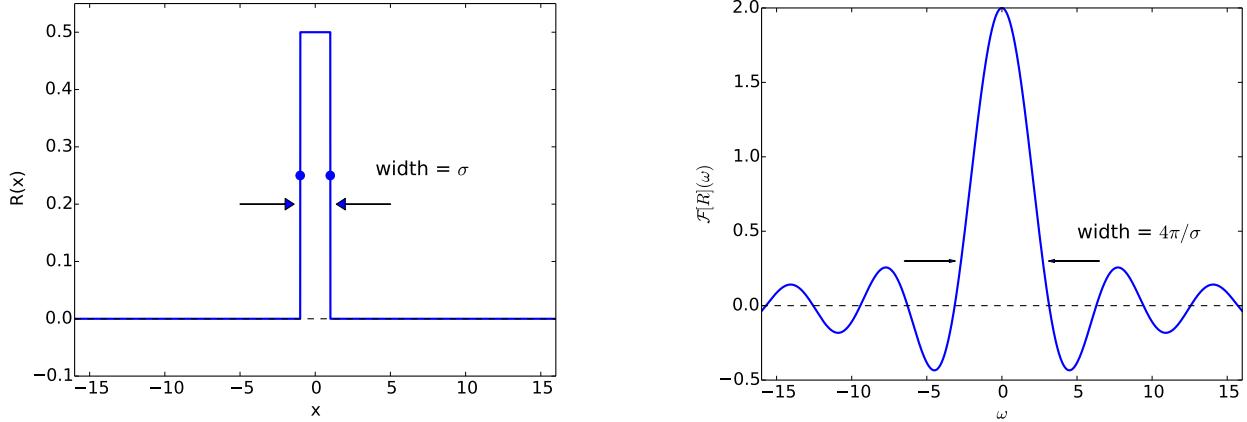


Figure 3.3: Rectangular pulse $R_\sigma(x)$ defined in (3.21) (left) and its Fourier-transform $\mathcal{F}[R_\sigma](\omega) = 2 \frac{\sin(\frac{\omega\sigma}{2})}{\omega\sigma}$ (right).

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 of 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

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

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

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

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

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

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

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

3.2 Elementary properties of the Fourier transform

Three obvious properties of the Fourier transform are:

- **Linearity:**

$$\mathcal{F}[\lambda u + \mu v] = \lambda \mathcal{F}[u] + \mu \mathcal{F}[v] \quad (3.25)$$

- **Symmetry:** If $v(x) \equiv u(-x)$ then

$$\mathcal{F}[u](-\omega) = \mathcal{F}[v](\omega) \quad (3.26)$$

- **Complex conjugation:**

$$\mathcal{F}[u^*](\omega) = \mathcal{F}[u]^*(-\omega) \quad (3.27)$$

Theorem 3.3 (First shift theorem).

$$\mathcal{F}[e^{i\bar{\omega}x}u](\omega) = \mathcal{F}[u](\omega - \bar{\omega}) \quad \text{for all real shifts } \bar{\omega} \in \mathbb{R}. \quad (3.28)$$

Proof.

$$\mathcal{F}[e^{i\bar{\omega}x}u](\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} e^{i\bar{\omega}x} u(x) dx = \int_{-\infty}^{\infty} e^{-i(\omega - \bar{\omega})x} u(x) dx = \mathcal{F}[u](\omega - \bar{\omega}) \quad (3.29)$$

□

Example 3.5. Consider

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

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

$$\mathcal{F}[v] = \frac{1}{2} \left(\frac{1}{\alpha + i(\omega - \beta)} + \frac{1}{\alpha + i(\omega + \beta)} \right). \quad (3.31)$$

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

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

Proof.

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

□

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 $|u(x)| \rightarrow 0$ as $|x| \rightarrow \infty$ then*

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

To paraphrase: “Differentiation in x -space \Leftrightarrow Multiplication in ω -space”. Note that there is a similar result for the Laplace transform.

Proof. Integrate by parts to obtain

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

□

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

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

Proof. Apply Theorem 3.5 twice. \square

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

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

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

Example 3.6. Let

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

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

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

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

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

and this is not well-defined since $e^{-i\omega x}$ has no meaningful value for $x \rightarrow \pm\infty$. 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

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

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

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

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

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

We now formally define the Dirac δ -function

$$\delta(\omega) = \lim_{\epsilon \rightarrow 0} \delta_\epsilon(\omega). \quad (3.43)$$

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*

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

with

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

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*

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

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

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

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 $\epsilon \rightarrow 0$ of the integral

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

For this, split the integral into three parts:

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

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:

$$|I_0| \leq \int_{-\sqrt{\epsilon}}^{\sqrt{\epsilon}} |\tilde{f}(y)|\delta_{\epsilon}(y) dy \leq L \int_{-\sqrt{\epsilon}}^{\sqrt{\epsilon}} |y|\delta_{\epsilon}(y) dy = 2L \int_0^{\sqrt{\epsilon}} y\delta_{\epsilon}(y) dy \quad (3.50)$$

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

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

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

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

But then:

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

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

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

□

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

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

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 , $i = 1, \dots, N$, such that $g(x_i) = 0$, $g'(x_i) \neq 0$ and g is continuously differentiable at x_i . Then we have:*

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

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

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

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 $\bar{\Omega} = [-\infty, \infty] \setminus \cup_{i=1}^N \Omega_i$, then we also have that $g(x) \neq 0$ for all $x \in \bar{\Omega}$. Split the integral in (3.57) as follows:

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

The rightmost integral vanishes since $g(x) \neq 0$ and hence $\delta(g(x)) = 0$ for all $x \in \bar{\Omega}$. 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 :

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

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

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

□

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

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

In this case we have

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

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

$$g'(0) = -4, \quad g'(-1) = 5, \quad g'(4) = 20. \quad (3.63)$$

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

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

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

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

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 $z = \alpha y$

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

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:

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

□

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*

$$\mathcal{F}[f \star g] = \mathcal{F}[f] \cdot \mathcal{F}[g]. \quad (3.68)$$

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

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

□

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

$$f \star g = \mathcal{F}^{-1} [\mathcal{F}[f] \cdot \mathcal{F}[g]]. \quad (3.70)$$

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

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

Expand the product using partial fractions

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

and we can read off

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

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 $h = f * g$. If we only have the blurred image, how can we reconstruct the original, unblurred image? This is covered by the following Corollary.

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

$$f = \mathcal{F}^{-1} \left[\frac{\mathcal{F}[h]}{\mathcal{F}[g]} \right] = \mathcal{F}^{-1} \left[\frac{\mathcal{F}[f * g]}{\mathcal{F}[g]} \right]. \quad (3.74)$$

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

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:

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

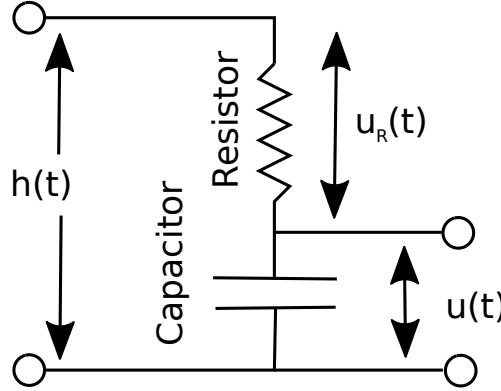


Figure 3.4: Electrical circuit which can be represented by (3.76)

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

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

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

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

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

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

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

$$u = h * g. \quad (3.79)$$

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). \square

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

Lemma 3.3. The transfer function of the first order ODE

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

with $a, b > 0$ is

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

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

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

and Example 3.2. \square

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

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

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

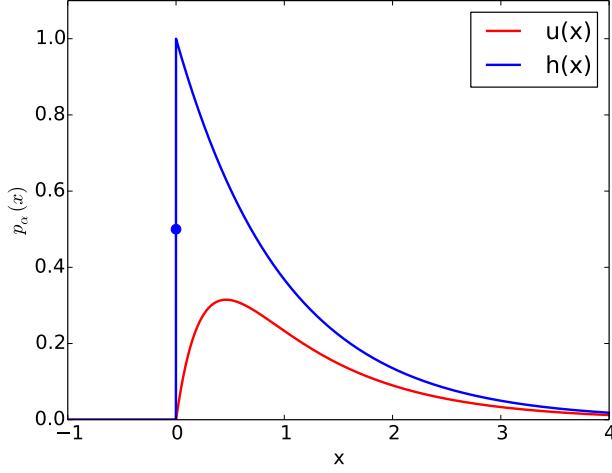


Figure 3.5: Solution of the first order ODE in (3.77).

Example 3.11. Let

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

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

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

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

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

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

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

3.5.2 Second order equations

We can study equations of the form

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

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

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

So, as before,

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

We now expand the Fourier-space transfer function as

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

where

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

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 $D \rightarrow 1$ and very weak damping ($b \rightarrow 0$) in $D \rightarrow -\infty$. 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,

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

Then

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

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

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

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,

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

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

Therefore we find with Theorem 3.3

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

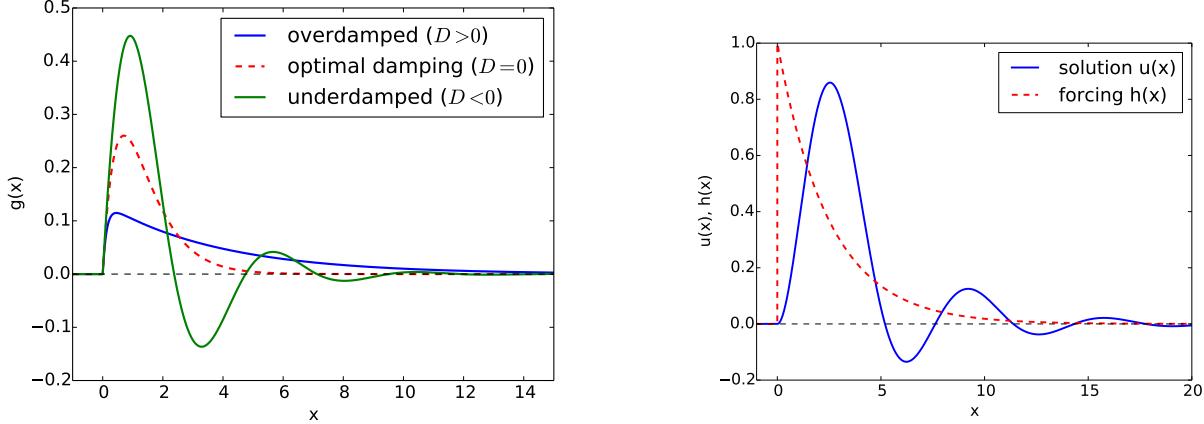


Figure 3.6: Transfer functions $g(x)$ for the damped oscillator (left) and solution $u(x)$ for the forcing $h(x)$ in the underdamped case in Example 3.12 (right).

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

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

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:

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

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

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

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

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

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)

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

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

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

This implies that

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

□

Example 3.13. Imagine we want to calculate the integral

$$\int_{-\infty}^{\infty} \text{sinc}^2(\omega) d\omega = \int_{-\infty}^{\infty} \frac{\sin^2(\omega)}{\omega^2} d\omega. \quad (3.111)$$

Recall from Example 3.4 that the sinc function is the Fourier-transform of the rectangular pulse R_2 with width 2, i.e. $\mathcal{F}[R_2](\omega) = \text{sinc}(\omega)$. 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:

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

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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + a \frac{\partial u}{\partial x} \quad (3.113)$$

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

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

$u(x, t)$ 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 $F(\omega, t) = \mathcal{F}[u](\omega)$

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

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

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

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

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

and hence

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

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

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

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

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

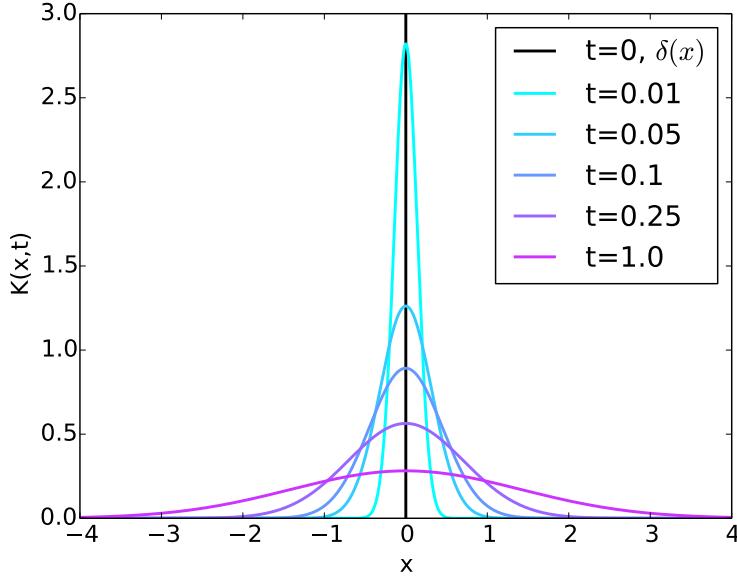


Figure 3.7: Heat kernel $K(x, t)$ defined by $g(x, t)$ in (3.117).

The most general solution of the first equation is

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

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

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

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

$$\begin{aligned} g(x, y) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} e^{-|\omega|y} d\omega = \frac{1}{2\pi} \left(\int_{-\infty}^0 e^{(ix+y)\omega} d\omega + \int_0^{\infty} e^{(ix-y)\omega} d\omega \right) \\ &= \frac{1}{2\pi} \left(\frac{1}{y+ix} + \frac{1}{y-ix} \right) = \frac{1}{\pi} \cdot \frac{y}{x^2+y^2} \end{aligned} \quad (3.123)$$

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

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

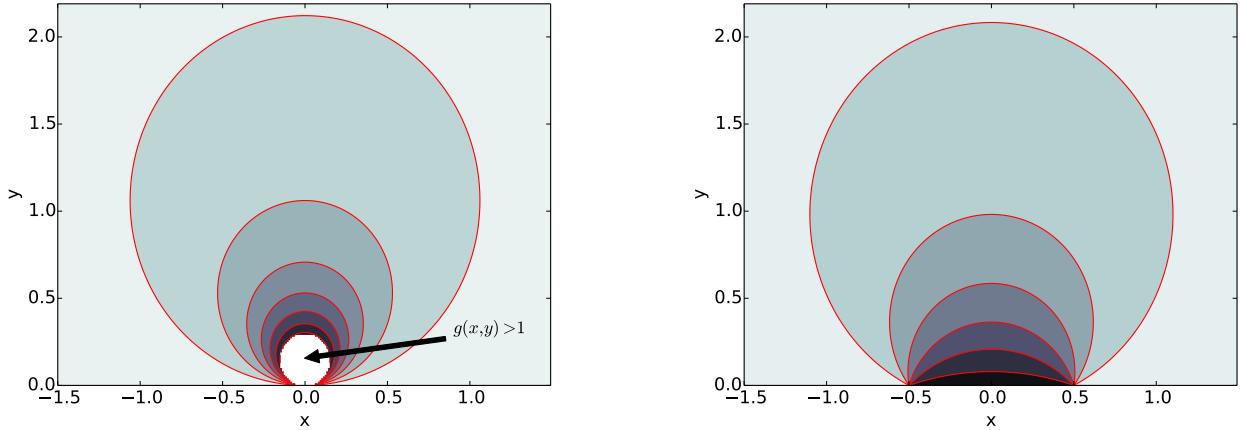


Figure 3.8: Contour plot of the fundamental solution $g(x, y)$ of the Laplace equation in a half-plane (left) and solution for $h(x)$ as defined in (3.125) for $a = 0.5$.

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

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

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

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

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 $\delta(\mathbf{x})$ 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

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y}, \quad u_t = \frac{\partial u}{\partial t} \quad \text{etc.} \quad (4.1)$$

to denote partial derivatives. Sometimes we will also use

$$\dot{u} = \frac{\partial u}{\partial t} \quad (4.2)$$

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

$$F(x, y, u, u_x, u_y) = 0 \quad (4.3)$$

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 :

$$F(x, y, u, u_x, u_y) = a(x, y)u_x + b(x, y)u_y + c(x, y)u - d(x, y) = 0 \quad (4.4)$$

2. *quasi-linear* if F is a linear functional of u_x and u_y (but F is not necessarily linear in u):

$$F(x, y, u, u_x, u_y) = P(x, y, u)u_x + Q(x, y, u)u_y - R(x, y, u) = 0 \quad (4.5)$$

3. *fully non-linear* otherwise.

A special case of a linear PDE is if the coefficients are constant,

$$au_x + bu_y + cu = d(x, y). \quad (4.6)$$

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:

$$u_x u_y = u^2, \quad u_x^2 + u_y^2 = 1. \quad (4.7)$$

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:

$$u_{xx} - c^2 u_{yy} = \frac{\partial^2 u}{\partial x^2} - c^2 \frac{\partial^2 u}{\partial y^2} = 0 \quad (4.8)$$

We will discuss this equation in Chapter 5.

Example 4.2. A famous quasi-linear equation is Burgers' equation

$$u_y + uu_x = c(x, y) \quad (4.9)$$

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

$$xu_y - yu_x = u^2. \quad (4.10)$$

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

- Fluid mechanics
- Meteorology

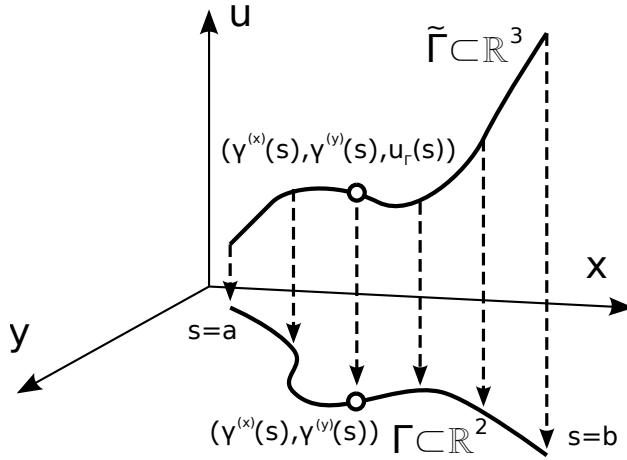


Figure 4.1: Data curve Γ and $\tilde{\Gamma}$

- Gas dynamics (detonations and explosions)

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

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

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

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

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)

$$u_y + uu_x = 0. \quad (4.13)$$

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}_{κ} along which $u(x, y)$ is constant:

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

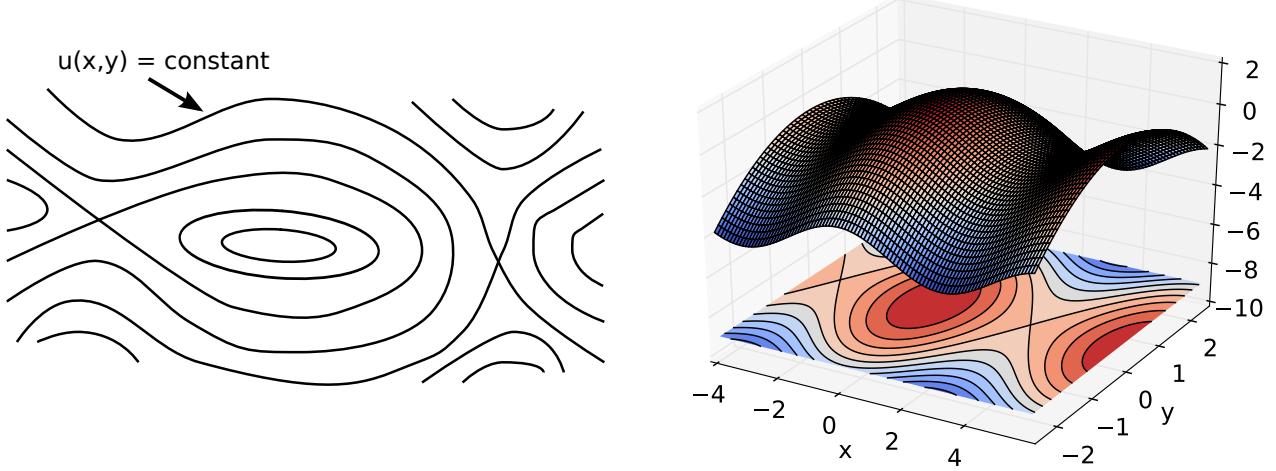


Figure 4.2: Schematic sketch of a family of contour lines (left) and contour lines of the function $u(x, y) = \cos(x) - \frac{1}{2}y^2$ (right).

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*

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

Then u satisfies the quasi-linear PDE

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

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:

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

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

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

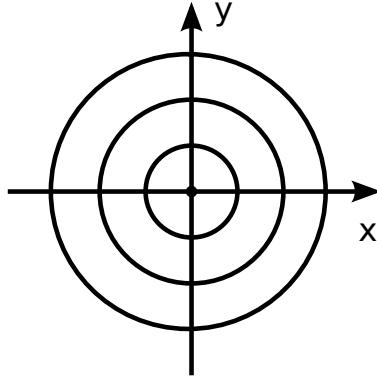


Figure 4.3: Concentric contour lines in Example 4.5.

where the last equality follows from (4.15). \square

Example 4.5. Suppose that we consider the equation

$$yu_x - xu_y = 0 \quad (4.19)$$

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

$$\dot{x} = \frac{dx}{dt} = y, \quad \dot{y} = \frac{dy}{dt} = -x. \quad (4.20)$$

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

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

Since $x^2 + y^2 = A^2 + B^2 = \text{const}$ 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

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

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 $u(x, y) = u(\gamma^{(x)}(s), \gamma^{(y)}(s)) = u_\Gamma(s)$.

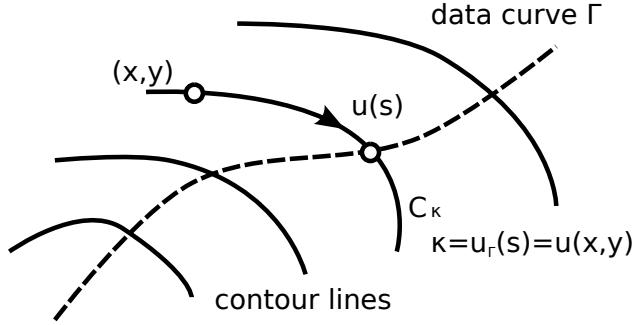


Figure 4.4: Calculating $u(x, y)$ by following contour lines $(x(t), y(t))$ and evaluating on the data curve.

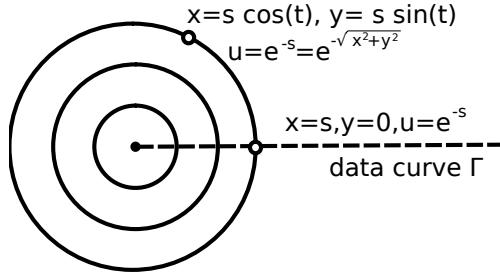


Figure 4.5: Calculation of $u(x, y)$ in Example 4.6

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

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

Note that Γ is the positive x -axis and it intersects 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

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

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

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

and therefore

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

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

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

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

$$x = x(t, s), \quad y = y(t, s) \quad (4.28)$$

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

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

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

$$\frac{dx}{dt} = -x^2, \quad \frac{dy}{dt} = \kappa \quad (4.30)$$

to obtain the curves

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

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

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

and are given explicitly as

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

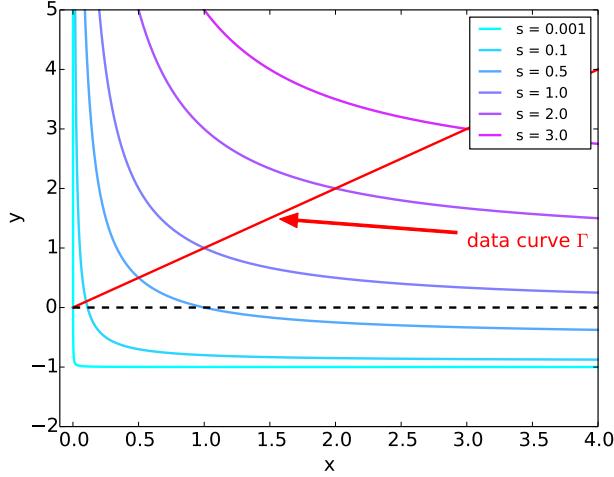


Figure 4.6: Contour lines and data curve Γ for Example 4.7.

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

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

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

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

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

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

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

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

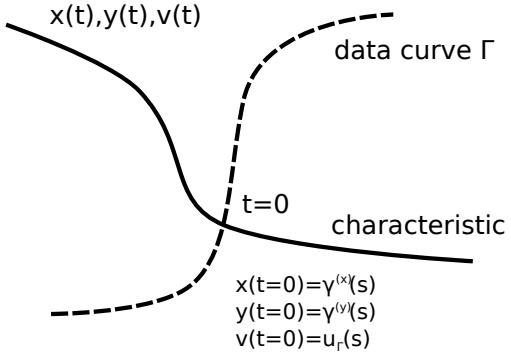


Figure 4.7: Solving equations with characteristics

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

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

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

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

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

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

$$x = x(t, s), \quad y = y(t, s) \quad (4.40)$$

to obtain

$$t = t(x, y), \quad s = s(x, y) \quad (4.41)$$

Then the solution for arbitrary x, y is given by

$$u(x, y) = v(t(x, y), s(x, y)). \quad (4.42)$$

Example 4.8. Consider the equation

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

Step 1: Solving

$$\frac{dx}{dt} = x, \quad \frac{dy}{dt} = y, \quad \frac{dv}{dt} = v \quad (4.44)$$

we obtain

$$x = Ae^t, \quad y = Be^t, \quad v = Ce^t. \quad (4.45)$$

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

$$x(0) = A = s, \quad y(0) = B = s^2, \quad v(0) = C = s^3 \quad (4.46)$$

and the characteristics can be written as

$$x(t, s) = se^t, \quad y(t, s) = s^2e^t = sx(t, s), \quad v(t, s) = s^3e^t = s^2x(t, s). \quad (4.47)$$

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

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

Therefore the solution is

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

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

Example 4.9 (harder). Consider

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

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

$$\frac{dx}{dt} = x + v, \quad \frac{dy}{dt} = y, \quad \frac{dv}{dt} = v + y^2. \quad (4.51)$$

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

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

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

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

and therefore

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

With this, we can finally solve the equation for x

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

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

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

Putting everything together, the characteristics are given by

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

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

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

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

$$x = y^2 - s^2 e^t \frac{1 + \log(y)}{1 + \log(s)}. \quad (4.59)$$

Similarly we find that

$$v = y^2 - s^2 e^t \frac{1}{1 + \log(s)} \quad (4.60)$$

$$\begin{aligned} v - y^2 &= -\frac{s^2 e^t}{1 + \log(s)} = \frac{x - y^2}{1 + \log(y)} \\ \Rightarrow \quad v(t, s) &= u(x(t, s), y(t, s)) = \frac{x(t, s) + y(t, s)^2 \log(y(t, s))}{1 + \log(y(t, s))} \\ \Rightarrow \quad u(x, y) &= \frac{x + y^2 \log(y)}{1 + \log(y)} \quad \text{for } y > e^{-1} \end{aligned} \quad (4.61)$$

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*

$$J = \frac{\partial(x, y)}{\partial(s, t)} = \begin{pmatrix} x_s & x_t \\ y_s & y_t \end{pmatrix} \quad (4.62)$$

is invertible, i.e.

$$E = \det(J) = x_s y_t - y_s x_t \neq 0. \quad (4.63)$$

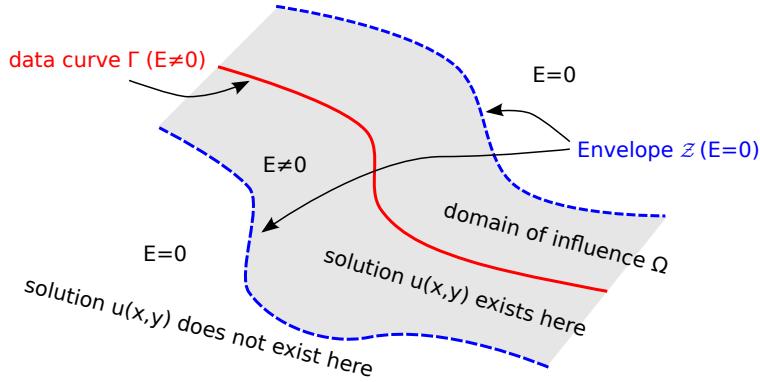


Figure 4.8: Domain of influence Ω and envelope Z .

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

$$Z = \{(x, y) : y_s x_t = y_t x_s\} \quad (4.64)$$

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

$$E = x_s y_t - y_s x_t \neq 0 \quad (4.65)$$

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 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 Z bounds Ω .

Example 4.10. Suppose the characteristics of a quasilinear PDE are given by

$$x = s + t, \quad y = st \quad (4.66)$$

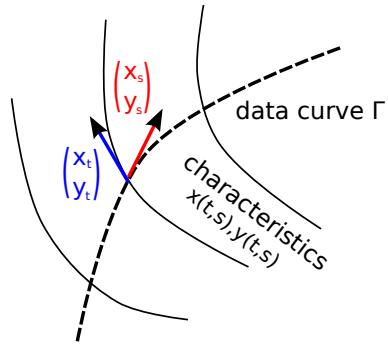


Figure 4.9: Tangents on characteristics and data curve.

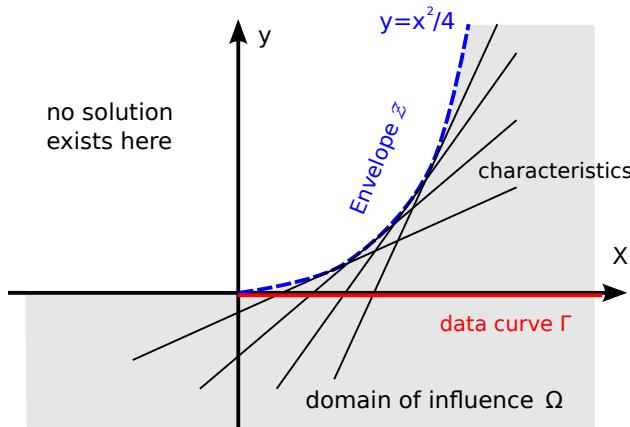


Figure 4.10: Envelope, characteristics lines and domain of influence for example 4.69.

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

$$J = \begin{pmatrix} 1 & 1 \\ t & s \end{pmatrix} \quad (4.67)$$

$$E = \det(J) = s - t$$

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

$$s = \frac{x}{2} \pm \frac{1}{2}\sqrt{x^2 - 4y}. \quad (4.68)$$

This confirms that we can find s if $y < x^2/4$, as predicted. In particular, if data is given for

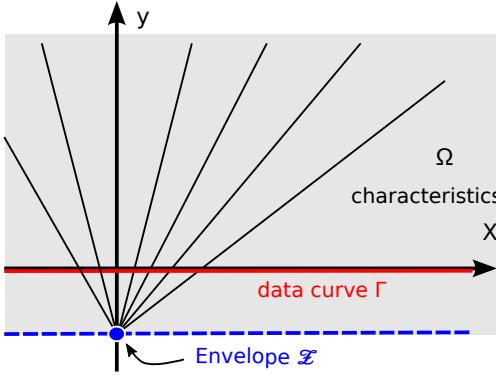


Figure 4.11: Characteristics for the Burgers' equation in (4.71).

$x = s > 0$, then for all $x > 0$ we can take

$$s = \frac{x}{2} + \frac{1}{2}\sqrt{x^2 - 4y}, \quad t = \frac{x}{2} - \frac{1}{2}\sqrt{x^2 - 4y}. \quad (4.69)$$

For $x \leq 0$ 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)

$$\Omega = \{(x, y) : (x > 0 \text{ and } y < x^2/4) \text{ or } (x \leq 0 \text{ and } y < 0)\} \subset \mathbb{R}^2. \quad (4.70)$$

Example 4.11 (Burgers' equation). Consider

$$uu_x + u_y = 0, \quad \Gamma = \{(x, y) : y = 0, x = s, -\infty < s < \infty\} \quad (4.71)$$

where y can be interpreted as time. We will look at two cases:

Case I: $u = s$ on Γ .

Calculate the characteristics by solving

$$\frac{dx}{dt} = v, \quad \frac{dy}{dt} = 1, \quad \frac{dv}{dt} = 0 \quad (4.72)$$

to obtain

$$x(t) = A + Ct, \quad y(t) = B + t, \quad v(t) = C. \quad (4.73)$$

Using the initial data we fix the constants A , B and C and find the characteristics

$$x(t, s) = s + st, \quad y(t, s) = t, \quad u(t, s) = s. \quad (4.74)$$

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

$$J = \begin{pmatrix} 1+t & s \\ 0 & 1 \end{pmatrix}, \quad E = \det(J) = 1 + t. \quad (4.75)$$

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

$$u(x, y) = s = \frac{x}{1+y} \quad (4.76)$$

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

$$x(t, s) = s + f(s)t, \quad y(t, s) = t, \quad u(t, s) = f(s) \quad (4.77)$$

and the Jacobian and its determinant are

$$J = \begin{pmatrix} 1 + f'(s)t & f(s) \\ 0 & 1 \end{pmatrix}, \quad E = \det(J) = 1 + f'(s)t = 1 + f'(s)y. \quad (4.78)$$

Hence $E = 0$ for $y = -1/f'(s)$ and we can not extend the solution for all x beyond

$$y_{\max} = \min_s \left\{ -\frac{1}{f'(s)} \right\} \quad (4.79)$$

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 $y \rightarrow 1$ 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.

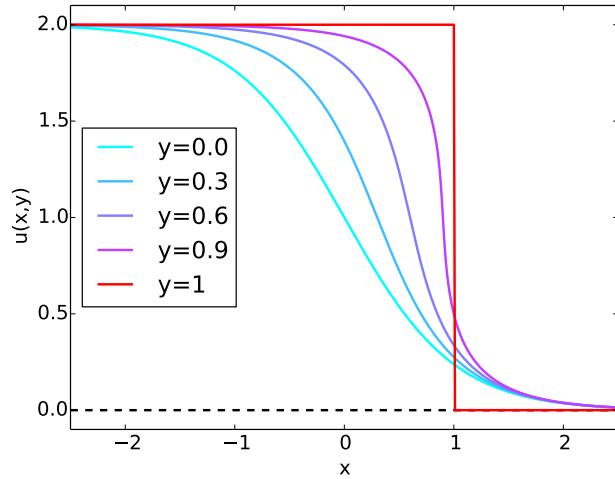


Figure 4.12: Solution of Burgers' equation for different values of y

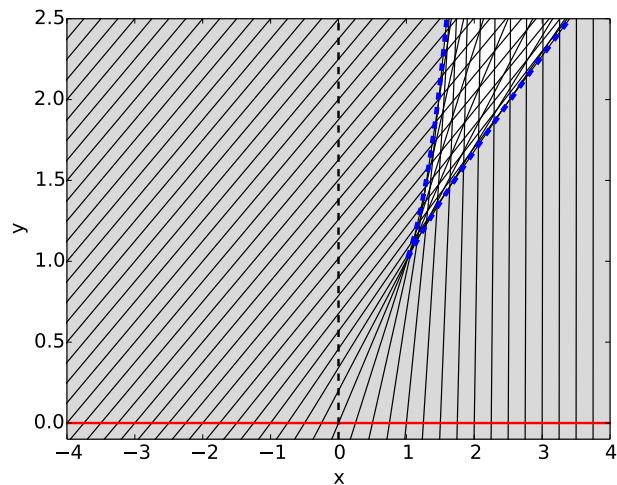


Figure 4.13: Characteristics and domain of influence of the Burgers' equation

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

$$a(x, y)u_{xx} + b(x, y)u_{xy} + c(x, y)u_{yy} = d(x, y, u, u_x, u_y) \quad (5.1)$$

where we use the obvious notation

$$u_{xx} \equiv \frac{\partial^2 u}{\partial x^2}, \quad u_{xy} \equiv \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}, \quad u_{yy} \equiv \frac{\partial^2 u}{\partial y^2} \quad (5.2)$$

for the second order derivatives. This defines a linear differential operator \mathcal{L}

$$\mathcal{L}u \equiv au_{xx} + bu_{xy} + cu_{yy} = d(x, y, u, u_x, u_y) \quad (5.3)$$

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**

$$D \equiv b^2 - 4ac. \quad (5.4)$$

The behaviour of the solution depends crucially on the sign of D .

Definition 5.1. If $D > 0$ then \mathcal{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 \mathcal{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 \mathcal{L} is *elliptic*.

Elliptic PDEs arise in electrostatics and potential flow/theory.

Example 5.1. The wave equation

$$u_{xx} - \lambda^2 u_{yy} = 0, \quad \lambda \in \mathbb{R} \quad (5.5)$$

is hyperbolic since $a = 1, b = 0, c = -\lambda^2 \Rightarrow D = 4\lambda^2 > 0$.

Example 5.2. The heat equation

$$u_{xx} - u_y = 0 \quad (5.6)$$

is parabolic since $a = 1, b = 0, c = 0 \Rightarrow D = 0$.

Example 5.3. The Laplace equation

$$\Delta u = u_{xx} + u_{yy} = 0 \quad (5.7)$$

is elliptic since $a = 1, b = 0, c = 1 \Rightarrow D = -4 < 0$.

Example 5.4. The canonical wave equation

$$u_{xy} - u_x - u_y = u \quad (5.8)$$

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 $a > 0$

$$\mathcal{L}u = au_{xx} + bu_{xy} + cu_{yy} = 0. \quad (5.9)$$

To solve this equation, make the following ansatz for the solution:

$$u(x, y) = f(y + \lambda x) \quad (5.10)$$

where f is a twice differentiable function. It follows that

$$u_{xx} = \lambda^2 f'', \quad u_{xy} = \lambda f'', \quad u_{yy} = f''. \quad (5.11)$$

Inserting this into (5.9) leads to an algebraic equation for λ

$$\mathcal{L}u = au_{xx} + bu_{xy} + cu_{yy} = (a\lambda^2 + b\lambda + c) f'' \quad (5.12)$$

Hence if

$$a\lambda^2 + b\lambda + c = 0 \quad (5.13)$$

then $\mathcal{L}u = 0$ for *any function* f . The solution of (5.13) is

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-b \pm \sqrt{D}}{2a}. \quad (5.14)$$

If \mathcal{L} is hyperbolic ($D > 0$) then \sqrt{D} is real and positive. We therefore have two distinct, nonzero real roots λ_+ and λ_- with

$$\lambda_+ = \frac{-b + \sqrt{D}}{2a} \neq 0, \quad \lambda_- = \frac{-b - \sqrt{D}}{2a} \neq 0 \quad (5.15)$$

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 $\mathcal{L}u = 0$ and \mathcal{L} defined by (5.9) is hyperbolic with $a > 0$, then*

$$u(x, y) = f_+(y + \lambda_+ x) + f_-(y + \lambda_- x) \quad (5.16)$$

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

$$s_+ \equiv y + \lambda_+ x, \quad s_- \equiv y + \lambda_- x \quad (5.17)$$

Note that since $\lambda_+ \neq \lambda_-$ and $\lambda_+ - \lambda_- = \sqrt{D}/a \neq 0$, this relationship can be inverted to obtain

$$x = x(s_+, s_-) = \frac{s_+ - s_-}{\lambda_+ - \lambda_-}, \quad y = y(s_+, s_-) = \frac{\lambda_+ s_- - \lambda_- s_+}{\lambda_+ - \lambda_-} \quad (5.18)$$

with λ_{\pm} defined by (5.15). Let $u(x, y)$ be the solution of $\mathcal{L}u = 0$ and introduce the function v defined by

$$v(s_+, s_-) \equiv u(x(s_+, s_-), y(s_+, s_-)). \quad (5.19)$$

Then by the chain rule

$$\begin{aligned} \frac{\partial v}{\partial s_+} &= \frac{\partial u}{\partial x} \frac{\partial x}{\partial s_+} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial s_+} = \frac{1}{\lambda_+ - \lambda_-} \left(\frac{\partial}{\partial x} - \lambda_- \frac{\partial}{\partial y} \right) u \\ \frac{\partial v}{\partial s_-} &= \frac{\partial u}{\partial x} \frac{\partial x}{\partial s_-} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial s_-} = \frac{1}{\lambda_+ - \lambda_-} \left(-\frac{\partial}{\partial x} + \lambda_+ \frac{\partial}{\partial y} \right) u \\ \frac{\partial^2 v}{\partial s_+ \partial s_-} &= \frac{1}{(\lambda_+ - \lambda_-)^2} \left(\frac{\partial}{\partial x} - \lambda_- \frac{\partial}{\partial y} \right) \left(-\frac{\partial}{\partial x} + \lambda_+ \frac{\partial}{\partial y} \right) u \\ &= \frac{1}{(\lambda_+ - \lambda_-)^2} \left(-\frac{\partial^2 u}{\partial x^2} + (\lambda_+ + \lambda_-) \frac{\partial^2 u}{\partial x \partial y} - \lambda_+ \lambda_- \frac{\partial^2 u}{\partial y^2} \right) \end{aligned} \quad (5.20)$$

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

$$\frac{\partial^2 v}{\partial s_+ \partial s_-} = \frac{1}{(\lambda_+ - \lambda_-)^2} \left(-u_{xx} - \frac{b}{a} u_{xy} - \frac{c}{a} u_{yy} \right) = -\frac{a}{D} (au_{xx} + bu_{xy} + cu_{yy}) = 0 \quad (5.21)$$

i.e. the function v has to satisfy

$$\frac{\partial^2 v}{\partial s_+ \partial s_-} = 0. \quad (5.22)$$

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 be a function of s_+ , i.e.

$$v_+(s_+, s_-) = \frac{\partial v}{\partial s_+} = f'_+(s_+) \quad (5.23)$$

Integrating this, we obtain

$$u(x, y) = v(s_+, s_-) = \int v_+(s_+, s_-) ds_+ = f_+(s_+) + f_-(s_-) = f_+(y + \lambda_+ x) + f_-(y + \lambda_- x). \quad (5.24)$$

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

$$u(x, y) = f_+(s_+) + f_-(s_-) = \tilde{f}_+(s_+) + \tilde{f}_-(s_-) = \tilde{f}_+(y + \lambda_+ x) + \tilde{f}_-(y + \lambda_- x). \quad (5.25)$$

□

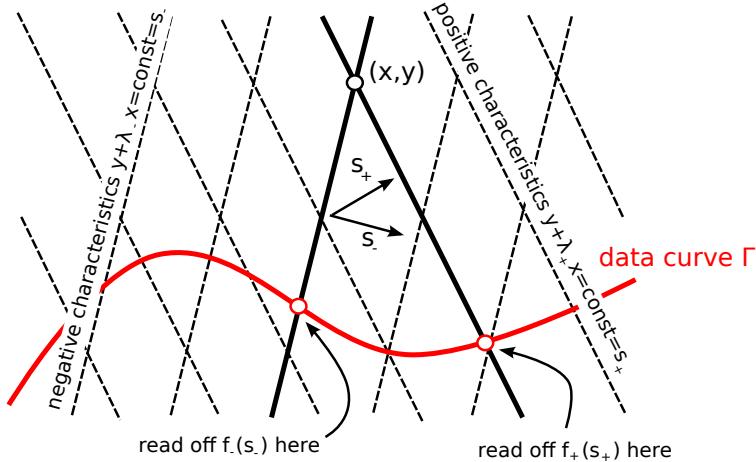


Figure 5.1: Characteristics and data curves for hyperbolic equations.

Definition 5.4. The form

$$\frac{\partial^2 v}{\partial s_+ \partial s_-} = 0 \quad (5.26)$$

is called the *canonical form* for the second order hyperbolic PDE in Eq. (5.9) and the lines

$$s_+ = y + \lambda_+ x = \text{const.}, \quad s_- = y + \lambda_- x = \text{const.} \quad (5.27)$$

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.

$$\Gamma = \{(x, y) : x = s, y = 0, -\infty < s < \infty\} \quad (5.28)$$

On Γ we specify

$$u(x, 0) = \psi(x), \quad u_y(x, 0) = \phi(x) \quad (5.29)$$

for some functions ψ and ϕ . We can use this information to deduce the functions f_+ and f_- since we know that

$$u(x, 0) = f_+(\lambda_+ x) + f_-(\lambda_- x) = \psi(x), \quad u_y(x, 0) = f'_+(\lambda_+ x) + f'_-(\lambda_- x) = \phi(x). \quad (5.30)$$

Integrating the second equation by x from a reference point x_0 we find that

$$\int_{x_0}^x \phi(z) dz = \frac{1}{\lambda_+} f_+(\lambda_+ z) + \frac{1}{\lambda_-} f_-(\lambda_- z) + C \quad (5.31)$$

Rearranging and solving for f_+ and f_- we find that

$$\begin{aligned} \psi(x) - \lambda_- \left(\int_{x_0}^x \phi(z) dz - C \right) &= \left(1 - \frac{\lambda_-}{\lambda_+} \right) f_+(\lambda_+ x) \\ \psi(x) - \lambda_+ \left(\int_{x_0}^x \phi(z) dz - C \right) &= \left(1 - \frac{\lambda_+}{\lambda_-} \right) f_-(\lambda_- x) \end{aligned} \quad (5.32)$$

Now use the fact that on the data curve $s_+ = \lambda_+ x$ and $s_- = \lambda_- x$ and hence

$$\begin{aligned} f_+(s_+) &= \frac{\lambda_+}{\lambda_+ - \lambda_-} \left(\psi(s_+/\lambda_+) - \lambda_- \int_{x_0}^{s_+/\lambda_+} \phi(z) dz + \lambda_- C \right) \\ f_-(s_-) &= \frac{\lambda_-}{\lambda_- - \lambda_+} \left(\psi(s_-/\lambda_-) - \lambda_+ \int_{x_0}^{s_-/\lambda_-} \phi(z) dz + \lambda_+ C \right) \\ u(x, y) &= f_+(y + \lambda_+ x) + f_-(y + \lambda_- x) \\ &= \frac{1}{\lambda_+ - \lambda_-} \left(\lambda_+ \psi(y/\lambda_+) - \lambda_- \psi(y/\lambda_-) - \lambda_+ \lambda_- \int_{x+y/\lambda_-}^{x+y/\lambda_+} \phi(z) dz \right). \end{aligned} \quad (5.33)$$

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

$$u_{xx} - 4u_{xy} + 3u_{yy} = 0, \quad \text{with } u(x, x) = x^3 \text{ and } u_x(x, x) = 0. \quad (5.34)$$

In this case we specified the solution and its derivative in the x -direction on the line $y = x$. The characteristic equation is

$$(\lambda - 1)(\lambda - 3) = 0 \quad (5.35)$$

which implies $\lambda_+ = 1$ and $\lambda_- = 3$. Hence the general solution is

$$u(x, y) = f_+(s_+) + f_-(s_-) = f_+(y + \lambda_+ x) + f_-(y + \lambda_- x) = f_+(y + x) + f_-(y + 3x). \quad (5.36)$$

To find the functions f_+ and f_- , use the initial conditions on the line $y = x$:

$$u(x, x) = f_+(2x) + f_-(4x) = x^3, \quad u_x(x, x) = f'_+(2x) + 3f'_-(4x) = 0. \quad (5.37)$$

Take the total derivative of the first equation¹ in Eq. (5.37) to obtain

$$\frac{d}{dx}u(x, x) = 2f'_+(2x) + 4f'_-(4x) = 3x^2. \quad (5.38)$$

Subtracting twice the second equation in Eq. (5.37) from Eq. (5.38) we find

$$\frac{d}{dx}u(x, x) - 2u_x(x, x) = -2f'_-(4x) = 3x^2. \quad (5.39)$$

Changing variables and introducing $z := 4x$ we conclude that

$$\begin{aligned} f'_-(z) &= -\frac{3}{32}z^2 \\ \Rightarrow f_-(z) &= -\frac{1}{32}z^3 + C. \end{aligned} \quad (5.40)$$

The function f_+ can be found by inserting this into the first equation of Eq. (5.37):

$$\begin{aligned} f_+(2x) &= x^3 - f_-(4x) = 3x^3 - C \quad \text{and hence with } \tilde{z} := 2x \\ \Rightarrow f_+(\tilde{z}) &= \frac{3}{8}\tilde{z}^3 - C \end{aligned} \quad (5.41)$$

Inserting the expressions for f_+ and f_- from Eqs (5.41) and (5.40) into the solution in Eq. (5.36), we obtain

$$u(x, y) = \frac{3}{8}(x+y)^3 - \frac{1}{32}(3x+y)^3. \quad (5.42)$$

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

$$\begin{aligned} u_x(x, y) &= \frac{9}{8}(x+y)^2 - \frac{9}{32}(3x+y)^2 \\ u_y(x, y) &= \frac{9}{8}(x+y)^2 - \frac{3}{32}(3x+y)^2 \\ u_{xx}(x, y) &= \frac{9}{4}(x+y) - \frac{27}{16}(3x+y) = \frac{9}{16}(-5x+y) \\ u_{xy}(x, y) &= \frac{9}{4}(x+y) - \frac{9}{16}(3x+y) = \frac{9}{16}(x+3y) \\ u_{yy}(x, y) &= \frac{9}{4}(x+y) - \frac{3}{16}(3x+y) = \frac{9}{16}\left(3x+\frac{11}{3}y\right) \end{aligned} \quad (5.43)$$

¹Make 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

$$u_{xx} - 4u_{xy} + 3u_{yy} = \frac{9}{16} \left(-5x + y - 4(x + 3y) + 3 \left(3x + \frac{11}{3}y \right) \right) = 0. \quad (5.44)$$

We also confirm that on the line $y = x$

$$\begin{aligned} u(x, x) &= \frac{3}{8}(2x)^3 - \frac{1}{32}(4x)^3 = x^3 \\ u_x(x, x) &= \frac{9}{8}(2x)^2 - \frac{9}{32}(4x)^2 = 0. \end{aligned} \quad (5.45)$$

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

$$u_{tt} - c^2 u_{xx} = 0, \quad \text{for } -\infty < x < \infty, t > 0 \quad (5.46)$$

with initial conditions

$$u(x, t=0) = \psi(x), \quad u_t(x, t=0) = \phi(x) \quad \text{for } -\infty < x < \infty. \quad (5.47)$$

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

$$s_+ = t + \frac{x}{c} = \text{const}, \quad s_- = t - \frac{x}{c} = \text{const}. \quad (5.48)$$

Hence the general solution is

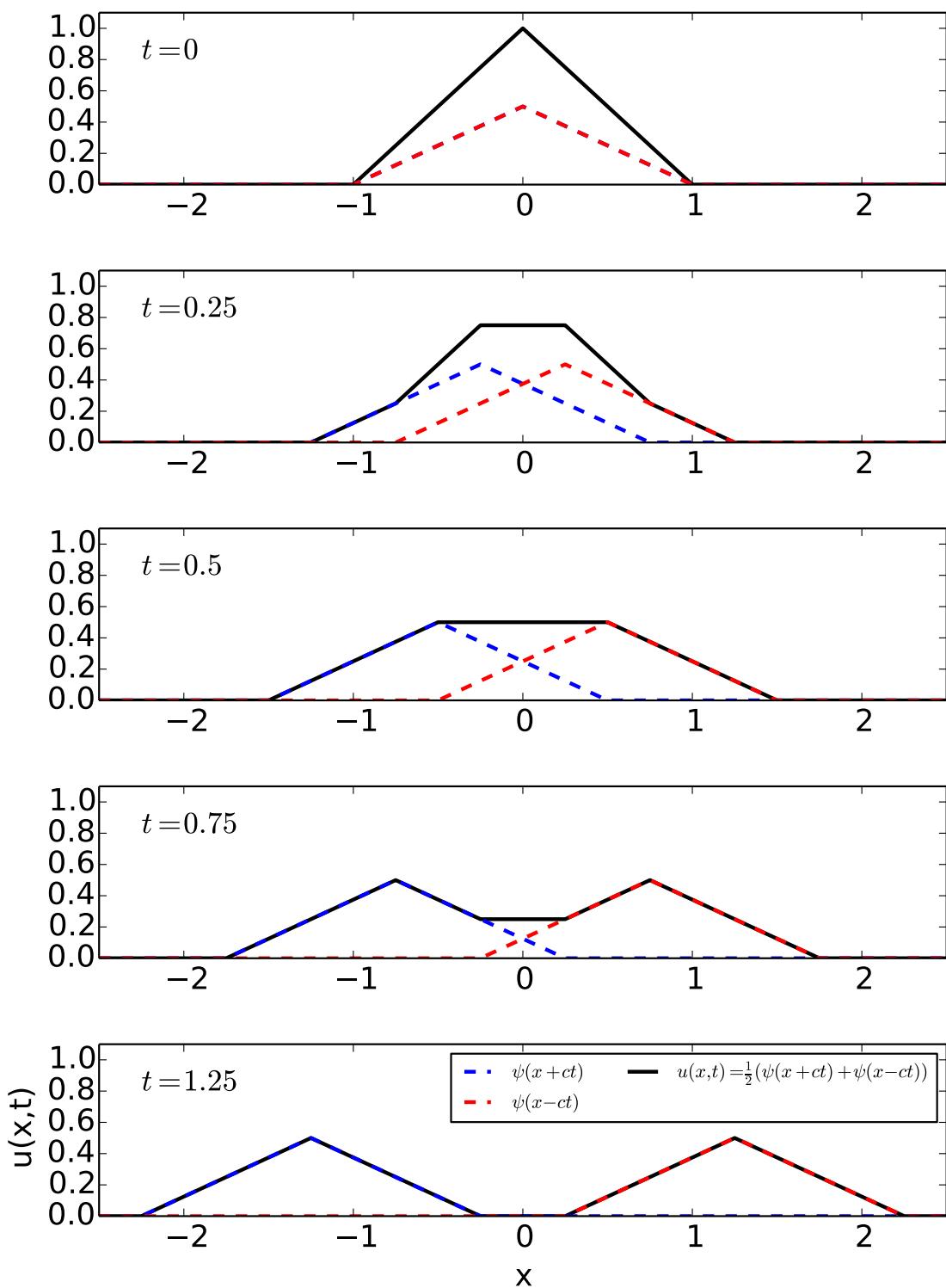
$$u(x, t) = f_+ \left(\frac{x}{c} + t \right) + f_- \left(\frac{x}{c} - t \right) \quad (5.49)$$

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

$$u(x, t) = \frac{1}{2} (\psi(x + ct) + \psi(x - ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} \phi(z) dz. \quad (5.50)$$

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

$$u(x, 0) = \psi(x) = \begin{cases} 1+x & \text{for } -1 \leq x < 0 \\ 1-x & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}, \quad u_t(x, 0) = \phi(x) = 0 \quad (5.51)$$

Figure 5.2: Solution of the wave equation $u(x, t)$ given in (5.50) for different times t .

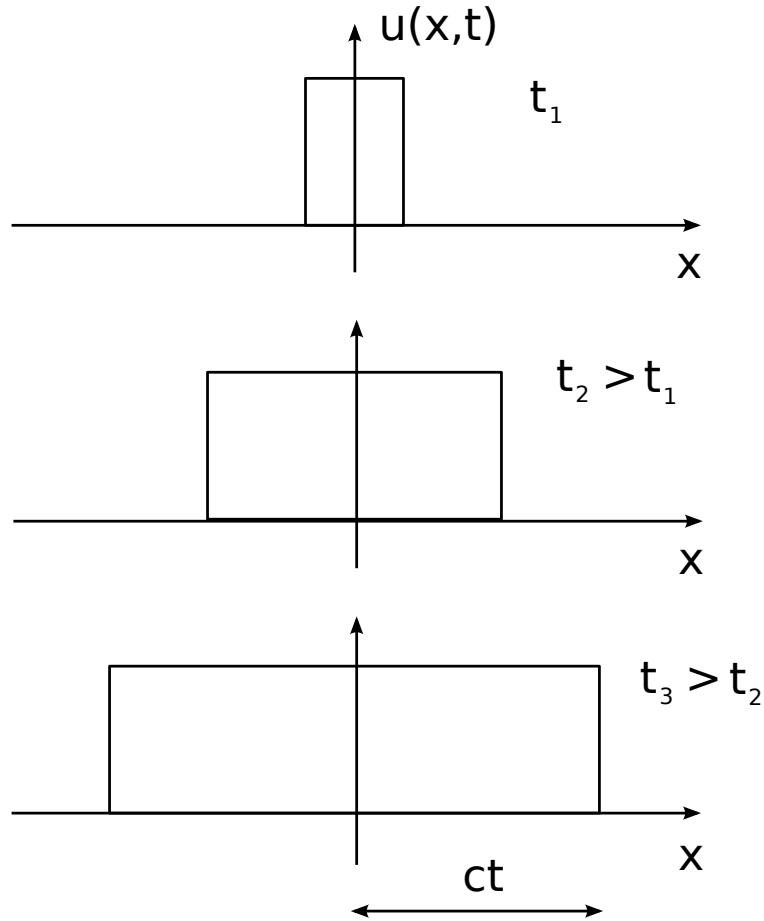


Figure 5.3: Solution of the wave equation $u(x, t)$ for the impulsively struck string in Example 5.7. The solution is shown for different times t .

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)

$$u(x, t) = \frac{1}{2c} \int_{x-ct}^{x+ct} \delta(z) dz = \frac{1}{2c} \quad \text{if } |x| < ct \text{ and } u(x, t) = 0 \text{ otherwise.} \quad (5.52)$$

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.
- $\mathbb{N} = 1, 2, 3, \dots$ 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

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y}, \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}, \quad u_{xy} = \frac{\partial^2 u}{\partial x \partial y}, \quad \dot{u} = u_t = \frac{\partial u}{\partial t} \quad (\text{A.1})$$

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

$$\frac{d^n(f \cdot g)}{dx^n} = \sum_{k=0}^n \binom{n}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} \quad (\text{B.1})$$

Proof. This can be proven by induction. Start with the base case $n = 1$, which is just the normal product rule

$$\frac{d(f \cdot g)}{dx} = \frac{df}{dx} \cdot g + f \cdot \frac{dg}{dx} = \sum_{k=0}^1 \frac{d^k f}{dx^k} \cdot \frac{d^{1-k} g}{dx^{1-k}} \quad (\text{B.2})$$

Assuming (B.1) is true for $n - 1$, we can obtain the case n by differentiating the equation (B.1) for $n - 1$:

$$\begin{aligned} \frac{d^n(f \cdot g)}{dx^n} &= \frac{d}{dx} \left(\frac{d^{n-1}(f \cdot g)}{dx^{n-1}} \right) = \frac{d}{dx} \left(\sum_{k=0}^{n-1} \binom{n-1}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-1-k} g}{dx^{n-1-k}} \right) \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} \left(\frac{d^{k+1} f}{dx^{k+1}} \cdot \frac{d^{n-1-k} g}{dx^{n-1-k}} + \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} \right) \\ &= \sum_{k=1}^n \binom{n-1}{k-1} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} + \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} \\ &= \frac{d^n f}{dx^n} \cdot g + \sum_{k=1}^{n-1} \left(\binom{n-1}{k-1} + \binom{n-1}{k} \right) \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} + f \cdot \frac{d^n g}{dx^n} \\ &= \sum_{k=0}^n \binom{n}{k} \frac{d^k f}{dx^k} \cdot \frac{d^{n-k} g}{dx^{n-k}} \end{aligned} \quad (\text{B.3})$$

APPENDIX B. GENERALISED PRODUCT RULE

since $\binom{n-1}{0} = \binom{n-1}{n-1} = \binom{n}{0} = \binom{n}{n} = 1$ and Pascal's rule

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \quad (\text{B.4})$$

□

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 $\langle \cdot, \cdot \rangle$ is defined as

$$\langle u, v \rangle \equiv \sum_{i=1}^d u_i^* v_i. \quad (\text{C.1})$$

Two vectors u, v are said to be **orthogonal** if $\langle u, v \rangle = 0$.

It is easy to see that

$$\langle u, v \rangle^* = \langle u^*, v^* \rangle = \langle v, u \rangle \quad (\text{C.2})$$

and that

$$\langle u, u \rangle \geq 0 \quad \text{for all } u \in \mathbb{C}^d \quad (\text{C.3})$$

with

$$\langle u, u \rangle > 0 \quad \text{for all } u \in \mathbb{C}^d \text{ with } u \neq 0. \quad (\text{C.4})$$

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-adjoint, then $H_{ij} = H_{ji}^*$.

Proof. Let $e^{(i)}$ be the unit vector in direction i , i.e. $e_k^{(i)} = \delta_{ik}$. Choose $u = e^{(i)}$ and $v = e^{(j)}$ and use the fact that H is self-adjoint to obtain

$$H_{ij} = \langle e^{(i)}, He^{(j)} \rangle = \langle He^{(i)}, e^{(j)} \rangle = H_{ji}^*. \quad (\text{C.5})$$

□

Definition C.3. A $d \times d$ matrix A is called positive semi-definite if for all $u \in \mathbb{C}^d$

$$\langle u, Au \rangle \geq 0. \quad (\text{C.6})$$

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 2.2.

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

$$\langle u_n, Hu_n \rangle = \lambda_n \langle u_n, u_n \rangle. \quad (\text{C.7})$$

Taking the complex conjugate of the left hand side of this equation, using (C.2) and the fact that H is Hermitian gives

$$\langle u_n, Hu_n \rangle^* = \langle Hu_n, u_n \rangle = \langle u_n, Hu_n \rangle = \lambda_n \langle u_n, u_n \rangle. \quad (\text{C.8})$$

On the other hand, taking the complex conjugate of the right hand side of (C.7) gives

$$\lambda_n^* \langle u_n, u_n \rangle^* = \lambda_n^* \langle u_n, u_n \rangle \quad (\text{C.9})$$

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:

$$\langle u_m, Hu_n \rangle = \lambda_n \langle u_m, u_n \rangle, \quad \langle u_n, Hu_m \rangle = \lambda_m \langle u_n, u_m \rangle \quad (\text{C.10})$$

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

$$\begin{aligned} \langle u_n, Hu_m \rangle^* &= \lambda_m \langle u_n, u_m \rangle^* \\ \Leftrightarrow \quad \langle Hu_m, u_n \rangle &= \langle u_m, Hu_n \rangle = \lambda_m \langle u_m, u_n \rangle \end{aligned} \quad (\text{C.11})$$

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 fulfilled if $\langle u_m, u_n \rangle = 0$, i.e. the eigenvectors u_n and u_m are orthogonal. \square