INTRODUCTION
TO
PARTIAL DIFFERENTIAL EQUATIONS

2WA90 COURSE NOTES, 1ST EDITION

Luc Florack

© May 19, 2020, Eindhoven University of Technology
# Contents

Preliminaries v

I PDE Systems: Concepts & Motivations 1

1 Introduction 3

1.1 Partial Differential Equations and Boundary Conditions 3

1.2 Examples 4

1.3 Issues 5

1.4 Ill-Posed versus Well-Posed Problems 6

1.5 Generalised Solutions 6

1.6 Classification of Partial Differential Equations 7

1.7 Overview 8

2 Existence and Uniqueness 11

2.1 Theorem of Cauchy-Kowalevski 11

2.2 Examples 12

3 Calculus of Variations 13

3.1 Origin 13

3.2 Basic Technique 13
3.3 Examples ......................................................... 15

4 Distribution Theory .................................................. 21
  4.1 Motivation .......................................................... 21
  4.2 Distributions Formalised ......................................... 23
  4.3 Reparametrisations of the Dirac Point Distribution .......... 26
  4.4 Examples .......................................................... 27

5 Fourier Transformation ................................................. 31
  5.1 Introduction ........................................................ 31
  5.2 The Fourier Transform on $\mathcal{S}(\mathbb{R}^n)$ ...................... 31
  5.3 The Fourier Transform on $\mathcal{S}'(\mathbb{R}^n)$ ...................... 33
  5.4 The Fourier Transform on $L^2(\mathbb{R}^n)$ ......................... 34
  5.5 Fourier Theorems .................................................. 35

6 Complex Analysis ....................................................... 39
  6.1 Introduction ........................................................ 39
  6.2 $\mathbb{C}$-Differentiability and Holomorphic Functions ............ 39

II PDE Systems: Techniques & Examples .............................. 43

7 The Fourier Method .................................................... 45
  7.1 Basic Technique .................................................... 45
  7.2 Examples .......................................................... 46

8 The Method of Characteristics ........................................ 49
  8.1 Basic Technique .................................................... 49
  8.2 Examples .......................................................... 51

9 Separation of Variables ............................................... 53
  9.1 Basic Technique .................................................... 53
## 9.2 Examples

### 10 Fundamental Solutions & Green’s Functions

10.1 Basic Technique

10.2 Construction

10.2.1 Via Classical Homogeneous Solutions

10.2.2 Via Separation of Variables

10.2.3 Via Fourier Transformation

### III PDE Systems: Common Cases

#### 11 First Order Systems

11.1 The Transport Equation

11.2 The Dirac Equation

#### 12 Second Order Hyperbolic Systems

12.1 The Wave Equation

12.2 Examples

#### 13 Second Order Parabolic Systems

13.1 The Heat Equation

13.2 Examples

#### 14 Second Order Elliptic Systems

14.1 The Laplace Equation

14.2 The Helmholtz Equation

14.3 The Poisson Equation
Preliminaries

Numbers

• \( \mathbb{N} = \{1, 2, 3, \ldots\} \): The set of natural numbers. If 0 is included we write \( \mathbb{N}_0 \).
• \( \mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\} \): The set of integer numbers.
• \( \mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z}, q \neq 0\} \): The set of rational numbers.
• \( \mathbb{R} \): The set of real numbers, i.e. the closure \( \mathbb{Q} \) containing all limit points of \( \mathbb{Q} \).
• \( \mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\} \): The set of complex numbers: \( i^2 = -1 \).
• \( \mathbb{K} \): Shorthand for either \( \mathbb{R} \) or \( \mathbb{C} \), depending on context.

Functions

It is tacitly understood that functions are \( \mathbb{K} \)-valued. If relevant, the context should reveal whether \( \mathbb{K} = \mathbb{R} \) or \( \mathbb{K} = \mathbb{C} \). We abbreviate \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) and \( \int_{\Omega} f(x) dx = \int_{\Omega} f(x_1, \ldots, x_n) dx_1 \ldots dx_n \) for \( \Omega \subset \mathbb{R}^n \). The essential supremum \( \text{ess} \sup_{x \in \Omega} f(x) \) of a function \( f \) is the smallest \( M \geq 0 \) s.t. \( f(x) \leq M \) for almost all \( x \in \Omega \).

• \( C(\Omega) \): The set of continuous functions with domain \( \Omega \).
• \( C^k(\Omega), k \in \mathbb{N}_0 \): The set of \( k \)-fold continuously differentiable functions on \( \Omega \), with \( C^0(\Omega) = C(\Omega) \).
• \( C^\infty(\Omega) \): The set of smooth, i.e. infinitely differentiable functions with domain \( \Omega \).
• \( C^\omega(\Omega) \): The set of analytical functions with domain \( \Omega \).
• \( L^p(\Omega), p \geq 1 \): The set of functions for which \( \int_{\Omega} |f(x)|^p \ dx \) exists. Notation: \( \|f\|_p = (\int_{\Omega} |f(x)|^p \ dx)^{1/p} \).
• \( L^\infty(\Omega) \): The set of functions \( f \) for which \( \text{ess} \sup_{x \in \Omega} |f(x)| \) exists. Notation: \( \|f\|_\infty = \text{ess} \sup_{x \in \Omega} |f(x)| \).
• \( \mathcal{S}(\mathbb{R}^n) \): The set of smooth functions of rapid decay on \( \mathbb{R}^n \).
Preliminaries

- $\mathcal{S}'(\mathbb{R}^n)$: The topological dual of $\mathcal{S}(\mathbb{R}^n)$, a.k.a. the set of tempered distributions on $\mathbb{R}^n$.

- $\mathcal{D}(\Omega)$: The set of smooth bump functions with compact support in $\Omega \subset \mathbb{R}^n$.

- $\mathcal{D}'(\Omega)$: The topological dual of $\mathcal{D}(\Omega)$, a.k.a. the set of distributions on $\Omega \subset \mathbb{R}^n$.

Some frequently used standard $C^\infty(\mathbb{R})$-functions:

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2} \quad (x \in \mathbb{R}).$$

Useful Inequalities

Hölder Inequality. Let $1 \leq p, q \leq \infty$ such that $1/p + 1/q = 1$, $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, then $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Young Inequality. Let $1 \leq p, q, r \leq \infty$ with $1/p + 1/q = 1 + 1/r$, $f \in L^p(\Omega)$, $\phi \in L^q(\Omega)$, then $\|f*\phi\|_r \leq \|f\|_p \|\phi\|_q$.

Multi-Indices

An $n$-dimensional multi-index $\alpha$ is an $n$-tuple of integers $(\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n$. Its order or norm is defined as $|\alpha| = \alpha_1 + \ldots + \alpha_n$. If $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n$, $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{Z}^n$ are multi-indices of dimension $n$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, and $f : \Omega \to \mathbb{K} : x \mapsto f(x)$ is a suitably defined function, then

- $\alpha + \beta \overset{def}{=} (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$
- $\nabla_{\alpha} f \overset{def}{=} \frac{\partial^{\alpha_1 + \ldots + \alpha_n} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}$
- $x^\alpha \overset{def}{=} x_1^{\alpha_1} \ldots x_n^{\alpha_n}$
- $\alpha! \overset{def}{=} \alpha_1! \ldots \alpha_n!$
- $\binom{\alpha}{\beta} \overset{def}{=} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \ldots \begin{pmatrix} \alpha_n \\ \beta_n \end{pmatrix}$
- $\sum_{\alpha} \overset{def}{=} \sum_{\alpha_1} \ldots \sum_{\alpha_n}$
Abbreviations

- a.e. almost everywhere
- b.c. boundary condition
- h.o.t. higher order term(s)
- i.c. initial condition
- iff if and only if
- l.h.s. left hand side
- o.d.e. ordinary differential equation
- p.d.e. partial differential equation
- p.d.o. partial differential operator
- r.h.s. right hand side
- s.t. such that

Oriented Region and its Boundary

An oriented region $\Omega$ (blue) with boundary $\partial \Omega$ (black) and outward normal $n$ (arrow). Closure of $\Omega$: $\overline{\Omega} = \Omega \cup \partial \Omega$. Complement of $\Omega$ (white background): $\Omega^c = U \setminus \Omega$, in which $U$ is the “universe” of all admissible sets, in this example $U = \mathbb{R}^2$.

Stokes’ Theorem

**Theorem** (Stokes’ Theorem). Let $\Omega \subset \mathbb{R}^n$ be an oriented region with boundary $\partial \Omega$ and $\omega$ a sufficiently smooth differential form. Then

$$\int_{\Omega} d\omega = \int_{\partial \Omega} \omega.$$

At this level of abstraction, Stokes’ Theorem requires knowledge of differential forms and exterior derivatives, cf. Spivak [5]. For our purposes the following special cases in standard vector calculus notation suffice.

**Corollary.** Cf. Stokes’ Theorem. Let $d\sigma$ denote the $(n-1)$-dimensional elementary surface measure of $\partial \Omega$ and $\hat{n}$ the unit outward normal. For sufficiently smooth $f, g : \Omega \to \mathbb{R}$ and $v, w : \Omega \to \mathbb{R}^n$, we have

- $\int_{\Omega} \text{div} w \, dx = \int_{\partial \Omega} w \cdot \hat{n} \, d\sigma.$
- $\int_{\Omega} (f \text{div} v + v \cdot \nabla f) \, dx = \int_{\partial \Omega} f \, v \cdot \hat{n} \, d\sigma.$
\[
\int_{\Omega} (f \Delta g - g \Delta f) \, dx = \int_{\partial \Omega} (f \nabla g - g \nabla f) \cdot \hat{n} \, d\sigma.
\]

Here \( \text{div} v = \sum_{i=1}^{n} \partial_{i} v_{i} \) denotes the divergence of \( v = (v_1, \ldots, v_n) \). Note that the last two identities follow from the first by taking \( w = f v \), respectively by taking \( w = f \nabla g \), and subsequently interchanging roles \( f \leftrightarrow g \).

**Common Coordinate Reparametrisations**

Polar coordinates for \((x, y) \in \mathbb{R}^2\), with \((r, \phi) \in \mathbb{R}^+ \times [0, 2\pi)\):

\[
\begin{align*}
  x(r, \phi) &= r \cos \phi, \\
  y(r, \phi) &= r \sin \phi.
\end{align*}
\]

Jacobian determinant:

\[
\det \frac{\partial (x, y)}{\partial (r, \phi)} = r.
\]

Cylindrical coordinates for \((x, y, z) \in \mathbb{R}^3\), with \((r, \phi, \zeta) \in \mathbb{R}^+ \times [0, 2\pi) \times \mathbb{R}\):

\[
\begin{align*}
  x(r, \phi, \zeta) &= r \cos \phi, \\
  y(r, \phi, \zeta) &= r \sin \phi, \\
  z(r, \phi, \zeta) &= \zeta.
\end{align*}
\]

Jacobian determinant:

\[
\det \frac{\partial (x, y, z)}{\partial (r, \phi, \zeta)} = r.
\]

Spherical coordinates (ISO convention) for \((x, y, z) \in \mathbb{R}^3\), with \((r, \theta, \phi) \in \mathbb{R}^+ \times [0, \pi] \times [0, 2\pi)\):

\[
\begin{align*}
  x(r, \phi, \theta) &= r \sin \theta \cos \phi, \\
  y(r, \phi, \theta) &= r \sin \theta \sin \phi, \\
  z(r, \phi, \theta) &= r \cos \theta.
\end{align*}
\]

Jacobian determinant:

\[
\det \frac{\partial (x, y, z)}{\partial (r, \theta, \phi)} = r^2 \sin \theta.
\]

Dirac function (minuscule/majuscule letters refer to free, respectively fixed point coordinates, l.h.s. and r.h.s. correspond to Cartesian, respectively polar/cylindrical/spherical representations):

\[
\begin{align*}
\delta(x - X)\delta(y - Y) &= \frac{\delta(r - R)\delta(\phi - \Phi)}{r}, \\
\delta(x - X)\delta(y - Y)\delta(z - Z) &= \frac{\delta(r - R)\delta(\phi - \Phi)\delta(\zeta - Z)}{r}, \\
\delta(x - X)\delta(y - Y)\delta(z - Z) &= \frac{\delta(r - R)\delta(\phi - \Phi)\delta(\theta - \Theta)}{r^2 \sin \theta}.
\end{align*}
\]

Note: \((X, Y)\), respectively \((X, Y, Z)\), do not coincide with the coordinate origin.
Unit Balls and Spheres

Let $B_n = \{ x \in \mathbb{R}^n \mid \|x\| < 1 \}$ be the unit ball in $\mathbb{R}^n$, and $S^{n-1}_n = \partial B_n = \{ x \in \mathbb{R}^n \mid \|x\| = 1 \}$ the corresponding unit $n$-sphere, then the volume, respectively surface area of $B_n$ and $S^{n-1}_n$ are given by

$$\text{vol}_n B^n = \frac{\sqrt{\pi^n}}{\Gamma\left(\frac{n}{2} + 1\right)},$$

$$\text{vol}_{n-1} S^{n-1} = \frac{2\sqrt{\pi^n}}{\Gamma\left(\frac{n}{2}\right)},$$

in which the so-called Gamma function, cf. the plot below, is given for any $z \in \mathbb{C}$ with Re $z > 0$ by

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \, dx.$$

Note that $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$.

Frequently encountered cases:

- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \Gamma(1) = 1, \Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt{\pi}, \Gamma(2) = 1, \Gamma\left(\frac{3}{2}\right) = \frac{3}{4}\sqrt{\pi}.$
- $\text{vol}_1 B^1 = 2, \text{vol}_2 B^2 = \pi, \text{vol}_3 B^3 = \frac{4}{3}\pi.$
- $\text{vol}_0 S^0 = 2, \text{vol}_1 S^1 = 2\pi, \text{vol}_2 S^2 = 4\pi.$

Recursions for $n \in \mathbb{N}$:

- $\text{vol}_{n+1} B^{n+1} = \frac{\text{vol}_n S^n}{n+1}$.
- $\text{vol}_{n+1} S^{n+1} = 2\pi \text{vol}_n B^n$.

If $B^n(r) = r^n B^n$ and $S^{n-1}(r) = \partial B^n(r) = r^{n-1} S^{n-1}$, then $\text{vol}_{n-1} S^{n-1}(r) = \frac{d}{dr} \text{vol}_n B^n(r)$.

![Gamma function](https://example.com/gamma.png)  

Gamma function $x \mapsto \Gamma(x)$ plotted for a real variable $x \in \mathbb{R}$. © Illustration by Alessio Damato.
I
PDE Systems: Concepts & Motivations
1 INTRODUCTION

1.1. Partial Differential Equations and Boundary Conditions

Recall the multi-index convention on page vi. A linear partial differential equation (p.d.e.) for a \( K \)-valued function \( u : \Omega \to K \) with domain \( \Omega \subset \mathbb{R}^n \) is an equation of the form

\[
Lu = f \quad \text{on } \Omega,
\]

in which \( f : \Omega \to K \) is a given function, and \( L \) is a linear partial differential operator (p.d.o.):

\[
L = \sum_{|\alpha| \leq N} c_\alpha \nabla^\alpha.
\]

The coefficients \( c_\alpha \in K \) may depend on the independent variable \( x \in \Omega \). If there exists a coefficient \( c_\alpha \neq 0 \) for some \( \alpha \) with \( |\alpha| = N \), then the p.d.e. is said to be of order \( N \). The function \( f \) is referred to as the inhomogeneous term. The p.d.e. is called inhomogeneous if \( f \neq 0 \), otherwise it is called homogeneous.

A p.d.e. is often furnished with boundary conditions (b.c.’s):

\[
Bu = g \quad \text{on } \partial \Omega,
\]

in which \( B \) is typically a linear p.d.o. and \( g : \partial \Omega \to K \) a given function on the boundary \( \partial \Omega \) of \( \Omega \):

\[
B = \sum_{|\alpha| \leq M} b_\alpha \nabla^\alpha.
\]

The coefficients \( b_\alpha \in K \) may vary along the boundary \( \partial \Omega \). If there is a distinguished entry in the \( n \)-tuple \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) that is interpreted as time (or some other evolution parameter), then boundary conditions specified for fixed initial time are usually referred to as initial conditions (i.c.’s).

The system Eqs. (1.1–1.4) is referred to as a boundary value (respectively initial value) problem, or simply as a (p.d.e.) “system”.

3
The following generalisations occur in practice:

- The p.d.e. system may be nonlinear, i.e. the coefficients $c_\alpha$ or $b_\alpha$ may depend on the dependent variable $u$.
- The boundary may be partitioned into parts, each with its own boundary condition.
- Instead of a scalar-valued function $u : \Omega \to \mathbb{K}$ we may have a multi-valued function $u : \Omega \to \mathbb{K}^k$.
- The boundary $\partial \Omega$ may not be a priori fixed (“free boundary”).
- We may want to admit “non-classical” solutions that are not in $C^N(\Omega)$, cf. Section 1.5.

### 1.2. Examples

Certain p.d.e.’s are frequently encountered in physics. Below are some of the most prevalent ones.

**Notation 1.2.1.** The following notational convention is in effect throughout these lecture notes:

- The Laplacian operator in $n$ spatial dimensions is defined as $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$.
- The d’Alembertian operator in $n+1$ spatiotemporal dimensions is defined as $\Box = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta$.

In relativistic theories the universal constant $c > 0$ represents the speed of light. By convention one often sets $c = 1$. This means that if one chooses to measure time intervals in the S.I. unit of one second, then the length unit for spatial distances is a light-second, i.e. $2.997924580 \times 10^8$ m in S.I. units. If relevant, the spatial dimension $n \in \mathbb{N}$ should be clear from the context. The d’Alembertian involves a distinguished time coordinate $t$.

**Caveat.** One needs to remain cautious about the above notation for $\Delta$, for several reasons:

- It is tacitly understood that the r.h.s. in the above notation pertains to Cartesian coordinates. A change of coordinates will generally result in different expressions. Moreover, on curved manifolds, such as the unit sphere, or general relativistic spacetime, Cartesian coordinates do not exist.
- A rigid motion, i.e. a coordinate transformation of the form $x = R y + a$, with $R \in \text{SO}(n)$ an $n \times n$ orthogonal matrix and $a \in \mathbb{R}^n$ a constant Euclidean vector, preserves the r.h.s. form of the Laplacian and d’Alembertian.
- A Lorentz transformation $x = Ly + a$ with $L \in \text{SO}(n, 1)$ and $a \in \mathbb{R}^n \times \mathbb{R}$, preserves the r.h.s. form of the d’Alembertian.
- In special relativity theory one often encounters the imaginary time coordinate $x_0 = i c t$. With this (potentially deceptive) convention, the d’Alembertian looks like an $(n+1)$-dimensional Laplacian, whereas a Lorentz transformation resembles an ordinary rotation.

**Example 1.2.1.** Laplace equation: $\Delta u = 0$.

**Example 1.2.2.** Poisson equation: $\Delta u = f$, i.e. the inhomogeneous Laplace equation, with data term $f \neq 0$.

**Example 1.2.3.** Helmholtz equation: $(\Delta + \nu^2 \omega^2) u = 0$. In relativistic theories $\omega$ represents angular temporal frequency for a propagating wave in a medium with index of refraction $\nu = 1/c$, i.e. inverse propagation speed.
Example 1.2.4. Wave equation: \( \Box u = 0 \).

Example 1.2.5. Heat or diffusion equation: \((\partial_t - \alpha \Delta)u = 0\). By convention one often sets \(\alpha = 1\), if constant. In heat diffusion problems \(\alpha > 0\) represents thermal diffusivity.

Example 1.2.6. Schrödinger equation: \((i\hbar \partial_t + \frac{\hbar^2}{2m} \Delta - V)u = 0\). This p.d.e. governs the evolution of the wave function for a particle with mass \(m\) and potential energy \(V\) in non-relativistic quantum mechanics. In particular we have \(V = 0\) for a free particle, and \(V(x) = \frac{1}{2}m\omega^2||x||^2\) for an isotropic harmonic oscillator with angular frequency \(\omega\). The universal constant \(\hbar = 1.054571800 \times 10^{-34}\) J s is called Planck’s constant. By convention one often sets \(\hbar = 1\) instead of using S.I. units.

Example 1.2.7. Klein-Gordon equation: \((\Box \mu^2) u = 0\). This is a relativistic counterpart of the free Schrödinger equation, with mass parameter \(\mu = mc/\hbar\), i.e. in natural units \(\mu = m\).

Example 1.2.8. Biharmonic equation: \(\Delta^2 u = 0\). This p.d.e. is encountered in elasticity theory.

Example 1.2.9. Advection or transport equation: \((\partial_t + v \cdot \nabla)u = 0\), in which \(v \cdot \nabla = \sum_{i=1}^{n} v_i \partial_i\). This p.d.e. plays a role in fluid dynamics, in which case the coefficients \(v_i\) represent flow velocity components. Note that if you consider the trajectory of a point particle comoving with the flow, with position \(x = x(t)\) at time \(t\), then \(v = \dot{x} = dx/dt\). Using this observation together with the chain rule it follows that along the particle’s trajectory, \(\partial_t + v \cdot \nabla\) acts as a total derivative \(d/dt\) on any function depending on \((x, t) = (x(t), t)\).

All of the above p.d.e.’s are linear. An example of a nonlinear p.d.e. is the following.

Example 1.2.10. Korteweg de Vries equation: \(u_t + 6uu_x - u_{xxx} = 0\). In this example, and henceforth, subscripts \(t\) and \(x\) denote partial derivatives w.r.t. \(t\) and \(x\), respectively. The Korteweg de Vries equation governs the propagation of water waves.

The physics of the underlying problems often suggest “natural” boundary conditions.

1.3. Issues

In p.d.e. theory various problems need attention, such as

- the genesis of a p.d.e. system (cf. Chapter 3),
- the mathematical concept of a solution (cf. Section 1.5 and Chapter 4),
- existence of a solution (cf. Chapter 2),
- uniqueness of the solution (cf. Chapter 2),
- regularity of the solution (cf. Chapter 2),
- bounds on the solution (pointwise or with respect to some norm),
- well-posedness of the system (cf. Section 1.4),
- construction of an exact solution,
- numerical approximation of a solution,
- other properties of the solution.
1.4. ILL-POSED VERSUS WELL-POSED PROBLEMS

Not all of these issues will be addressed in-depth. It is clear though that many of these problems are related. For instance, it makes no sense to try and approximate the solution if no solution exists, or if the solution does not depend continuously on the data (i.e. if the problem is ill-posed).

The situation for p.d.e.’s is more complicated than for ordinary differential equations (o.d.e.’s). For instance, whereas an \( N \)-th order o.d.e. typically has an \( N \)-dimensional solution space, a p.d.e. is, in general, not guaranteed to have a solution at all. If it does, the solution space tends to be infinite-dimensional. Moreover, uniqueness and well-posedness are issues more intricate for p.d.e.’s than for o.d.e.’s. The example in the next section illustrates this for the case of well-posedness.

1.4. Ill-Posed versus Well-Posed Problems

The problem of ill-posedness is best appreciated by a simple example.

**Example 1.4.1.** Consider the following Laplace equation with Dirichlet b.c.’s in \( n = 2 \) dimensions:

\[
\begin{align*}
\Delta u &= 0 \quad \text{on } \Omega = \mathbb{R}_0^+ \times \mathbb{R} \\
u(0, y) &= f_\epsilon(y) \\
u_x(0, y) &= 0
\end{align*}
\]

with data term \( f_\epsilon(y) = \epsilon \sin(y/\epsilon) \) for some \( \epsilon > 0 \). This problem has a unique solution (the proof relies on the Theorem of Cauchy-Kowalevski (a.k.a. Cauchy-Kowalevskaya, after Augustin Cauchy and Sophie Kowalevskaya), cf. Chapter 2), and is not difficult to establish, viz.

\[
u_\epsilon(x, y) = \epsilon \sin \left( \frac{y}{\epsilon} \right) \cosh \left( \frac{x}{\epsilon} \right) .
\]

You may verify that this is indeed a solution, that it is analytical on \( \Omega \), and that it has no upper bound. Moreover, the data term is bounded in the sense that \( \|f_\epsilon\|_\infty = \epsilon \), which can be made arbitrarily small by suitable choice of \( 0 < \epsilon \ll 1 \). Curiously, the solution \( u_\epsilon \) develops a singularity in the limit \( \epsilon \downarrow 0 \), showing that this problem is ill-posed. For a well-posed problem you would have expected the solution to converge to the null function on \( \Omega \) (the unique solution for vanishing data term) in any reasonable norm.

The ill-posedness of the problem in the example is not an inherent problem of the linear p.d.o. \( \Delta \) as such, but of the formulation of the overall system including b.c.’s. We will encounter b.c.’s that do make the problem well-posed.

1.5. Generalised Solutions

A classical solution to Eqs. (1.1–1.4) is one that satisfies the system in the strict sense of an \( N \)-fold continuously differentiable function fulfilling the stated differential conditions. The a priori regularity condition required may, however, be too restrictive, as it excludes “obvious” generalisations. Consider e.g. the system

\[
\frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x} \quad \text{for } (x, t) \in \mathbb{R} \times \mathbb{R},
\]

with constant parameter \( c \in \mathbb{R} \), subject to the initial condition

\[
u(x, 0) = f(x) ,
\]

\[1\]
1.6. Classification of Partial Differential Equations

P.d.e.’s can be classified according to different criteria, such as their order, their linear or nonlinear character, and—in case of linear p.d.e.’s—the constant or variable nature of the coefficients of the p.d.o.

For linear p.d.e.’s with constant coefficients a more detailed classification is obtained in terms of the so-called symbol. It is based on the observation that functions of the form $e^{ix \xi}$ are eigenfunctions of a linear p.d.o. for any parameter $\xi \in \mathbb{R}^n$:

$$L e(x; \xi) = \sum_{|\alpha| \leq N} c_{\alpha} \nabla e(x; \xi) = \sum_{|\alpha| \leq N} c_{\alpha} (i \xi)^\alpha e(x; \xi) \overset{\text{def}}{=} p(\xi) e(x; \xi).$$

**Definition 1.6.1.** Let $L = \sum_{|\alpha| \leq N} c_{\alpha} \nabla$ be a linear p.d.o.

* The polynomial $p(\xi) = \sum_{|\alpha| \leq N} c_{\alpha} (i \xi)^\alpha$ is called the symbol of $L$.

* The homogeneous polynomial of degree $N$, $p_N(\xi) = \sum_{|\alpha| = N} c_{\alpha} (i \xi)^\alpha$, is called the main symbol of $L$.

**Example 1.6.1.** In $n = 2$ dimensions a second order linear p.d.o. with constant coefficients has the form

$$L = a \frac{\partial^2}{\partial x^2} + 2b \frac{\partial^2}{\partial x \partial y} + c \frac{\partial^2}{\partial y^2} + d \frac{\partial}{\partial x} + e \frac{\partial}{\partial y} + f,$$

with $a, b, c, d, e, f \in \mathbb{R}$. Its main symbol defines a real quadratic form:

$$p_2(\xi, \eta) = a(i \xi)^2 + 2b(i \xi)(i \eta) + c(i \eta)^2 = -(a \xi^2 + 2b \xi \eta + c \eta^2) = -(\xi, \eta) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}. \quad (1.10)$$

The eigenvalues of the coefficient matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

are real by virtue of symmetry of the matrix. Three scenarios are possible:

* $\det A > 0$: $L$ is called elliptic (the eigenvalues of $A$ have equal signs);
The classification in this two-dimensional example can be extended to higher dimensions. However, the sign of the determinant \( \det A = \lambda_1 \ldots \lambda_n \) is then no longer conclusive, and one needs to inspect the individual eigenvalues \( \lambda_1, \ldots, \lambda_n \in \mathbb{R} \) of the \( n \times n \) symmetric coefficient matrix of the main symbol.

Example 1.6.2. In \( n > 2 \) dimensions a second order linear p.d.o. with constant coefficients has the form

\[
L = \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} + a_0.
\]

Its main symbol defines a real quadratic form:

\[
p_2(\xi) = \sum_{i,j=1}^{n} a_{ij} (i\xi_i)(i\xi_j) = -\sum_{i,j=1}^{n} a_{ij} \xi_i \xi_j.
\]

Example 1.6.2. In \( n > 2 \) dimensions a second order linear p.d.o. with constant coefficients has the form

The classification may then of course vary with the choice of \( x_0 \).

1.7. Overview

Part I

In the remainder of Part I we address some of the conceptual issues touched upon in this introductory chapter in more detail:

- Any attempt to solve a p.d.e. system should be accompanied by considerations of existence and uniqueness, recall Section 1.3. An important result in this context is formulated in Chapter 2.
- Chapter 3 discusses the variational principle often used in scientific models to generate p.d.e. systems, recall Section 1.3.
1. INTRODUCTION

- In Chapter 4 we consider the concept of generalised solutions of a p.d.e. system, a motivating example of which was given in Section 1.5.
- Chapter 5 introduces Fourier theory, providing important tools for linear shift invariant systems.
- Chapter 6 reviews some important results on differentiability from complex analysis.

Part II

Unfortunately, given existence and uniqueness, there is no single method of general applicability for solving p.d.e. systems. On the positive side, there do exist a few methods that are, to some extent, generic.

In Part II we focus on a number of generic techniques, de-emphasizing applications:

- Chapter 7 exploits the Fourier theory of Chapter 5 to solve p.d.e. systems.
- Chapter 8 introduces the method of characteristics in the context of first order p.d.e.’s.
- Chapter 9 explains the method of separation of variables.
- Chapter 10 introduces the method of Green’s functions, or fundamental solutions, using Chapter 4.

Part III

Finally, in Part III, we put things together, applying the concepts and techniques of the foregoing parts to study a selection of prevalent p.d.e. systems:

- Chapter 12 considers hyperbolic systems, illustrated by the wave equation.
- Chapter 13 considers parabolic systems, illustrated by the heat equation.
- Chapter 14 considers elliptic systems, illustrated by the Laplace equation.

We concentrate on analytical tools. Numerical and perturbative schemes constitute another important class of solution methods, especially for systems that cannot be handled analytically (the typical case). Such methods are, however, beyond the scope of this course. Nevertheless, the most powerful solution strategies tend to be of a “hybrid” nature, with perturbative expansions around an analytically known “zeroth order” system, and/or powerful numerical schemes constructed based on slick use of analytical properties of the system. In this sense one may view these lecture notes as an important stepping stone towards numerical and perturbative p.d.e. methods.
2

EXISTENCE AND UNIQUENESS

2.1. Theorem of Cauchy-Kowalevski

The Theorem of Cauchy-Kowalevski plays an important role in establishing existence and uniqueness of solutions for a broad class of p.d.e. systems. In this chapter we state the main result for later convenience. You are referred to the literature for a rigorous proof (the pivot of which is Taylor’s Theorem).

Definition 2.1.1. Let \( L = \sum_{|\alpha| \leq N} c_\alpha \nabla \alpha \) be a linear p.d.o. on the \( n \)-dimensional region \( \Omega \). An analytical \((n-1)\)-dimensional submanifold \( V \subset \Omega \) is called non-characteristic at \( x_0 \in V \) relative to \( L \) if there exist local coordinates \( y = (y_1, \ldots, y_n) \in \mathbb{R}^n \) in a neighbourhood of \( x_0 \), with \( x - x_0 = Ry \) for some rotation \( R \), s.t.

- \( y_n = 0 \) describes the tangent plane to \( V \) at \( y = 0 \) (i.e. at \( x = x_0 \)), and
- \( L \) can be written as \( L(y) = c(y) \frac{\partial^N}{\partial y_n^N} + \ldots \) with \( c(y) \neq 0 \) on a neighbourhood of \( y = 0 \), with the ellipsis denoting terms not involving \( \partial^N / \partial y_n^N \).

Thus \( V \) is non-characteristic at \( x_0 \in V \) if the p.d.o. has a nontrivial \( N \)-th order term transversal to \( V \) at that point.

Theorem 2.1.1 (Cauchy-Kowalevski). Recall Eqs. (1.1–1.2):

\[
Lu = f \quad \text{on } \Omega, \tag{2.1}
\]
with $L = \sum_{|\alpha| \leq N} c_\alpha \nabla^\alpha$. We assume that the b.c.'s take the form

$$u = g_0, \quad \frac{\partial u}{\partial n} = g_1, \ldots, \quad \frac{\partial u^{N-1}}{\partial n^{N-1}} = g_{N-1} \quad \text{on } V,$$

where $V \subset \Omega$ is an $(n-1)$-dimensional analytical manifold, and in which $\frac{\partial}{\partial n} = \hat{n} \cdot \nabla$ denotes derivation perpendicular to $V$. We furthermore assume that

- all coefficients $c_\alpha$, $|\alpha| \leq N$, as well as $f$ are analytical on $\Omega$,
- $g_k$ is analytical on a neighbourhood of $x_0 \in V$ for each $k = 0, \ldots, N-1$,
- $V$ is non-characteristic at $x_0$ relative to the p.d.o. $L$.

Then there exists a unique, analytical solution of Eqs. (2.1–2.2) on a neighbourhood of $x_0$.

It is important to keep in mind that Theorem 2.1.1 does not tell us anything about well-posedness of the system. Thus an “analytical solution” is not necessarily tantamount to a “nice solution”. Recall Example 1.4.1.

### 2.2. Examples

**Example 2.2.1.** Consider the following p.d.e. with i.c.'s (cf. Chapter 12, notably Section 12.1):

$$\begin{align*}
\Box u &= h \quad \text{for } (x,t) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^+,
 u\big|_{t=0} &= \phi , \\
 u_t\big|_{t=0} &= \psi ,
\end{align*}$$

The plane $V : t = 0$ on which the i.c.'s are specified has a normal in $t$-direction everywhere and the p.d.o. contains a corresponding second order term of the form $\partial^2 u/\partial t^2$. In other words, $V$ is non-characteristic relative to $\Box$. In case of analytical $h$, $\phi$ and $\psi$, Theorem 2.1.1 therefore guarantees a unique analytical solution near $t = 0$.

**Example 2.2.2.** Consider the following p.d.e. with i.c. (cf. Chapter 13, notably Section 13.1):

$$\begin{align*}
u_t - \Delta u &= 0 \quad \text{for } (x,t) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^+ ,
 u\big|_{t=0} &= f .
\end{align*}$$

The plane $V : t = 0$ on which the i.c. is specified has a normal in $t$-direction everywhere, but the p.d.o. does not contain a corresponding second order term $\propto \partial^2 / \partial t^2$ in that direction. In other words, $V$ is everywhere characteristic relative to $\partial_t - \Delta$. Thus in this case Theorem 2.1.1 is not applicable.

**Example 2.2.3.** Consider the following p.d.e. with b.c.'s (cf. Chapter 14, notably Section 14.1):

$$\begin{align*}
\Delta u &= 0 \quad \text{for } x \in \Omega \subset \mathbb{R}^n ,
 u\big|_V &= \phi , \\
 \frac{\partial u}{\partial n}\big|_V &= \psi ,
\end{align*}$$

in which $V : \hat{n} \cdot x = 0$ represents a plane with normal $\hat{n} \in \mathbb{R}^n$ on which the b.c.'s are specified and $\partial / \partial n$ denotes the derivative normal to $V$. By virtue of rotation invariance of $\Delta$ a rotation may be performed s.t. it contains an explicit second order normal derivative $\partial^2 / \partial n^2$. Therefore $V$ is non-characteristic relative to $\Delta$. In case of analytical $\phi$ and $\psi$, Theorem 2.1.1 thus guarantees a unique analytical solution near $x = 0$ (or any other $x \in V$).

**Example 2.2.4.** Recall Example 1.4.1. This special case is covered by the previous example and thus has a unique, analytical solution near $x = 0$, with explicit form given in that example. Clearly, Theorem 2.1.1 does not guarantee a well-posedness solution!
3

CALCULUS OF VARIATIONS

3.1. Origin

In science p.d.e. systems often emerge as necessary and/or sufficient conditions from physical laws originally stated in integral forms, collecting “stuff” inside some closed region. Variational problems constitute a particularly important class. These are often used to describe dynamical systems characterised by some conserved quantity, such as energy-momentum, or by some stationary quantity, such as the so-called action functional.

3.2. Basic Technique

Consider a fixed region $\Omega \subset \mathbb{R}^n$, together with a function $F : \Omega \times \mathbb{K} \times \mathbb{K}^n \to \mathbb{R} : (x, u, \nabla u) \mapsto F(x, u, \nabla u)$, in which $x$ is shorthand for the $n$ independent variables $x = (x_1, \ldots, x_n) \in \Omega$ as usual, and in which $u : \Omega \to \mathbb{K} : x \mapsto u(x)$ and $\nabla u : \Omega \to \mathbb{K}^n : x \mapsto \nabla u(x)$ denote some dependent variable and its first order partial derivatives, respectively. (Although it does cover a wide class of relevant problems, the restriction to first order is merely imposed to simplify the discussion and could be dropped.)

Next, consider the functional $J : \mathcal{W}(\Omega) \to \mathbb{R} : u \mapsto J[u]$, given by

$$J[u] = \int_{\Omega} F(x, u, \nabla u) \, dx,$$

(3.1)
in which $dx$ is shorthand for the elementary volume measure $dx = dx_1 \ldots dx_n$, and $\mathcal{V}(\Omega) \subset \mathcal{P}(\Omega)$ denotes some function subspace. In a nutshell, the variational principle entails the study of how this functional behaves under sufficiently small but otherwise arbitrary variations $\delta u$ of the function $u$, seeking those instances of $u$ for which the functional is stationary, i.e. for which the “infinitesimal” variations $\delta J[u] = J[u + \delta u] - J[u]$ vanish identically (to first order) for all $\delta u$.

To cope with the $\infty$-dimensional nature of the vector space $\mathcal{V}(\Omega)$, consider 1-dimensional variations around $u \in \mathcal{V}(\Omega)$ of the form $\delta u = \epsilon v$ along some arbitrarily fixed element $v \in \mathcal{V}(\Omega)$, with $\epsilon \in \mathbb{R}$ “infinitesimally” small. The variational problem can then be recast into a standard calculus problem of finding stationary points of the integrand $J_u$ to correspond to a (local) minimum of $J$.

For our purposes we require $F \in C^2(\Omega \times \mathbb{K} \times \mathbb{K}^n)$ as a necessary condition, de-emphasizing the details of the other abovementioned aspects.

Let us apply the recipe of Eq. (3.2) to Eq. (3.1). In that case we obtain

$$J[u + \epsilon v] = \int_\Omega F(x, u + \epsilon v, \nabla u + \epsilon \nabla v) \, dx.$$  \hspace{1cm} (3.3)

A straightforward expansion up to first order in $\epsilon$ produces

$$J[u + \epsilon v] = \int_\Omega F(x, u, \nabla u) \, dx + \epsilon \int_\Omega \left[ F_u(x, u, \nabla u) v + \sum_{i=1}^n F_{u_i}(x, u, \nabla u) \frac{\partial v}{\partial x_i} \right] \, dx + \text{h.o.t.}$$  \hspace{1cm} (3.4)

in which the subscripts $u$ and $u_i$ denote derivatives w.r.t. the scalar argument $u$ and the $i$-th component of the vector argument $\nabla u$ ($i = 1, \ldots, n$), respectively. The term “h.o.t.” stands for higher order terms in $\epsilon$. The $x$-dependence of the dependent variables $u = u(x)$ and $v = v(x)$ is suppressed for the sake of brevity.

In order to factor out the arbitrary function $v$ in the first order term of the integrand, apply partial integration, exploiting the chain rule

$$F_{u_i}(x, u, \nabla u) \frac{\partial v}{\partial x_i} = \frac{\partial}{\partial x_i} \left( F_{u_i}(x, u, \nabla u) v \right) - \frac{\partial F_{u_i}}{\partial x_i}(x, u, \nabla u) v.$$  \hspace{1cm} (3.5)

Together with the definition of $J[u]$, Eq. (3.1), this yields

$$J[u + \epsilon v] = J[u] + \epsilon \int_\Omega \left[ F_u(x, u, \nabla u) - \sum_{i=1}^n \frac{\partial F_{u_i}}{\partial x_i}(x, u, \nabla u) \right] v \, dx \quad \text{for all } \epsilon \in \mathbb{R}.$$  \hspace{1cm} (3.6)

The $\ast$-integral can be rewritten with a result from differential geometry known as Stokes’ Theorem, recall page vii. The specific instance actually needed here is known in vector calculus as Green’s Theorem or the Divergence Theorem. The divergence of a vector field $w : \Omega \to \mathbb{K}^n : x \mapsto w(x)$ is defined as

$$\text{div } w = \sum_{i=1}^n \frac{\partial w_i}{\partial x_i}.$$  \hspace{1cm} (3.7)
The consequence of Stokes’ Theorem is that the volumetric \( \ast \)-integral can be expressed as a surface integral:
\[
\int_{\Omega} \sum_{i=1}^{n} \frac{\partial}{\partial x_i} (F_u(x, u, \nabla u) \, v) \, dx = \int_{\partial \Omega} \sum_{i=1}^{n} F_{u_i}(x, u, \nabla u) \, n_i \, v \, d\sigma .
\] (3.8)

All in all we have obtained
\[
\lim_{\epsilon \to 0} \frac{J[u + \epsilon v] - J[u]}{\epsilon} = \int_{\Omega} \left[ F_u(x, u, \nabla u) - \sum_{i=1}^{n} \frac{\partial F_{u_i}}{\partial x_i} (x, u, \nabla u) \right] v \, dx + \int_{\partial \Omega} \sum_{i=1}^{n} F_{u_i}(x, u, \nabla u) \, n_i \, v \, d\sigma .
\] (3.9)

The condition of stationarity entails that this should vanish for all \( v \in \mathcal{U}(\Omega) \). This implies
\[
F_u(x, u, \nabla u) - \sum_{i=1}^{n} \frac{\partial F_{u_i}}{\partial x_i} (x, u, \nabla u) = 0 \quad \text{on} \ \Omega,
\] (3.10)
\[
\sum_{i=1}^{n} F_{u_i}(x, u, \nabla u) \, n_i = 0 \quad \text{on} \ \partial \Omega . \quad (3.11)
\]

To see this, first consider the subset of all modes \( v \in \mathcal{U}_0(\Omega) \) that vanish on the boundary \( \partial \Omega \). Then the second term on the r.h.s. of Eq. (3.9) is absent, implying that the part in-between square brackets in the integrand of the first term must be identically zero, which proves Eq. (3.10). Subsequently use this fact in Eq. (3.9), now also including all nontrivial modes \( v \in \mathcal{U}(\Omega) \) that do not vanish on the boundary. This establishes Eq. (3.11).

Eq. (3.10) is known as the Euler-Lagrange equation associated with the functional Eq. (3.1), furnished with a natural b.c., viz. Eq. (3.11).

### 3.3. Examples

We illustrate the variational principle by some examples taken from classical mechanics. The first example illustrates the principle in the simplest context of an action functional for the path of a single point particle, which takes the form of a time integral of a so-called Lagrangian function (kinetic minus potential energy associated with the particle’s trajectory \( x(t) \)). This case is extended to continuum mechanics in the second example, illustrating how a wave phenomenon for a field quantity \( u(x, t) \) emerges from a dynamical system described in terms of an action functional that takes the form of a spatiotemporal integral of a Lagrangian density (kinetic minus potential energy densities “stored” in a field configuration \( u(x, t) \)).

**Example 3.3.1.** For a single point particle the action functional over the time interval \([0, T] \) takes the form
\[
S[x] = \int_{0}^{T} L(t, x, \dot{x}) \, dt ,
\]
in which \( x = x(t) \in \mathbb{R}^n \) and \( \dot{x} = \dot{x}(t) \in \mathbb{R}^n \) are parametrisations of the particle’s position and velocity at time \( t \in [0, T] \).

The action principle entails that \( S[x] \) is stationary for a physical path \( x = x(t) \in \mathbb{R}^n \), i.e.
\[
\delta S[x] = \int_{0}^{T} \delta L(t, x, \dot{x}) \, dt = \int_{0}^{T} \left( \sum_{i=1}^{n} \frac{\partial L}{\partial x_i} \delta x_i + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}_i} \delta \dot{x}_i \right) \, dt = 0 .
\]

Via partial integration this is seen to be equivalent to
\[
\delta S[x] = \sum_{i=1}^{n} \int_{0}^{T} \left( \frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i \, dt + \sum_{i=1}^{n} \int_{0}^{T} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i \, dt = 0 .
\]
The last term is a boundary term:

\[
\int_0^T \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \delta x_i \right) dt = \left. \frac{\partial L}{\partial \dot{x}_i} \delta x_i \right|_{t=T} - \left. \frac{\partial L}{\partial \dot{x}_i} \delta x_i \right|_{t=0}.
\]

Since each instantaneous \( \delta x_i(t), t \in [0, T] \), is a priori independent, stationarity implies

\[
\left\{ \begin{array}{l}
\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = 0, \\
\left. \frac{\partial L}{\partial \dot{x}_i} \right|_{t=0} = 0, \\
\left. \frac{\partial L}{\partial \dot{x}_i} \right|_{t=T} = 0.
\end{array} \right.
\]

The o.d.e. obtained in this way is known as the “Euler-Lagrange equation” corresponding to the action \( S[x] \). The expression \( p_i = \frac{\partial L}{\partial \dot{x}_i} \) is known in the trade as the (\( i \)-th component of the) “momentum”. Thus the “natural” b.c.’s are provided by vanishing momenta.

In classical mechanics, the Lagrangian for a one-particle system represents the difference \( L = T - U \) of kinetic \( (T) \) and potential energy \( (U) \) of the particle. In particular, for a particle of mass \( m > 0 \) attached to an ideal spring with elastic modulus \( \kappa > 0 \) we have

\[
S[x] = \frac{1}{2} \int_0^T [m\dot{x}^2(t) - \kappa x^2(t)] dt.
\]

For \( N \)-particle systems one introduces a phase-space pair \((x_\nu(t), \dot{x}_\nu(t))\) for each particle, \( \nu = 1, \ldots, N \). One then adds up their joint kinetic and potential energy contributions to build the corresponding Lagrangian.

The next example illustrates the continuum limit for an \( N \)-particle system as \( N \to \infty \), replacing discrete point trajectories \( x_\nu(t) \) by a dense spatiotemporal field \( u(x,t) \).

**Example 3.3.2.** Consider the following functional on \((n+1)\)-dimensional flat spacetime \( \mathbb{R}^n \times \mathbb{R}^+ \), with \( \rho, \tau > 0 \):

\[
S[u] = \frac{1}{2} \int_{\mathbb{R}^+} \int_{\mathbb{R}^n} \left[ \rho u_t(x,t)^2 - \tau \| \nabla u(x,t) \|^2 \right] dx dt.
\]

The procedure of Section 3.2, notably Eqs. (3.10–3.11), produces the following Euler-Lagrange system (we disregard the boundary at infinity):

\[
\left\{ \begin{array}{l}
\rho u_{tt} - \tau \Delta u = 0 \quad \text{on} \ \mathbb{R}^n \times \mathbb{R}^+, \\
u_t = 0 \quad \text{for} \ (x,t) \in \mathbb{R}^n \times \{0\}.
\end{array} \right.
\]

This is the wave equation, recall the d’Alembertian operator introduced in Section 1.2 on page 4, with \( c^2 = \tau/\rho \).

If one takes the form of the Lagrangian density for granted, the above example explains the origin of the associated wave phenomenon, which the “deus ex machina” p.d.e. system as such fails to do. But where did the specific Lagrangian density come from? The next example, adapted from Gelfand and Fomin [3, Chapter 36], illustrates this for the simplified case \( n = 1 \), by making the explicit connection between a discrete \( N \)-particle system (in the appropriate limit \( N \to \infty \)) and its continuum counterpart formulated in terms of a physical field.
3. CALCULUS OF VARIATIONS

3.3. EXAMPLES

Figure 3.1: Vibrating string. Vertical displacement and stretching are highly exaggerated for the sake of clarity.

Example 3.3.3. We consider the transverse displacement of a homogeneous string, elastically fastened at the end points \( x = 0 \) and \( x = \ell \), meaning that the end points feel a restoring force proportional to their displacements \( u(0, t) \) and \( u(\ell, t) \). The setup is sketched in Fig. 3.1 for some fiducial moment \( t \in [0, T] \) in time.

More specifically, the string is attached at the end points \( x = 0 \) and \( x = \ell \) to ideal springs constrained to allow stretching only in vertical direction. Unstretched, these strings have zero length, while their elastic moduli are given by the constants \( \kappa_1 > 0 \) and \( \kappa_2 > 0 \), respectively. The string has horizontal length \( \ell \) and its equilibrium state is \( u(x, t) = 0 \) for all \( x \in [0, \ell] \) and \( t \in [0, T] \), say. During vibration it will be assumed that the overall stretching of the string remains small, so that the length of the string is approximately constant, i.e. roughly equal to \( \ell \).

In order to construct the Lagrangian for our continuous string (excluding the springs at the end points) we may divide it into \( N \) horizontal pieces of equal width \( \Delta x \) (so that \( \ell = N \Delta x \)), and consider each element as a particle with mass \( \Delta m_x = \rho \Delta x \), in which \( \rho > 0 \) is the string's (constant) linear mass density. Freezing position \( x \in [0, \ell] \) along the horizontal axis, the kinetic energy of the particle is determined by its velocity \( u_t(x, t) \), and equals, to lowest order in \( \Delta x \),

\[
\Delta T \approx \frac{1}{2} \rho u_t(x, t)^2 \Delta x .
\]

If the displacements \( u(x, t) \) and \( u(x + \Delta x, t) \) of the string at the ends of the interval \( (x, x + \Delta x) \) differ, then, by the Pythagorean rule, the difference \( \Delta u = u(x + \Delta x, t) - u(x, t) \) induces a (small) stretching of the string by an amount \( \Delta \ell = \sqrt{\Delta x^2 + \Delta u^2} - \Delta x = \Delta x \left( \sqrt{1 + \left( \frac{\Delta u}{\Delta x} \right)^2} - 1 \right) = \frac{1}{2} u_x(x, t)^2 \Delta x + \text{h.o.t.} \). Assuming zero potential energy of the string in equilibrium, the stretched piece of string then acquires a potential energy equal to the work needed for the stretching, viz. \( \Delta U = \tau \Delta \ell \), in which \( \tau \) is the string's (approximately) constant tension (magnitude of forces applied by the ends of the string once stretched). To lowest order in \( \Delta x \) we thus find:

\[
\Delta U \approx \frac{1}{2} \tau u_x(x, t)^2 \Delta x .
\]

Adding up all contributions \( \Delta L = \Delta T - \Delta U \) for the string pieces and taking the limit \( \Delta x \to 0 \) (and \( N \to \infty \) while keeping \( \ell = N \Delta x \) fixed), we obtain the action functional as a time-integral of the Lagrangian \( L \), which itself becomes a spatial integral of a Lagrangian density \( \mathcal{L} \):

\[
(*) \quad S[u] = \int_0^T \int_0^\ell \mathcal{L}(x, t, u, u_x, u_t) \, dx \, dt ,
\]

with

\[
(*) \quad \mathcal{L}(x, t, u, u_x, u_t) = \frac{1}{2} \left[ \rho u_t(x, t)^2 - \tau u_x(x, t)^2 \right] .
\]

Upon varying the function \( u \) by \( \delta u \) we obtain a lowest order variation \( \delta \mathcal{L} \) of \( \mathcal{L} \) in (\( * \)) given by

\[
\delta \mathcal{L} = \rho u_t \delta u_t - \tau u_x \delta u_x = - [\rho u_{tt} - \tau u_{xx}] \delta u + \frac{\partial}{\partial t} [\rho u_t \delta u] - \frac{\partial}{\partial x} [\tau u_x \delta u] ,
\]
3.3. EXAMPLES

The system’s natural b.c.’s are found by independent variation of the Euler-Lagrange equations, essentially already derived in the previous example (take $n = 1$), thus become:

\[ \rho u_{tt} - \tau u_{xx} = 0 \quad \text{on} \quad (0, \ell) \times (0, T). \] (••)

The system’s natural b.c.’s are found by independent variation of $\delta u(x, T) = \delta u(x, 0) = \delta u(\ell, t) = \delta u(0, t) = 0$. The Euler-Lagrange equations, essentially already derived in the previous example (take $n = 1$), thus become:

\[ \rho u_{tt} - \tau u_{xx} = 0 \quad \text{on} \quad (0, \ell) \times (0, T). \] (••)

The system’s natural b.c.’s are found by independent variation of $\delta u(x, T)$ and $\delta u(x, 0)$, considered as functions of $x \in (0, \ell)$, respectively of $\delta u(\ell, t)$ and $\delta u(0, t)$, considered as functions of $t \in (0, T)$:

\[ \begin{cases} u_1(0, 0) = u_1(x, T) = 0 & \text{on} \quad (0, \ell), \\ u_2(0, t) = u_2(\ell, t) = 0 & \text{on} \quad (0, T). \end{cases} \] (†)

These do not necessarily reflect realistic constraints, since the physical details of the fastening of the end points of the string have been ignored (so far).

In the final example we account for the contributions of the end points as well.

**Example 3.3.4.** The Euler-Lagrange equation (••)—or rather the analysis, still to be done, of its solutions—clarifies the potential string dynamics if no boundary or initial/final conditions, such as (†), are imposed. In realistic scenarios the b.c.’s must reflect the way in which the end points are fastened, so that we need to include the contribution of the springs in Fig. 3.1 into an extended action functional, which is the sum of (••) for the string itself and two additional terms, one for each end point. Moreover, unlike the b.c.’s, the temporal conditions are artifacts of our a priori arbitrary choice of temporal interval. For this reason $T \to \infty$ in Example 3.3.2, admitting only an initial condition, which may likewise be dispensed with by taking the time interval to be all of $\mathbb{R}$.

Let us investigate the natural physical boundary conditions by including the springs into the system. In order to pull the end points away from equilibrium ($u = 0$) towards their actual fixation points, $u(0, t)$ and $u(\ell, t)$, the work expended will be “remembered” in the form of potential energy. Recall that the springs are ideal, i.e. they have zero mass, and zero length if unloaded, with elastic moduli $\kappa_1 > 0$ and $\kappa_2 > 0$. Again assuming zero potential energy for unloaded springs, and defining $u_1(t) = u(0, t)$ and $u_2(t) = u(\ell, t)$, we thus have

\[ U_1 = \frac{1}{2} \kappa_1 u_1(t)^2 \quad \text{and} \quad U_2 = \frac{1}{2} \kappa_2 u_2(t)^2, \]

for their respective potential energies, with corresponding actions

\[ S_1[u_1] = -\frac{1}{2} \int_0^T \kappa_1 u_1(t)^2 \, dt \quad \text{and} \quad S_2[u_2] = -\frac{1}{2} \int_0^T \kappa_2 u_2(t)^2 \, dt. \]

Added to (••) this yields the action functional for the entire setup including b.c.’s:

\[ S_{\text{total}}[u] = S[u] + S_1[u_1] + S_2[u_2]. \]
Recall that \((\bullet\bullet)\) arises from the variational principle by considering stationarity of the action under all variations \(\delta u\) of \(u\) with \(\delta u(x, T) = \delta u(x, 0) = \delta u(\ell, t) = \delta u(0, t) = 0\). Thus when varying \(S_{\text{total}}[u]\) under variations \(\delta u\) including nontrivial variations \(\delta u(0, t) \neq 0\) and \(\delta u(\ell, t) \neq 0\), then \((\bullet\bullet)\) is a necessary but insufficient condition. Imposing this, we may restrict our attention to the variation of the boundary actions:

\[
\delta S_{\text{total}}[u] = (\circ) + (\circ) + \delta S_1[u_1] + \delta S_2[u_2].
\]

Discarding \((\circ)\) by imposing \(\delta u(x, T) = \delta u(x, 0) = 0\) (since we have not specified any a priori physical constraints for the initial and final configuration of the system), and observing that

\[
\delta S_1[u_1] = -\kappa_1 \int_0^T u_1(t) \delta u_1(t) \, dt = -\kappa_1 \int_0^T u(0, t) \delta u(0, t) \, dt,
\]

respectively

\[
\delta S_2[u_2] = -\kappa_2 \int_0^T u_2(t) \delta u_2(t) \, dt = -\kappa_2 \int_0^T u(\ell, t) \delta u(\ell, t) \, dt,
\]

we end up with

\[
\delta S_{\text{total}}[u] = -\int_0^T [\tau u_x(\ell, t) + \kappa_2 u(\ell, t)] \delta u(\ell, t) \, dt + \int_0^T [\tau u_x(0, t) - \kappa_1 u(0, t)] \delta u(0, t) \, dt.
\]

The action principle dictates that this should vanish for all \(\delta u(\ell, t)\) and \(\delta u(0, t)\), so that we finally obtain the following b.c.’s for \(t \in (0, T)\):

\[
\begin{align*}
-\kappa_1 u(0, t) + \tau u_x(0, t) &= 0, \\
\kappa_2 u(\ell, t) + \tau u_x(\ell, t) &= 0.
\end{align*}
\]

These “mixed” b.c.’s include the limiting cases of free end points \((\kappa_1, \kappa_2 \to 0, \text{implying } u_x(0, t) = u_x(\ell, t) = 0)\) as well as of rigidly fixed end points \((\kappa_1, \kappa_2 \to \infty, \text{implying } u(0, t) = u(\ell, t) = 0)\).
4 Distribution Theory

4.1. Motivation

A classical p.d.o., recall Eqs. (1.1–1.2), requires differentiability of its operand. If we wish to admit non-classical solutions, such as Eq. (1.8) for the p.d.e. system Eqs. (1.5–1.6) with \( f \notin C^1(\mathbb{R}) \), then this raises a paradox.

The resolution is provided by the theory of distributions. This theory employs auxiliary, so-called test functions, which are, by construction, differentiable in classical sense. The problem statement of Eqs. (1.1–1.2) is then reformulated into one in which the p.d.o. is “transposed” to these test functions.

To appreciate the idea, consider the simplest o.d.e. in \( n = 1 \):

\[
u' = f \quad \text{on } \mathbb{R}.
\]

(4.1)

For a classical solution to exist, we need to require \( u \in C^1(\mathbb{R}) \) and \( f \in C^0(\mathbb{R}) \). Assuming, for the moment, that this is the case, multiply left and right hand sides with a test function \( \phi \in \mathcal{S}(\mathbb{R}) \), and integrate:

\[
\int_{-\infty}^{\infty} u'(x)\phi(x)dx = \int_{-\infty}^{\infty} f(x)\phi(x)dx \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}).
\]

(4.2)

The class \( \mathcal{S}(\mathbb{R}) \) is explained in Section 4.2. For the present case all we need to know is that \( \phi \in C^1(\mathbb{R}) \) and \( \phi(x) \to 0 \) “sufficiently fast” as \( x \to \pm \infty \). Since we require Eq. (4.2) to hold for all \( \phi \in \mathcal{S}(\mathbb{R}) \), it is equivalent to the classical formulation Eq. (4.1).
Next, perform a partial integration step on the l.h.s. The result is

\[- \int_{-\infty}^{\infty} u(x)\phi'(x)dx = \int_{-\infty}^{\infty} f(x)\phi(x)dx \quad \text{for all } \phi \in \mathcal{D}(\mathbb{R}). \quad (4.3)\]

Note that there is no effective boundary term on the l.h.s., since by construction\(^1\) \(u(x)\phi(x) \to 0\) whenever \(x \to \pm \infty\).

Given the original regularity constraints on \(f\) and \(u\), Eq. (4.3) is again equivalent to Eq. (4.1), as well as to Eq. (4.2). However, if we take Eq. (4.3) as our point of departure, i.e. if we would have formulated the o.d.e. in terms of this equation \(a\ priori\), then there is apparently no need for insisting on differentiability or even continuity of the functions \(u\) and \(f\), as long as the integral expressions are well-defined.

We may mimic the original o.d.e. formulation, Eq. (4.1), by redefining the classical functions

\[f : \mathbb{R} \to \mathbb{K} : x \mapsto f(x), \quad (4.4)\]
\[u' : \mathbb{R} \to \mathbb{K} : x \mapsto u'(x), \quad (4.5)\]
in Eq. (4.1) through “function overloading”, as follows:

\[f : \mathcal{D}(\mathbb{R}) \to \mathbb{K} : \phi \mapsto f(\phi) = \int_{-\infty}^{\infty} f(x)\phi(x)dx, \quad (4.6)\]
\[u' : \mathcal{D}(\mathbb{R}) \to \mathbb{K} : \phi \mapsto u'(\phi) = -\int_{-\infty}^{\infty} u(x)\phi'(x)dx. \quad (4.7)\]

These are the distributional counterparts of the classical functions, Eqs. (4.4–4.5), originally intended in Eq. (4.1). Note that, \(by \ definition\), \(u'(\phi) = -u(\phi')\), so that no “classical derivative” \(u'(x)\) of \(u(x)\) is ever needed.

No confusion between a function and its corresponding distribution is likely to arise as long as one is clear about the applicable prototype, i.e. either \(\mathbb{R} \to \mathbb{K}\) (for a classical function) or \(\mathcal{D}(\mathbb{R}) \to \mathbb{K}\) (for a distribution). The distributional counterpart of Eq. (4.1) may thus be written as

\[u' = f \quad \text{on } \mathcal{D}(\mathbb{R}). \quad (4.8)\]

As a shorthand we say that \(f \in \mathcal{S}'(\mathbb{R})\) if \(f : \mathcal{D}(\mathbb{R}) \to \mathbb{K}\) is a continuous linear functional of the type Eq. (4.6). A few considerations arise:

- Can all continuous linear functionals of type \(\mathcal{S}'(\mathbb{R}) \to \mathbb{K}\) be represented by an integral formula such as Eq. (4.6)? (Lemma 4.2.1 and Theorem 4.2.2, v.i.)
- For those cases that can be represented by Eq. (4.6), is the correspondence with the classical representation one-to-one? (Notation 4.2.1 and Definition 4.2.5, v.i.)

These issues (for which the strict answer is negative in both cases) will be considered in the next section.

The example of “function overloading”, illustrated by Eqs. (4.1) and (4.8) via Eqs. (4.4–4.7), generalizes to arbitrary p.d.e.’s in any dimension \(n\), and to any domain of definition \(\Omega \subset \mathbb{R}^n\), provided the test functions are properly constructed. The formal requirements are given in Section 4.2.

**Observation 4.1.1.** The lesson is thus that we may, without formal notational distinction, interpret any p.d.e. either in the “classical” sense, in which operands and data are ordinary functions satisfying differentiability conditions compatible with the p.d.e., or in a “distributional” sense, in which operands and data represent linear functionals requiring only weak regularity conditions. The latter interpretation admits generalised solutions.

\(^1\)This explains why we need “sufficiently fast” convergence: the amplification factor \(u(x)\) might otherwise spoil convergence!
4. DISTRIBUTION THEORY

4.2. DISTRIBUTIONS FORMALISED

4.2. Distributions Formalised

We start by considering the multivariate case with unbounded domain $\mathbb{R}^n$.

**Definition 4.2.1.** The class $\mathcal{S}(\mathbb{R}^n)$ of smooth test functions of rapid decay, $\phi : \mathbb{R}^n \to \mathbb{K}$, a.k.a. Schwartz functions, is defined as follows:

$$\phi \in \mathcal{S}(\mathbb{R}^n) \iff \phi \in C^\infty(\mathbb{R}^n) \text{ and } \sup_{x \in \mathbb{R}^n} |x^\alpha \nabla^\beta \phi(x)| < \infty \text{ for all multi-indices } \alpha \text{ and } \beta.$$ 

Thus a Schwartz function is an infinitely differentiable function which is “essentially compact”, meaning that it decays rapidly in any direction towards infinity. Rapid decay applies to any of its derivatives and means that convergence of function values towards zero is faster than any monomial factor (or more generally, any function of polynomial growth) would be able to counteract.

**Definition 4.2.2.** An a.e. defined function $f : \mathbb{R}^n \to \mathbb{K}$ is called a function of polynomial growth, $f \in \mathcal{P}(\mathbb{R}^n)$, if there exist constants $c > 0$ and $m \geq 0$ such that $|f(x)| \leq c (1 + \|x\|^2)^m$ a.e. Equivalently, $f \in \mathcal{P}(\mathbb{R}^n)$, if there exists a real-valued polynomial $p$ such that $|f(x)| \leq p(x)$ a.e.

**Definition 4.2.3.** The topological dual $\mathcal{S}'(\mathbb{R}^n)$ of $\mathcal{S}(\mathbb{R}^n)$ is the class of continuous linear functionals $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{K}$, a.k.a. the class of tempered distributions.

**Theorem 4.2.1.** If $T \in \mathcal{S}'(\mathbb{R}^n)$, then there exist a constant $c > 0$ and multi-indices $\alpha, \beta$ such that

$$|T(\phi)| \leq c \sup_{x \in \mathbb{R}^n} |x^\alpha \nabla^\beta \phi(x)|.$$ 

Two tempered distributions $T_1, T_2 \in \mathcal{S}'(\mathbb{R}^n)$ are equal iff $T_1(\phi) = T_2(\phi)$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. $T \in \mathcal{S}'(\mathbb{R}^n)$ is called positive if $T(\phi) > 0$ for all positive test functions $\phi \in \mathcal{S}(\mathbb{R}^n)$. Of special interest are the so-called regular tempered distributions, as these can be associated (non-uniquely) with classical functions.

**Definition 4.2.4.** The regular tempered distribution $T_f \in \mathcal{S}'(\mathbb{R}^n)$ associated with a function of polynomial growth, $f \in \mathcal{P}(\mathbb{R}^n)$, is the tempered distribution

$$T_f : \mathcal{S}(\mathbb{R}^n) \to \mathbb{K} : \phi \mapsto \int_{\mathbb{R}^n} f(x) \phi(x) \, dx.$$ 

One says that the regular tempered distribution $T_f$ is associated with the “function $f$ under the integral”. A regular tempered distribution $T_f$ is often whimsically identified with the function $f$ despite ambiguity. This is just the “function overloading” principle introduced in Section 4.1. Although convenient, one needs to remain cautious.

**Notation 4.2.1.** If $T_f = T_g$, i.e. if $T_f(\phi) = T_g(\phi)$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$, we say that $f$ and $g$ are equivalent: $f \sim g$.

**Observation 4.2.1.** Equivalent functions may differ on a set of measure zero, i.e. w.r.t. physically void details.

**Definition 4.2.5.** It is customary to write $f = g$ instead of $f \sim g$ for two equivalent functions of polynomial growth, $f, g \in \mathcal{P}(\mathbb{R}^n)$, associated with the regular tempered distributions $T_f, T_g \in \mathcal{S}'(\mathbb{R}^n)$.

Recall that classically, i.e. outside the context of distributions, functions are identified, $f = g$, iff $f(x) = g(x)$ for all $x \in \mathbb{R}^n$. It should be clear from the context which of the two equivalence concepts are applicable.

The space of tempered distributions is strictly larger than the space of regular tempered distributions.

**Lemma 4.2.1.** Let $\phi \in \mathcal{S}(\mathbb{R}^n)$, and $a \in \mathbb{R}^n$ any point. Then

$$\delta_a : \mathcal{S}(\mathbb{R}^n) \to \mathbb{K} : \phi \mapsto \phi(a)$$

defines a tempered distribution, called the Dirac point distribution at point $a \in \mathbb{R}^n$. 

23
The Dirac point distribution $\delta_0$ is usually written as $\delta$.

**Definition 4.2.6.** One tends to abuse the notation of Definition 4.2.4, writing *any* tempered distribution $T$ in the form $T_j$, even if no such function $f \in \mathcal{D}(\mathbb{R}^n)$ exists. In particular, the Dirac point distribution $\delta_a$ is associated with a hypothetical “function under the integral” with the same symbol, $\delta_a : \mathbb{R}^n \to \mathbb{K} : x \mapsto \delta_a(x) = \delta(x - a)$, s.t. $\delta_a(\phi) \overset{\text{def}}{=} \int_{\mathbb{R}^n} \delta_a(x) \phi(x) \, dx \overset{\text{def}}{=} \phi(a)$. Notice, however, that, as a “function under the integral”, we have $\delta(x) = 0$ for all $x \neq 0$, and that it is meaningless to ask for the value $\delta(0)$.

**Observation 4.2.2.** Dimensional analysis reveals that, from a physicist’s point of view, one should interpret the function under the integral $\delta_a : \mathbb{R}^n \to \mathbb{K}$ as a (singular) density function (normalised to unity upon integration).

The deliberate malpractice of notational ambiguity is carried over to all tempered distributions, cf. Theorem 4.2.2, v.i.

**Notation 4.2.2.** One often declines to make a notational distinction between functions and distributions. Thus it should be clear from the context whether a symbol like $f$ denotes a classical function $f : \mathbb{R}^n \to \mathbb{K}$ as it might appear “under the integral” (i.e. $f \in \mathcal{D}(\mathbb{R}^n))$ or a proper distribution $f : \mathcal{D}'(\mathbb{R}^n) \to \mathbb{K}$ (or $f \in \mathcal{D}'(\mathbb{R}^n)$).

Except for a few instances where it might lead to confusion, we will adhere to this convenient abuse of notation, as it will allow us to interpret a p.d.e. in either classical or generalised sense, whichever is appropriate.

The core of distribution theory is that we can define arbitrary derivatives of tempered distributions in a well-posed way, independent of operationally meaningless “infinitesimal” details.

**Definition 4.2.7.** For any nonnegative multi-index $\alpha$ the derivative $\nabla^\alpha T \in \mathcal{D}'(\mathbb{R}^n)$ of $T \in \mathcal{D}'(\mathbb{R}^n)$ is the tempered distribution defined by $\nabla^\alpha T : \mathcal{D}(\mathbb{R}^n) \to \mathbb{K} : \phi \mapsto (\nabla^\alpha T)(\phi) \overset{\text{def}}{=} T(\nabla^\alpha_1 \phi)$, in which the transposed p.d.o. is given by $\nabla^\alpha_1 = (-1)^{\lvert \alpha \rvert} \nabla^\alpha$.

The reason for the minus signs is that in the subspace of regular tempered distributions $T_f$ for which $f$ is sufficiently smooth, the above definition boils down to the classical definition of differentiation. This can be verified by repeated partial integrations using the integral expression on the right hand side of Definition 4.2.4 and the fact that $f$ is smooth and of polynomial growth. For let $f \in C^\infty(\mathbb{R}^n) \cap \mathcal{D}(\mathbb{R}^n)$, then on the one hand we have

$$\nabla^\alpha T_f(\phi) \overset{\text{def}}{=} T_f(\nabla^\alpha_1 \phi) = (-1)^{\lvert \alpha \rvert} \int_{\mathbb{R}^n} f(x) \nabla^\alpha \phi(x) \, dx ,$$

by Definition 4.2.7. On the other hand, we may consider the regular tempered distribution $T_{\nabla^\alpha f}$ associated with the function $\nabla^\alpha f$, and apply repetitive partial integrations, i.e.

$$T_{\nabla^\alpha f}(\phi) = \int_{\mathbb{R}^n} \nabla^\alpha f(x) \phi(x) \, dx \overset{\text{p.i.}}{=} (-1)^{\lvert \alpha \rvert} \int_{\mathbb{R}^n} f(x) \nabla^\alpha \phi(x) \, dx .$$

All boundary terms that show up in this process vanish as a result of the rapid decay of the test function (recall the example in Section 4.1). Comparing the last two results we observe that $T_{\nabla^\alpha f} = \nabla^\alpha T_f$, a reasonable consistency requirement. This explains why Definition 4.2.7 is the natural way to define $\nabla^\alpha T$ for any tempered distribution. In particular, we may differentiate the Dirac point distribution.

**Result 4.2.1.** Recall Definition 4.2.7. Let $\phi \in \mathcal{D}(\mathbb{R}^n)$, $a \in \mathbb{R}^n$ any given point, and $\alpha$ a multi-index. Then $\nabla^\alpha \delta : \mathcal{D}(\mathbb{R}^n) \to \mathbb{K} : \phi \mapsto \nabla^\alpha_1 \phi(a) = (-1)^{\lvert \alpha \rvert} \nabla^\alpha \phi(a)$.
It can be shown that a distribution whose support is a single point is a finite linear combination of derivatives of the Dirac distribution at that point.

**Definition 4.2.8.** The support \( \operatorname{supp} f \) of a function \( f : \Omega \to \mathbb{K} \) is the closure of the set of all \( x \in \Omega \) for which \( f(x) \neq 0 \).

Despite the fact that \( \delta(a) \) is ill-defined, we also say that \( \operatorname{supp} \delta_a = \{ a \} \), because \( \delta_a(x) \) can be defined for \( x \neq a \), with \( \delta_a(x) = 0 \).

**Theorem 4.2.2.** If \( f \in \mathcal{S}'(\mathbb{R}^n) \) is a general point distribution, with \( \operatorname{supp} f = \{ a \} \) for some point \( a \in \mathbb{R}^n \), then there exists an order \( m \in \mathbb{Z}_+^0 \) such that

\[
    f = \sum_{|\alpha| \leq m} c_\alpha \nabla_\alpha \delta_a.
\]

Point distributions with \( m = 0 \) are also referred to as point measures.

Adaptations of the theory of Schwartz exist for arbitrary open subdomains \( \Omega \subset \mathbb{R}^n \). Such adaptations are useful for p.d.e. systems defined on \( \Omega \subset \mathbb{R}^n \).

**Definition 4.2.9.** The class \( \mathcal{D}(\Omega) \) of smooth compact test functions on the open set \( \Omega \subset \mathbb{R}^n \), \( \phi : \Omega \to \mathbb{K} \), is defined as follows:

\[
    \phi \in \mathcal{D}(\Omega) \iff \phi \in C^\infty(\Omega) \text{ and } \phi \text{ has compact support } \operatorname{supp} \phi \subset \Omega.
\]

Note that the supremum condition in Definition 4.2.1 is obsolete, its role (viz. the suppression of boundary terms) is automatically fulfilled by the condition of compact support: \( \nabla_\beta \phi|_{\partial \Omega} = 0 \) for all multi-indices \( \beta \).

A question that might arise is whether nonzero test functions \( \phi \in \mathcal{D}(\Omega) \) actually exist. If not, the stipulated equivalence of Eqs. (4.1) and (4.3) would fail for a general domain \( \Omega \). Example 4.4.1 in Section 4.4 illustrates how one may construct a smooth function of arbitrary compact support for the one-dimensional case.

**Definition 4.2.10.** The topological dual \( \mathcal{D}'(\Omega) \) of \( \mathcal{D}(\Omega) \) is the class of continuous linear functionals \( T : \mathcal{D}(\Omega) \to \mathbb{K} \), a.k.a. the class of distributions on \( \Omega \).

Results obtained for \( \mathcal{S}(\mathbb{R}^n) \) and \( \mathcal{S}'(\mathbb{R}^n) \) may be adapted to \( \mathcal{D}(\Omega) \) and \( \mathcal{D}'(\Omega) \) for any subset \( \Omega \subset \mathbb{R}^n \), including \( \Omega = \mathbb{R}^n \).

Limits on \( \mathcal{D}'(\Omega) \) (and a fortiori on \( \mathcal{S}'(\mathbb{R}^n) \), v.i. Theorem 4.2.3) are introduced as follows.

**Definition 4.2.11.** Let \( u, u_n \in \mathcal{D}'(\Omega), n \in \mathbb{N} \). We say that \( u = \lim_{n \to \infty} u_n \) if \( u(\phi) = \lim_{n \to \infty} u_n(\phi) \) for all \( \phi \in \mathcal{D}(\Omega) \).

It is not difficult to show that the acts of differentiation and of taking a limit can be interchanged.

**Lemma 4.2.2.** Recall Definition 4.2.11. If \( u = \lim_{n \to \infty} u_n \), then \( \nabla_\alpha u = \lim_{n \to \infty} \nabla_\alpha u_n \) for all multi-index orders \( \alpha \).

We end with a theorem that sheds light on how the various function spaces are nested. It clarifies in a most concise way why we may think of the elements of \( \mathcal{D}'(\Omega) \) as truly generalised functions.

**Theorem 4.2.3** (Inclusion Theorem). The following inclusions hold:

\[
    \mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n).
\]

Moreover, for compact \( \Omega \subset \mathbb{R}^n \) we have

\[
    L^q(\Omega) \subset L^p(\Omega) \text{ whenever } 1 \leq p \leq q.
\]
Note that the deliberate “confusion” of Notation 4.2.2 actually renders the first inclusion chain meaningful, as it relies on our identification of equivalence classes of functions with (regular) tempered distributions. The second inclusion chain tells us that, restricted to compact domains, \( L^1(\Omega) \) is the largest and \( L^\infty(\Omega) \) the smallest of all \( L^p \)-type spaces. A compact space is a normed space in which all points lie within some fixed distance of each other. A closed and bounded subset of \( \mathbb{R}^n \) (equipped with its standard norm) provides an example of a compact space. Beyond the realm of Euclidean geometry we may think of intrinsically compact universes \( U \), such as \( \mathbb{S}^{n-1} = \{ x \in \mathbb{R}^n \mid \| x \| = 1 \} \), the unit sphere (conceptually embedded in an \( n \)-dimensional Euclidean space).

To conclude, a p.d.e. and its solution may be interpreted either in classical or in distributional sense. The merit of the notational ambiguity introduced above is that the standard formulation of a linear p.d.e. needs no modification, as it admits both views simultaneously.

**Result 4.2.2.** The linear \( N \)th order p.d.e. \( Lu = f \) on \( \Omega \) may be interpreted either as \( Lu(x) = f(x) \) for all \( x \in \Omega \), or as \( Lu(\phi) = f(\phi) \) for all \( \phi \in \mathcal{D}(\Omega) \). In the former case it is assumed that \( u \in C^N(\Omega) \), \( f \in C(\Omega) \). The latter case requires only very weak assumptions, viz. \( u, f \in \mathcal{D}'(\Omega) \). In that case it is understood that, by definition, \( Lu(\phi) = u(L^1\phi) \). More specifically, if \( L = \sum_{|\alpha| \leq N} c_\alpha \nabla_\alpha \), then \( L^1 = \sum_{|\alpha| \leq N} c_\alpha \nabla_\alpha^1 = \sum_{|\alpha| \leq N} (-1)^{|\alpha|} c_\alpha \nabla_\alpha \).

### 4.3. Reparametrisations of the Dirac Point Distribution

A generalisation of Definition 4.2.6 is particularly useful in applications that call for curvilinear coordinates, such as polar, cylindrical, or spherical coordinates. Suppose \( x = x(\xi) \) represents a (diffeomorphic) reparametrisation of Cartesian coordinates \( x \in \mathbb{R}^n \) in terms of curvilinear coordinates \( \xi \in \mathbb{R}^n \), and \( a = x(\alpha) \in \mathbb{R}^n \), say.

**Lemma 4.3.1.** In Cartesian coordinates, the \( n \)-dimensional Dirac delta function, \( \delta_n(x) = \delta(x-a) \), is separable:

\[
\delta(x-a) = \prod_{i=1}^n \delta(x_i-a_i),
\]

in which each factor on the r.h.s. represents a one-dimensional Dirac delta function.

**Theorem 4.3.1.** We have the formal identity

\[
\delta(x-a) = \frac{\delta(\xi-\alpha)}{|J(\xi)|}
\]

in which \( J = \det \frac{\partial x}{\partial \xi} \) denotes the Jacobian determinant.

To see this, note that, on the one hand,

\[
\phi(a) = \int \delta(x-a) \phi(x) \, dx = \int \delta(x(\xi)-x(\alpha)) \phi(x(\xi)) \, |J(\xi)| \, d\xi,
\]

in which the r.h.s. is obtained by a change of variables, \( x = x(\xi) \). On the other hand we also have

\[
\phi(a) = \phi(x(\alpha)) = \int \delta(\xi-\alpha) \phi(x(\xi)) \, d\xi.
\]

Comparison with the previous representation establishes Theorem 4.3.1. With our conventions for polar, cylindrical and spherical coordinates (cf. Preliminaries) we obtain the following.

**Result 4.3.1.** In polar, cylindrical, respectively spherical coordinates in \( \mathbb{R}^2/\mathbb{R}^3 \) we have

\[
\begin{align*}
\delta(x-x_0)\delta(y-y_0) &= \frac{\delta(r-R)\delta(\phi-\Phi)}{r}, \\
\delta(x-x_0)\delta(y-y_0)\delta(z-z_0) &= \frac{\delta(r-R)\delta(\phi-\Phi)\delta(\zeta-Z)}{r^2 \sin \theta}.
\end{align*}
\]
The constant parameters \( R, \Phi, Z \) and \( \Theta \) pertain to the appropriate polar, cylindrical, respectively spherical coordinates \( r, \phi, z \) and \( \theta \) of the fixed points \((x_0, y_0) \in \mathbb{R}^2\), respectively \((x_0, y_0, z_0) \in \mathbb{R}^3\).

The proof is left as an exercise. Cf. Example 4.4.6 for the case of an affine reparametrisation.

### 4.4. Examples

**Example 4.4.1.** For \( c > 0 \) consider \( \psi : \mathbb{R} \to \mathbb{R} : x \mapsto \psi(x) \) given by

\[
\psi(x) = \begin{cases} 
\exp \left( \frac{1}{x^2 - c^2} \right) & \text{if } x \in (-c, c) \\
0 & \text{if } x \not\in (-c, c)
\end{cases}
\]

It is clear that \( \lim_{x \to \pm c} \psi(x) = 0 \), and in fact, \( \lim_{x \to \pm c} \psi^{(k)}(x) = 0 \) for any \( k \)-th order derivative of \( \psi \). Thus if we set \( \psi^{(k)}(\pm c) = 0 \), then \( \psi \in C^\infty(\mathbb{R}) \). The function \( \psi \) is said to be a smooth function of compact support, a.k.a. as a “bump function”. The support of \( \psi \) is the closure of the set of \( x \in \mathbb{R} \) for which \( \psi(x) \neq 0 \), i.e. the closed interval \([-c, c]\) in this example. Notation: \( \psi \in \mathcal{D}((-c, c)) \). We may generate new smooth functions of compact support via convolution with an integrable function \( \chi \) of compact support, say \( \text{supp } \chi = [a, b] \), with \( a < b \):

\[
\phi(x) = \int_{-\infty}^{\infty} \chi(y) \psi(x - y) \, dy.
\]

The function \( \phi \) inherits its smoothness from \( \psi \), and its compact support (viz. \([a - c, b + c]\)) from both \( \psi \) and \( \chi \). Slick choice of parameters will thus allow you to construct a smooth function of any desirable compact support.

**Example 4.4.2.** The following function fails to be differentiable at \( x = 0 \) in classical sense:

\[
f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x) = |x|
\]

However, interpreting it as a regular tempered distribution via the identification \( f \sim T_f \in \mathcal{S}'(\mathbb{R}) \), i.e.

\[
T_f(\phi) = \int_{\mathbb{R}} f(x) \phi(x) \, dx = - \int_{-\infty}^{0} x \phi(x) \, dx + \int_{0}^{\infty} x \phi(x) \, dx,
\]

for all \( \phi \in \mathcal{S}(\mathbb{R}) \), renders it infinitely differentiable in distributional sense. Indeed we have

\[
T_f'(\phi) = - \int_{\mathbb{R}} f(x) \phi'(x) \, dx = \int_{-\infty}^{0} x \phi'(x) \, dx - \int_{0}^{\infty} x \phi'(x) \, dx
\]

\[
= [x\phi(x)]_{x=-\infty}^{x=0} - [x\phi(x)]_{x=0}^{x=\infty} - \int_{-\infty}^{0} \phi(x) \, dx + \int_{0}^{\infty} \phi(x) \, dx = \int_{\mathbb{R}} \text{sgn } (x) \phi(x) \, dx = T_{\text{sgn }}(\phi),
\]

in which \( \text{sgn } : \mathbb{R} \to \mathbb{R} : x \mapsto \text{sgn } (x) \) is the sign function:

\[
\text{sgn } : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \text{sgn } (x) = \begin{cases} 
-1 & \text{als } x < 0, \\
0 & \text{als } x = 0, \\
1 & \text{als } x > 0.
\end{cases}
\]

Thus in a precise sense we have \( T_f' = T_{\text{sgn }} \). Identifying the r.h.s. with \( T_{f'} \), —which strictly holds only for classically differentiable \( f \)—amounts to the “pseudo-classical” statement \( f'(x) = \text{sgn } (x) \), i.e.

\[
\frac{d}{dx} |x| = \text{sgn } (x),
\]

27
but note that the (classically ill-defined) value of $f'(0) = \text{sgn}(0)$ is irrelevant and has been arbitrarily assigned here (to ensure antisymmetry). Moreover,

$$T''(\phi) = \int_{\mathbb{R}} f(x) \phi''(x) \, dx = - \int_{-\infty}^{0} x \phi''(x) \, dx + \int_{0}^{\infty} x \phi''(x) \, dx$$

$$= -\left[ x\phi'(x) \right]_{x=-\infty}^{0} + \left[ x\phi'(x) \right]_{x=0}^{\infty} + \int_{-\infty}^{0} \phi'(x) \, dx - \int_{0}^{\infty} \phi'(x) \, dx$$

$$= [\phi(x)]_{x=-\infty}^{0} - [\phi(x)]_{x=0}^{\infty} = 2 \phi(0) = 2\delta(\phi).$$

Thus $T'' = 2\delta$. Again adopting the interpretation $f \sim T_f$, we may state “pseudoclassically” that $f'' = 2\delta$. The factor 2 reflects the jump in slope from $-1$ to $+1$ of the graph of $f$ at the origin. Higher order derivatives are now straightforward: $T^{(k+1)} \sim f^{(k+1)} = 2\delta^{(k-1)}$, $k \in \mathbb{N}$.

**Example 4.4.3.** Let $U \in \mathcal{S}'(\mathbb{R})$ be a tempered distribution satisfying the following distributional o.d.e.:

$$U'' = \delta.$$

Clearly this o.d.e. does not admit a classical solution $U \in C^2(\mathbb{R})$, for then $U'' \in C^0(\mathbb{R})$, contradicting the r.h.s. Therefore, stipulate a solution of the form $U = T_u \in \mathcal{S}'(\mathbb{R})$ for some function $u : \mathbb{R} \to \mathbb{R}$. Example 4.4.2 shows that

$$u(x) = \frac{1}{2} |x|$$

is a generalised solution. Indeed, direct substitution of

$$U(\phi) = T_u(\phi) = \int_{-\infty}^{\infty} u(x)\phi(x) \, dx$$

in (1) yields

$$U''(\phi) = U'(\phi') = \int_{-\infty}^{\infty} u(x)\phi''(x) \, dx = \frac{1}{2} \int_{-\infty}^{\infty} |x|\phi''(x) \, dx = -\frac{1}{2} \int_{-\infty}^{0} x\phi''(x) \, dx + \frac{1}{2} \int_{0}^{\infty} x\phi''(x) \, dx$$

$$= -\frac{1}{2} [x\phi'(x)]_{-\infty}^{0} + \frac{1}{2} \int_{-\infty}^{0} \phi'(x) \, dx + \frac{1}{2} [x\phi'(x)]_{0}^{\infty} - \frac{1}{2} \int_{0}^{\infty} \phi'(x) \, dx$$

$$= \frac{1}{2} [\phi(x)]_{-\infty}^{0} + \frac{1}{2} [\phi(x)]_{0}^{\infty} = \phi(0) = \delta(\phi).$$

Since this holds for all $\phi \in \mathcal{S}(\mathbb{R})$ we indeed find $U'' = \delta$.

Note that this solution is not unique, since we may add a general solution of the corresponding homogeneous o.d.e. Thus the general solution of (1) is given by (the regular tempered distribution corresponding to)

$$U(x) = \frac{1}{2} |x| + ax + b.$$

**Example 4.4.4.** The staircase function $f : \mathbb{R} \to \mathbb{R} : x \mapsto f(x)$, given by $f(x) = |x|$, i.e. the entier of $x \in \mathbb{R}$, or the largest integer $k \in \mathbb{Z}$ such that $k \leq x$, has infinitely many discontinuities, cf. Figure 4.1.

Since $f$ has a discontinuity at each integer, and constant values in-between integers, its classical derivative is

$$f'(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R}\setminus\mathbb{Z} \\ \text{ill-defined} & \text{if } x \in \mathbb{Z}. \end{cases}$$

Thus

$$T_f' \neq T_{f'},$$
4. DISTRIBUTION THEORY

4.4. EXAMPLES

Figure 4.1: Graph of \( y = f(x) = |x| \). The red bullets at the discontinuities are not part of the graph.

the r.h.s. being trivial if \( f' : \mathbb{R} \setminus \mathbb{Z} \to \mathbb{R} \) is interpreted as the classical derivative of \( f \), with arbitrarily assigned values at the discontinuities:

\[
T_f'(\phi) = \int_{\mathbb{R}} f'(x) \phi(x) \, dx = \int_{\mathbb{R} \setminus \mathbb{Z}} f'(x) \phi(x) \, dx = 0.
\]

The last two equalities hold by virtue of the zero measure of the set \( \mathbb{Z} \). For the l.h.s., on the other hand, we find

\[
T_f'(\phi) = -T_f'(\phi') = -\int_{-\infty}^{\infty} f(x) \phi'(x) \, dx,
\]

which clearly does not vanish for all \( \phi \in \mathcal{S}(\mathbb{R}) \). Indeed, interpreted as a generalised function \( f \sim T_f \), we find

\[
T_f'(\phi) = -\sum_{k \in \mathbb{Z}} \delta_k \cdot \phi_k' = T_f(\phi') = -\sum_{k \in \mathbb{Z}} \int_{k}^{k+1} k \phi_k'(x) \, dx = -\sum_{k \in \mathbb{Z}} k \phi(k+1) - \phi(k) = \sum_{k \in \mathbb{Z}} \phi(k) = \sum_{k \in \mathbb{Z}} \delta_k(\phi),
\]

for any \( \phi \in \mathcal{S}(\mathbb{R}) \), whence

\[
T_f' = \sum_{k \in \mathbb{Z}} \delta_k.
\]

Thus, recalling Definition 4.2.6, we may loosely write

\[
\frac{d}{dx} |x| = \sum_{k \in \mathbb{Z}} \delta_k(x) = \sum_{k \in \mathbb{Z}} \delta(x-k) = \text{III}(x),
\]

where on the r.h.s. we have introduced the so-called “Dirac comb” \( \text{III} \), a.k.a. “impulse train” or “sampling function” in physics and engineering disciplines.

Example 4.4.5. Consider the o.d.e. \( u' + u = 0 \) for \( x \in \mathbb{R} \). The function \( f : \mathbb{R} \to \mathbb{R} : x \mapsto f(x) \), given by

\[
f(x) = \begin{cases} 
0 & x < 0 \\
e^{-x} & x \geq 0,
\end{cases}
\]

clearly provides a solution “almost everywhere”, viz. for \( x \in \mathbb{R} \setminus \{0\} \). Obviously \( f \not\in C^1(\mathbb{R}) \) is not a classical solution due to the discontinuity at the origin. Let us therefore verify whether its associated distribution \( T_f \in \mathcal{D}'(\mathbb{R}) \) satisfies the o.d.e. in distributional sense. (Note that \( f \) is not of polynomial growth, which is why we must look beyond \( \mathcal{S}'(\mathbb{R}) \).) This is in fact nearly the case, in the sense that the reservation “almost everywhere” for the classical validity domain is obsolete in the distributional view, and a certain “correction term” that has no classical counterpart enters the picture. Taking

\[
T_f(\phi) = \int_{-\infty}^{\infty} f(x) \phi(x) \, dx = \int_{0}^{\infty} e^{-x} \phi(x) \, dx,
\]

where
4.4. EXAMPLES

4. DISTRIBUTION THEORY

a direct verification shows that
\[ T'_f(\phi) \overset{*}{=} -T_f(\phi') = -\int_{-\infty}^{\infty} f(x) \phi'(x) \, dx = -\int_{0}^{\infty} e^{-x} \phi'(x) \, dx = \left[ -e^{-x} \phi(x) \right]_0^{\infty} - \int_{0}^{\infty} e^{-x} \phi(x) \, dx = \phi(0) - T_f(\phi). \]

The equality marked by \( * \) holds by definition of distributional differentiation, the one marked by \( \star \) follows by partial integration. (Note that for the last step to hold we must have suitably defined test functions; the choice \( \phi \in \mathcal{D}(\mathbb{R}^n) \) made here will do.) Using the definition of the Dirac point distribution, \( \delta(\phi) = \phi(0) \), we may rewrite the result as
\[ T'_f(\phi) = \delta(\phi) - T_f(\phi), \]
which shows that \( T_f \) in fact satisfies the inhomogeneous o.d.e. \( u' + u = \delta \) in \( \mathcal{D}'(\mathbb{R}) \)-distributional sense. Notice that no restrictions on the domain of definition need to be imposed, and that the result is consistent with the classical result given the reservation on the classical validity domain, since \( \delta(x) = 0 \) for \( x \in \mathbb{R} \setminus \{0\} \).

Example 4.4.6. Let \( A \) be a constant invertible \( n \times n \)-matrix and \( b \in \mathbb{R}^n \) a constant vector parameter. Then
\[ (\circ) \quad \delta_{A,b}(x) \overset{\text{def}}{=} \delta(Ax + b) = \frac{1}{|\det A|} \delta(x + A^{-1}b) = \frac{1}{|\det A|} \delta_{-A^{-1}b}(x). \]

The crucial thing to note is the appearance of the scaling factor. There is no “classical” view to see this (one sometimes sees physically void statements s.a. “\( \delta(0) = \infty \”). A true explanation requires a strict distributional approach. To this end, let \( \phi \in \mathcal{S}(\mathbb{R}^n) \), then
\[ \delta_{A,b}(\phi) \overset{\text{def}}{=} \int_{\mathbb{R}^n} \delta(Ax + b) \phi(x) \, dx \overset{*}{=} \frac{1}{|\det A|} \int_{\mathbb{R}^n} \delta(y) \phi(A^{-1}(y - b)) \, dy \]
\[ = \frac{1}{|\det A|} \phi(-A^{-1}b) \overset{\star}{=} \frac{1}{|\det A|} \int_{\mathbb{R}^n} \delta(x + A^{-1}b) \phi(x) \, dx \overset{\star}{=} \frac{1}{|\det A|} \int_{\mathbb{R}^n} \delta_{-A^{-1}b}(x) \phi(x) \, dx. \]

In \( * \) a change of variables has been carried out: \( y = Ax + b \), whence \( x = A^{-1}(y - b) \) and \( dy = |\det A| \, dx \) with invariant integration boundaries. In \( \star \) we have used the definition of the standard Dirac delta function, \( \delta(y) \), and in \( \star \) the one for a shifted Dirac delta function, \( \delta_a(x) \) with \( a = -A^{-1}b \), recall Definition 4.2.6. From these observations, using the fact that \( \phi \in \mathcal{S}(\mathbb{R}^n) \) is arbitrary, it follows that
\[ \delta_{A,b} = \frac{1}{|\det A|} \delta_{-A^{-1}b}, \]
for which \( (\circ) \) provides the “function under the integral”.
5

FOURIER TRANSFORMATION

5.1. Introduction

The spaces $\mathcal{S}(\mathbb{R}^n)$, $\mathcal{S}'(\mathbb{R}^n)$, and $L^2(\mathbb{R}^n)$ are closed under Fourier transformation\(^1\). These cases are covered in Sections 5.2–5.4, respectively. Section 5.5 summarizes some useful Fourier theorems.

Various definitions of the Fourier transform are encountered in the literature. To appreciate how these are related, we introduce the $(a,b)$-parametric convention, which covers all cases of practical interest. In this convention, $(a, b) \in \mathbb{R}\{0\} \times \mathbb{R}^+$ are fixed parameters.

Our default will be $(a, b) = (1, 1)$, which is fairly common in pure mathematics and systems engineering.

5.2. The Fourier Transform on $\mathcal{S}(\mathbb{R}^n)$

**Definition 5.2.1.** Fourier transformation is a continuous, linear, invertible mapping $\mathcal{F}_{(a,b)} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) : \phi \mapsto \mathcal{F}_{(a,b)}(\phi)$, defined as follows:

$$\mathcal{F}_{(a,b)}(\phi)(\omega) \overset{\text{def}}{=} b \int_{\mathbb{R}^n} e^{-ia\omega \cdot x} \phi(x) \, dx.$$

\(^1\)In particular, every tempered distribution $f \in \mathcal{S}'(\mathbb{R}^n)$ has a Fourier transform. This is not the case for every $f \in \mathcal{S}'(\mathbb{R}^n)$. 

31
5.2. THE FOURIER TRANSFORM ON $\mathcal{S}(\mathbb{R}^N)$

Notation 5.2.1. For our default convention, $(a, b) = (1, 1)$, we will write $\mathcal{F}(\phi)$ or $\hat{\phi}$ instead of $\mathcal{F}_{(1,1)}(\phi)$.

Conjecture 5.2.1. Cf. Definition 5.2.1. The inverse Fourier transformation is given by

$$\mathcal{F}^{-1}_{(a,b)}(\psi)(x) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{ia\omega \cdot x} \psi(\omega) \, d\omega.$$  

Observation 5.2.1. Forward and inverse Fourier transforms are formally related by an $(a, b)$-reparametrization:

$$\mathcal{T} : \mathbb{R} \setminus \{0\} \times \mathbb{R}^+ \longrightarrow \mathbb{R} \setminus \{0\} \times \mathbb{R}^+ : (a, b) \mapsto \mathcal{T} (a, b) \overset{\text{def}}{=} (a', b')$$

given by

$$\begin{cases} a' &= -a \\ b' &= \frac{|a|^n}{(2\pi)^n b}. \end{cases}$$

In other words:

$$\mathcal{F}^{-1}_{(a,b)} = \mathcal{F}_{(a',b')}.$$  

Note that the transformation $\mathcal{T} = \mathcal{T}^{-1}$ equals its own inverse, allowing us to “toggle” between forward and backward Fourier transforms.

The class of test functions is closed under differentiation. The question arises of how the Fourier transform of a derivative is related to that of the test function itself. This is expressed by the following theorem.

Theorem 5.2.1. Let $\phi \in \mathcal{S}(\mathbb{R}^n)$, and let $\mathcal{F}_{(a,b)}(\phi) \in \mathcal{S}(\mathbb{R}^n)$ be its Fourier transform according to Definition 5.2.1. Moreover, let $P(c, \nabla) = \sum_{c_\alpha \in \mathbb{R}, \in \mathbb{R}}$ be a $k$-th order linear p.d.o. with constant coefficients $c_\alpha \in \mathbb{K}$, in which the sum extends over all multi-indices $\alpha$ of order $|\alpha| = 0, \ldots, k$. Then we have

$$\mathcal{F}_{(a,b)} (P(c, \nabla) \phi)(\omega) = P(c, i\omega) \mathcal{F}_{(a,b)}(\phi)(\omega).$$

Proof. The basic observation underlying Theorem 5.2.1 is the fact that $\nabla \phi e^{i\omega \cdot x} = (i\omega)^\alpha e^{i\omega \cdot x}$, from which it follows by linearity that $P(c, \nabla) e^{i\omega \cdot x} = P(c, i\omega) e^{i\omega \cdot x}$. In the terminology of linear algebra this states that the complex “planar wave” function $x \mapsto e^{i\omega \cdot x}$—considered as a function of $x$ with parameter $\omega$—is an eigenfunction of the linear p.d.o. $P(c, \nabla)$ with eigenvalue $P(c, i\omega)$. Using this, it follows that

$$P(c, \nabla) \phi(x) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} P(c, \nabla) e^{i\omega \cdot x} \mathcal{F}_{(a,b)}(\phi)(\omega) \, d\omega = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{i\omega \cdot x} P(c, i\omega) \mathcal{F}_{(a,b)}(\phi)(\omega) \, d\omega.$$  

The first equality follows by Conjecture 5.2.1 upon interchanging differentiation and integration. Inspection of the result, again with the help of Conjecture 5.2.1, shows that the function $\omega \mapsto P(c, i\omega) \mathcal{F}_{(a,b)}(\phi)(\omega)$ is the Fourier transform of $x \mapsto P(c, \nabla) \phi(x)$.

Example 5.2.1. Let us consider a few simple cases for $n = 1$. According to Theorem 5.2.1 we have, for $\omega \in \mathbb{R}$,

$$\mathcal{F}_{(a,b)} \left( \frac{d^k \phi}{dx^k} \right)(\omega) = (i\omega)^k \mathcal{F}_{(a,b)}(\phi)(\omega).$$

In higher dimensions results are similar, only now we have a frequency vector $\omega \in \mathbb{R}^n$ rather than a scalar frequency, $\omega = (\omega_1, \ldots, \omega_n)$. Also, we may now consider any linear combination of partial derivatives with respect to the $n$ spatial coordinates $x = (x_1, \ldots, x^n)$.

Example 5.2.2. Recall Theorem 5.2.1, and suppose $n = 2$. Writing $\omega = (\omega_x, \omega_y) \in \mathbb{R}^2$, Theorem 5.2.1 yields:

$$\mathcal{F}_{(a,b)} \left( \frac{\partial^{p+q} \phi}{\partial x^p \partial y^q} \right)(\omega_x, \omega_y) = (i\omega_x)^p (i\omega_y)^q \mathcal{F}_{(a,b)}(\phi)(\omega_x, \omega_y).$$

Example 5.2.3. Linear combinations of partial derivatives map to corresponding linear combinations of their Fourier transforms. As an example, consider the Laplacian $\Delta \phi$ of $\phi$. Fourier transformation yields

$$\mathcal{F}_{(a,b)} \left( \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} \right)(\omega_x, \omega_y) = -a^2 (\omega_x^2 + \omega_y^2) \mathcal{F}_{(a,b)}(\phi)(\omega_x, \omega_y).$$

32
Example 5.2.4. It is beyond our scope to deal rigorously with so-called fractional powers of differential operators, but their significance is most intuitive in Fourier space. The previous example reveals that one may view $\mathcal{F}(-\Delta)$ as a multiplication operator in Fourier space, viz. $\mathcal{F}(-\Delta) = (\omega \mapsto f(\omega)) \mapsto (\omega \mapsto \|\omega\| f(\omega))$. By the same token, it appears reasonable to define the multiplication operator $\mathcal{F}(\sqrt{-\Delta}) = (\omega \mapsto f(\omega)) \mapsto (\omega \mapsto \|\omega\| f(\omega))$.

In this way a seemingly irreducible second order p.d.o. like the d’Alembertian may be factorised into two conjugate first order pseudo-p.d.o.’s:

$$\Box = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta = \left(\frac{1}{c} \frac{\partial}{\partial t} \pm i \sqrt{-\Delta}\right) \left(\frac{1}{c} \frac{\partial}{\partial t} \mp i \sqrt{-\Delta}\right).$$

Paul Dirac (1902–1984) was the first to realize that in four-dimensional spacetime the d’Alembertian (more generally, the Helmholtz operator) can in fact be decomposed into two linear first order (non-fractional) p.d.o.’s, stipulating a linear combination of the form

$$\sum_{\mu=0}^{3} c_{\mu} \partial_{\mu},$$

provided the constant coefficients $c_{\mu}$ (in the convention of physics: $c_{\mu} = i \hbar \gamma_{\mu}$) in this p.d.o. are interpreted as $4 \times 4$ matrices, and (thus) the operand as a 4-component so-called spinor instead of a scalar function. This cunning trick led to the discovery of antimatter (and earned him a Nobel Prize in 1933). For further details, cf. Section 11.2.

### 5.3. The Fourier Transform on $\mathcal{S}'(\mathbb{R}^n)$

Formally, distribution theory allows one to extend Fourier transformation to generalised functions, as follows.

**Definition 5.3.1.** Fourier transformation is a continuous, linear, invertible mapping $\mathcal{F}(a,b) : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$: $T \mapsto \mathcal{F}(a,b)(T)$, defined as follows:

$$\mathcal{F}(a,b)(T)(\phi) \overset{\text{def}}{=} T(\mathcal{F}(a,b)(\phi)) \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n),$$

recall Definition 5.2.1.

Informally, we have seen that tempered distributions may be loosely identified with “functions under the integral”, provided we adopt suitable notational conventions to deal with the Dirac point distribution and related non-regular tempered distributions, recall Definition 4.2.6 and Theorem 4.2.2. Accepting this level of rigor, which is what we shall do henceforth, the Fourier transform and its inverse take the same form as before.

As an exercise you may convince yourself of the fact that for a regular tempered distribution Fourier transformation in the formal sense of Definition 5.3.1 agrees with our informal definition below for the defining function under the integral, assuming that the Fourier integral expression for this function is well-defined. Do bear in mind, however, that functions involved should, strictly speaking, be interpreted as “functions under the integral”.

**Definition 5.3.2.** Fourier transformation is a continuous, linear, invertible mapping $\mathcal{F}(a,b) : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$: $f \mapsto \mathcal{F}(a,b)(f)$, defined as follows:

$$\mathcal{F}(a,b)(f)(\omega) \overset{\text{def}}{=} b \int_{\mathbb{R}^n} e^{-ia\omega \cdot x} f(x) \, dx.$$

**Conjecture 5.3.1.** Cf. Definition 5.3.2. The inverse Fourier transformation is given by

$$\mathcal{F}^{-1}(a,b)(g)(x) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{ia\omega \cdot x} g(\omega) \, d\omega.$$
5.4. THE FOURIER TRANSFORM ON $L^2(\mathbb{R}^n)$

Example 5.3.1. Let $\hat{u}_s : \mathbb{R} \rightarrow \mathbb{R} : \omega \mapsto \hat{u}_s(\omega) = e^{-s|\omega|}$, in which $s > 0$ is a positive constant. Then

$$u_s(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} \hat{u}_s(\omega) \, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-s|\omega|+i\omega x} \, d\omega = \frac{1}{2\pi} \left( \int_{-\infty}^{0} e^{(s+ix)\omega} \, d\omega + \int_{0}^{\infty} e^{-(s+ix)\omega} \, d\omega \right).$$

We have

$$\int_{-\infty}^{0} e^{(s+ix)\omega} \, d\omega = \left[ \frac{1}{s + i\varepsilon} e^{(s+ix)\omega} \right]_{\omega \to -\infty}^{\omega = 0} = \frac{1}{s + i\varepsilon}.$$ 

Similarly,

$$\int_{0}^{\infty} e^{-(s+ix)\omega} \, d\omega = \left[ \frac{1}{-s + i\varepsilon} e^{-(s+ix)\omega} \right]_{\omega = 0}^{\omega \to +\infty} = \frac{1}{-s + i\varepsilon},$$

so that we finally obtain

$$u_s(x) = \frac{1}{2\pi} \left( \frac{1}{s + i\varepsilon} + \frac{1}{-s + i\varepsilon} \right) = \frac{s}{\pi x^2 + s^2}.$$ 

That we actually obtain a real-valued solution is no coincidence. You should be able to figure out yourself how we might have anticipated this. (Hint: $\hat{u}_s(\omega) = \hat{u}_s(-\omega) \in \mathbb{R}$.)

**Lemma 5.3.1.** The Fourier transform of the Dirac point distribution at the origin, $x \mapsto \delta(x)$, is a constant function:

$$\mathcal{F}_{(a,b)}(\delta)(\omega) = b.$$ 

Note that the symbol $b$ here actually stands for $b : \mathbb{R}^n \rightarrow \mathbb{K} : \omega \mapsto b$.

**Proof.** Apply the definition of the Dirac point distribution$^3$, Lemma 4.2.1, Page 23.

**Lemma 5.3.2.** The inverse Fourier transform of a constant function is proportional to the Dirac point distribution at the origin, $\delta(x)$:

$$\mathcal{F}^{-1}_{(a,b)}(b)(x) = \delta(x).$$

We refrain from a proof of Lemma 5.3.2. Note that Conjectures 5.2.1 and 5.3.1 follow from this lemma.

The lemma is equivalent to the following result, which is useful to memorize, as it does not depend on a particular Fourier convention:

$$\int_{\mathbb{R}^n} e^{\pm i\omega \cdot x} \, d\omega = (2\pi)^n \delta(x). \quad (5.1)$$

You may verify the following formal scaling property, valid for any regular $n \times n$-matrix $A$ and vector $B \in \mathbb{R}^n$:

$$\delta(Ax + B) = \frac{1}{|\det A|} \delta(x + A^{-1}B). \quad (5.2)$$

### 5.4. The Fourier Transform on $L^2(\mathbb{R}^n)$

Among all $L^p(\mathbb{R}^n)$-spaces ($p \geq 1$ or $p = \infty$) the space $L^2(\mathbb{R}^n)$ is rather special in the sense that it is isomorphic to its own dual. In fact $L^2(\mathbb{R}^n)$ is a so-called Hilbert space, i.e. a complete inner product space. This means that it is equipped with an inner product $(\cdot, \cdot) : L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^n) \rightarrow \mathbb{R}$, which in turn induces a norm $\| \cdot \| : L^2(\mathbb{R}^n) \rightarrow \mathbb{R}$, and that relative to this norm every Cauchy sequence $\{f_n\}$ converges to an element of $L^2(\mathbb{R}^n)$. Recall that $\{f_n\}$

$^3$Note that $e^{i\omega \cdot x}$ is not a Schwartz function. Nevertheless, it is clear that the Dirac point distribution can be generalised to any function space in which point evaluation makes sense.
5. FOURIER TRANSFORMATION 5.5. FOURIER THEOREMS

is a Cauchy sequence if for all \( \epsilon > 0 \) there exists an \( N \in \mathbb{N} \) such that \( \| f_m - f_n \| < \epsilon \) whenever \( m, n > N \). This elegant property of \( L^2(\mathbb{R}^n) \) is also reflected in its closure under Fourier transformation. Technically we may use the same definitions as for the spaces \( \mathcal{S}(\mathbb{R}^n) \) and \( \mathcal{S}'(\mathbb{R}^n) \) (Definition 5.2.1 and Conjecture 5.2.1, respectively Definition 5.3.2 and Conjecture 5.3.1), but with the same caveat as for the latter, viz. that one should interpret elements \( f \in L^2(\mathbb{R}^n) \) as equivalence classes, with \( f \sim g \) iff \( f(x) = g(x) \) a.e.

**Definition 5.4.1.** Fourier transformation is a continuous, linear, invertible mapping \( \mathcal{F}(a, b) : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n) \): \( f \mapsto \mathcal{F}(a, b)(f) \), defined as follows:

\[
\mathcal{F}(a, b)(f)(\omega) \overset{\text{def}}{=} b \int_{\mathbb{R}^n} e^{-i\omega \cdot x} f(x) \, dx.
\]

**Conjecture 5.4.1.** Cf. Definition 5.4.1. The inverse Fourier transformation is given by

\[
\mathcal{F}^{-1}(a, b)(g)(x) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{i\omega \cdot x} g(\omega) \, d\omega.
\]

### 5.5. Fourier Theorems

Recall Definitions 5.2.1, 5.3.2, 5.4.1 and Conjectures 5.2.1, 5.3.1, 5.4.1. In what follows \( u_{1,2} : \mathbb{R}^n \rightarrow \mathbb{K} \) and \( \tilde{u}_{1,2} : \mathbb{R}^n \rightarrow \mathbb{K} \) are suitably defined functions, meaning that all stipulated products and convolutions are well-defined, with well-defined (inverse) Fourier transforms. (A \( \overset{\sim}{\text{ symbol}} \) indicates that the corresponding function is conceived of as living in Fourier space.)

**Theorem 5.5.1** (Fourier Transform of a Function Product),

\[
\mathcal{F}(a, b)(u_1 u_2) = \frac{|a|^n}{(2\pi)^n b} \mathcal{F}(a, b)(u_1) \ast \mathcal{F}(a, b)(u_2).
\]

*In particular, following our convention \((a, b) = (1, 1)\), we have*

\[
\mathcal{F}(u_1 u_2) = \frac{1}{(2\pi)^n} \mathcal{F}(u_1) \ast \mathcal{F}(u_2).
\]

This theorem explains the popularity of the Fourier convention in which \((a, b) = (\pm 2\pi, 1)\), or \((a, b) = (\pm 1, 1/(2\pi)^n)\).

**Proof.** For notational convenience we set \( u \equiv u_1 u_2 \) and \( \mathcal{F}(a, b)(u) \equiv \tilde{u} \) in this proof. We have

\[
\tilde{u}(\omega) \overset{\text{def}}{=} b \int_{\mathbb{R}^n} u_1(x) u_2(x) e^{-i\omega \cdot x} \, dx.
\]

Substituting the respective inverse Fourier transforms for \( u_{1,2} \), i.e.

\[
u_{1,2}(x) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{i\omega \cdot x} \tilde{u}_{1,2}(\omega) \, d\omega,
\]

yields

\[
\tilde{u}(\omega) = b \int_{\mathbb{R}^n} \left[ \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{i\omega_1 \cdot x} \tilde{u}_1(\omega_1) \, d\omega_1 \right] \left[ \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} e^{i\omega_2 \cdot x} \tilde{u}_2(\omega_2) \, d\omega_2 \right] e^{-i\omega \cdot x} \, dx.
\]

We may rewrite this expression as

\[
\tilde{u}(\omega) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} \tilde{u}_1(\omega_1) \tilde{u}_2(\omega_2) \left[ b \int_{\mathbb{R}^n} e^{ia(\omega_1 + \omega_2 - \omega) \cdot x} \, dx \right] d\omega_1 d\omega_2,
\]

35
provided we are allowed to interchange the order of integrations. (We take this for granted.) We recognize the innermost integral as a variant of Eq. (5.1). Thus we conclude that
\[ \tilde{u}(\omega) = \frac{|a|^n}{(2\pi)^n b} \int_{\mathbb{R}^n} \tilde{u}_1(\omega_1) \tilde{u}_2(\omega - \omega_1) d\omega_1, \]
in other words,
\[ \tilde{u}(\omega) = \frac{|a|^n}{(2\pi)^n b} (\tilde{u}_1 * \tilde{u}_2)(\omega). \]
The rest follows by inspection of this result. \( \square \)

An immediate result of Theorem 5.5.1 is the following.

**Theorem 5.5.2 (Fourier Transform of a Convolution Product).**
\[ \mathcal{F}_{(a, b)}(u_1 * u_2) = \frac{1}{b} \mathcal{F}_{(a, b)}(u_1) \mathcal{F}_{(a, b)}(u_2). \]

*In particular, following our convention \((a, b) = (1, 1)\), we have*
\[ \mathcal{F}(u_1 * u_2) = \mathcal{F}(u_1) \mathcal{F}(u_2). \]

**Proof.** Set \(v_{1,2} = \mathcal{F}_{(a, b)}(u_{1,2})\). Theorem 5.5.1 applied to the pair \(v_{1,2}\) states that
\[ \mathcal{F}_{(a', b')}(v_{1}v_{2}) = \frac{|a'|^n}{(2\pi)^n b'} \left( \mathcal{F}_{(a', b')}(v_1) * \mathcal{F}_{(a', b')}(v_2) \right) . \]
for all \((a', b') \in \mathbb{R}\{0\} \times \mathbb{R}^+\). By Observation 5.2.1, Page 32, together with the definitions of \(v_{1,2}\), we may rewrite this as
\[ \mathcal{F}^{-1}_{(a, b)} \left( \mathcal{F}_{(a, b)}(u_1) \mathcal{F}_{(a, b)}(u_2) \right) = b \left( \mathcal{F}^{-1}_{(a, b)}(\mathcal{F}_{(a, b)}(u_1)) * \mathcal{F}^{-1}_{(a, b)}(\mathcal{F}_{(a, b)}(u_2)) \right) = b(u_1 * u_2). \]
Application of \(\mathcal{F}_{(a, b)}\) to left and right hand sides completes the proof. \( \square \)

Theorems 5.5.1–5.5.2 are powerful results that are often used in the analysis of linear shift invariant systems. A diagrammatic representation of Theorem 5.5.1 is
\[ \mathcal{F}_{(a, b)} \downarrow \quad \begin{array}{c} (u_1, u_2) \quad \quad \quad \rightarrow \quad \quad \quad u_1 u_2 \\ \uparrow \mathcal{F}^{-1}_{(a, b)} \end{array} \]

\[ \left( \mathcal{F}_{(a, b)}(u_1), \mathcal{F}_{(a, b)}(u_2) \right) \quad \rightarrow \quad \frac{|a|^n}{(2\pi)^n b} \mathcal{F}_{(a, b)}(u_1) * \mathcal{F}_{(a, b)}(u_2) \]

Theorem 5.5.2 can be depicted as
\[ \mathcal{F}_{(a, b)} \downarrow \quad \begin{array}{c} (u_1, u_2) \quad \quad \quad \rightarrow \quad \quad \quad u_1 * u_2 \\ \uparrow \mathcal{F}^{-1}_{(a, b)} \end{array} \]

\[ \left( \mathcal{F}_{(a, b)}(u_1), \mathcal{F}_{(a, b)}(u_2) \right) \quad \rightarrow \quad \frac{1}{b} \mathcal{F}_{(a, b)}(u_1) \mathcal{F}_{(a, b)}(u_2) \]

In each diagram the upper horizontal route expresses an algebraic operation in the spatial domain (product, resp. convolution). The lower horizontal route expresses the equivalent operation in Fourier space. For each operation (moving from one corner in the diagram to another) there are always two distinct but equivalent alternatives. Note that you may reverse the vertical arrows if you replace \(\mathcal{F}_{(a, b)}\) by \(\mathcal{F}^{-1}_{(a, b)}\), vice versa.
The diagrams suggest how one could efficiently implement numerical schemes for convolution (either in the spatial domain or in Fourier space), viz. by avoiding the convolution corner in the applicable diagram. The reason for such a detour would be that whereas convolution is a multi-local operation, multiplication acts point-wise, which potentially saves us a significant amount of computation time. Of course, the concomitant cost of the forward and inverse Fourier transformations encountered along the detour has to be taken into consideration as well. Still, in certain cases this may be the most efficient way to compute a convolution product.

Other, equivalent formulations of Theorem 5.5.1 and 5.5.2 are stated in the following results.

Result 5.5.1 (Fourier Transform of a Function Product in Fourier Space).

\[
\mathcal{F}^{-1}_{(a,b)}(\tilde{u}_1) \cdot \mathcal{F}^{-1}_{(a,b)}(\tilde{u}_2) = \frac{|a|^n}{(2\pi)^n b} \mathcal{F}^{-1}_{(a,b)}(\tilde{u}_1 \ast \tilde{u}_2).
\]

In particular, following our convention \((a, b) = (1, 1)\), we have

\[
\mathcal{F}^{-1}(\tilde{u}_1) \cdot \mathcal{F}^{-1}(\tilde{u}_2) = \frac{1}{(2\pi)^n} \mathcal{F}^{-1}(\tilde{u}_1 \ast \tilde{u}_2).
\]

Proof. This is basically Theorem 5.5.1, with \(u_{1,2} = \mathcal{F}^{-1}_{(a,b)}(\tilde{u}_{1,2})\), and with \(\mathcal{F}^{-1}_{(a,b)}\) applied to both sides. □

Result 5.5.2 (Fourier Transform of a Convolution Product in Fourier Space).

\[
\mathcal{F}^{-1}_{(a,b)}(\tilde{u}_1) \ast \mathcal{F}^{-1}_{(a,b)}(\tilde{u}_2) = \frac{1}{b} \mathcal{F}^{-1}_{(a,b)}(\tilde{u}_1 \tilde{u}_2).
\]

In particular, following our convention \((a, b) = (1, 1)\), we have

\[
\mathcal{F}^{-1}(\tilde{u}_1) \ast \mathcal{F}^{-1}(\tilde{u}_2) = \mathcal{F}^{-1}(\tilde{u}_1 \tilde{u}_2).
\]

Proof. This is Theorem 5.5.2, subject to the same adaptation as in the previous proof. □

Diagrams similar to Diagrams (5.3–5.4) can be constructed for the latter two results.
6

COMPLEX
ANALYSIS

6.1. Introduction

Although complex numbers \( z = x + iy \in \mathbb{C} \) may be viewed as points \((x, y) \in \mathbb{R}^2\) in the plane, it is important to realize that \( \mathbb{C} \)-differentiability is a stronger requirement than \( \mathbb{R}^2 \)-differentiability. In this chapter we recall a few relevant definitions and results from complex analysis.

6.2. \( \mathbb{C} \)-Differentiability and Holomorphic Functions

**Definition 6.2.1.** Let \( \Omega \subset \mathbb{C} \) be an open subset, and \( f : \Omega \to \mathbb{C} \) a complex-valued function. Then \( f \) is said to be differentiable at \( z_0 \in \Omega \) if

\[
f'(z_0) \equiv \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \in \mathbb{C}
\]

exists, i.e. if

\[
|f(z) - f(z_0) - f'(z_0)(z - z_0)| = o(z - z_0) \quad (z \to z_0).
\]

**Theorem 6.2.1** (Cauchy-Riemann Equations). Let \( f : \Omega \to \mathbb{C} : z \mapsto f(z) = u(z) + iv(z) \), with \( z = x + iy \in \mathbb{C} \),

39
6.2. C-DIFFERENTIABILITY AND HOLOMORPHIC FUNCTIONS

Let \( u(z) = \text{Re} f(z) \in \mathbb{R}, v(z) = \text{Im} f(z) \in \mathbb{R} \), be differentiable at \( z_0 = x_0 + iy_0 \in \Omega \subset \mathbb{C} \). Then at \( z = z_0 \) we have

\[
\begin{align*}
\frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\
\frac{\partial v}{\partial x} &= -\frac{\partial u}{\partial y}
\end{align*}
\]

or, equivalently,

\[i \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y}.
\]

**Definition 6.2.2.** Let \( \Omega \subset \mathbb{C} \) be an open subset, and \( f : \Omega \to \mathbb{C} \) a complex-valued function. The function \( f \) is called holomorphic at \( z_0 \in \Omega \subset \mathbb{C} \) if it is \( C \)-differentiable at \( z_0 \), and holomorphic on \( \Omega \) if it is holomorphic at every point of \( \Omega \).

**Example 6.2.1.** Examples of holomorphic functions (with \( z = x + iy \)):

- \( z \mapsto z^{k-1}, k \in \mathbb{N} \) is holomorphic on \( \mathbb{C} \),
- \( z \mapsto \exp(z) \) is holomorphic on \( \mathbb{C} \),
- \( z \mapsto z^{-k}, k \in \mathbb{N} \), is holomorphic on \( \mathbb{C} \setminus \{0\} \).

**Example 6.2.2.** Counterexamples:

- \( z \mapsto z^* = x - iy \) is nowhere holomorphic,
- \( z \mapsto |z| = \sqrt{zz^*} = \sqrt{x^2 + y^2} \) is nowhere holomorphic,
- \( z \mapsto \text{Re} z = (z + z^*)/2 = x \) is nowhere holomorphic,
- \( z \mapsto \text{Im} z = (z - z^*)/(2i) = y \) is nowhere holomorphic.

Note that the counterexamples can all be written as functions of \( z \) and \( z^* \), treated formally as independent variables. As a rule of thumb, such functions are not holomorphic if they depend nontrivially on \( z^* \).

That \( \mathbb{C} \)-differentiability is a strong condition also becomes apparent from the following theorems.

**Theorem 6.2.2.** A holomorphic function on an open set \( \Omega \subset \mathbb{C} \) is infinitely differentiable on \( \Omega \).

**Theorem 6.2.3.** A holomorphic function \( f : \Omega \to \mathbb{C} \) on an open set \( \Omega \subset \mathbb{C} \) has a Taylor series expansion at every point \( z_0 \in \Omega \), which converges on any open ball \( \beta \subset \Omega \) centered at \( z_0 \). In particular, if \( \Omega = \mathbb{C} \), then \( f \) admits a globally convergent Taylor series expansion at every point.

Holomorphic functions of the latter type are called entire functions. An interesting result by Cauchy (but named after Liouville) states that if an entire function is bounded, then it is constant. Thus if an entire function is non-trivial in the sense of not being constant, then in some direction (but not necessarily in all directions) it must tend to infinity as \( |z| \to \infty \).

Another remarkable property of holomorphic functions is the following.

**Theorem 6.2.4** (Cauchy’s Integral Theorem). Let \( \Omega \subset \mathbb{C} \) be a simply connected open subset, and \( f : \Omega \to \mathbb{C} \) a holomorphic function. Then

\[
\oint_{\gamma} f(z) \, dz = 0
\]

for every closed loop \( \gamma \subset \Omega \).

Two types of singularities may occur at isolated points, “poles” and “essential singularities”.
**Theorem 6.2.5** (Cauchy’s Differentiation Formula). Let \( f : \Omega \to \mathbb{C} \) be a holomorphic function on a simply connected open subset \( \Omega \subset \mathbb{C} \), which is continuous on the closure \( \overline{\Omega} \) of \( \Omega \). Then for every \( a \in \Omega \) we have

\[
f^{(n)}(a) = \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z-a)^{n+1}} \, dz,
\]

in which the contour integral is taken counter-clockwise.

This theorem plays an important role in the following representation of a holomorphic function with singularities.

**Theorem 6.2.6** (Laurent Series Expansion). Let \( \Omega \subset \mathbb{C} \) be the annulus given by \( \Omega = \{ z \in \mathbb{C} \mid \rho_1 < |z-a| < \rho_2 \} \), with \( \rho_1, \rho_2 \in \mathbb{R}^+ \cup \{0, \infty\} \), and \( f : \Omega \to \mathbb{C} \) a holomorphic function. Then inside \( \Omega \), \( f \) admits a series expansion of the form

\[
f(z) = \sum_{n \in \mathbb{Z}} a_n (z-a)^n.
\]

Moreover, the coefficients are given by

\[
a_n = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z-a)^{n+1}} \, dz.
\]

recall Theorem 6.2.5.

Note that an infinite annulus, \( \Omega = \mathbb{C} \setminus \{ a \} \), is admitted. This theorem allows us to classify the singularity \( a \), and assign an “order” to it if it happens to be a pole:

- If \( a_k = 0 \) for all \( k < -m < 0 \), and \( a_{-m} \neq 0 \), then \( a \) is a pole of order \( m \in \mathbb{N} \).
- If for all \( m \in \mathbb{N} \) there exists a nontrivial \( a_k \neq 0 \) with \( k < -m \), then \( a \) is an essential singularity.

A pole of order \( m = 1 \) is called a simple pole.

**Definition 6.2.4.** Recall Theorem 6.2.6. The coefficient \( a_{-1} \in \mathbb{C} \) is called the residue of \( f \) at \( a \in \mathbb{C} \), notation: \( a_{-1} = \text{Res}(f; a) \).

**Theorem 6.2.7** (Residue Theorem). Let \( \Omega \subset \mathbb{C} \) be a simply connected open subset, and \( f : \Omega \setminus \{ z_1, \ldots, z_n \} \to \mathbb{C} \), in which all \( z_i \in \Omega \), \( i = 1, \ldots, n \), are distinct. Then

\[
\oint_{\gamma} f(z) \, dz = 2\pi i \sum_{i=1}^{n} \text{Res}(f; z_i) \text{Ind}_\gamma(z_i),
\]

in which \( \text{Ind}_\gamma(\zeta) \) is the winding number of the closed path \( \gamma \) around \( \zeta \in \mathbb{C} \setminus \gamma \):

\[
\text{Ind}_\gamma(\zeta) = \frac{1}{2\pi i} \oint_{\gamma} \frac{dz}{z-\zeta}.
\]
The winding number gives the number of times the curve $\gamma$ turns around the point $\zeta$, with a counter-clockwise turn counting as positive, and a clockwise turn as negative. This is best illustrated with a picture, cf. Fig. 6.1.

Figure 6.1: Illustration of the winding number. Source: Wolfram MathWorld.
II

PDE Systems: Techniques & Examples
7
THE FOURIER METHOD

7.1. Basic Technique

In this chapter we sketch the method of Fourier transformation to solve certain linear p.d.e. systems with constant coefficients. A tacit assumption made throughout is the existence of a well-defined Fourier representation of the solution \( u \) of a p.d.e. system, recall Eqs. (1.1–1.2) and the discussion in Sections 5.2–5.4. Furthermore, the domain of definition is either \( n \)-dimensional space, \( \Omega = \mathbb{R}^n \), or \((n+1)\)-dimensional spacetime, \( \Omega = I \times \mathbb{R}^n \), the latter in the case of a distinguished evolution coordinate \( t \in I \subset \mathbb{R} \) alongside the spatial coordinates \( x \in \mathbb{R}^n \). Finally, the coefficients \( c_\alpha \) will be assumed to be constants, and the b.c.’s trivial in the sense that the solution vanishes towards infinity in any direction, or is globally bounded.

Taking \( \Omega = \mathbb{R}^n \) and applying Theorem 5.2.1 to Eqs. (1.1–1.2) yields

\[
\sum_{|\alpha| \leq N} c_\alpha (i\omega)^\alpha \hat{u}(\omega) = \hat{f}(\omega) \quad \text{on } \mathbb{R}^n.
\]  

(7.1)

in which \( \hat{f} : \mathbb{R}^n \to \mathbb{K} \) is the Fourier transform of \( f : \mathbb{R}^n \to \mathbb{K} \). This is a simple algebraic equation, so that we obtain the explicit solution to our p.d.e. in Fourier space:

\[
\hat{u}(\omega) = \frac{\hat{f}(\omega)}{\sum_{|\alpha| \leq N} c_\alpha (i\omega)^\alpha} \quad \text{on } \mathbb{R}^n.
\]  

(7.2)
Fourier inversion then yields the solution in spatial representation,

\[ u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\omega \cdot x} \frac{\hat{f}(\omega)}{\sum_{|\alpha| \leq N} c_\alpha (i\omega)^\alpha} d\omega. \quad (7.3) \]

Explicit evaluation of this integral tends to be the hardest part (“there ain’t no such thing as a free lunch”).

If \( I = \mathbb{R} \) one may treat \( t \) essentially as a coordinate, so that Eqs. (7.1–7.3) apply with \( n \) replaced by \( n + 1 \). Often, however, the evolution coordinate \( t \in I \) is treated as a parameter, and Fourier transformation is applied only in the spatial domain. The result is an equation which is algebraic w.r.t. the spatial frequency parameters, but still containing \( t \)-derivatives, in other words, an o.d.e.

### 7.2. Examples

**Example 7.2.1.** Solve

\[
\begin{aligned}
\frac{\partial u}{\partial t} - \Delta u &= 0 & \text{for } (x, t) \in \mathbb{R}^n \times \mathbb{R}^+ \\
u(x, 0) &= f(x)
\end{aligned}
\]

in which \( f \in \mathcal{S}'(\mathbb{R}^n) \).

Solution: Substitute

\[ u(x, t) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\omega \cdot x} \hat{u}(\omega, t) d\omega. \quad (7.4) \]

Note that we treat \( t \) as a parameter. As a result we obtain

\[
\begin{aligned}
\frac{d\hat{u}(\omega, t)}{dt} + ||\omega||^2 \hat{u}(\omega, t) &= 0, \\
\hat{u}(\omega, 0) &= \hat{f}(\omega).
\end{aligned}
\]

This o.d.e. system is readily solved:

\[ \hat{u}(\omega, t) = \hat{f}(\omega) \hat{\phi}_t(\omega), \]

with

\[ \hat{\phi}_t(\omega) \overset{\text{def}}{=} e^{-t||\omega||^2}. \]

The final step is Fourier inversion. Things simplify a bit in view of the Fourier theorems of Section 5.5. For suppose, for the moment, that \( \hat{f}(\omega) = 1 \) for all \( \omega \in \mathbb{R}^n \). In that case, Eq. (7.4) yields

\[ u(x, t) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\omega \cdot x - t||\omega||^2} d\omega \overset{\text{def}}{=} \phi_t(x), \]

in which

\[ (\ast) \quad \phi(x, t) \overset{\text{def}}{=} \phi_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x|^2}{4t}}. \]

The details are left as an exercise. Hint: Rewrite

\[ i\omega \cdot x - t||\omega||^2 = -\left(\sqrt{7} - \frac{1}{2} t \right) \left(\frac{x}{\sqrt{t}}\right)^2 - \frac{||x||^2}{4t}. \]

and use the following theorem, a result from complex analysis:

**Lemma 7.2.1.** For all \( \xi \in \mathbb{R}^n \) we have

\[ \int_{\mathbb{R}^n} e^{-||x + i\xi||^2} d\omega = \sqrt{\pi^n}. \]
Finally, use Theorem 5.5.2 or Result 5.5.2, to obtain the solution for arbitrary data function \( f \in \mathcal{S}'(\mathbb{R}^n) \):

\[
    u(x,t) = (f \ast \phi_t)(x).
\]

In mathematics the function \( \phi_t \) is called the fundamental solution (i.c. of the heat equation). In linear systems theory \( \phi_t \) is usually referred to as the point spread function, and its Fourier counterpart \( \hat{\phi}_t \) as the transfer function.

**Example 7.2.2.** Cf. Example 13.2.2 in Chapter 13 for a comment.

**Example 7.2.3.** Consider the following inhomogeneous system with homogeneous i.c.

\[
\begin{cases}
    \frac{\partial u}{\partial t} - \Delta u &= h \quad \text{for } (x,t) \in \mathbb{R}^n \times \mathbb{R}^+ \\
    u(x,0) &= 0
\end{cases}
\]

in which \( h \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}) \) s.t. \( h_t(x) \overset{\text{def}}{=} h(x,t) = 0 \) for \( t < 0 \). Spatial Fourier transformation according to Eq. (7.4) yields

\[
\begin{cases}
    \frac{d\hat{u}(\omega,t)}{dt} + \|\omega\|^2 \hat{u}(\omega,t) &= \hat{h}_t(\omega) \\
    \hat{u}(\omega,0) &= 0
\end{cases}
\]

in which

\[
\hat{h}_t(\omega) = \int_{\mathbb{R}^n} e^{-i\omega \cdot x} h(x,t) \, dx.
\]

The homogeneous o.d.e. has been solved in Example 7.2.1. Let us write this solution as

\[
\hat{u}_0(\omega,t) = A(\omega) \hat{\phi}_t(\omega),
\]

with \( \hat{\phi}_t(\omega) \) as defined in Example 7.2.1. Stipulate a particular inhomogeneous solution of the form

\[
\hat{u}_p(\omega,t) = A(\omega,t) \hat{\phi}_t(\omega).
\]

Substitution into the inhomogeneous o.d.e. yields

\[
\frac{dA(\omega,t)}{dt} \hat{\phi}_t(\omega) = \hat{h}_t(\omega),
\]

from which we obtain a solution for \( A(\omega,t) \) in the form (note that \( 1/\hat{\phi}_t(\omega) = \hat{\phi}_{-t}(\omega) \))

\[
A(\omega,t) = \int_0^t \hat{h}_\tau(\omega) \hat{\phi}_{-\tau}(\omega) \, d\tau.
\]

With this choice, \( \hat{u}_p(\omega,t) \) fulfils the homogeneous i.c., thus

\[
\hat{u}(\omega,t) = \int_0^t \hat{h}_\tau(\omega) \hat{\phi}_{-\tau}(\omega) \, d\tau.
\]

Fourier inversion yields (using definition \((\ast)\) in Example 7.2.1)

\[
u(x,t) = \int_0^t (h_\tau \ast \phi_{1-\tau})(x) \, d\tau = \int_0^t \int_{\mathbb{R}^n} h(\xi,\tau) \phi(x-\xi,t-\tau) \, d\xi \, d\tau = \int_{\mathbb{R}^n+1} h(\xi,\tau) \theta(t-\tau) \phi(x-\xi,t-\tau) \, d\xi \, d\tau,
\]

in which \( \ast \) denotes the spatial convolution operator, recall Theorem 5.5.2, or equivalently, Result 5.5.2, and in which \( \theta \) is the indicator function on \( \mathbb{R}^+ \), i.e. \( \theta(t) = 0 \) for \( t < 0 \) and \( \theta(t) = 1 \) for \( t > 0 \). Note that the solution takes the form of a spatiotemporal convolution product, \( u(x,t) = (h \ast \psi)(x,t) \), cf. Section 10.2.3 for an interpretation in the context of fundamental solutions and Green’s functions.
8

THE METHOD OF CHARACTERISTICS

8.1. Basic Technique

The method of characteristics is generally applicable to first order p.d.e.’s. These take the following form:

\[ \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i} + cu = f, \quad \text{on } \Omega \subset \mathbb{R}^n. \tag{8.1} \]

in which \( a_i, c, f : \Omega \to \mathbb{K} \) are sufficiently smooth functions. Typically the b.c. is prescribed on an \((n - 1)\)-dimensional submanifold \( V \subset \Omega \) (a special case is \( V = \partial \Omega \)):

\[ u = \psi \text{ on } V, \tag{8.2} \]

with \( \psi : V \to \mathbb{K} \) sufficiently smooth.

To appreciate the method of characteristics, notice that the function \( a : \mathbb{R}^n \to \mathbb{K}^n \) defines a vector field, and that

\[ \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} = a \cdot \nabla \tag{8.3} \]

represents directional derivation along the vector field \( a \). For this reason it is natural to study the solution of Eqs. (8.1–8.2) along the integral curves of this vector field.

Definition 8.1.1. The characteristic \( \ell \) through \( x_0 \in \Omega \) is the integral curve \( x = x(t; x_0) \) uniquely determined by the o.d.e.

\[ \ell : \dot{x} = a(x) \quad \text{for } t \in I(x_0). \]
with initial condition \( x(0, x_0) = x_0 \). The existence interval of the solution is indicated here by \( I(x_0) \).

Below we are particularly interested in characteristics through points \( x_0 \in V \). If \( \ell : x = x(t) \) is any characteristic, and \( u_\ell(t) = u(x(t)), c_\ell(t) = c(x(t)), f_\ell(t) = f(x(t)) \), then

\[
\dot{u}_\ell + c_\ell u_\ell = f_\ell. \tag{8.4}
\]

If \( V \) is transversal to the characteristics, and \( \Omega_0 \subset \Omega \) is a region of interest such that each point lies on exactly one characteristic that intersects \( V \) at \( x(0) = x_0 \), say, then the initial condition for Eq. (8.4) becomes

\[
u_\ell(0) = \psi(x_0). \tag{8.5}\]

Eqs. (8.4–8.5) admit a unique solution.

In the case of constant \( a \in \mathbb{K}^n \), the characteristics form a family of parallel lines parametrised by \( x_0 \in \mathbb{R}^n \):

\[
\ell : x(t, x_0) = x_0 + at. \tag{8.6}
\]

The transversality assumption, \( n_V \cdot a \neq 0 \) at \( x_0 \in V \), implies that any neighbouring characteristic sufficiently close to \( \ell \) intersects \( V \) in a unique point close to \( x_0 \in V \). Let us denote the neighbourhood of \( V \) within the p.d.e.'s domain of definition \( \Omega \) on which the transversality condition holds by \( \Omega_0 \subset \Omega \), cf. Fig. 8.1. The heuristic principle of the method of characteristics is that any point \( x \in \Omega_0 \) is uniquely determined by a parameter pair \( (x_0, t) \), in which \( x_0 \) singles out a unique characteristic and \( t \) a unique point on that characteristic, viz. \( x = x(t, x_0) \). Therefore this equation can be inverted so as to express the parameters into the original coordinates of the problem, \( (t, x_0) = (t(x), x_0(x)) \) say. In this way the solution along each of the covering characteristics implies the solution of the p.d.e. problem, \( u(x) = u_\ell(t(x), x_0(x)) \), at least on \( \Omega_0 \subset \Omega \).

![Figure 8.1: The region \( \Omega_0 \subset \Omega \) is a union of non-intersecting characteristics \( \ell : x(t, x_0) = x_0 + at \), each uniquely defined by a transverse intersection \( x_0 \in V \). In other words, \( \forall x \in \Omega_0 \ \exists x_0 \in V, t \in I(x_0) \) s.t. \( x = x(t, x_0) \).](image)

For the sake of simplicity let us assume that also the coefficient \( c \) is a constant, then the solution of Eqs. (8.4–8.5) may be obtained by the method of “variation of constant”, as follows:

- Note that \( u_{\ell,0}(t) = A e^{-ct} \) solves the homogeneous o.d.e. \( \dot{u}_\ell + cu_\ell = 0 \), in which \( A \in \mathbb{K} \) is a constant.

- Replace \( A \) by \( A(t) \) and insert \( u_\ell(t) = A(t) e^{-ct} \) into the inhomogeneous o.d.e. \( \dot{u}_\ell + cu_\ell = f_\ell \) to find \( A(t) \).

- Disambiguate the solution by accounting for the initial condition.
The result is
\[ u_\ell(t, x_0) = \psi(x_0) e^{-ct} + \int_0^t f(\tau) e^{-c(t-\tau)} \, d\tau. \] (8.7)

Thus \( u_\ell \) is a function of the \( n \) independent degrees of freedom \((x_0, t) \in \Omega_0\), which can be re-expressed in terms of the \( n \) original coordinates \( x \in \Omega_0 \) to obtain the explicit solution \( u(x) \) on \( \Omega_0 \).

### 8.2. Examples

**Example 8.2.1.** Consider the following system:

\[
\begin{aligned}
\frac{d}{dt} u_\ell &= 0, \\
u_\ell(t, x_0) &= \psi(x_0) \quad \text{on } \Omega : y > 0,
\end{aligned}
\]

Consider the parametrised curve \( \ell : (x, y) = (\xi(t), \eta(t)) \). Denote the solution along this curve by \( u_\ell = u \circ \ell \). Take the curve \( \ell \) to be a characteristic:

\[
\ell : \begin{cases} 
\dot{\xi} = 1, \\
\dot{\eta} = -1.
\end{cases}
\]

In explicit form: \( \ell : (x, y) = (\xi_0 + t, \eta_0 - t) \). If we leave \((\xi_0, \eta_0)\) unconstrained, then the intersection points \( \ell \cap \partial \Omega \) are obtained at \( t_{\partial \Omega} = \eta_0 \), with corresponding points \((\xi_0 + \eta_0, 0) \in \partial \Omega\). Alternatively we may require \( t_{\partial \Omega} = 0 \) by assuming \( \eta_0 = 0 \), with corresponding intersection points \((\xi_0, 0) \in \partial \Omega\). We will adopt the latter choice henceforth:

\[ \ell : (x, y) = (\xi_0 + t, -t). \]

It is clear from its orientation that every characteristic \( \ell \) intersects the boundary \( \partial \Omega \) transversally.

The p.d.e. system (*) for \( u \) reduces to a simple o.d.e. system for \( u_\ell \), viz.

\[
\begin{aligned}
\frac{d}{dt} u_\ell &= 0, \\
u_\ell(0) &= \psi(\xi_0).
\end{aligned}
\]

The solution is the constant function

\[ u_\ell(t) = \psi(\xi_0). \]

In order to obtain the desired solution of (*) we must relate the parameters \((t, \xi_0)\) to \((x, y)\) using (square). Substituting \((x, y)\) for \((\xi_0 + t, -t)\) in \( u_\ell(t) = u(\xi_0 + t, -t) = \psi(\xi_0) \), using \( \xi_0 = x + y \), yields

\[ u(x, y) = \psi(x + y). \]

**Example 8.2.2.** Let us consider the special case of Eq. (8.1) with variable \( c = c(x) \), keeping all other coefficients constant. Then the same line of reasoning as in the previous section applies with respect to the characteristics. Also the procedure that led to Eq. (8.7) still holds as such, but produces a more general result, viz.

\[ u_\ell(t, x_0) = \psi(x_0) e^{-\int_0^t c_\ell(s) \, ds} + \int_0^t f_\ell(\tau) e^{-\int_\tau^t c_\ell(s) \, ds} \, d\tau. \] (8.8)

For constant \( c \) this is clearly consistent with Eq. (8.7).

The next example illustrates the case of a non-constant vector field \( a = a(x) \).

**Example 8.2.3.** Consider the following system:

\[
\begin{aligned}
 xu_x + yu_y + u &= 0 \quad \text{on } \Omega = \mathbb{R}^2 \setminus \{0, 0\}, \\
u(\cos \theta, \sin \theta) &= \psi(\theta) \quad \text{on } V : (x, y) = (\cos \theta, \sin \theta) \text{ with } \theta \in [0, 2\pi).
\end{aligned}
\]
Consider the parametrised curve \( \ell : (x, y) = (\xi(t), \eta(t)) \). Denote the solution along this curve by \( u_\ell = u \circ \ell \).

Take the curve \( \ell \) to be a characteristic:

\[
\ell : \begin{cases}
\dot{\xi} &= \xi, \\
\dot{\eta} &= \eta.
\end{cases}
\]

In explicit form: \( \ell : (x, y) = (\xi_0 e^t, \eta_0 e^t) \), i.e. a family of half lines emanating from the origin \((t \to -\infty)\) in radial direction through the point \((\xi_0, \eta_0)\). It is clear from its orientation that every characteristic \( \ell \) intersects the unit sphere transversally. The value \( t = 0 \) corresponds to intersection points \((\xi_0, \eta_0) \in V \) provided we assume \( \xi_0^2 + \eta_0^2 = 1 \), which we will do henceforth:

\( \square \quad \ell : (x, y) = (\xi_0 e^t, \eta_0 e^t) \quad \text{subject to} \quad \xi_0^2 + \eta_0^2 = 1. \)

The p.d.e. system for \( u \) reduces to the following o.d.e. system for \( u_\ell \):

\[
\begin{align*}
\dot{u}_\ell + u_\ell &= 0, \\
u_\ell(0) &= \psi(\text{arg}(\xi_0 + i\eta_0)).
\end{align*}
\]

The solution is

\[
u_\ell(t) = \psi(\text{arg}(\xi_0 + i\eta_0)) e^{-t}.
\]

To obtain the desired solution of \( \text{(**)} \) we must eliminate the parameters \((t, \xi_0, \eta_0)\)—two of which are independent—using \( \square \) in favour of \((x, y)\). Note that

- \( \text{arg}(\xi_0 + i\eta_0) = \text{arg}((\xi_0 + i\eta_0) e^t) = \text{arg}(\xi(t) + i\eta(t)) = \text{arg}(x + iy) \), and
- \( e^t = \sqrt{\xi_0^2 + \eta_0^2} e^{it} = \sqrt{\xi(t)^2 + \eta(t)^2} = \sqrt{x^2 + y^2}. \)

Therefore

\[
u(x, y) = \frac{\psi(\text{arg}(x + iy))}{\sqrt{x^2 + y^2}}.
\]

**Example 8.2.4.** There is a more direct way to reduce the p.d.e. system of the previous example to an o.d.e. system, obviating the intermediate step via characteristics. It is based on a slick choice of coordinates adapted to the symmetry of the problem. A close inspection reveals that \( \text{(**)} \) is symmetric under dilations (i.e. the form of the system does not change upon a transformation \((x', y') = (\lambda x, \lambda y) \) with \( \lambda > 0 \)), suggesting a polar coordinate transformation (note that a dilation amounts to a 1-dimensional radial scaling):

\[
\begin{cases}
x(r, \theta) &= r \cos \theta, \\
y(r, \theta) &= r \sin \theta,
\end{cases}
\]

with \((r, \theta) \in \mathbb{R}^+ \times [0, 2\pi)\). Setting \( v(r, \theta) = u(x(r, \theta), y(r, \theta)) \), and observing that

\[
\begin{align*}
\frac{\partial}{\partial x} &= \cos \theta \frac{\partial}{\partial r} - \sin \theta \frac{\partial}{\partial \theta}, \\
\frac{\partial}{\partial y} &= \sin \theta \frac{\partial}{\partial r} + \cos \theta \frac{\partial}{\partial \theta},
\end{align*}
\]

yields the following equivalent system, in which a ’ indicates \( d/dr \):

\[
\begin{cases}
rv' + v &= 0 \quad \text{on} \ \Omega = \mathbb{R}^2 \setminus \{0, 0\}, \\
v(1, \theta) &= \psi(\theta) \quad \text{with} \ \theta \in [0, 2\pi). \end{cases}
\]

This is an o.d.e. system in the variable \( r \), in which \( \theta \) enters as a parameter. The solution is

\[
v(r, \theta) = \frac{\psi(\theta)}{r}.
\]

Compare this to the result from the previous example.
9

SEPARATION OF VARIABLES

9.1. Basic Technique

In this chapter we sketch another generic method to solve p.d.e. systems, recall Eqs. (1.1-1.2), based on separation of variables. Certain tacit assumptions will be made concerning the general conditions under which this method is applicable, but again we will confine ourselves to specific examples to illustrate the main idea. An explicit assumption will be that the domain of definition \( \Omega_k \subset \mathbb{R}^{n-k} \) for a subset of independent variables \( x_1, \ldots, x_{n-k} \) \( (0 < k < n) \) is open and bounded, with boundary \( \partial \Omega_k \).

The method stipulates the existence of solutions of the form

\[
\begin{align*}
    u(x_1, \ldots, x_n) &= u_1(x_1, \ldots, x_{n-k}) u_2(x_{n-k+1}, \ldots, x_n),
\end{align*}
\]

with \( 0 < k < n \). In many cases of interest \( k = 1 \), e.g. in dynamic systems this ansatz is typically used to separate the system’s behaviour in space (with coordinates \( x_1, \ldots, x_{n-1} \), say) and time (\( x_n = t \)).

Applying this technique may result in a so-called Sturm-Liouville problem.

**Definition 9.1.1.** A (regular\(^1\)) Sturm-Liouville problem is an eigensystem expressed in terms of a second order o.d.e. with homogeneous b.c.’s of the following form:

\[
\begin{align*}
    \mathcal{L}u &= \lambda u \quad \text{on } (a, b) \subset \mathbb{R}, \\
    \alpha u(a) + \beta u'(a) &= 0 \quad \text{with } \alpha, \beta \in \mathbb{R}, \alpha^2 + \beta^2 \neq 0 \\
    \gamma u(b) + \delta u'(b) &= 0 \quad \text{with } \gamma, \delta \in \mathbb{R}, \gamma^2 + \delta^2 \neq 0
\end{align*}
\]

\(^1\)In a singular Sturm-Liouville problem we have \( a = -\infty \) and/or \( b = \infty \), or the coefficient of \( u'' \) diverges at \( a \) and/or \( b \).
in which the linear operator \( L \) has the form

\[
(\ast) \quad L u = -\frac{1}{w} \left( (pu')' + qu \right),
\]

for certain specified real-valued functions \( p, q \), and positive “weight” function \( w > 0 \).

The Sturm-Liouville problem consists of finding those values of \( \lambda \in \mathbb{R} \) for which nontrivial solutions exist. Under suitable conditions on the functions \( p, q \) and \( w \) (in particular \( p, p', q, w \in C([a, b]) \)), Sturm-Liouville theory provides us with the following result.

**Theorem 9.1.1.** Recall Definition 9.1.1. The eigenvalues \( \lambda_n, n \in \mathbb{N} \), are real and can be ordered s.t.

\[
\lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots \quad \text{with} \quad \lim_{n \to \infty} \lambda_n = \infty.
\]

The eigenfunction \( \phi_n \), corresponding to \( \lambda_n \), a.k.a. the \( n \)-th fundamental solution of the Sturm-Liouville problem, is unique up to normalization, and has exactly \( n-1 \) zero crossings in the interval \((a, b)\). Moreover, the (appropriately normalized) eigenfunctions form an orthonormal basis for the Hilbert space \( L^2([a, b], w) \), in the sense that

\[
(\phi_m, \phi_n)_w \overset{\text{def}}{=} \int_a^b \phi_m^*(x) \phi_n(x) w(x) \, dx = \delta_{mn}.
\]

**Remark.** An important property of the Sturm-Liouville operator in Definition 9.1.1 is

\[
(L u, v)_w = (u, L^* v)_w.
\]

Among others, this implies \( \lambda \in \mathbb{R} \) as well as orthogonality of the eigenfunctions \( \{\phi_n\}_{n \in \mathbb{N}} \). The proof is left as an exercise.

### 9.2. Examples

**Example 9.2.1.** Consider the following system, describing a vibrating string without a priori given initialization:

\[
(\ast) \begin{cases}
    u_{tt} - u_{xx} = 0 & \text{for } (x, t) \in (0, 1) \times \mathbb{R}^+,
    \\
    u(0, t) = u(1, t) = 0.
\end{cases}
\]

Suppose

\[
(\hat{\ast}) \quad u(x, t) = \phi(x) \chi(t).
\]

For the p.d.e. this yields \( \ddot{\chi} + c^2 \chi = 0 \), with dots and primes indicating temporal, respectively spatial differentiation. The b.c. reduces to \( \phi(0) = \phi(1) = 0 \), assuming a nontrivial solution. Ignoring denominator problems that might occur, this implies that \( \ddot{\chi} / \chi = \phi'' / \phi = -c^2 \) constant, since l.h.s. and r.h.s. are functions of different independent variables \( t \) and \( x \). Note that the arbitrary constant \(-c^2\) (the notation betrays a modest amount of foresight) may be non-positive (in which case take \( c \geq 0 \)) or non-negative (in which case take \( c = i\kappa \) with \( \kappa \geq 0 \) and \( i^2 = -1 \)). Setting

\[
\ddot{\chi} + c^2 \chi = 0,
\]

we obtain

\[
(\ast) \quad \chi(t) = S_c \sin(ct) + C_c \cos(ct).
\]

Here, \( S_c \) and \( C_c \) are constants that depend on the parameter \( c \), which, at this stage, is left undetermined. Note that \( \sin(i\kappa t) = i \sinh(\kappa t) \) and \( \cos(i\kappa t) = \cosh(\kappa t) \).

As a consequence we have the following Sturm-Liouville problem for \( \phi \) (recall Definition 9.1.1 with \( p = 1, q = 0, w = 1, a = 0, b = 1, \lambda = c^2 \)):

\[
\begin{cases}
    \phi'' + c^2 \phi = 0 & \text{for } x \in (0, 1),
    \\
    \phi(0) = \phi(1) = 0.
\end{cases}
\]
The solution of the o.d.e. is formally similar to that of $\chi$,

$$\phi(x) = A_c \sin(cx) + B_c \cos(cx).$$

The thing to note is that the spatial b.c.’s prevent $c$ from taking just any value, at least for a nontrivial solution. To admit a nontrivial solution satisfying the b.c.’s we must apparently set $B_c = 0$ and take a real-valued $c \in \mathbb{R}$ (consistent with Theorem 9.1.1), viz.

$$c = k\pi \quad \text{with } k \in \mathbb{N}.$$ 

Finally, without loss of generality, we may normalize the solution, since the integration constants $s_k = S_{k\pi}$ and $c_k = C_{k\pi}$ in the factor $\chi(t)$ (†) can still be freely chosen, e.g.

$$(\circ) \quad \phi_k(x) = \sqrt{2} \sin(k\pi x) \quad k \in \mathbb{N}.$$ 

The normalization is such that

$$\phi_k \cdot \phi_\ell = \delta_{k\ell}$$

for the standard inner product on $L^2((0, 1))$. All in all

$$(\bullet) \quad u(x, t) = \sum_{k=1}^{\infty} (s_k \sin(k\pi t) + c_k \cos(k\pi t)) \phi_k(x).$$

Whether this formal series represents a classical or a generalised solution depends on the convergence properties of the particular sequences $\{s_k\}_{k \in \mathbb{N}}$ and $\{c_k\}_{k \in \mathbb{N}}$. Note that, generically, it is not separable, unlike the ansatz (†).

In the next example we exploit the freedom of choice of the coefficients to further restrict the system.

**Example 9.2.2.** Consider the p.d.e. system (†) of Example 9.2.1 under the additional initial condition

$$(\text{i.c.)} \quad \left\{ \begin{array}{l} u(x, 0) = f(x), \\
                      u_t(x, 0) = g(x). \end{array} \right. $$

We expect this to induce constraints on the coefficients $s_k$ and $c_k$. Inspection of (•) for $t = 0$ and comparison with the i.c.’s reveals that

$$\left\{ \begin{array}{l} f(x) = \sum_{k=1}^{\infty} c_k \phi_k(x), \\
                      g(x) = \pi \sum_{k=1}^{\infty} k s_k \phi_k(x), \end{array} \right. $$

in terms of the basis functions given by (◦) in the previous example. To retrieve the coefficients, take inner products of $f$ and $g$ with $\phi_\ell$, using orthonormality, $\phi_k \cdot \phi_\ell = \delta_{k\ell}$:

$$\left\{ \begin{array}{l} c_\ell = \int_0^1 f(x) \phi_\ell(x) \, dx = f \cdot \phi_\ell, \\
                      s_\ell = \frac{1}{\pi \ell} \int_0^1 g(x) \phi_\ell(x) \, dx = \frac{1}{\pi \ell} g \cdot \phi_\ell. \end{array} \right. $$

All in all, recalling (◦) and (•) from the previous problem,

$$(\star \star) \quad u(x, t) = \sum_{k=1}^{\infty} \left( \frac{g \cdot \phi_k}{\pi k} \sin(k\pi t) + f \cdot \phi_k \cos(k\pi t) \right) \phi_k(x).$$

In general, the result (\star \star) above represents a generalised solution, thus one should take point evaluation with a grain of salt. To appreciate this, let us write $u(x, t) = \lim_{N \to \infty} U_N(x, t)$ with $U_N(x, t) = \sum_{k=1}^{N} u_k(x, t)$, with self-explanatory definition of $u_k(x, t)$. This limit generally holds in the weak sense of distribution theory, recall Definition 4.2.11:

$$u = \lim_{N \to \infty} U_N \quad \text{iff} \quad u(\phi) = \lim_{N \to \infty} U_N(\phi) \quad \text{for all } \phi \in \mathcal{D}((0, 1) \times \mathbb{R}^+).$$

(9.2)

If $f$ and $g$ are sufficiently regular, stronger convergence results can be obtained that render $u$ a classical solution.
Example 9.2.3. Cf. Examples 12.2.5–12.2.6 in Chapter 12 for more examples in the context of hyperbolic systems.

Example 9.2.4. Cf. Example 13.2.1 in Chapter 13 for an application in the context of parabolic systems.

Example 9.2.5. Consider the Laplace equation with inhomogeneous b.c.’s on the unit square:

\[
\begin{cases}
  u_{xx} + u_{yy} = 0 & \text{for } (x, y) \in (0, 1) \times (0, 1), \\
  u(x, 0) = u(x, 1) = u(0, y) = 0, \\
  u(1, y) = 1.
\end{cases}
\]

Suppose

\[
\begin{cases}
  u(x, y) = X(x) Y(y).
\end{cases}
\]

The result is

\[
\frac{X''}{X} = -\frac{Y''}{Y} = c^2.
\]

In view of the homogeneous b.c.’s on the horizontal edges, consider the system for \(Y\) first. As in Example 9.2.1 we obtain a Sturm-Liouville problem:

\[
\begin{cases}
  Y'' + c^2 Y = 0 & \text{for } y \in (0, 1), \\
  Y(0) = Y(1) = 0
\end{cases}
\]

which implies \(Y(y) = A \sin(cy) + B \cos(cy)\) for some constant \(c \in \mathbb{C}\). Imposing the homogeneous b.c.’s yields \(B = 0\) and (given that \(A \neq 0\)) \(c = k\pi, k \in \mathbb{N}\). Recall that the normalised family \(\{Y_k(y) = \sqrt{2} \sin(k\pi y)\}_{k \in \mathbb{N}}\) is orthonormal relative to the standard inner product on \(L^2([0, 1])\).

Setting \(c = k\pi\), the resulting system for \(X\) is \(X'' - k^2 \pi^2 X = 0\). The single homogeneous b.c. \(X(0) = 0\) implies \(X(x) \propto \sinh(k\pi x)\). All in all we may stipulate a solution of the original problem the form

\[
u(x, y) = \sum_{k=1}^{\infty} c_k \sinh(k\pi x) \sin(k\pi y).
\]

The coefficients should follow by enforcing the final, inhomogeneous b.c., viz. \(u(1, y) = 1\), or

\[
\sum_{k=1}^{\infty} c_k \sinh(k\pi) \sin(k\pi y) = 1.
\]

To isolate the coefficients, take the standard inner product with \(\phi_\ell(y) = \sqrt{2} \sin(\ell\pi y)\), using \(\phi_k \cdot \phi_\ell = \delta_{k\ell}\). For the l.h.s. we then obtain \(c_\ell \sinh(\ell\pi)/\sqrt{2}\), for the r.h.s. we find

\[
1 \cdot \phi_\ell = \int_0^1 \phi_\ell(y) dy = \left[ -\frac{\sqrt{2}}{\ell\pi} \cos(\ell\pi y) \right]_0^1 = \begin{cases} 0 & \text{if } \ell \text{ is even}, \\
\frac{2\sqrt{2}}{\ell\pi} & \text{if } \ell \text{ is odd}.
\end{cases}
\]

All in all

\[
u(x, y) = \sum_{k=1, k \text{ odd}}^{\infty} \frac{4}{k\pi \sinh(k\pi)} \sinh(k\pi x) \sin(k\pi y) = \sum_{k=1}^{\infty} \frac{4}{(2k-1)\pi \sinh((2k-1)\pi)} \sinh((2k-1)\pi x) \sin((2k-1)\pi y).
\]
10.1. Basic Technique

Solving the inhomogeneous linear p.d.e. $Lu = f$ on $\Omega$ subject to $Bu = g$ on $\partial \Omega$ defined by Eqs. (1.1–1.4) formally amounts to finding a left-inverse $L^{-1}$ of $L$ incorporating the b.c.’s. This whimsical observation can be made more rigorous using the concept of a Green’s function, a.k.a. fundamental solution or, in systems engineering, impulse response.

Throughout this chapter we will assume that the coefficients $c_{\alpha}(x) = c_{\alpha}$ in Eqs. (1.1–1.2) are constant, although this is not a necessary condition per se for the proposed method.
The concept of a fundamental solution relies on the fact that the domain of definition of the Dirac point measure, recall Lemma 4.2.1 and Definition 4.2.6, can be extended so as to include all functions admitting point evaluation. With Eqs. (1.1–1.2) in mind, suppose we are in possession of a (generalised!) symmetric function \( \psi \in \mathcal{D}'(\Omega) \), s.t.
\[
L \psi = \delta \quad \text{on } \Omega. 
\]
(10.1)

Let us define the shifted function \( \psi_{\xi} \) in analogy with the shifted Dirac function \( \delta_{\xi} \) for \( \xi \in \Omega \), Definition 4.2.6:
\[
\psi_{\xi}(x) = \psi(x - \xi). 
\]
(10.2)

Symmetry entails \( \psi(x) = \psi(-x) \) for all \( x \in \Omega \), whence \( \psi_{\xi}(x) = \psi_{-\xi}(\xi) \). Note that \( \psi_{\xi} \) is a solution of
\[
L \psi_{\xi} = \delta_{\xi} \quad \text{on } \Omega. 
\]
(10.3)

For this reason \( L \) is said to be linear shift invariant. This implies that
\[
u(x) = \int_{\Omega} \psi_{\xi}(x) f(\xi) \, d\xi 
\]
(10.4)
satisfies \( Lu = f \). Note that in this shift invariant case (linear p.d.o. with constant coefficients) we may write this as a convolution by virtue of Eq. (10.2):
\[
u(x) = (\psi *_{\Omega} f)(x). 
\]
(10.5)

**Proof.** We have \( Lu(x) = L \int_{\Omega} \psi_{\xi}(x) f(\xi) \, d\xi = \int_{\Omega} L \psi_{\xi}(x) f(\xi) \, d\xi = \int_{\Omega} \delta_{\xi}(x) f(\xi) \, d\xi = \int_{\Omega} \delta_{\xi}(\xi) f(\xi) \, d\xi = f(x) \).

Alternatively, if we interpret Eq. (10.4) in \( \mathcal{D}'(\Omega) \)-distributional sense, i.e. if \( u(\phi) = \int_{\Omega} \int_{\Omega} \psi_{\xi}(x) f(\xi) \phi(x) \, dx \)
then \( Lu(\phi) = u(L^1 \phi) = \int_{\Omega} \int_{\Omega} \psi_{\xi}(x) f(\xi) \phi(x) \, dx = \int_{\Omega} f(\xi) \int_{\Omega} \psi_{\xi}(x) L^1 \phi(x) \, dx \, d\xi = \int_{\Omega} f(\xi) L \psi_{\xi}(\phi) \, d\xi = \int_{\Omega} f(\xi) \delta_{\xi}(\phi) \, d\xi = \int_{\Omega} f(\xi) \phi(\xi) \, d\xi = f(\phi) \) for all \( \phi \in \mathcal{D}(\Omega) \), so that \( Lu = f \in \mathcal{D}'(\Omega) \).

Any generalised function \( \psi \in \mathcal{D}'(\Omega) \) that obeys Eq. (10.1) is called a fundamental solution of the p.d.e. Via Eq. (10.2) a suitably chosen fundamental solution gives rise to a so-called Green’s function, a symmetric bivariate function \( \Psi : \Omega \times \Omega \rightarrow \mathbb{R} : (x, \xi) \mapsto \Psi(x, \xi) = \psi_{\xi}(x) \), adapted to certain b.c.’s. Obviously \( \psi \) is non-differentiable at the origin, which implies that \( \Psi \) is non-differentiable at ‘diagonal’ points \( (x, x) \in \Omega \times \Omega \). Fundamental solutions, and a fortiori Green’s functions, are not unique. One may add classical solutions of the homogeneous p.d.e., which may be exploited to enforce specific boundary conditions.

### 10.2. Construction

#### 10.2.1. Via Classical Homogeneous Solutions

To construct a Green’s function a useful observation is that outside the singularity it must have the same form as a classical homogeneous solution of the p.d.e. As an illustration of this principle we consider the construction of a Green’s function for the Laplace/Poisson equation.

**Theorem 10.2.1.** A Green’s function \( \Psi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} : (x, \xi) \mapsto \Psi(x, \xi) \) for the Laplacian \( \Delta \) is given by
\[
\Psi(x, \xi) = \begin{cases} 
\frac{1}{2} |x - \xi| & \text{if } n = 1, \\
\frac{1}{2\pi} \ln \|x - \xi\| & \text{if } n = 2, \\
\frac{-1}{(n - 2) \text{vol}_{n-1} \mathbb{S}^{n-1} \|x - \xi\|^{n-2}} & \text{if } n = 3,
\end{cases}
\]
in which \( \text{vol}_{n-1} \mathbb{S}^{n-1} \) denotes the surface area of the unit sphere \( \mathbb{S}^{n-1} \subset \mathbb{R}^n \) in \( n \) dimensions. That is, if \( \psi_{\xi}(x) = \Psi(x, \xi) \), then \( \Delta \psi_{\xi} = \delta_{\xi} \).

58
Note that the case \( n = 1 \) has already been handled in Example 4.4.3.

**Proof.** The proof for general \( n \in \mathbb{N} \) requires a lemma.

**Lemma 10.2.1.** If \( u : \mathbb{R}^n \to \mathbb{R} : x \mapsto u(x) \) is rotationally symmetric w.r.t. the origin, then \( \Delta u = \frac{1}{r^{n-1}} \partial_r \left( r^{n-1} \partial_r u \right) \), in terms of the radial coordinate \( r = \| x \| = \sqrt{x_1^2 + \ldots + x_n^2} \).

**Proof of Theorem 10.2.1.** With the lemma it is straightforward to prove that, for fixed \( \xi \in \mathbb{R}^n \), \( \psi \) as conjectured in Theorem 10.2.1 satisfies \( \Delta \psi = 0 \) for \( \| x - \xi \| > 0 \) (set \( r = \| x - \xi \| \)). The point \( x = \xi \) needs special care. To this end, consider the region \( \Omega_\epsilon = \{ x \in \mathbb{R}^n \mid \epsilon < \| x - \xi \| < R \} \subset \Omega \) with \( R > \epsilon > 0 \) and \( \Omega = \{ x \in \mathbb{R}^n \mid \| x - \xi \| < R \} \).

Then for any test function \( \phi \in \mathcal{D}(\Omega) \) we have, by Stokes’ Theorem,

\[
\int_{\Omega_{\epsilon}} (\psi \Delta \phi - \phi \Delta \psi) \, dx = \int_{\partial \Omega_{\epsilon}} (\psi \nabla \phi - \phi \nabla \psi) \cdot \quad \hat{n} \, d\sigma.
\]

We have just argued that II = 0 identically on \( \Omega_\epsilon \). For III and IV only the innermost boundary with \( \| x - \xi \| = \epsilon \) yields potential contributions by definition of \( \mathcal{D}(\Omega) \). Let us consider them one by one.

For term III we observe that \( \| \psi \| \nabla \phi \cdot \hat{n} \leq C(\epsilon) \max \| \nabla \phi \| \), with constant \( C(\epsilon) = \| \psi(x) \|_{x - \xi = \epsilon} = O(\epsilon^{2-n}) \) if \( n \neq 2 \), and \( O(\ln \epsilon) \) if \( n = 2 \), and \( \int_{\| x - \xi \| = \epsilon} d\sigma = c^{n-1} \text{vol}_{n-1} \), so that

\[
\| III \| = \int_{\| x - \xi \| = \epsilon} \| \psi \| \nabla \phi \cdot \hat{n} \\ d\sigma \leq C(\epsilon) \max \| \nabla \phi \| \int_{\| x - \xi \| = \epsilon} d\sigma = \begin{cases} O(\epsilon) & \text{if } n \neq 2, \\ O(\ln \epsilon) & \text{if } n = 2. \end{cases}
\]

That is, \( \lim_{\epsilon \to 0} III = 0 \).

As for term IV, we explicitly compute \( \nabla \psi \cdot \hat{n} = \frac{\partial \psi}{\partial r} \bigg|_{r=\epsilon} = -\frac{\partial \psi}{\partial r} \bigg|_{r=\epsilon} \) on the \( \epsilon \)-sphere \( \| x - \xi \| = \epsilon \) in terms of the radial coordinate \( r = \| x - \xi \| : \)

\[
\frac{\partial \psi}{\partial r} \bigg|_{r=\epsilon} = \begin{cases} 1 & \text{if } n = 1, \\ 2\pi \epsilon & \text{if } n = 2, \\ \frac{1}{\text{vol}_{n-1} \epsilon^{n-1}} & \text{if } n \geq 3. \end{cases}
\]

(The latter formula thus in fact covers all cases including \( n = 1, 2 \).) Noting that \( |\phi(x) - \phi(\xi)| \leq \epsilon \max \| \nabla \phi \| \), and writing \( \phi(x) = \phi(\xi) + (\phi(x) - \phi(\xi)) \), we find

\[
IV = -\int_{\| x - \xi \| = \epsilon} \phi \nabla \psi \cdot \hat{n} \\ d\sigma = -\left( \phi(\xi) + O(\epsilon) \right) \int_{\| x - \xi \| = \epsilon} \nabla \psi \cdot \hat{n} \\ d\sigma = \left( \phi(\xi) + O(\epsilon) \right) \int_{\| x - \xi \| = \epsilon} \frac{\partial \psi}{\partial r} \bigg|_{r=\epsilon} \\ d\sigma.
\]

We conclude that \( \lim_{\epsilon \to 0} IV = \phi(\xi) = \delta_\xi(\phi) \), and therefore, combining all of the above observations,

\[
\int_{\Omega} \psi \Delta \phi \\ dx = \lim_{\epsilon \to 0} \int_{\Omega_{\epsilon}} \psi \Delta \phi \\ dx = \delta_\xi(\phi),
\]

i.e. \( \Delta \psi = \delta_\xi \) in distributional sense. This concludes the proof of Theorem 10.2.1.

**10.2.2. Via Separation of Variables**

If, via the method of separation of variables (Chapter 9, Eq. (10.1)), furnished with homogeneous b.c.’s, can be reduced to a Sturm-Liouville problem (recall Definition 9.1.1 and Theorem 9.1.1), then there is a systematic...
way to construct a Green’s function that satisfies the b.c.’s. This construction also works in higher dimensions, but let us consider the Sturm-Liouville problem \((*)\) of Definition 9.1.1 for the sake of simplicity: Let \(\{\phi_n\}_{n \in \mathbb{N}}\) be a complete orthonormal family for the Sturm-Liouville problem in Definition 9.1.1, then we may stipulate a Green’s function of the form

\[
\psi_{\xi}(x) = \Psi(x, \xi) = \sum_{n \in \mathbb{N}} a_n(\xi) \phi_n(x).
\]

Insertion into Eq. (10.3) yields the following result.

**Result 10.2.1.** Recall Definition 9.1.1 for the definition of \(\mathcal{L}\). A Green’s function for the system

\[
\begin{align*}
\mathcal{L}\psi_{\xi} &= \delta_{\xi} \text{ on } (a, b), \\
\alpha \psi_{\xi}(a) + \beta \psi'_{\xi}(a) &= 0 \text{ with } \alpha, \beta \in \mathbb{R}, \alpha^2 + \beta^2 \neq 0, \\
\gamma \psi_{\xi}(b) + \delta \psi'_{\xi}(b) &= 0 \text{ with } \gamma, \delta \in \mathbb{R}, \gamma^2 + \delta^2 \neq 0
\end{align*}
\]

is given by

\[
\psi_{\xi}(x) = \Psi(x, \xi) = \sum_{n \in \mathbb{N}} w(\xi) \frac{\phi^*_n(\xi) \phi_n(x)}{\lambda_n},
\]

in which \(\{\phi_n\}_{n \in \mathbb{N}}\) is an orthonormal basis of \(L^2([a, b], w)\) induced by the Sturm-Liouville problem in Definition 9.1.1, recall Theorem 9.1.1.

**Proof.** Expand \(\psi_{\xi}(x) = \sum_{n \in \mathbb{N}} a_n(\xi) \phi_n(x)\) and substitute this into \(\mathcal{L}\psi_{\xi} = \delta_{\xi}\) to determine the coefficients \(a_n(\xi)\). Using \(\mathcal{L}\phi_n = \lambda_n \phi_n\) we find

\[
\sum_{n \in \mathbb{N}} a_n(\xi) \lambda_n \phi_n(x) \overset{\text{def}}{=} \delta_{\xi}(x).
\]

Multiplying both sides with \(\phi^*_m(x) w(x)\) and integrating over \(x \in (a, b)\) yields for the l.h.s.

\[
\sum_{n \in \mathbb{N}} a_n(\xi) \lambda_n \int_a^b \phi^*_m(x) \phi_n(x) w(x) \, dx = \sum_{n \in \mathbb{N}} a_n(\xi) \lambda_n \langle \phi_m, \phi_n \rangle_w = \lambda_m a_m(\xi),
\]

by virtue of orthonormality \(\langle \phi_n, \phi_m \rangle_w = \delta_{nm}\), and for the r.h.s.

\[
\int_a^b \delta(x - \xi) \phi^*_m(x) w(x) \, dx = w(\xi) \phi^*_m(\xi),
\]

Thus

\[
a_m(\xi) = w(\xi) \frac{\phi^*_m(\xi)}{\lambda_m}.
\]

The extension to genuine p.d.e.’s in \(n\) dimensions leads to the following result.

**Result 10.2.2.** Let \(L\psi_{\xi} = \delta_{\xi}\) on \(\Omega\) with homogeneous b.c.’s on \(\partial\Omega\). Let \(\{\phi_\alpha\}_{\alpha \in \mathbb{N}^n}\) be a complete family of orthonormal eigenfunctions on \(L^2(\Omega, w)\) subject to given b.c.’s, i.e. \(L\phi_\alpha = \lambda_\alpha \phi_\alpha\), with

\[
\langle \phi_\alpha, \phi_\beta \rangle_w \overset{\text{def}}{=} \int_\Omega \phi^*_\alpha(x) \phi_\beta(x) w(x) \, dx = \delta_{\alpha \beta}.
\]

Then

\[
\psi_{\xi}(x) = \Psi(x, \xi) = \sum_{\alpha \in \mathbb{N}^n} w(\xi) \frac{\phi^*_\alpha(\xi) \phi_\alpha(x)}{\lambda_\alpha}.
\]
10.2.3. Via Fourier Transformation

In Example 7.2.3 we have effectively constructed a fundamental solution and a fortiori Green’s function for the heat operator (with homogeneous i.c.) on \( \mathbb{R}^n \times \mathbb{R} \) via the (spatial\(^1\)) Fourier route:

\[
\begin{aligned}
\frac{\partial u}{\partial t} - \Delta u &= \delta \quad \text{for } (x, t) \in \mathbb{R}^n \times \mathbb{R} \\
u(x, 0) &= 0
\end{aligned}
\]

with a spatiotemporal Dirac point distribution \( \delta \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}) \) on the r.h.s. Taking \( h(x, t) = \delta(x)\delta(t) \) in that example we obtain a fundamental solution

\[\psi(x, t) = \theta(t) \phi(x, t) = \frac{\theta(t)}{\sqrt{4\pi t^n}} \exp\left(-\frac{\|x\|^2}{4t}\right),\]

in which \( \theta \in \mathcal{S}'(\mathbb{R}) \) is the Heaviside function (or indicator function on \( \mathbb{R}^+ \)), defined by \( \theta(t) = 0 \) for \( t < 0 \) and \( \theta(t) = 1 \) for \( t > 0 \). The Green’s function for the above system including homogeneous b.c. is a shifted copy:

\[\Psi(x, t, \xi, \tau) = \psi_{\xi, \tau}(x, t) = \psi(x - \xi, t - \tau) = \theta(t - \tau) \phi(x - \xi, t - \tau) = \frac{\theta(t - \tau)}{\sqrt{4\pi(t - \tau)^n}} \exp\left(-\frac{\|x - \xi\|^2}{4(t - \tau)}\right).\]

Recall that the solution of the inhomogeneous system of Example 7.2.3 is obtained by spatiotemporal convolution of the data term \( h : \mathbb{R}^n \times \mathbb{R} \rightarrow h(x, t) \) (with \( h(x, t) = 0 \) identically if \( t < 0 \)) and the fundamental solution:

\[u(x, t) = (h * \psi)(x, t) = \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} h(\xi, \tau) \Psi(x, t, \xi, \tau) \, d\xi \, d\tau,
\]

in which * now denotes spatiotemporal convolution.

\(^1\)Alternatively one may apply spatiotemporal Fourier transformation on \( \mathbb{R}^n \times \mathbb{R} \).
III

PDE Systems: Common Cases

63
11

FIRST ORDER SYSTEMS

11.1. The Transport Equation

In this chapter we consider the inhomogeneous transport equation with i.c., recall Example 1.2.9 on page 5:

$$\begin{cases}
  u_t - a \cdot \nabla u + a_0 u &= f \quad \text{for } (x,t) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^+, \\
  u|_{t=0} &= g,
\end{cases} \quad (11.1)$$

with

$$a \cdot \nabla = \sum_{i=1}^n a_i \partial_i. \quad (11.2)$$

11.2. The Dirac Equation

A famous first order p.d.e. in quantum physics is the so-called Dirac equation, which was introduced as one of the attempts to replace the non-relativistic Schrödinger equation by a relativistic counterpart. The Klein-Gordon equation may be seen as another attempt, recall Example 1.2.7: $(\Box + m^2) u = 0$ (in natural units). Aiming for a first order time-evolution equation, Paul Dirac stipulated (using different but equivalent notation) a first order
Recall the relativistic Klein-Gordon operator:

\[ \Box + m^2 = \partial_{tt} - \Delta + m^2. \]

Dirac imposed the constraint that the Dirac equation for a physical field must imply the Klein-Gordon equation.

To see how this comes about, consider a field \( \psi \) satisfying Eq. (11.3). We then have

\[
0 = \left( -i \sum_{\mu=0}^{3} \gamma^\mu \partial_\mu - m \right) \left( i \sum_{\nu=0}^{3} \gamma^\nu \partial_\nu - m \right) \psi = \left( \frac{1}{2} \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \{\gamma^\mu, \gamma^\nu\} \partial_\mu \partial_\nu + m^2 \right) \psi,
\]

in which \( \{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu \). Consistency with the Klein-Gordon equation thus produces the following formal condition

\[
\{\gamma^\mu, \gamma^\nu\} = 2 \eta^{\mu\nu},
\]

in which \( \eta^{00} = 1, \eta^{11} = \eta^{22} = \eta^{33} = -1 \), and all other \( \eta^{ij} = 0 \) (Minkowski metric). The conditions (●) cannot be fulfilled if the coefficients are ordinary (commuting) numbers \( \gamma^\mu \in \mathbb{K} \). The “anti-commutator” introduced here, however, concisely captures Dirac’s brilliant idea, viz. that one can satisfy (●) by interpreting each \( \gamma^\mu \) as a particular \( 4 \times 4 \)-matrix, and, thus, \( \psi \) as a \( 4 \)-component so-called “spinor” instead of a scalar field. You may verify that the \( 4 \times 4 \)-matrices

\[
\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & \sigma^1 \\ -\sigma^1 & 0 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & \sigma^3 \\ -\sigma^3 & 0 \end{pmatrix},
\]

with \( 2 \times 2 \)-submatrices

\[
\emptyset = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
\]

do fulfill the stated conditions (●). The three matrices \( \sigma^i, i = 1, 2, 3 \), are known as the Pauli spin matrices. These were already known from Wolfgang Pauli’s phenomenological theory of spin (Nobel Prize in physics, 1945, “for the discovery of the Exclusion Principle, also called the Pauli Principle”). However, this theory was constructed on the basis of 2-component wave functions. Dirac’s \( 4 \)-component spinor formalism introduced a new paradigm in physics, for which he received the Nobel Prize in physics in 1933 (jointly with Erwin Schrödinger “for the discovery of new productive forms of atomic theory”). Among others it led to the prediction of antimatter to explain the theoretical possibility of electrons with negative energy states in his formalism, which would effectively behave in electromagnetic interactions as if positively charged. These are nowadays known as anti-electrons or positrons and have been experimentally confirmed.

All in all, the Dirac equation takes the form

\[
(i \not{\psi} - m) \psi = 0,
\]

in which \( \not{\psi} = \sum_{\mu=0}^{3} \gamma^\mu \partial_\mu \) is a commonly used abbreviation in physics. In this notation we have set \( \hbar = c = 1 \). Reintroducing these fundamental constants, the Dirac equation takes the form \( (i \hbar \not{\psi} - mc) \psi = 0 \). Note that the Klein-Gordon equation holds component-wise for the spinor, whereas the Dirac equation couples the components through the \( \gamma^\mu \)-matrices.

The Dirac equation can be derived from a variational principle, involving besides the Dirac spinor \( \psi \) a second, so-called conjugate Dirac spinor \( \bar{\psi} = \psi^\dagger \gamma^0 \). The Dirac action functional is given by

\[
S[\psi, \bar{\psi}] = \int \bar{\psi}(x)(i \sum_{\mu=0}^{3} \gamma^\mu \partial_\mu - m)\psi(x) \, dx.
\]
Variation with respect to $\bar{\psi}$ (leaving $\psi$ fixed) gives the Dirac equation in the form of Eq. (11.3). Variation with respect to $\psi$ (leaving $\bar{\psi}$ fixed) yields the conjugate Dirac equation for $\bar{\psi} = \psi^0\gamma^0$:

$$i \sum_{\mu=0}^{3} \partial_{\mu} \bar{\psi} \gamma^\mu + m \bar{\psi} = 0.$$ 

Note that $\bar{\psi}$ is still to the left of $\gamma^\mu$, as in the action functional.
12
SECOND ORDER
HYPERBOLIC SYSTEMS

12.1. The Wave Equation

In this chapter we consider the inhomogeneous wave equation with i.c.’s, recall Section 1.2:

\[
\begin{align*}
\Box u &= h \quad \text{for } (x,t) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^+, \\
\big|_t=0 u &= \phi , \\
\big|_t=0 u_t &= \psi , 
\end{align*}
\]

(12.1)

with

\[
\Box = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta .
\]

(12.2)

Recall from Section 1.6 that this is an example of a hyperbolic system.

We focus explicitly on spatial dimensions \( n = 1, 2, 3 \). In all examples below, tacit assumptions are in effect concerning such matters as regularity of the data, existence of well-defined Fourier transforms, interchangeability of differentiation and integration, et cetera, so as to justify all manipulations. Recall that generalised solutions may still exist if some of these assumptions fail to hold.
Examples 12.2.1–12.2.6 illustrate the case $n = 1$. Examples 12.2.7–12.2.11 deal with the case $n = 3$. Example 12.2.12 considers the case $n = 2$. The reason for this particular order will become apparent in due course.

### 12.2. Examples

We start with the homogeneous case in $n = 1$ dimension.

**Example 12.2.1.** Consider the $(1+1)$-dimensional homogeneous wave equation with i.c.’s,

\[
\begin{align*}
(\circ) & \quad \begin{cases} 
u_{tt} - c^2 \nu_{xx} = 0 \quad \text{for } (x, t) \in \mathbb{R} \times \mathbb{R}^+. \\
\nu|_{t=0} = \phi \\
\nu_t|_{t=0} = \psi.
\end{cases}
\end{align*}
\]

As an ansatz we try the method proposed in Chapter 7. Since spatiotemporal Fourier transformation would scramble the time-localised i.c.’s, we consider spatial Fourier transformation, ignoring the time coordinate $t$:

\[
u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\nu}(\omega, t) \, d\omega,
\]

Substituting into $(\circ)$ yields the following o.d.e. system (with $\omega \in \mathbb{R}$ interpreted as a parameter):

\[
(\ast) \quad \begin{cases} rac{d^2\hat{\nu}}{dt^2} + c^2 \omega^2 \hat{\nu} = 0 \quad \text{for } (\omega, t) \in \mathbb{R} \times \mathbb{R}^+. \\
\hat{\nu}|_{t=0} = \hat{\phi} \\
\frac{d\hat{\nu}}{dt}|_{t=0} = \hat{\psi}.
\end{cases}
\]

The general solution of the o.d.e. in $(\ast)$ takes the form

\[
\hat{\nu}(\omega, t) = A_+ (\omega) e^{i\omega ct} + A_- (\omega) e^{-i\omega ct}.
\]

Imposing the i.c.’s in $(\ast)$ fixes the integration constants:

\[
A_\pm (\omega) = \frac{1}{2} \left( \hat{\phi}(\omega) \pm \frac{1}{i\omega c} \hat{\psi}(\omega) \right)
\]

Fourier inversion yields

\[
(\bullet) \quad u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega) \left( \frac{e^{i\omega(x+ct)} + e^{i\omega(x-ct)}}{2} \right) d\omega + \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\psi}(\omega) \left( \frac{e^{i\omega(x+ct)} - e^{i\omega(x-ct)}}{2i\omega c} \right) d\omega.
\]

In order to evaluate the second integral we need to recall a basic result from calculus. Setting

\[
\Psi(x) \overset{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\psi}(\omega) \frac{e^{i\omega x}}{i\omega} \, d\omega,
\]

we observe that

\[
\Psi'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\psi}(\omega) e^{i\omega x} \, d\omega = \psi(x),
\]

provided differentiation and integration may be interchanged, implying

\[
\Psi(x) = \int_{x_0}^{x} \psi(\xi) \, d\xi.
\]
for some arbitrary constant $x_0 \in \mathbb{R}$. It follows that

$$u(x, t) = \frac{1}{2} \phi(x + ct) + \frac{1}{2} \phi(x - ct) + \frac{1}{2c} \int_{x_0}^{x+ct} \psi(\xi) \, d\xi - \frac{1}{2c} \int_{x_0}^{x-ct} \psi(\xi) \, d\xi.$$  

Note that the $x_0$-dependence actually cancels out, since

$$\int_{x_0}^{b} \psi(\xi) \, d\xi - \int_{x_0}^{a} \psi(\xi) \, d\xi = \int_{a}^{b} \psi(\xi) \, d\xi.$$  

All in all, the solution of $(\circ)$ is given in semi-closed form by

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

Cf. Fig. 12.1 for an illustration.

Figure 12.1: Travelling wave solutions of $(\circ)$ in one spatial dimension. Top row: With i.c.’s $\phi(x) = \delta(x)$, $\psi(x) = 0$, the initial spike splits into two half-weight spikes, one travelling to the left and the other to the right. Each disturbance moves with constant velocity $|c|$ without altering shape, and remains locally confined to a single point. Bottom row: With i.c.’s $\phi(x) = 0$, $\psi(x) = \delta(x)$, the initial (unmeasurable) disturbance develops two wave fronts, one travelling to the left and the other to the right. In-between the fronts, moving with constant velocity $|c|$, the solution has a constant amplitude.

Example 12.2.2. Another option for solving the system $(\circ)$ in Example 12.2.1 via the Fourier route is to ignore, momentarily, the i.c.’s, extending the domain of definition to all of spacetime. That is, we first look for a globally defined solution $u : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. In this case, the genuine two-dimensional spacetime Fourier transform, 

$$(*) \quad u(x, t) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\omega x + ivt} \hat{u}(\omega, \nu) \, d\nu d\omega,$$  

71
can be applied to transform the p.d.e. in (⊙) into the following equivalent algebraic equation (cf. the o.d.e. in (⋆)):

\((-\nu^2 + c^2 \omega^2) \hat{u}(\omega, \nu) = 0\) for \((\omega, \nu) \in \mathbb{R}^2\).

This implies \(\hat{u}(\omega, \nu) = 0\) for all \((\omega, \nu) \in \Omega = \{\mathbb{R}^2 | \nu \pm c \omega \neq 0\}\), and, a fortiori, that we must expect nontrivial solutions to exist only in strict distributional sense, i.e. \(\hat{u} \in \mathcal{S}(\mathbb{R}^2)\). In physics, a constraint like \(C: \nu^2 - c^2 \omega^2 = 0\) is known as a dispersion relation, whereas the support set \(\Omega \subset \mathbb{R}^2\) is sometimes referred to as a “shell”.

In view of the Theorem 4.2.2 we take as an ansatz

\[ \hat{u}(\omega, \nu) = \sum_{\pm} A_{\pm}(\omega, \nu) \delta(\nu \mp c \omega). \]

Substitution into (⋆), and evaluation of the time-integral, yields

\[ (⋆⋆) \quad u(x, t) = \sum_{\pm} \frac{1}{2\pi} \int_{-\infty}^{\infty} \alpha_{\pm}(\omega) e^{i\omega(x \pm ct)} d\omega, \]

in which we have redefined the coefficients:

\[ \alpha_{\pm}(\omega) \overset{\text{def}}{=} \frac{1}{2\pi} A_{\pm}(\omega, \pm c \omega). \]

With this general solution we may finally turn to the i.c.’s. By considering

\[ \phi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega) e^{i\omega x} d\omega \quad \text{and} \quad \psi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\psi}(\omega) e^{i\omega x} d\omega, \]

the i.c.’s take the form

\[ \hat{\phi}(\omega) = \alpha_{+}(\omega) + \alpha_{-}(\omega) \quad \text{respectively} \quad \hat{\psi}(\omega) = i\omega(c(\alpha_{+}(\omega) - \alpha_{-}(\omega))), \]

so that

\[ \alpha_{\pm}(\omega) = \frac{1}{2} \hat{\phi}(\omega) \pm \frac{1}{2i\omega c} \hat{\psi}(\omega). \]

Insertion into (⋆⋆) produces exactly the same result as (●) in Example 12.2.1, from which the solution (●●) follows.

Example 12.2.3. Again, recall Example 12.2.1. Instead of the Fourier route we now employ another solution strategy, based on the following observation:

\[ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} = \left( \frac{1}{c} \frac{\partial}{\partial t} - \frac{\partial}{\partial x} \right) \left( \frac{1}{c} \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right), \]

i.e. the d’Alembertian in \(n = 1\) spatial dimension is a composition of two commuting first order (“forward” and “backward”) advection p.d.o.’s (cf. Example 1.2.9 in Chapter 1). The form of the associated characteristics (recall Chapter 8, notably Definition 8.1.1) argues for a strategy similar to the one outlined in Section 1.5, as follows.

Consider a coordinate transformation adapted to the characteristics of the two advection p.d.o.’s (“light cone coordinates”):

\[ (●●) \quad \left\{ \begin{array}{l}
\xi = x + ct \\
\eta = x - ct
\end{array} \right. \]

and define \(v(\xi(x, t), \eta(x, t)) = u(x, t)\). Note that

\[ \frac{\partial}{\partial x} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \quad \text{and} \quad \frac{\partial}{\partial t} = c \frac{\partial}{\partial \xi} - c \frac{\partial}{\partial \eta}. \]

Now \(v(\xi, \eta)\) satisfies the p.d.e.

\[ \frac{\partial^2 v}{\partial \eta \partial \xi} = 0. \]
The solution is given by
\[ (\star) \quad v(\xi, \eta) = f(\xi) + g(\eta), \]
in which \( f, g : \mathbb{R} \to \mathbb{K} \) are a priori arbitrary functions. The i.c.'s originally formulated for \( u(x, t) \) on the \( x \)-axis, \( t = 0 \), translate to conditions for \( v(\xi, \eta) \) along the line \( \eta = \xi \), viz.
\[ v(\xi, \xi) = \phi(\xi) \quad \text{respectively} \quad v_\xi(\xi, \xi) - v_\eta(\xi, \xi) = \frac{1}{c} \psi(\xi). \]
Combined with the solution (\( \star \)) this yields
\[ f(\xi) + g(\xi) = \phi(\xi) \quad \text{respectively} \quad f'(\xi) - g'(\xi) = \frac{1}{c} \psi(\xi), \]
whence
\[ f(\xi) = \frac{1}{2} \phi(\xi) + \frac{1}{2c} \int_\xi^\infty \psi(\alpha) \, d\alpha \quad \text{respectively} \quad g(\xi) = \frac{1}{2} \phi(\xi) - \frac{1}{2c} \int_\xi^\infty \psi(\alpha) \, d\alpha. \]
Thus
\[ v(\xi, \eta) = \frac{1}{2} \phi(\xi) + \frac{1}{2} \phi(\eta) + \frac{1}{2c} \int_\eta^\xi \psi(\alpha) \, d\alpha. \]
Transformed back from light cone to Cartesian coordinates using (\( \bullet \bullet \)), \( u(x, t) = v(\xi(x, t), \eta(x, t)) \), we once again find the same result (\( \infty \)) as in Example 12.2.1.

**Example 12.2.4.** Consider the analogue of the system in Example 12.2.1 on a finite spatial interval:
\[
(\star) \quad \begin{cases} 
  u_{tt} - u_{xx} = 0 & \text{for } (x, t) \in (0, 1) \times \mathbb{R}^+, \\
  u(0, t) = u(1, t) = 0, \\
  u(x, 0) = f(x), \\
  u_t(x, 0) = g(x).
\end{cases}
\]
This system has been scrutinised in Examples 9.2.1 and 9.2.2 on pages 54ff. In particular, its solution is given by (\( \star \star \)) at the end of Example 9.2.2.

**Example 12.2.5.** Consider the inhomogeneous extension of the system in Example 12.2.4:
\[
(\star \star) \quad \begin{cases} 
  u_{tt} - u_{xx} = h & \text{for } (x, t) \in (0, 1) \times \mathbb{R}^+, \\
  u(0, t) = u(1, t) = 0, \\
  u(x, 0) = f(x), \\
  u_t(x, 0) = g(x).
\end{cases}
\]
with nontrivial inhomogeneous term \( h : (0, 1) \times \mathbb{R}^+ \to \mathbb{K} \). A naive ansatz of the form (\( \dag \)) in Example 9.2.1 won’t work, since the function \( h \) cannot be assumed to be separable. Instead, we take inspiration from the homogeneous problem to solve the inhomogeneous problem, as follows. Denoting the solution to the homogeneous problem by \( u_0 \) for the sake of clarity, we have (recall (\( \star \star \)) at the end of Example 9.2.2)
\[ u_0(x, t) = \sum_{k=1}^{\infty} \chi_k(t) \phi_k(x). \]
For this homogeneous solution we have \( \chi_k(t) = \frac{g \cdot \phi_k}{\pi k} \sin(k\pi t) + f \cdot \phi_k \cos(k\pi t) \) and \( \phi_k(x) = \sqrt{2} \sin(k\pi x) \). According to the Sturm-Liouville Theorem, recall Theorem 9.1.1, the spatial eigenfunctions \( \{ \phi_k \}_{k \in \mathbb{N}} \) span \( L^2((0, 1)) \). For this reason we stipulate a solution of the form
\[ (\dag) \quad u(x, t) = \sum_{k=1}^{\infty} \psi_k(t) \phi_k(x), \]
i.e. we hypothesize that the effect of the inhomogeneous term can be accounted for by a suitable adaptation of the time-dependent coefficients, a trick of the trade reminiscent of the “variation of constant” method used in the
context of inhomogeneous linear first order o.d.e.’s as outlined on page 50. Inserting (‡) into (∗∗) and using the
eigenfunction properties,
\[ \phi_k''(x) = -k^2 \pi^2 \phi_k(x), \]
yields
\[ \sum_{k=1}^{\infty} \left( \ddot{\psi}_k + k^2 \pi^2 \psi_k \right) \phi_k = h. \]

Multiplication by \( \phi_\ell \) and integration over \( x \in (0,1) \), using orthonormality of \( \{ \phi_k \}_{k \in \mathbb{N}} \) relative to the standard
inner product on \( L^2((0,1)) \), shows that
\[ (\diamond) \quad \ddot{\psi}_\ell = h_\ell, \]
in which
\[ h(x,t) = \sum_{\ell=1}^{\infty} h_\ell(t) \phi_\ell(x), \]
with
\[ h_\ell(t) = \int_0^1 h(x,t) \phi_\ell(x) \, dx. \]

The o.d.e. (\diamond) governs the time-dependent coefficients \( \psi_\ell(t) \) in (‡) up to a pair of integration constants. Writing
\[ f(x) = \sum_{k=1}^{\infty} f_k \phi_k(x) \quad \text{and} \quad g(x) = \sum_{k=1}^{\infty} g_k \phi_k(x), \]
with
\[ f_k = \int_0^1 f(x) \phi_k(x) \, dx \quad \text{respectively} \quad g_k = \int_0^1 g(x) \phi_k(x) \, dx, \]
we find, by taking \( t=0 \) in (‡) and imposing the i.c.’s in (∗∗):
\[ (\diamond\diamond) \quad \psi_\ell(0) = f_\ell \quad \text{respectively} \quad \dot{\psi}_\ell(0) = g_\ell. \]

Taken together, (‡), (\diamond) and (\diamond\diamond) admit a unique solution for \( \psi_\ell(t) \) and thus for \( u(x,t) \) in (∗∗) in terms of the
inhomogeneous term \( h(x,t) \) and the initial data \( f(x) \) and \( g(x) \).

Example 12.2.6. Consider the system (∗∗) in Example 12.2.5 for the special case in which the inhomogeneous
term is given by a time-periodic function, say
\[ h(x,t) = \alpha(x) \cos(\beta t). \]

The i.c.’s are immaterial to the following discussion, as we will not be interested in a specific choice of integration
constants, so we consider the following system:
\[ \left\{ \begin{array}{l}
\quad u_{tt} - u_{xx} = h \quad \text{for} \ (x,t) \in (0,1) \times \mathbb{R}^+, \\
\quad u(0,t) = u(1,t) = 0,
\end{array} \right. \]

Recall (‡) and (\diamond) in Example 12.2.5, in this particular case:
\[ (i) \quad \ddot{\psi}_\ell + \ell^2 \pi^2 \psi_\ell = h_\ell, \]
with
\[ h_\ell(t) = \alpha_\ell \cos(\beta t) \quad \text{and} \quad \alpha_\ell = \int_0^1 \alpha(x) \phi_\ell(x) \, dx. \]

The solution of o.d.e. (i) is the sum of its homogeneous solution \( \psi_{0\ell}(t) \) and a particular solution \( \psi_{p\ell}(t) \). We have
\[ \psi_{0\ell}(t) = A_\ell \cos(\ell \pi t) + B_\ell \sin(\ell \pi t). \]
To find a particular solution, stipulate one of the form $\psi_{lp}(t) = a_\ell \cos(\beta t) + b_\ell \sin(\beta t)$. Substitution into (i) shows that one can indeed find one with nontrivial coefficients, provided $\beta \neq \ell \pi$, viz.

$$\psi_{lp}(t) = \frac{a_\ell}{\ell^2 \pi^2 - \beta^2} \cos(\beta t).$$

If $\beta = \ell \pi$, the eigenfrequency of the system (i), then a nontrivial particular solution can be found by stipulating a form $\psi_{lp}(t) = (a_\ell + b_\ell t) \sin(\beta t)$, viz.

$$\psi_{lp}(t) = \frac{a_\ell t}{\ell \pi} \sin(\ell \pi t).$$

Thus $u(x, t) = \sum_{\ell=1}^{\infty} \psi_{\ell}(t) \phi_{\ell}(x)$, with $\phi_{\ell}(x) = \sqrt{2} \sin(\ell \pi x)$ and

$$\psi_{\ell}(t) = A_\ell \cos(\ell \pi t) + B_\ell \sin(\ell \pi t) + \alpha_\ell \times \begin{cases} \cos(\beta t) \ell^2 \pi^2 - \beta^2 t \sin(\ell \pi t) \\ \ell \pi \end{cases}$$

off resonance, i.e. for $\beta \neq \ell \pi$,

on resonance, i.e. for $\beta = \ell \pi$.

If desired, i.e.’s may be imposed to fix $A_\ell$ and $B_\ell$, recall ($\infty$) in Example 12.2.5. A remarkable phenomenon becomes apparent, known as resonance. If the “driving force” $h(x, t)$ contains a mode ($\alpha_\ell \neq 0$ for some $\ell \in \mathbb{N}$) with a temporal frequency $\beta \approx L \pi$, then the corresponding mode $u_{L}(x, t) = \psi_{L}(t) \phi_{L}(x)$ of the solution has an amplitude $\propto 1/|L^2 \pi^2 - \beta^2| \gg 1$, exhibiting a singularity as $\beta \to L \pi$. In the singular case, $\beta = L \pi$, we observe an “envelope” that linearly increases without bound as $t \to \infty$. The system is said to resonate at the eigenfrequency $\beta = L \pi$, triggered by the inhomogeneous term $h$.

**Example 12.2.7.** Consider the $(3+1)$-dimensional homogeneous wave equation with i.e.’s,

\[
\begin{aligned}
(\dddot{)} & \quad \begin{align*}
\dddot{u} - c^2 (\dddot{u}_{xx} + \dddot{u}_{yy} + \dddot{u}_{zz}) &= 0 & \text{for } (x, y, z, t) & \in \mathbb{R}^3 \times \mathbb{R}^+, \\
\dddot{u}_{t=0} &= f, \\
\dddot{u}_{t=0} &= g.
\end{align*}
\end{aligned}
\]

We are interested in spherically symmetric solutions around some fiducial point $P$ with Cartesian coordinates $(\xi, \eta, \zeta)$, i.e. solutions that take the form $u(x, y, z, t) = v(r, t)$ with $r = \sqrt{(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2}$, assuming, for the moment, that also the i.e.’s are spherically symmetric, with, by abuse of notation, $f(x, y, z) = f(r), g(x, y, z) = g(r)$, say. This suggests the introduction of spherical coordinates around $P$:

\[
(x) \quad \begin{aligned}
x &= \xi + r \sin \theta \cos \phi, \\
y &= \eta + r \sin \theta \sin \phi \\
z &= \zeta + r \cos \theta.
\end{aligned}
\]

With this transformation and under the given symmetry assumption, (\dddot{)} is transformed into

\[
(\dddot{)} \quad \begin{aligned}
(rv)_{tt} - c^2 (rv)_{rr} &= 0 & \text{for } (r, t) & \in \mathbb{R}^+ \times [0, 2\pi) \times [0, \pi], \\
rv_{t=0} &= f, \\
rv_{t=0} &= g.
\end{aligned}
\]

By an argument similar to that used in Example 12.2.3 we then find

$$v(r, t) = \frac{1}{2r} F(r + ct) + \frac{1}{2r} F(r - ct) + \frac{1}{2rc} \int_{r-ct}^{r+ct} G(\alpha) \, d\alpha,$$

with $F(r) = rf(r)$ and $G(r) = rg(r)$, i.e. a superposition of incoming and outgoing waves. The singularity at $r = 0$ may be removed by slick choice of $f$ and $g$.

**Example 12.2.8.** From a spherically symmetric solution of the three-dimensional wave equation we may construct new (non-symmetric) solutions, as follows. Let us consider an outgoing wave emanating from the fiducial origin of the following form (in distributional sense, recall Chapter 4):

$$v(x, y, z, t) = \frac{\delta(r - ct)}{4\pi r},$$
with \( r \equiv \sqrt{x^2 + y^2 + z^2} \), the distance from the point of evaluation \((x, y, z)\) relative to the origin. For sufficiently smooth \( g : \mathbb{R}^3 \to \mathbb{K} \),

\[
h(x, y, z, t) = (g *_{\mathbb{R}^3} v)(x, y, z, t) = \int_{\mathbb{R}^3} g(\xi, \eta, \zeta) v(x - \xi, y - \eta, z - \zeta, t) \, d\xi d\eta d\zeta,
\]

also satisfies the wave equation. We may interpret this solution as the constructive interference pattern caused by point disturbances propagating from their source points \((\xi, \eta, \zeta) \in \mathbb{R}^3\) at \( t = 0 \) to \((x, y, z) \in \mathbb{R}^3\) at time \( t \geq 0 \), where each point source contributes with an amplitude \( g(\xi, \eta, \zeta) \) (“Huygens’ Principle”), cf. Fig. 12.2.

Figure 12.2: The solution at \((x, y, z, t)\) is only affected by source beacons that have been instantaneously active at \( t = 0 \) at points \((\xi, \eta, \zeta)\) a distance \( R = ct \) away from \((x, y, z)\). Thus an observer at \((x, y, z)\) “sees” the history of the (aggregated) source activities delayed by \( t \) time units. Vice versa, a point source active at time \( t = 0 \) in \((x, y, z)\) will only affect those points \((\xi, \eta, \zeta, t)\) in spacetime that lie on the “light cone” with tip \((x, y, z, t)\) and radius \( R = ct \). The propagation phenomenon of local disturbances is thus qualitatively similar to that in the one-dimensional case, recall Fig. 12.1 (top row).

Using spherical integration variables \((R, \Theta, \Phi)\) relative to the fixed destination point \((x, y, z)\),

\[
\begin{align*}
\xi &= x + R \sin \Theta \cos \Phi, \\
\eta &= y + R \sin \Theta \sin \Phi, \\
\zeta &= z + R \cos \Theta,
\end{align*}
\]

we obtain

\[
h(x, y, z, t) = \frac{ct}{4\pi} \int_0^{2\pi} \int_0^\pi g(x + ct \sin \Theta \cos \Phi, y + ct \sin \Theta \sin \Phi, z + ct \cos \Theta) \sin \Theta \, d\Theta d\Phi.
\]

In other words,

\[
h(x, y, z, t) = ct \, \mu_{(x, y, z, ct)}[g],
\]

in which \(\mu_{(x, y, z, R)}\) returns the average of its operand over a sphere of radius \( R \) centered at \((x, y, z)\). Note that

\[
\lim_{t \to 0} \mu_{(x, y, z, ct)}[g] = g(x, y, z).
\]

The rescaled function \(u(x, y, z, t) = h(x, y, z, t)/c = t \mu_{(x, y, z, ct)}[g]\) apparently solves the following system:

\[
\begin{align*}
(\dagger) \quad u_{tt} - c^2 (u_{xx} + u_{yy} + u_{zz}) &= 0 \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+.
\end{align*}
\]

\[
\begin{align*}
||_{t=0} &= 0, \\
\hspace{0.5cm} u_{t} \big|_{t=0} &= g.
\end{align*}
\]
Example 12.2.9. Consider now the homogeneous system with general i.c.'s:

$$
\begin{align*}
\{ & u_{tt} - c^2 (u_{xx} + u_{yy} + u_{zz}) = 0 \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\
& u|_{t=0} = f \\
& u|_{t=0} = g,
\end{align*}
$$

Using the fact that any t-derivative of $u$ will be a solution of the homogeneous linear p.d.e. if $u$ is itself a solution, one readily verifies that the solution to (†) including i.c.'s is given by

$$
u(x, y, z, t) = t\mu(x,y,z,ct)[g] + \frac{\partial}{\partial t} (t\mu(x,y,z,ct)[f]).$$

The proof relies on elementary calculus (such as the product rule and chain rule for differentiation) in addition to the following observations, in which $\chi: \mathbb{R}^3 \to \mathbb{R}$ denotes a continuous function (either $f$ or $g$):

- $\mu(x,y,z,0)[\chi] = \chi(x,y,z)$;
- $\partial_t \mu(x,y,z,ct)[\chi]|_{t=0} = c\nabla \chi(x,y,z) \cdot \frac{1}{4\pi} \int_0^\pi \hat{r}(\Theta, \Phi) \sin \Theta d\Theta d\Phi = 0.$

Here $\hat{r}(\Theta, \Phi) = (\sin \Theta \cos \Phi, \sin \Theta \sin \Phi, \cos \Theta)$ is the unit vector in the direction specified by $(\Theta, \Phi)$. The crucial step (⋆) follows by evaluating the three components of the vector-valued integral, but should also be immediately obvious once you recognize that it represents the average of all unit vectors over all directions.

Example 12.2.10. Consider the inhomogeneous system with homogeneous i.c.'s:

$$
\begin{align*}
\{ & u_{tt} - c^2 (u_{xx} + u_{yy} + u_{zz}) = h \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\
& u|_{t=0} = 0 \\
& u|_{t=0} = 0,
\end{align*}
$$

for some sufficiently smooth data term $h: \mathbb{R}^3 \times \mathbb{R}^+ \to \mathbb{K}$. Its solution can be obtained with the help of Duhamel's principle, which provides an explicit solution of (●) in terms of the already obtained solution of the system (†) with slick choice of i.c.'s formulated for a shifted initial moment. Namely, consider

$$
\begin{align*}
\{ & v_{tt} - c^2 (v_{xx} + v_{yy} + v_{zz}) = 0 \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\
& v|_{t=\tau} = 0 \\
& v|_{t=\tau} = h|_{t=\tau}.
\end{align*}
$$

This is basically system (††) in Example 12.2.8, subject to a time shift. As the solution depends on the auxiliary parameter $\tau$, let us write it as a pentavariate function $v: \mathbb{R}^3 \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{K}: (x, y, z, t, \tau) \mapsto v(x, y, z, t; \tau).$ Then it follows straightforwardly by substitution that

$$
u(x, y, z, t) = \int_0^t v(x, y, z, t; \tau) d\tau$$

solves (●). To see this, note that the i.c.'s are trivially fulfilled:

$$
\begin{align*}
u(x, y, z, 0) &= 0, \\
u(x, y, z, 0) &= v(x, y, z, t; t) = 0,
\end{align*}
$$

as a result of the first i.c. for $v$. Furthermore,

$$
u_t(x, y, z, t) = v_t(x, y, z, t; t) + \int_0^t v_t(x, y, z, t; \tau) d\tau = \int_0^t v_t(x, y, z, t; \tau) d\tau,$$

v.s., and

$$
u_{tt}(x, y, z, t) = v_{tt}(x, y, z, t; t) + \int_0^t v_{tt}(x, y, z, t; \tau) d\tau = h(x, y, z, t) + c^2 \Delta \int_0^t v(x, y, z, t; \tau) d\tau$$

as a result of the second i.c., respectively the homogeneous wave equation for $v$. Also note that we did not need any details about the specific form of the solution, thus this argument generalises to any dimension $n$. 77
Example 12.2.11. Now consider the general, inhomogeneous problem:

\[
\begin{align*}
\{ \ & u_{tt} - c^2 (u_{xx} + u_{yy} + u_{zz}) = h \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\
& u_{t|t=0} = f , \\
& u_{t|t=0} = g ,
\end{align*}
\]

for some sufficiently smooth data term \( h : \mathbb{R}^3 \times \mathbb{R}^+ \to \mathbb{K} \). The solution is a sum of the solution of the homogeneous system with i.c.’s as obtained in Example 12.2.9 and a particular solution of the inhomogeneous system with homogeneous i.c.’s considered in Example 12.2.10. In turn, recall from Example 12.2.10 that the latter can be reduced to a homogeneous system of the form studied in Examples 12.2.8–12.2.9 by exploiting Duhamel’s principle. Thus we have

\[
(\dot{\tau}) \quad u = \dot{u} + U ,
\]

with

\[
\begin{align*}
\{ \ & \dot{u}_{tt} - c^2 (\dot{u}_{xx} + \dot{u}_{yy} + \dot{u}_{zz}) = 0 \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\
& \dot{u}_{t|t=0} = f , \\
& \dot{u}_{t|t=0} = g ,
\end{align*}
\]

and

\[
\begin{align*}
\{ \ & U_{tt} - c^2 (U_{xx} + U_{yy} + U_{zz}) = h \quad \text{for } (x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\
& U_{t|t=0} = 0 , \\
& U_{t|t=0} = 0 .
\end{align*}
\]

Thus the solution (\( \dot{\tau} \)) is the sum of

\[
\dot{u}(x, y, z, t) = t\mu(x,y,z,ct)[g] + \frac{\partial}{\partial t} \left( t\mu(x,y,z,ct)\{f\} \right)
\]

and, with \( v(x, y, z, t; \tau) \) as constructed in Example 12.2.10,

\[
U(x, y, z, t) = \int_0^t (t-\tau) v(x, y, z, t; \tau) \, d\tau = \int_0^t (t-\tau) \mu(x,y,z,(t-\tau))[h|_{t=\tau}] \, d\tau
\]

\[
= \int_0^t \frac{(t-\tau)}{4\pi} \int_0^{2\pi} \int_0^\pi h(x + c(t-\tau) \sin \Theta \cos \Phi, y + c(t-\tau) \sin \Theta \sin \Phi, z + c(t-\tau) \cos \Theta, \tau) \sin \Theta \, d\Theta \, d\Phi \, d\tau.
\]

In the following example \( z \)-invariance is imposed as a ploy to reduce the three-dimensional system to a two-dimensional one.

Example 12.2.12. In this example we consider the homogeneous two-dimensional system

\[
\begin{align*}
\{ \ & u_{tt} - c^2 (u_{xx} + u_{yy}) = 0 \quad \text{for } (x, y, t) \in \mathbb{R}^2 \times \mathbb{R}^+, \\
& u_{t|t=0} = f , \\
& u_{t|t=0} = g ,
\end{align*}
\]

The solution can be derived from that of the three-dimensional system in Example 12.2.9, viz. by assuming the i.c.’s \( f, g : \mathbb{R}^3 \to \mathbb{K} \) in that example to be independent of \( z \), so that, by abuse of notation, we may equivalently consider these as functions of type \( f, g : \mathbb{R}^2 \to \mathbb{K} \) as they enter in the i.c.’s of the problem at hand. With such \( z \)-independent i.c.’s we may simplify the pseudo-three-dimensional solution found in Example 12.2.9:

\[
u(x, y, t) = t\mu(x,y,0,ct)[g] + \frac{\partial}{\partial t} \left( t\mu(x,y,0,ct)\{f\} \right) ,
\]

i.e.

\[
u(x, y, t) = \frac{t}{4\pi} \int_0^{2\pi} \int_0^\pi g(x + ct \sin \Theta \cos \Phi, y + ct \sin \Theta \sin \Phi) \sin \Theta \, d\Theta \, d\Phi + \\
+ \frac{\partial}{\partial t} \left[ \frac{t}{4\pi} \int_0^{2\pi} \int_0^\pi f(x + ct \sin \Theta \cos \Phi, y + ct \sin \Theta \sin \Phi) \sin \Theta \, d\Theta \, d\Phi \right] .
\]

78
A change of integration variables, \( r = ct \sin \Theta \) (the radial coordinate obtained after projecting points \((\xi, \eta, \zeta)\) on the active 3-sphere with radius \(ct\) around \((x, y, 0)\) onto the “shadow” 2-disk in the \(xy\)-plane), turns this expression into the following form (to see this, split the \(\Theta\)-integral into two parts, one for the upper and one for the lower hemisphere, i.e. \(\Theta \in [0, \pi/2]\) respectively \(\Theta \in [\pi/2, \pi]\)): 

\[
 u(x, y, t) = \frac{1}{2\pi c} \int_0^{2\pi} \int_0^{ct} g(x + r \cos \Phi, y + r \sin \Phi) \frac{r}{\sqrt{c^2t^2 - r^2}} \, dr \, d\Phi + \\
 + \frac{\partial}{\partial t} \left[ \frac{1}{2\pi c} \int_0^{2\pi} \int_0^{ct} f(x + r \cos \Phi, y + r \sin \Phi) \frac{r}{\sqrt{c^2t^2 - r^2}} \, dr \, d\Phi \right].
\]

That there are two equal contributions from each hemisphere should be obvious by mirror symmetry relative to the \(xy\)-plane, which explains the factor \(1/(2\pi) = 2 \times 1/(4\pi)\). This geometric projection also explains why residual “echoes” of the sources \(f(\xi, \eta)\) and \(g(\xi, \eta)\) originating from all points \((\xi, \eta)\) in the interior of the 2-disk, \(0 \leq r = \sqrt{(x-\xi)^2 + (y-\eta)^2} \leq ct\), affect the signal \(u(x, y, t)\) in point \((x, y)\) at time \(t\), in marked contrast with the cases \(n = 1\) and \(n = 3\), in which only the boundaries of the respective \(n\)-spheres play a role. The appearance of an attenuation factor \(1/\sqrt{c^2t^2 - r^2}\) in the integrand shows that the dominant contribution still comes from the “wave front” \(r = ct\). Sources at distances \(r > ct\) do not affect the signal at time \(t\), a common property for all dimensions \(n\).
13

Second Order
Parabolic Systems

13.1. The Heat Equation

In this chapter we consider the homogeneous heat equation with i.c.'s, recall Section 1.2:

\[
\begin{aligned}
\begin{cases}
\frac{\partial u}{\partial t} - \Delta u &= 0 \quad \text{for } (x,t) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^+, \\
\left. u \right|_{t=0} &= f,
\end{cases}
\end{aligned}
\]  

(13.1)

with, in Cartesian coordinates,

\[
\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}.
\]  

(13.2)

Recall from Section 1.6 that this is an example of a parabolic system.
13.2. Examples

Example 13.2.1. Consider the heat equation with i.c. on a finite interval for a real-valued function:

\[
\begin{align*}
\tag{*}
\frac{\partial u}{\partial t} - u_{xx} &= 0 & \text{for } (x, t) \in (0, 1) \times \mathbb{R}^+, \\
u(x, 0) &= f(x), \\
u(0, t) = \nu(1, t) &= 0.
\end{align*}
\]

This example resembles that of Example 9.2.1 in Chapter 9, so let us likewise stipulate a nontrivial separable solution:

\[
\tag{†} u(x, t) = \phi(x) \chi(t).
\]

For the p.d.e. this yields \(\phi' \chi = \phi'' \chi\), with dots and primes indicating temporal, respectively spatial differentiation. This implies

\[
\phi'' + c^2 \phi = 0,
\]

for some constant \(c \in \mathbb{C}\), with b.c.’s \(\phi(0) = \phi(1) = 0\), and

\[
\chi + c^2 \chi = 0.
\]

The latter has a general solution of the form

\[
\tag{*} \chi(t) = Ce^{-c^2 t}.
\]

The former has a general solution of the form

\[
\phi(x) = A \sin(cx) + B \cos(cx),
\]

with \(B = 0\), \(A \in \mathbb{R}\) arbitrary, and with a restriction on \(c\), viz. \(c = k\pi, k \in \mathbb{N}\). Adopting the standard normalization on the unit interval we thus find a sequence of solutions, one for every \(k \in \mathbb{N}\):

\[
\phi_k(x) = \sqrt{2} \sin(k\pi x) \quad (k \in \mathbb{N}).
\]

Taking the restriction on \(c\) into account, \(\phi\) takes the form

\[
\chi_k(t) = C_k e^{-k^2\pi^2 t},
\]

with \(C_k \in \mathbb{R}\). A general solution of \(\phi\) is a linear superposition

\[
u(x, t) = \sum_{k \in \mathbb{N}} C_k \phi_k(x) e^{-k^2\pi^2 t}.
\]

The coefficients are determined by the initial data \(u(x, 0) = f(x)\). By setting \(t = 0\) in the above expression and taking the standard inner product with \(\phi_m\) we obtain

\[
f(x) = \sum_{k \in \mathbb{N}} C_k \phi_k(x) \quad \text{with} \quad C_k = \int_0^1 f(x) \phi_k(x) \, dx.
\]

Putting things together, we may thus write

\[
u(x, t) = \sum_{k \in \mathbb{N}} \left( \int_0^1 f(y) \phi_k(y) \, dy \right) \phi_k(x) e^{-k^2\pi^2 t}.
\]

Example 13.2.2. Consider the heat equation with i.c. without spatial boundaries:

\[
\begin{align*}
\tag{*}
\frac{\partial u}{\partial t} - \Delta u &= 0 & (x, t) \in \mathbb{R}^n \times \mathbb{R}^+, \\
u(x, 0) &= f(x).
\end{align*}
\]
in which \( f \in \mathcal{S}'(\mathbb{R}^n) \). This problem has been solved in Example 7.2.1, recall page 46:

\[
\begin{align*}
  u(x, t) &= \frac{1}{(4\pi t)^n} \int_{\mathbb{R}^n} f(y) e^{-\frac{4\pi^2}{t} y^2} dy.
\end{align*}
\]

Note that this solution may also be written as

\[
\begin{align*}
  u(x, t) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) e^{i\omega \cdot (x - y)} e^{-t\omega^2} dy d\omega.
\end{align*}
\]

In this form one recognizes the analogy with the “sum/integral” expression at the end of Example 13.2.1.
13.2. EXAMPLES  

13. SECOND ORDER PARABOLIC SYSTEMS
14
SECOND ORDER ELLIPTIC SYSTEMS

14.1. The Laplace Equation

14.2. The Helmholtz Equation

14.3. The Poisson Equation
BIBLIOGRAPHY


