Modelling and Perturbation Methods

R.M.M. Mattheij, S.W. Rienstra and J.H.M. ten Thije Boonkkamp
Eindhoven University of Technology

Chapters from a coming book on Partial Differential Equations

7 Oct 2003
## Contents

1 Differential and difference equations  
   1 Introduction ........................................ 5  
   2 Nomenclature ........................................... 10  
   3 Difference equations .................................... 12  
   4 Discussion ............................................. 14  
   Exercises .................................................. 14

2 Characterisation and classification  
   1 First order scalar partial differential equations .......... 17  
   2 First order systems ...................................... 21  
   3 Second order scalar partial differential equations .......... 23  
   4 Linear second order equations in more space variables ...... 27  
   5 Reduction to ODE; similarity solutions ...................... 29  
   6 Initial and boundary conditions; well-posedness ............ 31  
   7 Discussion ............................................. 33  
   Exercises .................................................. 34

3 Fourier theory  
   1 Fourier series ........................................... 37  
   2 Fourier transforms ....................................... 42  
   3 Discrete Fourier transforms ................................ 46  
   4 Fourier analysis applied to PDEs .......................... 48  
   5 Discussion ............................................. 51  
   Exercises .................................................. 51
## 4 Distributions and fundamental solutions

1. Introduction ........................................... 55
2. Distributions in one variable .......................... 57
3. Distributions in more variables ....................... 61
4. Strong and weak solutions ............................ 62
5. Fundamental solutions ................................. 65
6. Initial (boundary) value problems; Duhamel integrals . 67
7. Discussion ............................................ 72

### Exercises ............................................ 72

## 6 Continuum mechanics and electromagnetics

1. Introduction ........................................... 75
2. Eulerian and Lagrangian coordinates ................... 76
3. The transport theorem .................................. 77
4. Conservation equations ................................. 78
5. Conservation of mass .................................... 79
6. Conservation of momentum .............................. 79
7. Conservation of energy .................................. 81
8. Constitutive relations and thermodynamic relations .. 82
   8.1 Heat conduction and mass diffusion ................ 84
   8.2 Newtonian viscous fluid ............................ 84
   8.3 Linear elastic and viscoelastic deformations ....... 86
9. Maxwell’s equations .................................... 88
   9.1 Constitutive relations .............................. 89
   9.2 Energy conservation and Poynting’s theorem ....... 90
   9.3 Electromagnetic waves and Lorentz’s force ........ 90
10. Discussion ............................................ 91

### Exercises ............................................ 92
## Contents

### 7 The art of modelling

1. Introduction ........................................ 95
2. Models ............................................. 97
3. Non-dimensionalisation and scaling ................. 102
   3.1 General concepts .................................. 102
   3.2 Dimensional Analysis ............................. 106
   3.3 Similarity solutions. ............................. 109
4. Scaling and reduction of the Navier-Stokes equations .... 112
   4.1 Scaling, non-dimensionalisation. ............... 113
   4.2 Some dimensionless groups with their common names . 114
   4.3 Asymptotic reductions of the Navier-Stokes equations . 115
5. Discussion ........................................ 118

Exercises ............................................. 118

### 8 The analysis of elliptic equations

1. The Laplace operator .................................. 121
   1.1 Problem types .................................... 122
   1.2 Uniqueness ...................................... 125
2. Eigenvalues and eigenfunctions ....................... 125
   2.1 The 1-D eigenvalue problem ...................... 126
   2.2 Eigenvalue problems in more dimensions ....... 128
3. Separation of variables ................................ 130
4. Fundamental solutions ................................ 133
5. Green’s functions; superposition ..................... 136
6. The maximum principle ................................ 139
7. The Stokes equations .................................. 141
8. Discussion ......................................... 145

Exercises ............................................. 145

17:22 7 Oct 2003 v version: 26-09-2003
## Contents

### 10 Analysis of parabolic equations

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cauchy problems</td>
<td>149-154</td>
</tr>
<tr>
<td>1.1</td>
<td>The heat equation in one space dimension</td>
<td>149</td>
</tr>
<tr>
<td>1.2</td>
<td>The heat equation in ( d ) space dimensions</td>
<td>153</td>
</tr>
<tr>
<td>1.3</td>
<td>Problems on half spaces</td>
<td>154</td>
</tr>
<tr>
<td>2</td>
<td>The heat equation with spatial symmetries</td>
<td>156</td>
</tr>
<tr>
<td>3</td>
<td>Similarity solutions</td>
<td>157</td>
</tr>
<tr>
<td>4</td>
<td>Initial boundary value problems</td>
<td>159</td>
</tr>
<tr>
<td>5</td>
<td>Moving boundaries; Stefan problems</td>
<td>162</td>
</tr>
<tr>
<td>6</td>
<td>Long-time behaviour of solutions</td>
<td>164</td>
</tr>
<tr>
<td>6.1</td>
<td>Linear initial boundary value problem</td>
<td>165</td>
</tr>
<tr>
<td>6.2</td>
<td>Equilibrium and travelling-wave solutions for nonlinear problems</td>
<td>168</td>
</tr>
</tbody>
</table>

**Exercises** | 171 |

### 12 Analysis of hyperbolic equations

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>First-order scalar equations</td>
<td>175-182</td>
</tr>
<tr>
<td>1.1</td>
<td>Semilinear equations</td>
<td>176</td>
</tr>
<tr>
<td>1.2</td>
<td>Quasilinear equations</td>
<td>178</td>
</tr>
<tr>
<td>1.3</td>
<td>Nonlinear equations</td>
<td>182</td>
</tr>
<tr>
<td>2</td>
<td>Weak formulation of first-order scalar equations</td>
<td>183</td>
</tr>
<tr>
<td>2.1</td>
<td>Weak solutions</td>
<td>183</td>
</tr>
<tr>
<td>2.2</td>
<td>The Riemann problem</td>
<td>187</td>
</tr>
<tr>
<td>3</td>
<td>First order systems</td>
<td>193-199</td>
</tr>
<tr>
<td>3.1</td>
<td>Linear systems</td>
<td>194</td>
</tr>
<tr>
<td>3.2</td>
<td>Quasilinear systems</td>
<td>196</td>
</tr>
<tr>
<td>3.3</td>
<td>Method of characteristics</td>
<td>199</td>
</tr>
<tr>
<td>4</td>
<td>Weak formulation of first order systems</td>
<td>201-203</td>
</tr>
<tr>
<td>4.1</td>
<td>Weak solutions</td>
<td>201</td>
</tr>
<tr>
<td>4.2</td>
<td>The Riemann problem</td>
<td>203</td>
</tr>
<tr>
<td>5</td>
<td>The shallow water equations</td>
<td>208</td>
</tr>
<tr>
<td>6</td>
<td>The wave equation</td>
<td>216-218</td>
</tr>
<tr>
<td>6.1</td>
<td>One-dimensional problems</td>
<td>216</td>
</tr>
<tr>
<td>6.2</td>
<td>Solutions in more dimensions</td>
<td>218</td>
</tr>
<tr>
<td>7</td>
<td>Boundary conditions</td>
<td>221</td>
</tr>
<tr>
<td>8</td>
<td>Discussion</td>
<td>226</td>
</tr>
</tbody>
</table>

**Exercises** | 226 |

---

17:22 7 Oct 2003     vi      version: 26-09-2003
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>15 Perturbation methods</strong></td>
<td>229</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>229</td>
</tr>
<tr>
<td>2 Asymptotic approximations and expansions</td>
<td>231</td>
</tr>
<tr>
<td>2.1 Asymptotic approximations</td>
<td>231</td>
</tr>
<tr>
<td>2.2 Asymptotic expansions</td>
<td>232</td>
</tr>
<tr>
<td>2.3 Perturbation problems</td>
<td>234</td>
</tr>
<tr>
<td>2.4 Asymptotic expansions of Poincaré type</td>
<td>236</td>
</tr>
<tr>
<td>3 Regular perturbation problems</td>
<td>237</td>
</tr>
<tr>
<td>3.1 Method of slow variation</td>
<td>238</td>
</tr>
<tr>
<td>3.2 Lindstedt-Poincaré method</td>
<td>243</td>
</tr>
<tr>
<td>4 Singular perturbation problems</td>
<td>245</td>
</tr>
<tr>
<td>4.1 Matched asymptotic expansions</td>
<td>245</td>
</tr>
<tr>
<td>4.2 Multiple scales</td>
<td>253</td>
</tr>
<tr>
<td>5 Discussion</td>
<td>263</td>
</tr>
<tr>
<td>Exercises</td>
<td>263</td>
</tr>
</tbody>
</table>

**Useful definitions and properties**

| A Asymptotic order symbols                   | 267  |
| B Trigonometric relations                    | 268  |
| C Convergence of series                      | 269  |
| D Multistep formulae                         | 270  |
| E Solution of recursions                     | 273  |
| F Eigenvalues and eigenvectors of a tridiagonal matrix | 274  |
| G Norms                                      | 276  |
| H Similarity                                 | 277  |
| I Estimates of eigenvalues and consequences  | 279  |
| J Theorems from vector calculus              | 280  |
| K Properties of cartesian, cylindrical and spherical co-ordinate systems | 281  |
| L Tensors                                    | 282  |
| M Dimensionless numbers                      | 284  |

**Bibliography**

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>287</td>
</tr>
</tbody>
</table>

**Index**

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>294</td>
</tr>
</tbody>
</table>
Glossary of Notation

Variables and operators

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>time</td>
</tr>
<tr>
<td>$x$</td>
<td>space coordinate in $\mathbb{R}$</td>
</tr>
<tr>
<td>$x = (x, y)^T, x = (x, y, z)^T$</td>
<td>space coordinate in $\mathbb{R}^d$ ($d = 2, 3$)</td>
</tr>
<tr>
<td>$v(x, t), v(x, t)$</td>
<td>scalar function</td>
</tr>
<tr>
<td>$\mathbf{v}(x, t), \mathbf{v}(x, t)$</td>
<td>vector function</td>
</tr>
<tr>
<td>$(u, v)$</td>
<td>inner product of the scalar functions $u$ and $v$</td>
</tr>
<tr>
<td>$\mathbf{u} \cdot \mathbf{v}$</td>
<td>inner product of the vectors $\mathbf{u}$ and $\mathbf{v}$</td>
</tr>
<tr>
<td>$\mathbf{u} \times \mathbf{v}$</td>
<td>vector product of the vectors $\mathbf{u}$ and $\mathbf{v}$</td>
</tr>
<tr>
<td>$\mathcal{L}(x, t)[v]$</td>
<td>differential operator $\mathcal{L}(x, t)$ applied to $v$</td>
</tr>
<tr>
<td>$\mathcal{L}^*(x, t)$</td>
<td>adjoint operator of $\mathcal{L}(x, t)$</td>
</tr>
<tr>
<td>$\frac{\partial}{\partial t} v, v'$</td>
<td>derivative of $v(x)$</td>
</tr>
<tr>
<td>$\frac{\partial}{\partial x} v_x$</td>
<td>partial derivative of $v(x, t)$</td>
</tr>
<tr>
<td>$\frac{\partial}{\partial n} n \cdot \nabla v$</td>
<td>directional derivative of $v$ in the direction of $n$</td>
</tr>
<tr>
<td>$\nabla v$</td>
<td>gradient of $v$</td>
</tr>
<tr>
<td>$\nabla \cdot v$</td>
<td>divergence of $v$</td>
</tr>
<tr>
<td>$\nabla \times v$</td>
<td>curl of $v$</td>
</tr>
<tr>
<td>$\nabla^2 v$</td>
<td>Laplace operator applied to $v$</td>
</tr>
<tr>
<td>$\int_\Omega v , dV$</td>
<td>generic integral over a domain $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$)</td>
</tr>
<tr>
<td>$\iint_\Omega v(x, t) , dx , dy$</td>
<td>integral over a domain $\Omega \subset \mathbb{R}^2$</td>
</tr>
<tr>
<td>$\int_{\partial \Omega} \mathbf{v} \cdot \mathbf{n} , dS$</td>
<td>integral over a closed surface $\partial \Omega \subset \mathbb{R}^3$</td>
</tr>
<tr>
<td>$\int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{l}$</td>
<td>integral over a closed contour $\mathcal{C} \subset \mathbb{R}^2$</td>
</tr>
<tr>
<td>$[v]^+$</td>
<td>jump of $v$ across a discontinuity</td>
</tr>
<tr>
<td>$\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$</td>
<td>unit vectors in the cartesian coordinate system $(x, y, z)$</td>
</tr>
<tr>
<td>$\mathbf{e}<em>r, \mathbf{e}</em>\theta, \mathbf{e}_z$</td>
<td>unit vectors in the cylindrical coordinate system $(r, \phi, z)$</td>
</tr>
<tr>
<td>$\mathbf{e}<em>r, \mathbf{e}</em>\theta, \mathbf{e}_\phi$</td>
<td>unit vectors in the spherical coordinate system $(r, \theta, \phi)$</td>
</tr>
</tbody>
</table>
GLOSSARY OF NOTATION

Numerical parameters and variables

$\Delta t$ \hspace{1cm} step size
$\Delta x$ \hspace{1cm} grid size in $x$-direction
$h$ \hspace{1cm} generic grid size
$x_j$ \hspace{1cm} $j$th grid point
$x_{j+\frac{1}{2}}$ \hspace{1cm} location of a control volume boundary
$t^n$ \hspace{1cm} time level $n\Delta t$
$v^n_j$ \hspace{1cm} numerical approximation of $v(x, t)$ at grid point $x_j$ and time level $t^n$
$F^n_{j+1/2}$ \hspace{1cm} numerical approximation of flux $f(x, t)$ at control volume boundary $x_{j+1/2}$ and time level $t^n$

C, E, etc. \hspace{1cm} location names of grid points

$x_c$ \hspace{1cm} generic grid point
$v_c$ \hspace{1cm} numerical approximation of $v(x_c)$
$v_\Delta$ \hspace{1cm} grid function
$L_\Delta[v^n_j]$ \hspace{1cm} difference operator $L_\Delta$ applied to $v^n_j$
$L_\Delta^{-1}$ \hspace{1cm} inverse of $L_\Delta$

$d^n_j = O(\Delta x^2)$ \hspace{1cm} $d^n_j$ is of the order $\Delta x^2$ for $\Delta x \to 0$

Vectors and matrices

$v = (v_1, v_2, \cdots , v_n)^T$ \hspace{1cm} column vector in $\mathbb{R}^n$
$v^T = (v_1, v_2, \cdots , v_n)$ \hspace{1cm} row vector in $\mathbb{R}^n$
$v^k$ \hspace{1cm} $k$th vector in a sequence
$\|v\|_p$ \hspace{1cm} $p$-norm of $v$
$A = (a_{ij})$ \hspace{1cm} matrix with $a_{ij}$ in $i$th row and $j$th column
$A = (a_1, a_2, \cdots , a_n)$ \hspace{1cm} matrix with $a_i$ as the $i$th column
$A^T$ \hspace{1cm} the transpose of $A$
$A^{-1}$ \hspace{1cm} the inverse of $A$
$I$ \hspace{1cm} identity matrix
diag$(a_1, a_2, \cdots , a_n)$ \hspace{1cm} diagonal matrix with $a_i$ in $i$th row and column
det$A$ \hspace{1cm} determinant of $A$
$\|A\|_p$ \hspace{1cm} $p$-norm of $A$
$\rho(A)$ \hspace{1cm} spectral radius of $A$
GLOSSARY OF NOTATION

Miscellaneous

\[ :=, =: \quad \text{is defined as, defines} \]
\[ \doteq \quad \text{is equal to when neglecting terms of higher order} \]
\[ \sim \quad \text{is asymptotically equal to} \]
\[ \mathcal{O} \quad \text{asymptotic order symbol ("big } \mathcal{O} \text{")} \]
\[ o \quad \text{asymptotic order symbol ("small } o \text{")} \]
\[ e \quad \text{base of the natural logarithm (} e = 2.71828 \ldots \text{)} \]
\[ i \quad \text{imaginary unit} \]
\[ |z| \quad \text{absolute value of } z \in \mathbb{C} \]
\[ \bar{z} \quad \text{complex conjugate of } z \in \mathbb{C} \]
\[ [v] \quad \text{dimension of variable/constant } v \text{ (e.g. in SI-units)} \]
\[ \mathcal{C} \quad \text{characteristic/curve in } \mathbb{R}^d \text{ (} d = 2, 3 \text{)} \]
\[ \partial \Omega \quad \text{boundary of a domain } \Omega \subset \mathbb{R}^d \text{ (} d = 2, 3 \text{)} \]
Chapter 1
Differential and difference equations

In this chapter we give a brief introduction to partial differential equations. In Section 1 some simple problems are derived to show that they may arise in simple daily life phenomena (a more detailed derivation of such problems will follow in later chapters). We show by a number of examples how they often may be seen as continuous analogues of discrete formulations (i.e. based on difference equations). In Section 2 we briefly overview the terminology used to describe various partial differential equations. Thus concepts like order and linearity are introduced. In Chapter 2 we shall discuss the classification of the various types of partial differentials in more detail. Finally, we introduce difference equations and notions like scheme and stencil, which play a rôle in numerical approximation, in Section 3.

1 Introduction

Many phenomena in nature may be described mathematically by functions of a small number of independent variables and parameters. In particular, if such a phenomenon is given by a function of spatial position and time, their description gives rise to a wealth of (mathematical) models, which often result in equations, usually containing a large variety of derivatives with respect to these variables. Apart from the spatial variable(s), which are essential for the problems to be considered, the time variable will play a special rôle. Indeed, many events exhibit gradual or rapid changes as time proceeds. They are said to have an evolutionary character and an essential part of their modelling is therefore based on causality, i.e. the situation at any time is dependent on the past. As far as (mathematical) modelling leads to partial differential equations (PDE), the latter will therefore be called evolutionary, i.e. involve the time $t$ as a variable. The other type of problems is often referred to as steady state. We will give some examples illustrating this background.

A typical PDE arises if one studies the flow of quantities like density, concentration, heat, etc. If there are no restoring forces, they usually have a tendency to spread out. In particular, one may e.g. think of particles with higher velocities (or rather energy), colliding with particles having lower velocities. The former are initially rather clustered. The energy
1. INTRODUCTION

will gradually spread out, mainly because they collide with other particles, thereby trans-
fering some of the energy. This is called dissipation. A similar effect can be observed for
mass dissolved in a fluid with concentrations varying in space. Brownian motion (again)
will gradually spread out the material over the entire domain. This is called diffusion.

Example 1.1 Consider a long tube of cross section \( A \) filled with water and a dye. Initially the
dye is concentrated in the middle. Let \( u(x, t) \) denote the concentration or density (mass per unit
length) of the dye at position \( x \) and time \( t \); then we see that in a small volume \( A \Delta x \), positioned
between \( x - \frac{1}{2} \Delta x \) and \( x + \frac{1}{2} \Delta x \) (Fig. 1.1), the total amount of dye equals approximately
\( u(x, t) \Delta x \).

Figure 1.1. Sketch of dye diffusion

with a corresponding dye concentration \( u(x + \Delta x, t) \). The mass that flows per unit time through
a cross section is called the mass flux. From the physics of solutions it is known that the dye
will move from the volume with the higher concentration to one with the lower concentration,
such that the mass flux \( f \) between the respective volumes is proportional to the difference in
concentration between both volumes, and thus given by

\[
f(x + \frac{1}{2} \Delta x, t) = \alpha (u(x + \frac{1}{2} \Delta x, t) - u(x, t)) \frac{u(x + \Delta x, t) - u(x, t)}{\Delta x},
\]

where \( \alpha \), the diffusion coefficient, is usually depending on \( u \). This relation is called Fick’s law
for mass transport by diffusion, which is the analogue of Fourier’s law for heat transport by
conduction.

As there is a similar flux between the center volume and its left neighbour, we have a rate
of change of total amount of mass in the center volume equal to the difference between both
fluxes, given by

\[
\frac{\partial}{\partial t} u(x, t) \Delta x = f(x + \frac{1}{2} \Delta x, t) - f(x - \frac{1}{2} \Delta x, t).
\]

If the diffusion coefficient \( \alpha \) is a constant, we have

\[
\frac{\partial}{\partial t} u(x, t) = \frac{\partial^2}{\partial x^2} u(x, t) - 2u(x, t) + u(x - \Delta x, t).
\]

(*)

By taking the limits for small volumes (i.e. \( \Delta x \to 0 \)) we find

\[
\frac{\partial}{\partial t} u(x, t) = \alpha \frac{\partial^2}{\partial x^2} u(x, t),
\]

which is called the one-dimensional diffusion equation. As heat conduction satisfies the same
equation, it is also called the heat equation if \( u \) denotes temperature.

Another kind of PDE is met in transport of particles. Here a flow typically has a
dominant direction; mutual collision of particles (which is felt globally as a kind of internal
friction, or viscosity) is neglected.

Example 1.2 Consider a road with heavy traffic moving in one direction, say $x$-direction
(Fig. 1.2). Let the number of cars at time $t$ on a stretch $[x, x + \Delta x]$ be denoted by $N(x,t)$. Furthermore, let the number of cars, passing at a point $x$ per time period $\Delta t$, be given by $f(x, t)\Delta t$. In that period, the number of cars, $N(x, t + \Delta t)$, can only be changed from a
difference between inflow at $x$ and outflow at $x + \Delta x$, i.e.
\[ \Delta N(x, t + \Delta t) = \Delta N(x, t) = (f(x + \Delta x, t) - f(x, t))\Delta t. \]

Rather than the number of cars $\Delta N$ per interval of length $\Delta x$, it is convenient to consider a 

\[ \text{car density} \ n(x, t), \text{which is defined by} \]
\[ \Delta N(x, t) = n(x, t)\Delta x. \]

Hence we obtain the relation
\[ \frac{n(x, t + \Delta t) - n(x, t)}{\Delta t} = -\frac{f(x + \Delta x, t) - f(x, t)}{\Delta x}. \]

Assuming sufficient smoothness (which implies that we have to allow for fractions of cars ...),
this leads in the limit of $\Delta t, \Delta x \to 0$ to
\[ \frac{\partial n}{\partial t} + \frac{\partial f}{\partial x} = 0. \]

which takes the form of a so-called conservation law. We may recognise $f$ again as a flux. If
this flux is only dependent of the local car density, i.e. $f = f(n)$, and $f$ is sufficiently smooth,
we obtain
\[ \frac{\partial n}{\partial t} + f'(n) \frac{\partial n}{\partial x} = 0, \]

also known as the transport equation. \hfill \square

An important class of problems arises from classical mechanics, i.e. Newtonian
systems.

Example 1.3 Consider a chain consisting of elements, each with mass $m$, and springs, with
spring constant $\beta > 0$ and length $\Delta x$, see Fig. 1.3. Denote the elements by $V_1, V_2, \ldots$ with
position of the masses $x = u_1, u_2, \ldots$ Assuming linear springs, the force necessary to increase
the original length $\Delta x$ of the spring of element $V_i$ by an amount $\delta u = u_i - u_{i-1} - \Delta x$ is equal to
$F_i = \beta \delta u$. Apart from the end points, all masses are free to move in the $x$-direction, their inertia
being balanced by the reaction forces of the springs. Noting that each element $V_i$ (except for
the end points) experiences a spring force from the neighbouring $i$-th and $i + 1$-th spring, we
have from Newton’s law for the $i$-th element
\[ m \frac{\partial^2 u_i}{\partial t^2} = F_i + F_{i+1} = \beta(u_i+1 + u_{i-1} - u_i), \quad i = 1, 2, \ldots . \]

\[ \text{(9)} \]
If the chain elements increase in number, while the springs and masses decrease in size, it is natural and indeed more convenient not to distinguish between each individual element, but to blend the discrete description of (\(\ast\)) into a continuous analogue. The small masses are conveniently described by a density \(\rho\) such that \(m = \rho \Delta x\), while the large spring constants are best described by a stiffness \(\sigma = \beta \Delta x\). Then we obtain from (\(\ast\)) for the position function \(u(x, t)\) the partial differential equation

\[
\frac{\partial^2 u}{\partial t^2} = \frac{\sigma}{\rho} \frac{\partial^2 u}{\partial x^2}.
\]

As solutions of this equation are typically wave-like, it is known as the wave equation, with a wave velocity equal to \(\sqrt{\sigma/\rho}\). In our example it describes longitudinal waves along the suspended chain of masses. In the context of pressure-density perturbations of a compressible fluid like air, the equation describes one-dimensional sound waves, for example as they occur in organ pipes. In that case the air stiffness is equal to \(\sigma = \gamma p\), where \(\gamma = 1.4\) is a gas constant and \(p\) is the atmospheric pressure (see Section 8.2).

As a last example we mention the analogue in electrical circuits of the motion of coupled spring-dashpot elements.

**Example 1.4** The time-behaviour of electric currents in a network may be described by the variables potential \(V\), current \(I\), and charge \(Q\). If the network is made of simple wires connecting isolated nodes, resistances, capacities and coils, and the frequencies are low, the network may be modelled (a posteriori confirmed by analysis of the Maxwell equations) one-dimensionally by a series of elements with the material properties resistance \(R\), capacitance \(C\), and inductance \(L\). Such a model is called an electrical circuit. If the frequencies are high, such that the wavelength is comparable with the length of conductors, we have to be more precise. As the signal cannot change instantaneously at all locations, it propagates as a wave of voltage and current along the line. In such a case we cannot neglect the resistance and inductance properties of the wires. By considering the wires as being built up from a series of (infinitesimally) small elements, the system can be modelled by what is called a transmission line, leading to partial differential equations in time and space.

In or across each element we have the following relations. The current is defined as the change of charge in time, \(I = \frac{d}{dt}Q\). The capacitance of a pair of conductors is given by \(C = Q/V\), where \(V\) is the potential difference and \(Q\) the charge difference between the conductors (Coulomb’s law). The resistance between two points is given by \(R = V/I\) where \(V\) is the potential difference between these points and \(I\) is the corresponding current (Ohm’s law). A changing electromagnetic current in a coil with inductance \(L\) induces a counter-acting...
potential, given by \( V = -L \frac{d^2 I}{dt^2} \) (Faraday’s law). At a junction, no charge can accumulate, and we have the condition \( \sum I = 0 \), while around a loop the summed potential vanishes \( \sum V = 0 \) (Kirchhoff’s laws). With these building blocks we can construct transmission line models.

A famous example is the \textit{telegraph equation} where an infinitesimal piece of telegraph wire is modelled (Fig. 1.4) as an electrical circuit, consisting of a resistance \( R \Delta x \) and an inductance \( L \Delta x \), while it is connected to the ground via a resistance \( (G \Delta x)^{-1} \) and a capacitance \( C \Delta x \).

Let \( i(x, t) \) and \( u(x, t) \) denote the current and voltage through the wire at position \( x \) and time \( t \).

\[
\begin{align*}
&x \quad R \Delta x \quad L \Delta x \quad x + \Delta x \\
&\quad C \Delta x \quad (G \Delta x)^{-1}
\end{align*}
\]

\textbf{Figure 1.4. A transmission line model of a telegraph wire}

The change of voltage across the piece of wire is now given by

\[
u(x + \Delta x, t) - u(x, t) = \left[ -i R \Delta x - \frac{\partial i}{\partial t} L \Delta x \right]_{x+\Delta x}.
\]

The amount of current that disappears via the ground is

\[
i(x + \Delta x, t) - i(x, t) = \left[ -u G \Delta x - \frac{\partial u}{\partial t} C \Delta x \right].
\]

By taking the limit \( \Delta x \to 0 \) we get

\[
\frac{\partial u}{\partial x} = -Ri - L \frac{\partial i}{\partial t}, \quad \frac{\partial i}{\partial x} = -Gu - C \frac{\partial u}{\partial t}.
\]

By eliminating \( i \) this may be combined into the telegraph equation for \( u \)

\[
\frac{\partial^2 u}{\partial x^2} = LC \frac{\partial^2 u}{\partial t^2} + (LG + RC) \frac{\partial u}{\partial t} + RGu.
\]

\(\star\)

\textbf{Example 1.5} Consider the following crowd of \( N^2 \) very accommodating people, for convenience ordered in a rectangular square of size \( L \times L \), while each person, labelled by \((i, j)\), is positioned at \( x_i = ih, y_j = jh \) with \( h = L/N \). Each person has an opinion given by the (scalar) number \( p_{ij} \) and can only communicate with its immediate neighbours. Assume that each person tries to minimize any conflict with its neighbours and is willing to take an opinion which is the average of its neighbours’ opinions. So we have

\[
p_{ij} = \frac{1}{4} (p_{i+1,j} + p_{i-1,j} + p_{i,j+1} + p_{i,j-1}).
\]

\(\star\)

Only at the borders of the rectangle the individuals are provided with information such that \( p \) is fixed.
2. Nomenclature

In the previous section we met a number of equations with derivatives with respect to more than one variable. In general, such an equation is called a partial differential equation. Let \( x \) and \( t \) be two independent variables and let \( u(x, t) \) denote a quantity depending on \( x \) and \( t \). Furthermore, let

\[
t \in [0, T], \quad 0 \leq T \leq \infty, \quad x \in [a, b] \subset \mathbb{R}. \tag{2.1}
\]

For an integer \( n \), a general form for a scalar PDE (in two independent variables) reads

\[
F \left( \frac{\partial^n u}{\partial t^n}, \frac{\partial^n u}{\partial t \partial x^{n-1}}, \ldots, \frac{\partial^{n-1} u}{\partial x^{n-1}}, \ldots, \frac{\partial^{n-1} u}{\partial t^{n-1}}, \ldots, \frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, u, x, t \right) = 0. \tag{2.2}
\]
CHAPTER 1. DIFFERENTIAL AND DIFFERENCE EQUATIONS

The highest order derivative is called the order of the PDE; not all partial derivatives (except the highest of at least one variable) need to be present. The form (2.2) is an implicit formulation, i.e. the highest order derivative(s), the so-called principal part, do(es) not appear explicitly. If the latter is the case we call it an explicit PDE. The generalization to more than two independent variables is obvious.

Example 1.6 Some important examples of PDE’s are:

(i) \[
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{1}{6} \frac{\partial^3 u}{\partial x^3} = 0
\] (Korteweg-de Vries equation).

This is a third order PDE.

(ii) \[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0
\] (nonlinear transport equation).

If \( f \) is differentiable we see that this is a first order PDE in \( u \).

(iii) \[
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \varepsilon \frac{\partial^2 u}{\partial x^2}
\] (Burgers’ equation).

If \( \varepsilon = 0 \) this may be referred to as the inviscid Burgers’ equation, which is a special case of the transport equation.

(iv) \[
\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} + \frac{1}{3} \varepsilon^2 \frac{\partial^4 u}{\partial x^2 \partial t^2} = 0
\] (linearized Boussinesq equation).

(v) \[
EI \frac{\partial^4 u}{\partial x^4} - T \frac{\partial^2 u}{\partial x^2} + m \frac{\partial^2 u}{\partial t^2} = 0
\] (vibrating beam equation).

(vi) \[
\frac{\partial u}{\partial y \partial y} - \frac{\partial}{\partial x} D(u) \frac{\partial u}{\partial x} = f(x)
\] (Prandtl’s boundary layer equation).

In quite a few cases the order can only be deduced after some (though trivial) manipulation.

Example 1.7

\[
\frac{\partial u}{\partial t} - \frac{\partial}{\partial x} \left( D(u) \frac{\partial u}{\partial x} \right) = f(x)
\] (nonlinear diffusion equation).

It is clear that this PDE is second order. There is no analytical, numerical or practical need to rework this and have \( \frac{\partial^2 u}{\partial x^2} \) appear explicitly.

Usually, the variables are space and/or time. Although the variables in (2.2) are generic, we shall use the symbol \( t \) to indicate the time variable in general. The variable \( x \) will refer to space. There are major differences between problems where time does and where it does not play a rôle. If the time is not explicitly there, the problem is referred to as a steady state problem. If the PDE possesses solutions which evolve explicitly with \( t \) we call it an evolutionary problem, i.e. there is causality. Most of the theory will be devoted to problems in one space variable. However, occasionally we shall encounter more than one such space variable. Fortunately, problems in more such variables often have many analogues of the one-dimensional case. We shall indicate vectors by boldface characters. So in higher dimensional space the space variable is denoted by \( x \), or by \( (x, y, z) \). The PDE can still be scalar. We have obvious analogues for vectorial dependent variables of the foregoing.
3. DIFFERENCE EQUATIONS

Example 1.8 A few other examples are as follows.

(i) \( \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0 \) (heat equation in 3-D).

We prefer to write this as \( \frac{\partial u}{\partial t} - \nabla^2 u = 0 \). \( \nabla^2 \) is referred to as the Laplace operator.

(ii) \( \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = 0 \) (wave equation in 3-D).

(iii) \( \nabla^2 u + k^2 u = 0 \) (Helmholtz or reduced wave equation).

(iv) \( (1 - M^2) \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \) (equation for small perturbations in steady subsonic \( (M^2 < 1) \) or supersonic \( (M^2 > 1) \) flow).

Sometimes one also denotes a partial derivative of a certain variable by an index, like

\[ u_t := \frac{\partial u}{\partial t}, \quad u_{tx} := \frac{\partial^2 u}{\partial t \partial x}. \]  

(2.3)

If we can write (2.2) as a linear combination of derivatives of \( u \) with respect to \( x \) and \( t \), and with coefficients only depending on \( x \) and \( t \), the PDE is called linear. Moreover, it is called homogeneous if it does not depend explicitly on \( x \) and/or \( t \). If the PDE is a linear combination of derivatives but the coefficients of the highest derivative, say \( n \), depend on \((n-1)\)-st order derivatives at most, then we call it quasilinear [26].

Like any differential equation we have to prescribe certain initial conditions (IC) and boundary conditions (BC) for the time and space variable(s) respectively. In evolutionary problems they often appear both as initial boundary conditions (IBC). We shall encounter various types and combinations in later chapters.

We finally remark that we may look for solutions that satisfy the PDE in a weak sense. In particular the derivatives may not exist everywhere on the domain of interest. Again we refer to later chapters for further details.

3 Difference equations

Initially, the actual form of the equations we derived in the examples in Section 1 was of a difference equation. Like a partial differential equation we may define a partial difference equation as any relation between values of \( u(x, t) \), where \((x, t) \in \mathcal{F} \subset [a, b] \times [0, T) \), \( \mathcal{F} \) being a finite set of points of the domain \([a, b] \times [0, T) \). We shall encounter difference equations when solving a PDE numerically and therefore it should approximate the PDE in some well defined way. The simplest way to describe the latter is by defining a scheme, i.e. a discrete analogue of the (continuous) PDE. Since we shall mainly deal with finite difference approximations in this book, we perceive a scheme as the result of replacing the differentials by finite differences. To this end we have to indicate some (generic) points in the domain \([a, b] \times [0, T) \), at which the function values \( u(x, t) \) are taken. The latter set of points is called a stencil. We shall clarify this by some examples.

Example 1.9
(i) Consider Example 1.1 again. If we replace in equation (\(\ast\)) \[\frac{\partial}{\partial t} u(x, t)\] by a straightforward discretisation, then we obtain the scheme
\[
\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = \alpha \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{\Delta x^2},
\]
and the stencil is the set of bullets (\(\bullet\)) in Fig. 1.6.

![Figure 1.6. Stencil of Example 1.9 (i)](image)

(ii) Consider the wave equation (\(\ast\ast\)) of Example 1.3. From equation (\(\ast\)) a discrete version may be found to be
\[
\frac{u(x, t + \Delta t) - 2u(x, t) + u(x, t - \Delta t)}{\Delta t^2} = \frac{\sigma}{\rho} \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{\Delta x^2}.
\]
The stencil is given in Fig. 1.7.

![Figure 1.7. Stencil of Example 1.9 (ii)](image)

Given the special rôle of time and the implication it has for the actual computation which should be based on the causality of the problem, we may distinguish schemes according to the number of time levels involved. If \((k + 1)\) such time levels are involved, we call the scheme a \(k\)-step scheme. If it involves only spatial differences at earlier time levels, it is called explicit, otherwise implicit.
4. DISCUSSION

Example 1.10

(i) The schemes in Example 1.9 are both explicit, the first one being a one-step and the second one a two-step scheme.

(ii) We may as well approximate the $u_{xx}$-term in the heat equation at time-level $t + \Delta t$, and obtain the scheme

$$\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = \frac{u(x + \Delta x, t + \Delta t) - 2u(x, t + \Delta t) + u(x - \Delta x, t + \Delta t)}{\Delta x^2}.$$  

This scheme has the stencil given in Fig. 1.8. Clearly, it is an implicit one-step scheme.

![Stencil of Example 1.10 (ii)](image)

Figure 1.8. Stencil of Example 1.10 (ii)

4 Discussion

- The use of the variables $x$ and $y$ in an equation does not mean that the partial differential equation cannot have an evolutionary character. There are some cases where they refer to spatial coordinates, yet the corresponding equation may be hyperbolic, a type of equation we will encounter in the next chapter as an instance of evolutionary type.

- If in a system of time-dependent partial differential equations all spatial derivatives are replaced by suitable difference approximations, we obtain a system of ordinary differential equations in time. If one of the partial differential equations is independent of time, we obtain a so-called differential-algebraic system. A typical example is the condition that an incompressible flow is divergence free (equivalent to conservation of mass), like in the Stokes equations. This problem will be discussed in Sections 7 and 4.52.

Exercises

1.1. Show that a nonconstant diffusivity $\alpha(u)$ leads to the equation

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left( \alpha(u) \frac{\partial u}{\partial x} \right).$$
1.2. Determine the order of the *eikonal* equation
\[
\left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial u}{\partial z} \right)^2 = c^2.
\]

1.3. Determine the order of the PDE
\[
\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial y^2}.
\]
Derive a first order system by writing \( p := \frac{\partial u}{\partial x}, \ q := \frac{\partial u}{\partial y}. \)

1.4. Determine the order of the PDE (where \( a \) and \( b \) are parameters)
\[
\frac{\partial u}{\partial t} = a \nabla^2 u + b \frac{\partial u}{\partial x} + c(u).
\]

1.5. Verify that the solution \( u = u(x, t) \) of the transport equation (cf. Example 1.2 or 1.6 (ii))
\[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0, \quad u(x, 0) = U(x),
\]
for sufficiently smooth \( f \) is implicitly given by
\[
u = U(x - f'(u)t).
\]
Chapter 2

Characterisation and classification

The study of PDEs is quite diverse. Therefore it makes sense to first characterise them according to certain properties that will provide us guidelines to investigate them further. It turns out to be useful to start with first order equations in two independent variables. Therefore we start in Section 1 with describing scalar first order equations, thereby introducing the important notion of characteristics. This is generalised for first order systems in Section 2, leading to the definition of hyperbolicity. A well-known class of PDEs consists of second order scalar equations. In Section 3 we reformulate them as a first order system of equations and then discuss the classification in hyperbolic parabolic and elliptic equations. Quite naturally, this can be generalised to more (space) dimensions as is shown in Section 4. Sometimes the underlying structure of a problem is simpler than suggested, and after a suitable transformation the PDE may be transformed into an ODE. Examples are given in Section 5. PDEs need further conditions to make their solutions meaningfully exist and (hopefully) unique. In Section 6 we briefly deal with the question how to properly choose the initial and boundary conditions from a more theoretical point of view. For this we have the Hadamard condition, which states the conditions for a problem to be well-posed.

1 First order scalar partial differential equations in two independent variables

Consider the quasilinear (explicit) first order PDE

\[ a(x, t, u) \frac{\partial u}{\partial t} + b(x, t, u) \frac{\partial u}{\partial x} = c(x, t, u). \]  

(1.1)

Usually, the independent variables \( x \) and \( t \) denote a space coordinate and time, respectively, although strictly speaking, \( t \) might denote a space coordinate as well. Let \( u = \varphi(x, t) \) be a solution of (1.1). A geometrical interpretation of this solution is as follows. The independent variables \( x \) and \( t \) and the dependent variable \( u \) constitute a two parameter family of vectors \((x, t, u)^T\), which is lying on a surface \( S \subset \mathbb{R}^3 \). This surface \( S \), given by...
1. FIRST ORDER SCALAR PARTIAL DIFFERENTIAL EQUATIONS

\[ F(x, t, u) := \varphi(x, t) - u = 0, \]

is called the integral surface of (1.1). A normal \( n \) on \( S \) is given by

\[ n := \nabla F = \left( \frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial t}, -1 \right)^T. \]

Hence, for an infinitesimal displacement \( du := (dx, dt, du)^T \) along the surface, we find

\[ n \cdot du = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial t} dt - du = 0. \]

Comparing (1.1) and (1.3), we conclude that for a solution \( u = \varphi(x, t) \) of (1.1) the following should hold on the integral surface \( S \)

\[
\begin{pmatrix}
a & b \\
dt & dx \\
\frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial t}
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c \\
du
\end{pmatrix} = 0.
\]

The solution of this system is unique if and only if \( a \, dx - b \, dt \neq 0 \). In the following we will simply write \( u = u(x, t) \) instead of \( u = \varphi(x, t) \) to indicate the solution of (1.1).

This result can be interpreted as follows (see Figure 2.1): suppose we have a smooth, one-parameter curve \( J = \{ (x(\sigma), t(\sigma), u(\sigma) \mid \sigma \in I \subset \mathbb{R} \} \) on \( S \) where the condition \( a \, dx - b \, dt \neq 0 \) holds. Then the derivatives \( u_x \) and \( u_t \) are uniquely determined on \( J \) through (1.4). If, moreover, \( u \) is given along \( J \), then the solution \( u = u(x, t) \) exists and is unique, at least in some neighbourhood of \( J \). The curve \( J \) is referred to as a curve of initial values, or briefly, an initial curve.

The actual construction of the solution proceeds as follows. Suppose \( u \) is given along an initial curve \( J \). Consider a curve \( C \) on \( S \) for which \( a \, dx - b \, dt = 0 \). Then system (1.4) has either no solution or infinitely many. In the latter case the relations

\[ \frac{dt}{a} = \frac{dx}{b} = \frac{du}{c}, \]

should hold along \( C \). Clearly, the vector \( (a, b, c)^T \) is everywhere tangent to \( C \). We can now introduce a parametrization \( C = \{ (x(s), t(s), u(s)) \mid s \in I \subset \mathbb{R} \} \), such that \( ds = dt/a = dx/b = du/c \) and \( s = 0 \) on the initial curve \( J \). This way we obtain the following set of ODEs

\[ \frac{dt}{ds} = a, \quad \frac{dx}{ds} = b, \quad \frac{du}{ds} = c, \]

(1.6a)

coupled with an initial condition of the form

\[ u(0; \sigma) = v(x(\sigma), t(\sigma)), \quad \text{for} \quad (x(\sigma), t(\sigma)) \in J', \]

(1.6b)

where \( J' \) is the projection of \( J \) on the \((x, t)\)-plane. The set of ODEs (1.6a) is referred to as the characteristic equations. Consequently, the curve \( C \) is a solution of (1.6). \( C \) is called a characteristic and its projection on the \((x, t)\)-plane a base characteristic. Note however, that there is no uniformity in the nomenclature in literature; usually, no distinction is made between characteristics and base characteristics. In order to construct the integral surface,
we compute for each point on the initial curve $\mathcal{J}$ the characteristic passing through that point from (1.6). We formally obtain the solution

$$t = t(s; \sigma), \quad x = x(s; \sigma), \quad u = u(s; \sigma).$$

Inverting the first two relations, we find $s = s(t, x), \sigma = \sigma(t, x)$ and substitution of these in the expression for $u$ gives $u(x, t) := u(s(t, x); \sigma(t, x))$. This inversion is only possible if the Jacobian $t_\sigma x_s - t_s x_\sigma \neq 0$. Thus, the integral surface $\mathcal{S}$ is generated by a one-parameter family of characteristics $\mathcal{C}$ all passing through an initial curve $\mathcal{J}$. We will demonstrate this by an example.

**Example 2.1** Consider the following initial value problem for the inviscid Burgers’ equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = v(x) := \begin{cases} 1 & \text{if } x \leq 0, \\ 1 - x & \text{if } 0 < x \leq 1, \\ 0 & \text{if } x \geq 1. \end{cases}$$

The characteristic equations and the corresponding initial conditions read

$$\frac{dr}{ds} = 1, \quad \frac{dx}{ds} = u, \quad \frac{du}{ds} = 0, \quad t(0; \sigma) = 0, \quad x(0; \sigma) = \sigma, \quad u(0; \sigma) = v(\sigma).$$

![Figure 2.1. Initial curve $\mathcal{J}$ and a characteristic $\mathcal{C}$ on the integral surface $\mathcal{S}$](image)
1. FIRST ORDER SCALAR PARTIAL DIFFERENTIAL EQUATIONS

The solution of this set of ODE is given by

\[ t(s; \sigma) = s, \quad x(s; \sigma) = \sigma + v(\sigma)s, \quad u(s; \sigma) = v(\sigma). \]

We can easily invert the first two relations provided the Jacobian

\[ t_\sigma x_s - ts_x = -1 - v'(\sigma) \neq 0. \]

This way we obtain

\[ s(x, t) = t \quad \text{and} \quad \sigma(x, t) = x - t \quad \text{for} \quad x \leq t, \quad \sigma(x, t) = (x - t)/(1 - t) \quad \text{for} \quad t < x \leq 1 \quad \text{and} \quad \sigma(x, t) = x \quad \text{for} \quad x \geq 1. \]

Consequently, the solution is defined for \( 0 < t < 1 \) and is given by

\[ u(x, t) = \begin{cases} 
1 & \text{if} \quad x \leq t, \\
1 - x & \text{if} \quad t < x \leq 1, \\
0 & \text{if} \quad x \geq 1.
\end{cases} \]

Instead of \( s \), we can parametrize the characteristics by the variable \( t \), provided \( a \neq 0 \). We thus obtain the ODEs

\[
\frac{dx}{dt} = \frac{b}{a}, \quad \frac{du}{dt} = \frac{c}{a}, \tag{1.7}
\]

We also refer to (1.7) as the characteristic equations. The first equation gives the location of the base characteristics, possibly depending on the solution \( u \), and the second gives \( u \) along the base characteristics. We see that existence of solutions of (1.1) can be established from studying the ODEs (1.7). It is easy to see that such a solution is composed of solutions of initial value problems defined along base characteristics. If the right hand side of (1.7) satisfies a Lipschitz condition, then (1.7) together with an initial value for \( u \) determines the solution \( u \) on \( C \).

**Example 2.2** Consider the transport equation

\[
\frac{\partial u}{\partial t} + b(u)\frac{\partial u}{\partial x} = 0, \tag{*}
\]

subject to an initial condition of the form \( u(x, 0) = v(x) \). For this equation, system (1.7) reduces to

\[
\frac{dx}{dt} = b(u), \quad \frac{du}{dt} = 0,
\]

implying that \( u(x, t) = \text{Const} \) along the base characteristics, which however depend on the solution we seek. Since \( u(x, t) = \text{Const} \), also \( x - b(u)t = \text{Const} \), and we obtain the following (implicit) representation of the solution

\[ u(x, t) = v(x - b(u(x, t))t). \]

In the special case of the linear advection equation, for which \( b(u) = b = \text{Const} \), this representation reduces to \( u(x, t) = v(x - bt) \), i.e. the initial profile is propagated undisturbed with speed \( b \) along the base characteristics; see Figure 2.2.
CHAPTER 2. CHARACTERISATION AND CLASSIFICATION

2 First order linear systems in two independent variables

The previous theory for scalar equations can be extended to systems. In the sequel of this text we will often encounter systems of the form

\[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = c, \tag{2.8}
\]

where \( u : I_1 \times I_2 \to \mathbb{R}^n, I_1 \subset \mathbb{R} \) and \( I_2 \subset [0, \infty) \). Introducing the Jacobi matrix

\[
B := \frac{\partial f}{\partial u} = \left( \frac{\partial f_i(u)}{\partial u_j} \right),
\]

we find the quasilinear first order system of PDEs

\[
\frac{\partial u}{\partial t} + B \frac{\partial u}{\partial x} = c. \tag{2.9}
\]

More generally a quasilinear, first order system of PDEs may read

\[
A \frac{\partial u}{\partial t} + B \frac{\partial u}{\partial x} = c. \tag{2.10}
\]

In this section we assume that the coefficient matrices \( A \) and \( B \) are constant and that \( c = c(x, t, u) \). The quasilinear case where \( A = A(x, t, u) \) and \( B = B(x, t, u) \) is discussed in Chapter 12.

In order to employ the theory of the previous section, we try to decouple system (2.10) in \( n \) scalar equations. For the sake of simplicity we further assume that the matrix \( A \) is nonsingular. Then we look for a nonsingular transformation matrix \( S \) such that

\[
S^{-1}(BA^{-1})S = A, \tag{2.11}
\]

where \( A \) is a diagonal matrix. If such \( S \) exists we can introduce the characteristic variable \( \tilde{u} \) defined by

\[
\tilde{u}(x, t) := S^{-1}Au(x, t), \tag{2.12}
\]
satisfying the decoupled system
\[
\frac{\partial \tilde{u}}{\partial t} + A \frac{\partial \tilde{u}}{\partial x} = \tilde{c} := S^{-1} c.
\] (2.13a)

Written componentwise we have the scalar equations
\[
\frac{\partial \tilde{u}_k}{\partial t} + \lambda_k \frac{\partial \tilde{u}_k}{\partial x} = \tilde{c}_k, \quad k = 1, 2, \ldots, n.
\] (2.13b)

The previous derivation gives rise to the following

**Definition 2.3.** The linear system (2.10) with nonsingular matrix \( A \) is called hyperbolic, if \( BA^{-1} \) has \( n \) real eigenvalues and \( n \) linearly independent eigenvectors.

Consequently, when system (2.10) is hyperbolic, the matrices \( A \) and \( S \) do exist and are given by
\[
A := \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \quad S := (s_1, s_2, \ldots, s_n),
\] (2.14)
where \( \lambda_k \) and \( s_k \) are the eigenvalues and corresponding eigenvectors of \( BA^{-1} \), respectively. Thus the \( k \)th column of \( S \) is the eigenvector \( s_k \). Note that (2.10) is always hyperbolic if \( BA^{-1} \) is symmetric; for a general matrix hyperbolicity is assured if all eigenvalues are real and distinct.

Each equation in (2.13b) induces a characteristic \( C_k \) corresponding to the eigenvalue \( \lambda_k \) and eigenvector \( s_k \). The characteristic equations (1.7) in this case read
\[
\frac{dx}{dt} = \lambda_k, \quad \frac{d\tilde{u}_k}{dt} = \tilde{c}_k.
\] (2.15)

The solution of (2.10) can be composed of solutions of (2.15). This is demonstrated in the following.

**Example 2.4** Consider the initial value problem
\[
\begin{pmatrix}
1 & 0 \\
0 & 2
\end{pmatrix} \frac{\partial u}{\partial t} + \begin{pmatrix}
1 & 4 \\
1 & 0
\end{pmatrix} \frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R}, t > 0,
\]
\[
u(x, 0) = v(x), \quad x \in \mathbb{R}.
\]

We can easily verify that the eigenvalues and corresponding eigenvectors of \( BA^{-1} \) are given by
\[
\lambda_1 = -1, \lambda_2 = 2, \quad s_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.
\]

The characteristic variable \( \tilde{u} \) defined in (2.12) is now given by
\[
\tilde{u}_1 = \frac{1}{3}(u_1 - 4u_2), \quad \tilde{u}_2 = \frac{1}{3}(u_1 + 2u_2).
\]
These variables can be computed from (2.15) and we find \( \tilde{u}_1(x, t) = \tilde{u}_1(x + t, 0) \) and \( \tilde{u}_2(x, t) = \tilde{u}_2(x - 2t, 0) \). Combining this result with the above relations for the characteristic variables, we obtain the the final solution

\[
\begin{align*}
    u_1(x, t) &= \frac{1}{3}(v_1(x + t) - 4v_2(x + t) + 2v_1(x - 2t) + 4v_2(x - 2t)), \\
    u_2(x, t) &= \frac{1}{6}(-v_1(x + t) + 4v_2(x + t) + v_1(x - 2t) + 2v_2(x - 2t)).
\end{align*}
\]

Clearly, the solution contains waves propagating along base characteristics \( x + t = \text{Const} \) and \( x - 2t = \text{Const} \), respectively. \( \square \)

In general we conclude that (2.10) should not be subject to an initial condition prescribed on a characteristic. In fact, one should prescribe \( u \) on a curve \( \mathcal{J} \) that does not intersect any of these characteristics twice.

Finally, we introduce the following notions; see Figure 2.3. The domain of dependence of a point \( (x_0, t_0) \) is the region in the \( (x, t) \)-plane such that \( u(x_0, t_0) \) depends on all values \( u(x, t) \) with \( (x, t) \) in this domain. It is bounded by the two extreme characteristics through \( (x_0, t_0) \) facing back to the initial line \( t = 0 \). On the other hand, the domain of influence of \( (x_0, t_0) \) is the region in the \( (x, t) \)-space where the solution is influenced by \( u(x_0, t_0) \).

In the next section we shall consider the special case of systems arising from scalar second order problems.

![Figure 2.3. Region of influence and region of dependence. In case of constant coefficients these characteristics are straight lines.](image)

3 Second order scalar partial differential equations in two independent variables

For PDEs with higher order derivatives the classification may be reduced to first order systems met before, at least if they are scalar. In particular, we consider the second order
linear equation
\[ a \frac{\partial^2 u}{\partial t^2} + b \frac{\partial^2 u}{\partial t \partial x} + c \frac{\partial^2 u}{\partial x^2} + d \frac{\partial u}{\partial t} + e \frac{\partial u}{\partial x} = f, \]  
(3.16)

where the coefficients \( a, b, \ldots, e \) are assumed constant and where the right hand side \( f \) possibly depends on \( x, t \) and \( u \). The independent variable \( x \) is a space coordinate, whereas \( t \) is either time or a space coordinate. Introducing the variables

\[ p := \frac{\partial u}{\partial t}, \quad q := \frac{\partial u}{\partial x}, \]  
(3.17)

we obtain the following linear system

\[ \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} p \\ q \end{pmatrix} + \begin{pmatrix} b & c \\ -1 & 0 \end{pmatrix} \frac{\partial}{\partial x} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} f - d p - e q \\ 0 \end{pmatrix}. \]  
(3.18)

Note that this form is not unique. Clearly, this system is of the form (2.10), with the coefficient matrices given by

\[ A := \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \quad B := \begin{pmatrix} b & c \\ -1 & 0 \end{pmatrix}. \]  
(3.19)

We now investigate the eigenvalues and eigenvectors of this system.

In the following, we assume that \( a \neq 0 \), so that \( A \) is nonsingular. Like in the previous section, we look for a transformation matrix \( S \) that can diagonalise the matrix

\[ BA^{-1} = \begin{pmatrix} b/a & c/a \\ 1/a & 0 \end{pmatrix}. \]  
(3.20)

This will succeed if the characteristic equation

\[ \det(BA^{-1} - \lambda I) = \lambda^2 - \frac{b}{a} \lambda + \frac{c}{a} = 0 \]  
(3.21)

has two different, real roots, since then the corresponding eigenvectors are linearly independent. Consequently, system (3.18) is hyperbolic if \( b^2 - 4ac > 0 \). If \( b^2 - 4ac = 0 \) we have a degeneracy of the eigensystem and so only one “double” characteristic exists; in fact no \( S \) and \( A \), as required in (2.11), can be found. If, finally, \( b^2 < 4ac \) there are no real characteristic values at all. This leads to the following definition.

**Definition 2.5.** The partial differential equation (3.16) is called

(i) **hyperbolic**, if \( b^2 - 4ac > 0 \);
(ii) **parabolic**, if \( b^2 - 4ac = 0 \);
(iii) **elliptic**, if \( b^2 - 4ac < 0 \).
The nomenclature in this definition is adopted from the theory of quadratic forms. In particular, the corresponding quadratic curve \( at^2 + btx + cx^2 + dt + ex = \text{Const} \) is a hyperbola, parabola or ellipse, depending on the value of \( b^2 - 4ac \).

Next, we will derive the normal, or canonical, form of equation (3.16) in these three different cases, which only depends on the principal part of the equation, i.e. the first three terms containing the second order derivatives.

In the hyperbolic case we find a transformation matrix
\[
S = \begin{pmatrix}
\lambda_1 & \lambda_2 \\
\frac{1}{a} & \frac{1}{a}
\end{pmatrix},
\]
where \( \lambda_1 < \lambda_2 \) are the eigenvalues of \( BA^{-1} \). The characteristic variables \( \tilde{u} \) defined in (2.12) are now given by
\[
\tilde{u} = a \frac{\lambda_2 - \lambda_1}{p - p + \lambda_1 q - \lambda_2 q}.
\]

Inserting the first component \( \tilde{u}_1 \) into (2.13b), we obtain an equation of the form
\[
\left( \frac{\partial}{\partial t} + \lambda_1 \frac{\partial}{\partial x} \right) \left( \frac{\partial}{\partial t} + \lambda_2 \frac{\partial}{\partial x} \right) u = \tilde{f}.
\]

Taking into account the characteristic equations (2.15), we see that the first and second differential operator in (3.24) are just internal derivatives along characteristics of the \( C_1 \) and \( C_2 \)-family, respectively. Introducing coordinates \( \xi \) and \( \eta \) along these characteristics, we obtain the canonical form
\[
\frac{\partial^2 u}{\partial \xi \partial \eta} = \tilde{f}.
\]

Note that we would obtain the same equation starting from the equation for the second characteristic variable \( \tilde{u}_2 \).

In the parabolic case there is no transformation matrix \( S \) possible that can diagonalise \( BA^{-1} \). However, we have the Jordan normal form (see Appendix H)
\[
S^{-1}(BA^{-1})S = J := \begin{pmatrix}
\lambda & 1 \\
0 & \lambda
\end{pmatrix},
\]
with \( \lambda = b/(2a) \) the eigenvalue of \( BA^{-1} \), having algebraic multiplicity 2 and geometric multiplicity 1. A possible matrix \( S \) is given by
\[
S = \begin{pmatrix}
\lambda & 1 \\
\frac{1}{a} & 0
\end{pmatrix}.
\]

We can now reduce system (3.18) to
\[
\frac{\partial \tilde{u}}{\partial t} + \frac{\partial \tilde{u}}{\partial x} = \tilde{f},
\]
for the variable $\tilde{u}$, given by

$$\tilde{u} = a \begin{pmatrix} -q \\ p + \lambda q \end{pmatrix}. \quad (3.29)$$

From equation (2.13b) for $\tilde{u}$ we easily deduce

$$\left( 2 \frac{\partial}{\partial t} + \lambda \frac{\partial}{\partial x} \right)^2 u = f, \quad (3.30)$$

in which we recognize an internal differentiation along a characteristic. The canonical form for (3.16) is thus given by

$$\frac{d^2 u}{d\xi^2} = f, \quad (3.31)$$

where $\xi$ is the coordinate along the characteristic.

Finally, in the elliptic case, we have complex characteristics and the transformation matrix $S$ is also complex. Completely analogous to the hyperbolic case, we obtain the form

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = f, \quad (3.32)$$

with $\xi$ and $\eta$ the (complex) coordinates along characteristics of the $C_1$ and $C_2$-family, respectively. One can prove that $\eta = \overline{\xi}$; see e.g. [26]. Introducing the new coordinates

$$\mu := \frac{1}{2}(\xi + \eta) = \text{Re}(\xi), \quad \nu := \frac{1}{2i}(\xi - \eta) = \text{Im}(\xi), \quad (3.33)$$

we obtain the canonical form

$$\frac{\partial^2 u}{\partial \mu^2} + \frac{\partial^2 u}{\partial \nu^2} = 4f. \quad (3.34)$$

**Example 2.6** The standard examples of hyperbolic, parabolic and elliptic equations are, respectively,

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad \text{wave equation},$$

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad \text{heat equation},$$

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \text{Laplace equation}.$$

The classification as given in definition 2.5 also holds for linear equations with coefficients depending on $x$ and $t$ and even for quasilinear equations. The definition should then be applied pointwise as is demonstrated in the next example.
Example 2.7 Two well-studied equations in the theory of transonic flow are the Tricomi equation and the isentropic potential flow equation. The **Tricomi equation** (see e.g. [9]) is given by

\[
\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0.
\]

For \( y > 0 \) we apparently have a hyperbolic equation (related to supersonic flow), whereas for \( y < 0 \) the equation is elliptic (related to subsonic flow)! The **isentropic potential flow equation** for the velocity potential \( \psi \) reads (see e.g. [53])

\[
\left[ c^2 - \left( \frac{\partial \psi}{\partial x} \right)^2 \right] \frac{\partial^2 \psi}{\partial x^2} + \left[ c^2 - \left( \frac{\partial \psi}{\partial y} \right)^2 \right] \frac{\partial^2 \psi}{\partial y^2} - 2 \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial y} \frac{\partial^2 \psi}{\partial x \partial y} = 0,
\]

with the speed of sound \( c \) related to the velocity via Bernoulli’s equation for compressible flow (see Eq. 7.4.12)

\[
\left( \frac{\partial \psi}{\partial x} \right)^2 + \left( \frac{\partial \psi}{\partial y} \right)^2 + \frac{2c^2}{\gamma - 1} = \frac{2c_0^2}{\gamma - 1},
\]

where \( \gamma \) is the specific-heat ratio and \( c_0 \) is the sound speed for stagnant flow. Evidently, the equation is hyperbolic for supersonic flow (\( \varphi^2_x + \varphi^2_y > c^2 \)) and elliptic for subsonic flow (\( \varphi^2_x + \varphi^2_y < c^2 \)). □

There is a distinct difference between hyperbolic, parabolic and elliptic equations. Solutions of each type of equation show an entirely different behaviour, which is also reflected in the solution methods, either analytically or numerically. In the sequel of this book we will extensively address these three types of equations.

### 4 Linear second order equations in more space variables

The general linear second order PDE with variables \( x_1, x_2, \ldots, x_n \) is given by

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} + cu = f. \tag{4.35}
\]

Here \( x_1, x_2, \ldots, x_n \) can be time and/or any number of space variables. More precisely, for time dependent problems we have \( n = d + 1 \) and \( x_n = t \), whereas for stationary problems \( n = d \) and all variables are space coordinates. In Section 3 we used characteristics to define new variables and obtained a normal form. Here we shall simply consider just transformations of the variables without such a theory, and look for simplified forms of (4.35) from a geometrical point of view.

To start with, we may associate to (4.35) a symmetric matrix \( A \), where

\[
A := \frac{1}{2} \left( (a_{i,j})_{j \geq i} \right) + \frac{1}{2} \left( (a_{j,i})_{j \geq i} \right)^T; \tag{4.36}
\]

actually, we have distributed the coefficient \( a_{i,j} \) symmetrically between the entries \((i, j)\) and \((j, i)\) of the matrix \( A \). Now also define

\[
x := (x_1, \ldots, x_n)^T, \quad b := (b_1, \ldots, b_n)^T. \tag{4.37}
\]
Then (4.35) may be formulated as

$$A \nabla_x u \cdot \nabla_x u + b \cdot \nabla_x u + cu = f,$$

(4.38)

where $\nabla_x$ denotes the gradient with respect to $x$. From this we see that the PDE has been reformulated as a quadratic form. Quadratic forms can be simplified by diagonalising $A$. This is possible through an orthogonal similarity transformation. So let $Q$ be such that

$$Q^T A Q = \Lambda,$$

(4.39)

where $\Lambda$ is a real diagonal matrix. Next, we introduce a new set of variables $y := (y_1, \ldots, y_n)^T$ by

$$y := Qx.$$

(4.40)

Using the relation $\nabla_x = Q \nabla_y$, we then obtain the desired simplified form

$$\Lambda \nabla_y u \cdot \nabla_y u + Q^T b \cdot \nabla_y u + cu = f.$$

(4.41)

Viewing $\nabla_y u$ as a vector in $\mathbb{R}^n$ the quadratic form can now be classified as a (generalised)

(i) **ellipse**, if all eigenvalues have the same sign.
(ii) **parabola**, if at least one of the eigenvalues is zero.
(iii) **hyperbola**, if all eigenvalues are nonzero and have the same sign, except for one.

These geometric descriptions make sense in $\mathbb{R}^2$ at least. The corresponding PDEs are classified similarly: elliptic, parabolic and hyperbolic. If there are at least two positive and negative eigenvalues (and the others are all nonzero), one sometimes calls the PDE ultrahyperbolic.

If we scale the variables $y_1, \ldots, y_n$ by multiplying them by $\sqrt{|\lambda_1|}, \ldots, \sqrt{|\lambda_n|}$ respectively (unless $\lambda_i = 0$), we would obtain a quadratic form with $\pm 1, 0$ as eigenvalues. Hence it is not restrictive to assume this has already been done. It then follows that the multi-dimensional Laplace operator $\nabla^2$ becomes an important symbol to describe second order PDEs in more dimensions. In particular we have:

the **(elliptic) Laplace equation**

$$\nabla^2 u = 0,$$

(4.42a)

the **(parabolic) heat equation**

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

(4.42b)

and the **(hyperbolic) wave equation**

$$\frac{\partial^2 u}{\partial t^2} = \nabla^2 u.$$

(4.42c)
5 Reduction to ODE; similarity solutions

In the forgoing we have seen that it is sometimes possible to reformulate a (set of) PDE(s) as a set of ODEs, which describe the solution along characteristics. In this section we will reiterate this in a slightly different way. To that purpose, consider the following initial value problem

\[ \frac{\partial u}{\partial t} + b(u) \frac{\partial u}{\partial x} = c(u), \quad x \in \mathbb{R}, \; t > 0, \]  
\[ u(x, 0) = v(x), \quad x \in \mathbb{R}. \]  

(5.43a)  
(5.43b)

For the special case \( c(u) \equiv 0 \), we have derived in example 2.2 from the corresponding characteristic equations the following (implicit) representation of the solution

\[ u(x, t) = v(x - b(u(x, t))). \]

If furthermore \( b(u) = b = \text{Const} \) this solution reduces to \( u(x, t) = v(x - bt) \), which is a wave propagating undisturbed with speed \( b \).

Based on these observations, we try a solution of the following form

\[ u(x, t) = \hat{u}(\xi), \quad \xi := x - st. \]  

(5.44)

This solution is called a travelling wave solution with wave speed \( s \). Substituting (5.44) in (5.43a), we obtain the ordinary differential equation

\[ (b(\hat{u}) - s) \frac{d\hat{u}}{d\xi} = c(\hat{u}). \]  

(5.45)

The travelling wave \( \hat{u}(\xi) \) and its wave speed \( s \) have to be determined from (5.45) and (5.43b). We will illustrate this in the next example.

**Example 2.8** An example from combustion theory reads [73]

\[ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = c(u) := u(1 - u)(u - \beta) / \tau, \]  

(\*)

with parameters \( \tau > 0 \) and \( 0 < \beta < 1 \). The corresponding ODE \( u'(t) = c(u(t)) \) has the stable equilibrium solutions \( u(x, t) \equiv 0, 1 \), and the unstable one \( u(x, t) \equiv \beta \). This means that any initial value \( v(x) \neq \beta \) approaches either one of the stable solutions for \( t \to \infty \). Substituting (5.44) in (\*) we obtain the ODE

\[ (\hat{u} - s) \frac{d\hat{u}}{d\xi} = \frac{1}{\tau} \hat{u}(1 - \hat{u})(\hat{u} - \beta). \]  

(**)

Suppose that the initial solution \( v(x) \) increases monotonically from 0 to 1, then the appropriate boundary conditions for (**) read

\[ \lim_{\xi \to -\infty} \hat{u}(\xi) = 0, \quad \lim_{\xi \to \infty} \hat{u}(\xi) = 1. \]

Equation (**) can only satisfy these boundary conditions if \( s = \beta \), otherwise \( \hat{u}(\xi) \) cannot cross the unstable solution \( u(x, t) \equiv \beta \). The resulting solution then reads

\[ \hat{u}(\xi) = \left(1 + e^{-\xi/\tau}\right)^{-1}. \]

\[ \square \]
Travelling waves occur naturally in transport equations, but also in a number of other equations. We return to this in Chapter 10, where we consider travelling waves for a parabolic equation.

Another class of solutions consists of the so-called similarity solutions, which are functions of a (dimensionless) combination of $x$ and $t$. We will introduce these solutions for the homogeneous transport equation (5.43a), i.e. $c(u) \equiv 0$. A more systematic discussion of similarity solutions, based on dimension analysis, is presented in Chapter 7.

Ignoring initial and boundary conditions, we see that if $u(x, t)$ is a solution of (5.43a), then $u_{\alpha}(x, t) := u(\alpha x, \alpha t)$ is a solution as well, for any $\alpha > 0$. Therefore, we may try a similarity solution of the form

$$u(x, t) = \hat{u}(\eta), \quad \eta := \frac{x}{t},$$

i.e. $u(x, t) = \text{Const}$ along rays $x/t = \text{Const}$ through the origin of the $(x, t)$-plane. In Chapter 12 we will use this formulation to compute solutions of hyperbolic equations. Substituting (5.46) in (5.43a), we have

$$(b(\hat{u}) - \eta) \frac{d\hat{u}}{d\eta} = 0,$$

implying that either $\hat{u}(\eta) = \text{Const}$, resulting in the trivial solution $u(x, t) \equiv \text{Const}$, or $b(\hat{u}) = \eta$. In the latter case we obtain the solution

$$u(x, t) = b^{-1}(x/t),$$

with $b^{-1}(u)$ the inverse of $b(u)$, assuming it exists.

Example 2.9 Consider the traffic flow problem of Example 1.1.2, given by the transport equation

$$\frac{\partial n}{\partial t} + \frac{\partial f(n)}{\partial x} = 0.$$ 

A model for the flux $f(n)$ proposed in [73] reads

$$f(n) := u_m n \left(1 - \frac{n}{n_m}\right),$$

with $u_m$ the maximum speed of vehicles and $n_m$ the maximum density of cars. We may verify that $\hat{u}$ satisfies the equation

$$b(n) = f'(n) = u_m \left(1 - \frac{2n}{n_m}\right) = \eta,$$

resulting in the similarity solution

$$u(x, t) = \frac{1}{t} n_m \left(1 - \frac{x}{u_m t}\right).$$

See also Example 12.12.14.

Similarity solutions are also of importance for parabolic equations. In Chapter 10 we give a full account; here we restrict ourselves to an example.
CHAPTER 2. CHARACTERISATION AND CLASSIFICATION

Example 2.10 Consider the heat equation

\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \]  

Note, that if \( u(x, t) \) is a solution of (\( \ast \)), then also \( u_\alpha(x, t) := u(\alpha x, \alpha^2 t) \) for any \( \alpha > 0 \). Therefore, an obvious choice for a similarity solution is

\[ u(x, t) = \hat{u}(\eta), \quad \eta := \frac{x}{\sqrt{t}}. \]

Substituting \( \hat{u}(\eta) \) in (\( \ast \)), we obtain the ODE

\[ \frac{d^2 \hat{u}}{d\eta^2} + \frac{1}{2} \eta \frac{d\hat{u}}{d\eta} = 0. \]

This equation can be solved, having as solution

\[ \hat{u}(\eta) = C_1 \int_0^{\eta} e^{-\tau^2} \, d\tau + C_2, \]

where \( C_1, C_2 \) are constants to be determined from the initial and boundary conditions. \( \square \)

6 Initial and boundary conditions; well-posedness

For any differential equation one needs to specify the solution somewhere and somehow. The actual problem is then to find a solution of the PDE subject to certain conditions. If the latter are given at a time point onward from which the evolution takes place, we call them initial values. For the space domain we have (possibly) boundary conditions. For time dependent problems we usually have both: initial and boundary conditions. If we have purely initial conditions, we call the problem a Cauchy problem.

For determining whether the problem is meaningful we use the following

Definition 2.11 (Hadamard’s Well-posedness Conditions). A problem is well-posed if

(i) a solution exists,

(ii) the solution is unique,

(iii) the solution depends continuously on the data, in particular the initial and boundary values.

Note that (i) implies that one should not have too many (conflicting) initial and boundary conditions, (ii) not too few and (iii) that the effect of small perturbations is also small. The latter may be interpreted either in a local (small strip, interval, etc.) or global (infinite strip, interval, etc.) region.

In an elliptic problem the interaction between the coordinate points, as described by the equation, is in all directions. In time, this is of course not possible. Therefore, it is physically very unlikely that a time dependent equation is of elliptic type. This is made precise by the following example.
6. INITIAL AND BOUNDARY CONDITIONS; WELL-POSEDNESS

Example 2.12 Consider the elliptic equation
\[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad x \in \mathbb{R}, \ y > 0, \]
where \( u = u(x, y; n) \) is subject to \( u(x, 0; n) = 0, \)
\[ \frac{\partial u}{\partial y}(x, 0; n) = \frac{1}{n} \sin nx, \quad n \in \mathbb{N}. \]
This initial value problem was used by Hadamard to show that a Cauchy problem setting is not appropriate for elliptic problems. This can be seen as follows. One easily checks that \( u(x, y; n) := \frac{1}{n} \sin nx \sin ny, \)
is a solution on \( \mathbb{R} \times [0, \infty) \times \mathbb{N}. \) For \( y > 0, \) \( x = \frac{1}{n} \pi \) and \( n \) odd, we note that \( u(x, y; n) \to \infty, \) however small \( y. \) On the other hand, the initial conditions \( u(x, 0) = \frac{\partial u}{\partial y}(x, 0) = 0 \) give the solution \( u(x, y) = 0. \) This shows that \( u \) is not continuously depending on the initial data, i.e. it violates criterion (iii). □

Elliptic partial differential equations give typically rise to boundary value problems. We remark that elliptic operators will play a rôle by itself as well as part of parabolic or hyperbolic problems. Now consider a hyperbolic problem. Let \( \mathcal{D} \) be a smooth curve and \( n \) denote the normal direction. Then a Cauchy problem has as initial values
\[ u = A(x, t), \quad \frac{\partial u}{\partial n} = B(x, t), \quad (x, t)^T \in \mathcal{D}. \]

Theorem 2.13. Let \( \mathcal{D} \) be a curve in \( \mathbb{R}^2 \) such that \( \mathcal{D} \) intersects the characteristics only once. Then the Cauchy problem (3.16) and (6.49) is well posed.

Proof. Let \( \mathcal{D} \) have a parameter representation \( \varphi(x, t) = 0. \) Let \( \psi(x, t) \) be such that \( \varphi, \psi \) is a genuine coordinate transformation (i.e. the Jacobian is nonzero). Then we can reformulate (3.16) in terms of \( \varphi, \psi; \) giving
\[ \alpha \frac{\partial^2 u}{\partial \varphi^2} + \beta \frac{\partial^2 u}{\partial \varphi \partial \psi} + \gamma \frac{\partial^2 u}{\partial \psi^2} + \delta \frac{\partial u}{\partial \varphi} + \epsilon \frac{\partial u}{\partial \psi} = \zeta, \quad (*) \]
where
\[ \alpha = a \left( \frac{\partial \varphi}{\partial x} \right)^2 + b \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial x} + c \left( \frac{\partial \psi}{\partial x} \right)^2 \text{ etc.} \]
For \( u = u(\varphi, \psi) \) we then have the initial values \( u(0, \psi) = \hat{A}(\psi), \ u_{\psi} = \hat{B}(\psi). \) Hence we can find \( u_{\psi}(0, \psi), \ u_{\psi\psi}(0, \psi) \) and \( u_{\psi\psi\psi}(0, \psi) \) but not \( u_{\psi\psi\psi}(0, \psi). \) If \( \alpha \neq 0 \) we can also find \( u_{\psi\psi}(0, \psi) \) and thus all higher order derivatives from the transformed PDE (*). A formal solution, away from \( \mathcal{D} \) can now be found through a Taylor expansion
\[ u(\tilde{\varphi}, \tilde{\psi}) = \sum_{i=0}^{\infty} \sum_{j=0}^{i} \frac{1}{j! (i - j)!} \frac{\partial^i u}{\partial \tilde{\varphi}^i \partial \tilde{\psi}^j}(\varphi, \psi)(\tilde{\varphi} - \varphi)^i (\tilde{\psi} - \psi)^{i-j}. \]
CHAPTER 2. CHARACTERISATION AND CLASSIFICATION

So at least the solution exists in a neighbourhood of $D$, i.e. the problem is well-posed. Thus we only need to prove $\alpha \neq 0$. However, this follows from the fact that if $\xi(t, x) = 0$ denotes a characteristic, then $\text{det}(A\xi_t + B\xi_x) = a\xi^2_x + b\xi_x\xi_t + c\xi^2_t = 0$, while $D$ was assumed not to be a characteristic.

The preceding theorem also applies to parabolic problems.

Example 2.14 Consider the partial differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in I \subset \mathbb{R}, \quad t > 0,$$

subject to the initial conditions

$$u(x, 0) = \alpha(x), \quad \frac{\partial u}{\partial t}(x, 0) = \beta(x).$$

We trivially see that $u_{xx}(x, 0)$ should be equal to $\beta(x)$, which is not true in general. If, on the other hand, we would have a Cauchy problem in the $x$-variable, i.e.

$$u(0, t) = \tilde{\alpha}(t), \quad \frac{\partial u}{\partial x}(0, t) = \tilde{\beta}(t),$$

then one can still show the solution to exist, cf. Theorem 2.13.

Theorem 2.15. If (3.16) is parabolic, i.e. $a = b = 0$, and $u(x, 0) = \alpha(x)$ is given, then this defines a well-posed Cauchy problem, at least locally.

Proof. As in the proof of Theorem 2.13, one finds that $u_x(x, 0)$ and $u_{xx}(x, 0)$ are well defined. Hence $u_t(x, 0)$ is well defined and so are then their higher derivatives. Hence we conclude the existence of $u(x, t)$ in a neighbourhood of $t = 0$, whence we obtain local existence.

7 Discussion

- The classification into elliptic, parabolic and hyperbolic equations is more tradition than always natural. The most important distinction is between boundary value problems and initial (boundary) value problems, which have an evolutionary character. In the latter the information is travelling with a “finite speed”, while for the former everything happens with “infinite speed”. As remarked in the discussion of the previous chapter the usage of spatial coordinates does not exclude the problem to have an evolution. The boundary data are then typically needed on a part of the boundary only.

- A possible classification of evolutionary partial differential equations is whether or not they allow for a wave-like solutions; by this we mean to have solutions of the type $f(x - st)$. Hyperbolic equations clearly have such solutions. But there also exist so called dispersive waves, that do not necessarily fit into the definition of hyperbolicity that was given here and will be used later in Chapter 12. For more details see [127]
Exercises

2.1. Consider the following partial differential equation
\[
\frac{\partial^2 u}{\partial x^2} + 4 \frac{\partial^2 u}{\partial x \partial y} + 3 \frac{\partial^2 u}{\partial y^2} + 3 \frac{\partial u}{\partial x} - \frac{\partial u}{\partial y} + 2u = 0.
\]
(a) Show that the partial differential equation is hyperbolic.
(b) Find the characteristics and bring it to normal form.
(c) Find a coordinate transformation such that the first order terms vanish in the resulting equation.

2.2. Consider the partial differential equation
\[
\frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} + 5 \frac{\partial u}{\partial x} + 5 \frac{\partial u}{\partial y} + u = 0.
\]
(a) Show that the partial differential equation is parabolic.
(b) Find the normal form.
(c) Find a coordinate transformation such that the first order terms vanish in the resulting equation.

2.3. Consider the partial differential equation
\[
\frac{\partial^2 u}{\partial x^2} - 6 \frac{\partial^2 u}{\partial x \partial y} + 12 \frac{\partial^2 u}{\partial y^2} + 4 \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0.
\]
(a) Show that the partial differential equation is elliptic.
(b) Find the normal form.
(c) Find a coordinate transformation such that the first order terms vanish in the resulting equation.

2.4. Classify the partial differential equation
\[
\frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial^2 u}{\partial y \partial z} + \cos x \frac{\partial u}{\partial z} - e^y u = \cosh z.
\]

2.5. Show that in an \(d\)-dimensional space any second order elliptic partial differential equation with constant coefficients can be brought to the following form
\[
\sum_{i=1}^{d} \frac{\partial^2 u}{\partial x_i^2} + cu = f.
\]

2.6. Show that in an \(n\)-dimensional space any second order hyperbolic partial differential equation with constant coefficients can be brought to the following form
\[
\frac{\partial^2 u}{\partial x_n^2} = \sum_{i=1}^{n-1} \frac{\partial^2 u}{\partial x_i^2} + cu + f.
\]
2.7. Can you classify the partial differential equation
\[ x^\alpha y^\beta \frac{\partial^2 u}{\partial x^2} + x^\gamma y^\delta \frac{\partial^2 u}{\partial y^2} + \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0 \]?

2.8. Consider the hyperbolic equation
\[ \frac{\partial^2 u}{\partial x \partial y} = 0, \]
on the unit square, whereas \( u \) is given on the boundary. Show that this problem is not well-posed.

2.9. Consider the parabolic equation
\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \]
for \((x, t)\) in the positive \((x, t)\)-plane. Let \( u(x, 0) \) be given. Show that this Cauchy problem is not well-posed.

2.10. Determine the solution of
\[ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = u, \quad x \in \mathbb{R}, \ t > 0, \]
\[ u(x, 0) = e^x, \quad x \in \mathbb{R}. \]

2.11. Find the characteristics of the equation
\[ \frac{\partial u}{\partial t} + x \frac{\partial u}{\partial x} = 0. \]

2.12. Show that all travelling wave solutions of the wave equation
\[ \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \]
are of the form \( u(x, t) = \hat{u}_1(x - ct) \) and \( u(x, t) = \hat{u}_2(x + ct) \). Consequently, the general solution (see Chapter 12) is given by
\[ u(x, t) = \hat{u}_1(x - ct) + \hat{u}_2(x + ct). \]
Fourier theory plays an important rôle in applied analysis. In this chapter we give an overview of the most important aspects as they are needed in this book. First we introduce an inner product and (orthonormal) basis functions in Section 1. Here we also define Fourier series, consider its convergence and have Parseval’s identity. We give both the complex and the trigonometric representation. Finally the integral analogue of a series is introduced and exemplified. Next, in Section 3, the discrete form of the Fourier transform is considered, derived from the continuous version. Again, convergence and Parseval’s identity are studied. Also important phenomena like aliasing, which show the restrictions of the discrete Fourier transform are treated. One very important application of this Fourier transform is its use in analysing linear equations with periodic boundary values. Despite the limitations of this problem class, it turns out that many physically meaningful concepts, like stability, dissipation and dispersion can be studies quite fruitfully for the transformed equation, both in the continuous and in the discrete case. In Section 4, therefore, the use of these transformations is demonstrated, leading to the important concept of dispersion relation.

1 Fourier series

A powerful tool in analysis is the expansion of a function \( f \) in terms of suitably chosen functions, that form a basis. There are several ways to find such expansions. The most elegant way to describe it mathematically is to use projections, for which we need the concept of “orthogonality”. This is provided by using an inner product, much like the one met in linear algebra (the “natural inner product”). Since this involves integration, we shall restrict ourselves to a finite interval, which is typically chosen as \((0, L)\). The functions we consider are square-integrable, i.e. \( L^2 \)-functions. Moreover, we shall assume that they are periodically extended to the full real axis. We then define the inner product for two such functions as

\[
(f, g) := \int_0^L f(x) \overline{g}(x) \, dx, \quad (1.1)
\]
1. FOURIER SERIES

where the overbar denotes the complex conjugate (i.e. \( \overline{a + ib} = a - ib \)). Note that from periodicity it follows that any integral over an interval of length \( L \) is equivalent. As can simply be verified the functions

\[ p_j(x) := e^{i\alpha_j x}, \quad j = 0, \pm 1, \pm 2, \ldots, \quad (1.2a) \]

where

\[ \alpha_j = \frac{2\pi j}{L}, \quad (1.2b) \]

are orthogonal. By changing (1.1) into

\[ (f, g)_L := \frac{1}{L} \int_0^L f(x) \overline{g(x)} \, dx, \quad (1.1') \]

they are even orthonormal, i.e. \((p_j, p_k) = \delta_{jk}\). We call \( \alpha_j \) a wave number, if \( x \) denotes a spatial variable and frequency, if the independent parameter is time. In physical texts, a wave number is usually denoted by the letter \( k \) and a frequency by \( \omega \).

Introducing

\[ c_j := \frac{1}{L} \int_0^L f(y) e^{-i\alpha_j y} \, dy \quad (1.3) \]

we can form a so-called Fourier series

\[ \tilde{f}(x) := \sum_{j=-\infty}^{\infty} c_j e^{i\alpha_j x}. \quad (1.4) \]

The function \( p_j(x) = e^{i\alpha_j x} \) is called a Fourier mode with wave number (frequency) \( \alpha_j \). The important question is whether \( \tilde{f} \) can be identified with \( f \) (and of course whether \( \tilde{f} \) makes sense at all). We then say that the Fourier series of \( f \) converges to \( f \). In fact, no simple test is known that is both necessary and sufficient to relate a periodic function with its Fourier coefficients [24]. There is, however, a vast amount of partial results.

We have the following theorems.

**Theorem 3.1.** If for all points \( x \in (0, L) \) the left and right derivative of \( f \), i.e.

\[ \lim_{h \downarrow 0} \frac{f(x+h) - f(x)}{h} \quad \text{and} \quad \lim_{h \downarrow 0} \frac{f(x) - f(x-h)}{h} \]

exist, then the Fourier series (1.4) of \( f \) at \( x \) converges to \( f(x) \).

**Theorem 3.2.** If \( f \) is continuous in \( x \) and \( f(0) = f(L) \), then the Fourier series (1.4) of \( f \) converges uniformly to \( f \) (Appendix C), i.e.

\[ \lim_{N \to \infty} \sup_{x \in [0, L]} \left| \sum_{j=-N}^{N} c_j e^{i\alpha_j x} - f(x) \right| = 0. \]
Definition 3.3. The function $f$ is piecewise continuous on $[0, L]$ if there are a finite number of open subintervals $0 < x < x_1, \ldots, x_{N-1} < x < L$ on which $f$ is continuous, while the limits $f(0+), f(x_1), \ldots, f(L-)$ exist. The function $f$ is piecewise smooth on $[0, L]$ if $f$ and its derivative $f'$ are both piecewise continuous.

For such a function $f$ we have the following theorem.

Theorem 3.4 (Existence of Fourier series). If a function $f$ is piecewise smooth on the interval $[0, L]$, while $f(x) = \frac{1}{2}[f(x+) + f(x-)]$, then the Fourier series of $f$ converges for every $x$ to the $L$-periodic continuation of $f$.

For a given Fourier series $\sum c_j e^{i\alpha_j x}$ we have the following theorem.

Theorem 3.5 (Continuity of Fourier series). If a Fourier series is absolutely convergent, i.e. $\sum |c_j| < \infty$, then it converges absolutely and uniformly to a continuous periodic function $f$, such that $c_j$ are just the Fourier coefficients of $f$.

If $f$ is only a function in $L^2$ then we still have the following identity to hold at least.

Theorem 3.6 (Parseval’s identity.). Let $f \in L^2_2(0, L)$ with Fourier coefficients $c_j$. Then

$$ (f, f)_L = \sum_{j=-\infty}^{\infty} c_j^2. $$

**Proof.** For a proof of Theorems 3.1, 3.2, 3.6: see [24, 84].

Corollary 3.7. If $f$ and $f'$ are piecewise smooth, the Fourier coefficients $c_j$ of $f$ behave asymptotically for $j \to \infty$ like $c_j = O(j^{-1})$.

**Proof.** By partial integration it follows that

$$ c_j = \frac{1}{i\alpha_j} \sum_{x=x_d} \left[ f(x) e^{-i\alpha_j x} \right]_{x_d}^{x_d+} - \frac{1}{i\alpha_j} c'_j $$

where the summation runs over all points $x_d$ of discontinuity of $f$ (possibly including the end points), and $c'_j$ is the $j$-th Fourier coefficient of $f'$. As $c'_j \to 0$, the result follows.

In a suitable context, the inner product $(f, f)$ may be interpreted as an energy content and thus $(f, f)_L$ as a mean energy. Therefore Theorem 3.6 is sometimes referred to as Parseval’s energy theorem.
Often it is useful to rewrite the Fourier series in terms of trigonometric functions. Using the well-known relation
\[ e^{iz} = \cos z + i \sin z, \]
we thus find from (1.4)
\[ \tilde{f}(x) = a_0 + \sum_{j=1}^{\infty} a_j \cos(\alpha_j x) + b_j \sin(\alpha_j x), \]
where
\[ a_0 = c_0, \quad a_j = c_j + c_{-j}, \quad b_j = i(c_j - c_{-j}). \]
We can more directly write
\[ a_0 = \frac{1}{L} \int_{-L}^{L} f(y) \, dy, \quad a_j = \frac{2}{L} \int_{0}^{L} f(y) \cos(\alpha_j y) \, dy, \]
\[ b_j = \frac{2}{L} \int_{0}^{L} f(y) \sin(\alpha_j y) \, dy, \quad j = 1, 2, \ldots. \]
Theorems 3.1 and 3.2 carry over to the trigonometric representation and for Parseval’s identity (Theorem 3.6) we have
\[ (f, f)_L = a_0^2 + \frac{1}{2} \sum_{j=1}^{\infty} (a_j^2 + b_j^2). \]

**Example 3.8** Consider the function
\[ f(x) = x, \quad x \in (-\pi, \pi]. \]
In order to make this periodic, we extend this function periodically beyond \((-\pi, \pi]\) to obtain the saw-tooth function as in Fig. 3.1. It is then straightforward to see that \( f \in L_2(-\pi, \pi) \). Since
\[ a_j \propto \int_{-\pi}^{\pi} y \cos(jy) \, dy = 0, \quad \text{for all } j, \]
and
\[
b_j = \frac{1}{\pi} \int_{-\pi}^{\pi} y \sin(jy) \, dy = \frac{1}{\pi} \left[ -\frac{\gamma \cos jy}{j} + \frac{\sin jy}{j^2} \right]_{-\pi}^{\pi} = \frac{2}{j} (-1)^{j+1}
\]
we deduce, on account of Theorem 3.1 (we take \( L = 2\pi \) and the interval \((-\pi, \pi]\) in equation 1.3), that the resulting series
\[
\hat{f}(x) = 2 \left[ \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots \right]
\]
converges to \( f(x) \) for any \( x \in (-\pi, \pi) \).

However, we cannot guarantee uniform convergence (Appendix C) on the whole interval. At \( x = \pi \) the series does not converge to \( f(\pi) = \pi \) but rather to 0, the average between the left- and right limits, as is shown graphically by Fig. 3.2. We see an interesting phenomenon at the discontinuities of \( f \): there is an overshoot to the left and an “undershoot” to the right. This is known as Gibbs phenomenon.

\[\square\]

![Gibbs phenomenon](image)

**Figure 3.2.** Example of Gibbs’ phenomenon (50 terms)

The function in Example 3.8 was clearly odd and so were the Fourier terms. This is generally true. If \( f \) is odd (i.e. \( f(-x) = -f(x) \)) we have a Fourier sine series and if \( f \) is even (i.e. \( f(-x) = f(x) \)) we have a Fourier cosine series.

**Example 3.9** The following Fourier sine and cosine series define periodic functions with period 1.
\[
\sum_{n=1}^{\infty} \frac{\sin(2\pi nx)}{\pi n} = \left[ \frac{1}{2} - x \right], \quad \sum_{n=1}^{\infty} \frac{\cos(2\pi nx)}{n} = -\log |2 \sin \pi x|,
\]
\[
\sum_{n=1}^{\infty} \frac{\cos(2\pi nx)}{\pi^2 n^2} = \left[ x^2 - x + \frac{1}{6} \right], \quad \sum_{n=1}^{\infty} \frac{\cos(2\pi nx)}{n^2 - \frac{1}{4}} = 2 - \pi |\sin \pi x|.
\]

[\cdot, \cdot] denotes a function originally defined on \([0, 1]\) and continued periodically. Another interesting example is the block-wave function, defined along \([-1, 1]\) by
\[
4 \sum_{n=0}^{\infty} \frac{\sin(2n+1)\pi x}{(2n+1)\pi} = \text{sign}(x),
\]
and 2-periodically continued otherwise. \[\square\]
2. FOURIER TRANSFORMS

In a straightforward way, Fourier series are generalised to more dimensions. Suppose that the square-integrable function \( f : \mathbb{R}^d \to \mathbb{R} \) is periodic in every argument \( x_j, j = 1, \ldots, d \). Then

\[
f(x_1, \ldots, x_d) := \sum_{m_1 = -\infty}^{\infty} \cdots \sum_{m_d = -\infty}^{\infty} C_{m_1, \ldots, m_d} e^{i\alpha_{m_1} x_1 + \cdots + i\alpha_{m_d} x_d},
\]

(1.9)

where

\[
\alpha_{m_j} = \frac{2\pi}{L_j} m_j
\]

and

\[
C_{m_1, \ldots, m_d} = \frac{1}{L_1 \cdots L_d} \int_0^{L_1} \cdots \int_0^{L_d} f(y_1, \ldots, y_d) e^{-i\alpha_{m_1} y_1 - \cdots - i\alpha_{m_d} y_d} \, dy_1 \cdots dy_d
\]

(1.10)

If \( f \) is scaled such that it is \( 2\pi \)-periodic in each independent variable, this may be written more compactly as

\[
f(x) := \sum_{m = -\infty}^{\infty} C_m e^{im \cdot x}, \quad C_m = \frac{1}{(2\pi)^d} \int_0^{2\pi} f(y) e^{-im \cdot y} \, dy.
\]

(1.11)

where \( m \in \mathbb{Z}^d \) denotes the index vector \( m = [m_1, \ldots, m_d] \), and the sum and integral sign are to be interpreted \( d \)-fold.

2. Fourier transforms

There exists an integral analogue to the Fourier series. Recalling (1.2) we may let \( L \to \infty \), i.e. \( \frac{1}{L} \to 0 \). Writing

\[
\Delta \alpha := \frac{1}{L},
\]

(2.1)

we have for an \( L \)-periodic, square integrable function \( f \)

\[
c_j = \Delta \alpha \int_0^{(\Delta \alpha)^{-1}} f(y) e^{-2\pi ij \Delta \alpha y} \, dy
\]

(2.2)

whence

\[
f(x) = \sum_{j = -\infty}^{\infty} e^{2\pi ij \Delta \alpha x} \Delta \alpha \int_0^{(\Delta \alpha)^{-1}} f(y) e^{-2\pi ij \Delta \alpha y} \, dy.
\]

(2.3)

Here we recognize a Riemann sum of the function \( g(z) \), where

\[
g(z) := \frac{1}{2\pi} \int_0^{(\Delta \alpha)^{-1}} f(y) e^{-izy} \, dy.
\]

(2.4)

One should just take piecewise constant approximations of \( g \) at the points \( 2\pi j \Delta \alpha, j = 0, \pm 1, \pm 2, \ldots \), multiplied by the interval width \( 2\pi \Delta \alpha \). Hence by a limit argument we find

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(y) e^{-i\alpha y} \, dy \right] e^{ix \alpha} \, d\alpha,
\]

(2.5)
which leads to \( \hat{f} \), the Fourier transform of \( f \), which is together with its inversion given by

\[
\hat{f}(\alpha) := \int_{-\infty}^{\infty} f(x) e^{-i\alpha x} \, dx, \tag{2.6a}
\]

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\alpha) e^{i\alpha x} \, d\alpha. \tag{2.6b}
\]

(Note that other, equivalent, definitions frequently occur, causing a lot of confusion.) The Fourier transform (or also called: Fourier integral) plays an important rôle in the analysis of problems where we have a continuous spectrum of wave numbers (or frequencies). One can show that \(|f|\) and \(|\hat{f}|\) cannot vanish simultaneously outside a finite domain. Note that it is sufficient for the existence of \( \hat{f} \) that \( f \in L_1 \), but if \( f \in L_2 \) then also \( \hat{f} \in L_2 \) \cite{24} and therefore both (2.6a) and (2.6b) exist.

**Example 3.10**

(i) The “top hat” function, given by

\[
f(x) = \begin{cases} 
1, & -1 \leq x \leq 1, \\
0, & |x| > 1,
\end{cases}
\]

has a Fourier transform, closely related to the sinc-function:

\[
\hat{f}(\alpha) = \int_{-1}^{1} e^{-i\alpha x} \, dx = \frac{e^{i\alpha} - e^{-i\alpha}}{i\alpha} = \frac{2\sin\alpha}{\alpha} = 2 \text{sinc} \left( \frac{\alpha}{\pi} \right).
\]

(ii) The decaying exponential, vanishing for \( x < 0 \), defined by

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

where \( p > 0 \), has a Fourier transform consisting of a single pole in the upper complex \( \alpha \)-plane

\[
\hat{f}(\alpha) = \int_{0}^{\infty} e^{-px} e^{-i\alpha x} \, dx = \left[ -\frac{e^{-(p+i\alpha)x}}{p + i\alpha} \right]_{0}^{\infty} = \frac{1}{p + i\alpha}.
\]

If \( \alpha^n \hat{f}(\alpha) \) is Fourier transformable, where \( n \in \mathbb{N} \) and \( \hat{f}(\alpha) \) is the Fourier transform of \( f(x) \), then

\[
\frac{d^n}{dx^n} f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (i\alpha)^n \hat{f}(\alpha) e^{i\alpha x} \, d\alpha. \tag{2.7}
\]

**Theorem 3.11 (Convolution Theorem).**

If \( \hat{f}(\alpha), \hat{g}(\alpha) \) are the Fourier transforms of \( f(x), g(x) \in L_2 \), then the inverse Fourier transform of \( \hat{f}(\alpha)\hat{g}(\alpha) \) is given by

\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\alpha)\hat{g}(\alpha) e^{i\alpha x} \, d\alpha = \int_{-\infty}^{\infty} f(x - y)g(y) \, dy.
\]
2. FOURIER TRANSFORMS

**Proof.** As \( f \) and \( g \) are square integrable, we may change the order of integration to get

\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\alpha) \hat{g}(\alpha) e^{i\alpha x} \, d\alpha = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\alpha) \int_{-\infty}^{\infty} g(y) e^{-i\alpha y} \, dy \, e^{i\alpha x} \, d\alpha
\]

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} g(y) \int_{-\infty}^{\infty} \hat{f}(\alpha) e^{i\alpha(x-y)} \, d\alpha \, dy = \int_{-\infty}^{\infty} f(x-y)g(y) \, dy,
\]

\( \square \)

If we consider Theorem 3.11 for \( x = 0 \) and take \( g(y) = f(-y) \) with \( \hat{g}(\alpha) = \hat{f}(\alpha) \), we obtain the analogue of Parseval’s identity (Theorem 3.6) for integrals

\[
\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\alpha)|^2 \, d\alpha.
\]

This is sometimes referred to as the *energy theorem* for the continuous case.

**Example 3.12** Consider Example 3.10 (ii) again. We obtain indeed

\[
\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{0}^{\infty} e^{-2px} \, dx = \frac{1}{2p},
\]

\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\alpha)|^2 \, d\alpha = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{p^2 + \alpha^2} \, d\alpha = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{|\alpha|^2} \, d\alpha = \frac{1}{2p}.
\]

\( \square \)

Let \( f \) be a Fourier transformable function on \( \mathbb{R} \). Noting that the function

\[
\sum_{m=-\infty}^{\infty} f(mL + x)
\]

is periodic in \( x \) with period \( L \), we can write (following [21])

\[
\sum_{m=-\infty}^{\infty} f(mL + x) = \sum_{j=-\infty}^{\infty} \left[ \frac{1}{L} \int_{0}^{L} \sum_{m=-\infty}^{\infty} f(mL + y) \, e^{-iyj} \, dy \right] e^{i\alpha j x} = \frac{1}{L} \sum_{j=-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(\eta) \, e^{-i\eta j} \, d\eta \right] e^{i\alpha j x} = \frac{1}{L} \sum_{j=-\infty}^{\infty} \hat{f}(\alpha j) \, e^{i\alpha j x},
\]

(with \( \alpha_j \) as defined in Eq. (1.2b)), which leads for \( x = 0 \) to Poisson’s Formula

\[
\sum_{m=-\infty}^{\infty} f(mL) = \frac{1}{L} \sum_{m=-\infty}^{\infty} \hat{f}\left(\frac{2\pi}{L} m\right),
\]

where the average of the left and right limit is to be taken at any discontinuities.
Example 3.13 Poisson’s Formula is an excellent tool to accelerate slowly converging series. From Example 3.10 (ii) we find for $p > 0$ and $L = 2\pi$

$$\frac{1}{2\pi p} + \frac{p}{\pi} \sum_{m=1}^{\infty} \frac{1}{m^2 + p^2} = \frac{1}{2} + \sum_{m=1}^{\infty} e^{-2\pi pm}.$$ 

The left-hand side converges algebraically slow, in contrast to the fast, exponential convergence of the right-hand side. As a bonus, we have in this case even an explicit expression if we recognize the geometric series with common ratio $e^{-2\pi p}$.

We are interested to know when a given $\hat{f}(\omega)$ is a time-Fourier transform of a physical signal $f(t)$. First, in order for $f$ to be real, $\hat{f}(\omega)$ has to satisfy the reality condition

$$\hat{f}(\omega) = \overline{\hat{f}(-\omega)}.$$ 

(2.11)

No physical process can exist for all time. A process $f(t)$ that starts by some cause at some finite time $t = t_0$, while it vanishes before $t_0$, is called causal. The corresponding Fourier transform

$$\hat{f}(\omega) = \int_{t_0}^{\infty} f(t) e^{-i\omega t} \, dt$$

has the property that $\hat{f}(\omega)$ is analytic in the lower complex half-plane $\text{Im}(\omega) < 0$. So this is a necessary condition on $\hat{f}$ for $f$ to be causal. A sufficient condition is the following causality condition [84].

**Theorem 3.14 (Causality Condition).**

If $\hat{f}(\omega)$ is analytic in $\text{Im}(\omega) \leq 0$, $|\hat{f}(\omega)|^2$ is integrable along the real axis, and there is a real $t_0$ such that $\hat{f}(\omega) e^{i\omega t_0} \to 0$ uniformly with regard to $\arg(\omega)$ for $|\omega| \to \infty$ in the lower complex half plane, then $f(t)$ is causal, and vanishes for $t < t_0$.

**Proof.** It is no restriction for the proof to assume $t_0 = 0$. Consider, for $t < 0$, the integral $\int_{t}^{\infty} \hat{f}(\omega) e^{i\omega t} \, d\omega$ along the real contour $[-R, R]$ closed via a semi-circle of radius $R$ in the lower complex half plane. As the integrand is analytic the integral is zero. Let $R \to \infty$. The contribution $I_R$ of the integral along the semi-circle tends to zero, because

$$|I_R| \leq \int_{0}^{\pi} |\hat{f}(\omega)| e^{-i|\omega| R \sin \theta} R \, d\theta \leq 2R \max_{\theta} |\hat{f}(\omega)| \int_{0}^{\pi} e^{-i|\omega| R |\sin \theta|} \, d\theta \to 0,$$

where $\omega = Re^{-i\theta}$. So the contribution from the real axis, being equal to $2\pi \hat{f}(t)$, is also zero.

Note that the lower complex half-space becomes the upper half-space if the opposite Fourier sign convention is taken.

**Example 3.15** The Fourier transform $\hat{f}(\omega) = (p + i\omega)^{-1}$ is causal if $p > 0$, as may be confirmed by the inverse transform

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{p + i\omega} \, d\omega = \begin{cases} e^{-pt} & \text{if } t > 0, \\ 0 & \text{if } t < 0. \end{cases}$$

3. DISCRETE FOURIER TRANSFORMS

In the limit of no damping ($p \downarrow 0$) the singularity at $\omega = ip$ moves to $\omega = 0$, which is on the real axis. The integral is to be interpreted via a suitable indentation of the contour under the pole in order to retain causality.

In a straightforward way, Fourier transforms are generalised to more dimensions. For the square-integrable function $f : \mathbb{R}^d \to \mathbb{R}$ we have the couple

$$\hat{f}(\alpha) := \int_{-\infty}^{\infty} f(x) e^{-i\alpha \cdot x} \, dx,$$

(2.12a)

$$f(x) = \frac{1}{(2\pi)^d} \int_{-\infty}^{\infty} \hat{f}(\alpha) e^{i\alpha \cdot x} \, d\alpha,$$

(2.12b)

where $\alpha \in \mathbb{R}^d$ denotes the Fourier wave number vector and the integral signs are to be interpreted $d$-fold.

3. Discrete Fourier transforms

Often a function is known at a certain finite number of points only. For instance, $f$ may be measured at a few values of the argument ("sampled"). If one uses a numerical approximation method based on discretisation of the argument, as in the case with finite differences, there is also a natural set of points on which $f$ is "monitored" only. Given the grid points $x_k$, $k = 0, \ldots, N - 1$, we can give straightforward analogues of the continuous case by considering the discrete inner product

$$\langle f, g \rangle := \sum_{k=0}^{N-1} f(x_k) g(x_k),$$

(3.1)

where $x_k \in [0, L]$. Because of the special properties exhibited by $e^{i\alpha x}$ when the points $x_k$ are chosen equispaced, i.e. $x_{k+1} - x_k$ is constant ($k = 0 \ldots N - 1$ and $x_N = x_0 + L$), we shall take them as

$$x_k := \frac{k}{N} L \quad k = 0, \ldots, N - 1.$$

(3.2)

It can be verified that the functions

$$p_j(x) := e^{i\alpha_j x},$$

(3.3)

with

$$\alpha_j = \frac{2\pi j}{L}$$

are orthogonal. This is stated in the following theorem.

**Theorem 3.16.** The polynomials $p_j$, defined in (3.3), are orthogonal with respect to (3.1) in the sense that $\langle p_j, p_l \rangle = 0$ for all $j, l$ with $j - l$ not a multiple of $N$. 

Proof. Substituting (3.3) in (3.1), we have
\[
\langle p_j, p_l \rangle = \sum_{k=0}^{N-1} \exp\left[2\pi i (j - l) \frac{k}{N}\right] = \sum_{k=0}^{N-1} \sigma^k = \begin{cases} 
\frac{\sigma^N - 1}{\sigma - 1} & \text{if } \sigma \neq 1, \\
\frac{N}{\sigma} & \text{if } \sigma = 1,
\end{cases}
\]
with \( \sigma := \exp(2\pi i \frac{j - l}{N}) \). Apparently, \( \sigma^N = 1 \) while \( \sigma \neq 1 \) if \( (j - l) \mod N \neq 0 \), which proves the result.

Corollary 3.17. \( \langle p_j, p_j \rangle = N \).

We therefore may as well consider the inner product
\[
\langle f, g \rangle_N := \frac{1}{N} \sum_{k=0}^{N-1} f(x_k) g(x_k),
\]
which makes \( \{ p_j(x) \}_{j=0}^{N-1} \) an orthonormal set of basis functions. We can now give a discrete Fourier transform (DFT) of a function \( f \)
\[
c_j := \langle f, p_j \rangle_N = \frac{1}{N} \sum_{k=0}^{N-1} f(x_k) e^{-2\pi i j k/N}.
\]
Due to the special choice of the grid we immediately find that \( \exp(i\alpha_j x_k) = 1 \) when \( j \) is a multiple of \( N \). Hence for \( j \geq N \) we have
\[
e^{i\alpha_j x_k} = e^{2\pi i j k/N}, \quad \hat{j} = j \mod N.
\]
This relation tells us that the basis functions \( p_j(x) \), for \( j \geq N \), will not provide additional information to represent (an approximation of) \( f \). The phenomenon that we cannot distinguish discrete Fourier components of \( p_j \) and \( p_{j+lN}, l \in \mathbb{Z} \), is called aliasing. Hence we will be satisfied to have the finite series
\[
\hat{f}(x) := \sum_{j=0}^{N-1} c_j e^{2\pi i j x}, \quad x \in [0, L].
\]
Of particular interest is \( \hat{f}(x) \) at the points \( x = x_k \), because here the original values of \( f \) are exactly recovered, i.e. \( f(x_k) = \hat{f}(x_k) \). So at these points \( f \) is completely defined by the coefficients \( c_j \), and vice versa. If we define
\[
\hat{c}_k := \frac{1}{\sqrt{N}} f(x_k)
\]
the following reciprocity relation between \( c_j \) and \( \hat{c}_k \) may be shown
Property 3.18 (Reciprocity of DFT).

\[
\hat{c}_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} c_j e^{2\pi i j k / N}, \quad c_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{c}_k e^{-2\pi i j k / N}.
\]

We remark that the DFT has very important applications, in particular through its efficient implementation, the so-called Fast Fourier Transform (FFT, see [108]).

From an approximation point of view, aliasing implies that we cannot obtain more information about a function \( f \) than the sampling rate (the density of the grid) is allowing us. See Example 3.19. In particular we thus conclude that the discrete wave numbers generated by the grid will limit the accuracy of the approximation \( \hat{f}(x) - f(x) \).

Example 3.19 In Fig. 3.3 we have drawn two sinus functions, one with wave number 1, sampled with a rate of \( 20/2\pi \) and one with wave number 21. As \( \sin x \) coincide with \( \sin(21x) \) right at the sampling points, \( \sin(21x) \) cannot be represented with this sampling rate. \( \square \)

Figure 3.3. Aliasing

4 Fourier analysis applied to PDEs

Consider now the linear partial differential equation with constant coefficients

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + cu. \tag{4.1}
\]

We may look for solutions of this equation by using a Fourier series (on a finite domain) or Fourier integrals (on an infinite domain). Of course this does not make much sense in general if we do not specify the initial and boundary conditions. Yet one may hope that an analysis based on Fourier expansions – if successful – might give insights for the general case. If for instance such an analysis leads to prediction of instabilities or other undesirable phenomena, it may imply such a result for a more general case.

Since we have at least two variables, space and time, our Ansatz will be

\[
u(x, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{u}(\kappa, \omega) e^{i\kappa x - i\omega t} \, d\kappa \, d\omega, \tag{4.2}
\]

where \( \omega \) denotes frequency and \( \kappa \) wave number. (For a natural definition of phase and group velocity below, these Fourier variables are defined with opposite signs.) This expression

CHAPTER 3. FOURIER THEORY

is a very general one, valid for any function \( u(x, t) \). We are not dealing with any \( u \), but with a solution of (4.1) and usually a simpler form is possible, for example, if we require each Fourier mode to be a solution of the defining equation (4.1). In order to analyse the solution we consider the single mode

\[
 u(x, t) = A e^{ixx - i\omega t}.
\]  

(4.3)

(Such a planar wave solution is even useful for analyzing nonlinear evolution equations; see [127]). Upon substituting (4.3) in (4.1) we obtain that the mode is a solution if

\[
 -i\omega = -a\kappa^2 + ib\kappa + c
\]  

(4.4)

Since this relation gives information about the propagation properties of the various modes it is often referred to as the dispersion relation (dispersion being explained below).

If we follow a modal wave crest, i.e. such that the phase \( \kappa x - \omega t = \text{constant} \), we move with the phase velocity or wave speed

\[
 v_p := \frac{\omega}{\kappa}.
\]  

(4.5)

As a mode has an infinite extension in \( x \), it is hard to tell by which velocity any associated properties, like energy, propagate. Therefore, we consider a “localised mode”, or wave packet, that decays slowly to zero for large \( |x| \). This is not exactly one mode any more (with a single frequency and wave number), but a superposition of modes near a main frequency and wave number. To be more precise, let \( f(x) \) be an absolute integrable smooth function and \( \varepsilon \) is small compared to \( \kappa_0 \), such that

\[
 f(\varepsilon x) = \hat{f}(\kappa - \kappa_0 \varepsilon^{1/2}) e^{ixx - i\omega_0 t}.
\]  

(4.6a)

The Fourier transform of (4.6a) is given by

\[
 \int_{-\infty}^{\infty} e^{-i(\kappa - \kappa_0)x} f(\varepsilon x) dx = \varepsilon^{-1} \hat{f}(\frac{x-\varepsilon}{\varepsilon}).
\]  

(4.6b)

Let the modes be defined by a dispersion relation \( \omega = \omega(\kappa) \). Hence we can write the wave packet as

\[
 u(x, t) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\kappa x} \hat{f}(\eta) e^{i\omega t} e^{i\omega(\kappa - \kappa_0) t} d\kappa
\]  

\[
 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\eta) e^{i\kappa x + i\omega_0 t - i\omega(\kappa_0 + \varepsilon \eta)t} d\eta.
\]  

By Taylor-expanding \( \omega(\kappa_0 + \varepsilon \eta) = \omega_0 + \varepsilon \eta \omega_0' + O(\varepsilon^2) \) for small \( \varepsilon \) - where \( \omega_0 = \omega(\kappa_0) \) and \( \omega_0' = \omega'(\kappa_0) \) - and rearranging terms we find in good approximation

\[
 u(x, t) = \frac{1}{2\pi} e^{i\omega_0 x - i\omega_0 t} \int_{-\infty}^{\infty} \hat{f}(\eta) e^{i\omega_0 t} e^{i\omega(\kappa_0 + \varepsilon \eta)t} d\eta
\]  

\[
 = e^{i\omega_0 x - i\omega_0 t} f(\varepsilon|x - \omega_0'|t).  
\]  

(4.7)
4. FOURIER ANALYSIS APPLIED TO PDES

We see that the wave crests indeed propagate with the phase velocity \( \omega_0/\kappa_0 \), whereas the group as a whole propagates with a velocity \( \omega'(\kappa_0) \). This velocity is called the group or energy or signal velocity, and is thus given by

\[
v_g := \frac{d\omega}{d\kappa}.
\] (4.8)

In general, all modes propagate with their own speed, and a group superposed of many different modes gets dispersed. The shape of the group remains intact, i.e. shows no dispersion, if all modes propagate with the same speed, in other words if

\[
\frac{dv_g}{d\kappa} = \frac{d^2\omega}{d\kappa^2} = 0.
\] (4.9)

Note that a mode like \( e^{-i\omega t} e^{i\kappa(x-bt)} \) is clearly not dispersive, so condition (4.9) is not exactly equivalent to a condition of constant phase speed.

We return to equation (4.1) and consider two special cases. First let \( b \) and \( c \) be zero, i.e. we have the standard heat equation. We then find upon substituting \(-i\omega = -a\kappa^2\) that

\[
u(x, t) = e^{i\kappa x} e^{-a\kappa^2 t}.
\] (4.10)

The first factor is just the spatial Fourier component, but the second is an exponentially growing/decaying quantity. Hence we see that we need \( a > 0 \) to have a stable mode, and we may call solution (4.10) dissipative.

The other special case is when \( a = c = 0 \). Then we have

\[
\omega = -b\kappa.
\] (4.11)

This means that each mode is propagating with the same wave speed \( b \). In particular, we find

\[
u(x, t) = e^{i(x+bt)\kappa}.
\] (4.12)

Hence on a line in the \((x, t)\) plane where \( x + bt \) is constant, i.e. on a characteristic (cf. Chapter 2), we note that \( \nu \) is constant. We may therefore call solutions like \( \nu \) in (4.11) conservative.

In the context of numerical methods for partial differential equations based on finite differences, the analysis of error propagation leads to typically linearized equations with slowly varying coefficients. Take for example the “true” solution \( y \) satisfying

\[
y_t = (A(y)y_x)_x + C(y),
\] (4.13)

the numerical solution \( \tilde{y} \), and the error \( u := y - \tilde{y} \). The error is by assumption small. Its typical fluctuations, due to the small grid and time steps, vary over a much shorter time and length scale than the “true” solution. As we are only interested in the error, i.e. the behaviour along the short time and length scales, the linearized equation for \( u \) in the neighbourhood of \( x = x_0 \) and \( t = t_0 \) may look like

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + cu,
\] (4.14)
where \(a\), \(b\), and \(c\) are assumed constant, depending on \(y(x_0, t_0)\). (Note that this equation is just an example, and has no other purpose than to illustrate.) As before, the behaviour of \(u\) in \(x\) and \(t\) may be analysed by Fourier analysis. Assume for \(u\) a single mode of the form (4.3), i.e. with constant amplitude and phase. Substitution in equation (4.14) yields dispersion relation (4.4).

Fourier analysis in difference equations will turn out to be a powerful tool to determine necessary conditions for a numerical method to be practically useful. We shall defer details to the specific chapters where we assess the numerical methods. Refinements based on the method of multiple scales (Chapter 15, Section 4.2) allows including the variation of amplitude and phase with the (relatively) slowly varying \(x_0, t_0\).

5 Discussion

- Fourier theory is an essential tool in many applications, far beyond the goals of this book. Traditionally the Fourier series are used to approximate functions. There is a host of other choices for this, depending on the application, see e.g. books on special functions, like [83]. Quite another application is the efficient solution of systems through the so-called Fast Fourier Transform (FFT), see [87]

- Dispersion is a very important concept for the theory of waves. As it will turn out in Chapters 12, ??, ?? both in analytical and numerical study of hyperbolic problems, the actual behaviour of the solution is critically depending on properties like dissipation or dispersion. In fact one of the major problems in numerically solving hyperbolic problems is to capture the physical behaviour, i.e. not introducing too much “numerical” dissipation or dispersion. Finally we remark that there are many equations having wave-like solutions, not being of hyperbolic type, so called dispersive waves, [127].

Exercises

3.1. Let \(L\) be some positive number. Show that the functions \(\frac{1}{2} \sqrt{2} s, s \cos(j \pi s^2 x), s \sin(j \pi s^2 x), \) with \(s := L^{-\frac{1}{4}}, j = 1, 2, \ldots,\) form an orthonormal basis on \((-L, L)\).

3.2. Show that an even function on \((-L, L)\) is orthogonal to an odd.

3.3. Prove the reciprocity relation for DFT (property 3.18).

3.4. (a) Find the Fourier series of \(\cos^2 x\).

(b) Find the Fourier series of \(x^2\), defined on \([0, 1]\).

(c) Consider the function \(f(x) = x(x-1)\), defined on \([0, 1]\). Let \(\sum_{j=1}^{\infty} \gamma_j \sin(j \pi x)\) be the Fourier sine-series of \(f\). Show that \(\gamma_j = 0\) for \(j\) even, and \(\gamma_j = -8/(j \pi)^3\) for \(j\) odd.

3.5. Given the Fourier coefficients of \(f(x)\), determine the Fourier coefficients of the first and second derivative.
Exercises

3.6. If $f(x)$ is a periodic $L_2$-function with a jump condition at $a$ say, then the Fourier series at that point converges to $\frac{1}{2}(f(a+) + f(a-))$, where $a+$ denotes the right limit and $a-$ the left limit. Show this for $f(x) = x$ on $(-\pi, \pi)$.

3.7. Find the Fourier series of the following function, defined on $[0, 1]$, 

$$f(x) = \frac{2}{3}x^3 - x^2 + \frac{1}{3}x.$$ 

3.8. Show that the functions given by

$$f_j(x) = \sin((j + \frac{1}{2})\pi x), \quad j \in \mathbb{Z},$$

are orthonormal on $[-1, 1]$.

3.9. Show by utilizing Poisson’s Formula that 

$$\sum_{m=1}^{\infty} \frac{\sin m}{m} = \frac{\pi - 1}{2}.$$

3.10. Determine 

$$\sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1}$$

from the Fourier series of $\frac{1}{2} - x$ on $[0, 1]$.

3.11. (a) Determine the dispersion relation for the beam equation

$$\frac{\partial^2 u}{\partial t^2} + c^2 \frac{\partial^4 u}{\partial x^4} = 0.$$ 

(b) The same for the Korteweg-de Vries equation

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} + d \frac{\partial^3 u}{\partial x^3} = 0.$$ 

(c) The same for the Boussinesq equation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = b^2 \frac{\partial^4 u}{\partial x^2 \partial t^2}.$$ 

3.12. The dispersion relation for water waves is given by

$$\omega^2 = \left( 1 + \frac{T}{\rho g} \kappa^2 \right) g \kappa \tanh(\kappa h),$$

where $\omega$ is the frequency, $\kappa$ the wave number, and $h$ the undisturbed water height, while $g = 9.8 \text{ m/s}^2$ is the gravitational acceleration, $\rho = 1000 \text{ kg/m}^3$ the water density and $T = 0.074 \text{ N/m}$ the surface tension. Waves controlled by surface tension ($T \kappa^2 / \rho g$ is not small) are called ripples. Waves controlled by just gravity are called gravity waves.

(a) Verify that for deep water the phase velocity of gravity waves is twice the group velocity, so the waves are dispersive.
(b) Verify that for long waves \((\kappa \to 0)\) group and phase velocity become the same (and equal to \(\sqrt{gh}\)), so long waves propagate without dispersion.

(c) A practical parameter to maximize the wave number range of dispersionless waves is the water depth \(h\). A device that uses these waves to model physically sound waves (which are dispersionless) is called a ripple tank [74].

Consider \(V(z, \beta) = \frac{v_g}{\sqrt{gh}}\) as a function of \(z = \kappa h\) and \(\beta = T/\rho gh^2\). Our aim is to select \(\beta\) such that \(V\) remains close to 1 for a considerable interval \(0 < z < z_0\). Verify that \(V(0, \beta) = 1\) and \(V_z(0, \beta) = V_{zzz}(0, \beta) = 0\). For what value of \(\beta\) is \(V_z(0, \beta) = 0\)? This value produces practically dispersionless waves for \(z\) between 0 and 0.5, i.e. any wave length larger than \(4\pi h\).

By decreasing \(\beta\) slightly, the range in \(z\) may be increased with an acceptable deviation of \(V\). In terms of \(h\), suitable values are found for 5 – 8 mm, at a wave speed of 22 – 28 cm/s.

3.13. Show that for higher dimensional waves the group velocity, implied by the dispersion relation \(\omega = \omega(\kappa)\), is given by

\[
v_g = \frac{\delta \omega}{\delta \kappa}.
\]

\(\frac{\delta}{\delta \kappa}\) denotes the gradient with respect to \(\kappa\).
Chapter 4

Distributions and fundamental solutions

This chapter is devoted to rather fundamental concepts. In Section 1 we first sketch the idea behind a so-called fundamental solution. For a number of properties and phenomena of partial differential equations the concept of distribution is needed. Important functions like Dirac delta functions or Heaviside functions are an instance of this. In Section 2 we first consider distributions in one dimension, and define what we mean by convergence in a distributional sense. The extension to higher dimensions, which is rather straightforward is treated in Section 3. Distributions play a crucial rôle in problems that do not possess solutions “in a classical sense”, i.e. that are non-smooth. This then leads to a notion of solutions in a so-called “weak form”, which are solutions in distributional sense. They are discussed in Section 4. Another use of distributions is describing particular solutions of linear partial differential equations, so-called fundamental solutions, see Section 5. These fundamental solutions are defined on the whole \( \mathbb{R}^d \). A special form of such a solution in distributional sense is the Green’s function, which is such a solution that moreover satisfies the homogenous boundary condition. In fact the latter lead to expressions of the solution in terms of the source term of the equation. A more classical approach is to use a Duhamel integral, giving an expression for the solution of a partial differential equation by superposition of elementary solutions that represent the source term. These Duhamel integrals are discussed in Section 6. In fact it turns out that there is a natural relationship between these two forms.

1 Introduction

Consider the Cauchy problem for the ODE

\[
\begin{align*}
\frac{du}{dt} &= \lambda u + f(t), \quad t > 0, \\
u(0) &= u_0.
\end{align*}
\]

If \( f(t) \equiv 0 \), the solution, \( v \) say, is simply given by

\[ v(t) = e^{\lambda t} u_0. \]
In order to find the general solution of (1.1) one can use the so-called variation of constant method, i.e. substitute

\[ u(t) = c(t) v(t), \]

and determine \( c(t) \) from (1.1a). The well-known result is

\[ u(t) = e^{\lambda t} u_0 + \int_0^t v(t)[v(\tau)]^{-1} f(\tau) \, d\tau. \]  

(1.3)

Here \( w(t; \tau) \), defined as

\[ w(t; \tau) := v(t)[v(\tau)]^{-1}, \quad t \geq \tau, \]  

(1.4)

is sometimes called a \textit{fundamental solution}. For \( \tau \) fixed it satisfies the homogeneous part of (1.1a). One can view the integral in (1.3) as a superposition of initial value solutions propagating values \( f(\tau) \, d\tau \) until \( t \). One can also describe this as follows. Let

\[ \xi := t - \tau, \]  

(1.5a)

\[ \overline{w}(\xi) := w(t; \tau). \]  

(1.5b)

Clearly \( \overline{w}(\xi) = e^{\lambda \xi} \) for \( \xi \geq 0 \). From (1.3) we conclude that it makes sense to define \( \overline{w}(\xi) = 0 \) for \( \xi < 0 \). But beyond the point \( \xi = 0 \), \( \overline{w}(\xi) \) satisfies the ODE

\[ \frac{d\overline{w}}{d\xi} = \lambda \overline{w}, \quad \xi \neq 0, \]  

(1.6)

with solution \( \propto e^{\lambda \xi} \), so at the point \( \xi = 0 \) a jump occurs. This may be described by adding a (for the moment hypothetical) source term \( \delta(\xi) \) to the right hand side of (1.6), in such a way that

\[ \overline{w}(\xi) = \int_{-\infty}^\xi e^{\lambda(\xi-s)} \delta(s) \, ds = e^{\lambda \xi}, \quad \xi > 0. \]  

(1.7)

This is a non-trivial matter (even more: it is a condition impossible for any normal function to satisfy!) but in loose terms we apparently require \( \delta(\xi) \) to be a spike-like function which integrates to zero along any interval, except for an infinitesimally small interval around \( \xi = 0 \) where its integral contribution adds up to 1, i.e.

\[ \int_{-\infty}^\infty \delta(\xi) \, d\xi = 1. \]  

(1.8)

This function \( \delta \) is sometimes called the \textit{(Dirac) delta function}. In the next section, we shall give a more rigorous definition. Summarizing, we may view the fundamental solution \( w(t; \tau) \) to satisfy the Cauchy problem

\[ \frac{d}{dt} w(t; \tau) = \lambda w(t; \tau) + \delta(t-\tau), \quad t > \tau, \]  

(1.9a)

\[ w(\tau; \tau) = 0. \]  

(1.9b)

This notion can simply be extended to vector valued ODE and to boundary value problems as well. We finally remark that for Cauchy problems the value \( \tau = 0 \) needs special consideration; here, one just has a solution of the homogeneous problem, satisfying \( w(0, 0) = 1 \).
2 Distributions in one variable

In Section 1 we saw that we needed a notion of function of which some properties only made sense after integration. It took mathematicians a while before they had the proper formulation for the delta function. It shows up as a special instance of a distribution, to be discussed below.

Let \( D \) be some class of functions \( \mathbb{R} \to \mathbb{R} \), to be called test functions, and consider mappings, or functionals, from \( D \) to \( \mathbb{R} \). A simple but important class of examples is generated by the “inner” product of a test function \( \varphi \in D \) and a given integrable real function \( f : \mathbb{R} \to \mathbb{R} \), i.e.

\[
(f, \varphi) := \int_{-\infty}^{\infty} f(x) \varphi(x) \, dx.
\]  

(2.10)

If the functional can be written in this way, the functional is identified with the function \( f \), and we call it “the functional \( f \)”. We see that this functional is even linear, i.e.

\[
(f, \alpha \varphi_1 + \beta \varphi_2) = \alpha (f, \varphi_1) + \beta (f, \varphi_2).
\]  

(2.11)

A suitable test space can be found as follows. First we define the support of a function \( \varphi : \mathbb{R} \to \mathbb{R} \) as the closure of the set of all points \( x \) such that \( \varphi(x) \neq 0 \), i.e.

\[
\text{supp}(\varphi) = \{ x \in \mathbb{R} \mid \varphi(x) \neq 0 \}.
\]

Example 4.1 The infinitely many times differentiable real function \( \varphi \), defined by

\[
\varphi(x) = \begin{cases} 
\exp[(x^2 - 1)^{-1}] & \text{for } |x| < 1, \\
0 & \text{for } |x| \geq 1,
\end{cases}
\]

clearly has \( \text{supp}(\varphi) = [-1, 1] \). \( \square \)

If \( \text{supp}(\varphi) \) is a bounded set, then \( \varphi \) is said to have a compact support. Now define the test function space

\[
\mathcal{D} = C_0^\infty(\mathbb{R}) := \{ \varphi \in C^\infty(\mathbb{R}) \mid \text{\varphi has compact support} \}.
\]  

(2.12)

This space of test functions \( \mathcal{D} \) will be used here throughout, unless indicated otherwise.

In order to have a practically meaningful linear functional we like it to be continuous. For this we need a convergence concept. In view of the compact support property, the following makes sense.

**Definition 4.2.** A sequence of test functions \( \{ \varphi_i \}_{i \geq 0} \), where \( \varphi_i \in C_0^\infty(\mathbb{R}) \) is called convergent to 0 if

(i) there is a closed and finite interval \( I \) such that \( \text{supp}(\varphi_i) \subset I \);

(ii) \( \lim_{i \to \infty} \frac{d^k}{dx^k} \varphi_i(x) = 0 \), for \( k = 0, 1, 2, \ldots \), uniformly on \( \mathbb{R} \).

17:22 7 Oct 2003 57 version: 26-09-2003
From now on we shall identify $\mathcal{D}$ with $C_0^\infty(\mathbb{R})$, equipped with the convergence notion as defined in Definition 4.2.

**Definition 4.3.** A linear functional $f$ is continuous if for any sequence $\{\varphi_i\}_{i \geq 0} \subset \mathcal{D}$, convergent to 0, also $(f, \varphi_i) \to 0$.

(Because of linearity it suffices to consider continuity at zero.) The space of continuous linear functionals on $\mathcal{D}$ is the dual space $\mathcal{D}'$. A continuous linear functional is called a distribution. Actually we see that $\mathcal{D}'$ is a linear space by writing for $f, g \in \mathcal{D}'$

\[(\alpha f + \beta g, \varphi) = \alpha(f, \varphi) + \beta(g, \varphi) \quad \text{for all } \varphi \in \mathcal{D}. \quad (2.13)\]

If $f$ is locally integrable (i.e. the integral exists on any finite interval) then by (2.10) $f$ generates a distribution. Such a distribution is called regular, and may be identified with $f$. The delta function encountered in Section 1 can now be defined by

\[(\delta, \varphi) = \varphi(0), \quad \text{for all } \varphi \in \mathcal{D}. \quad (2.14)\]

This is not a regular distribution (see e.g. [106]). However, we will write this, as tradition does, like (2.10), i.e.

\[(\delta, \varphi) = \int_{-\infty}^{\infty} \delta(x) \varphi(x) \, dx. \quad (2.15)\]

Take due note that this is just symbolism. The delta function $\delta(x)$ should be interpreted via its definition (2.14). It is not a function in the classical sense.

If $f, g \in \mathcal{D}'$ are regular distributions, it follows from (2.10) that if $(f, \varphi) = (g, \varphi)$ for any $\varphi \in \mathcal{D}$, then $f(x) = g(x)$ almost everywhere. In general we call two distributions $f$ and $g$ identical if

\[(f, \varphi) = (g, \varphi) \quad \text{for all } \varphi \in \mathcal{D}. \quad (2.16)\]

If we shift the argument in a distribution $f$ by $\xi$, i.e.

\[f_{\xi}(x) := f(x - \xi), \quad (2.17)\]

we find

\[(f_{\xi}, \varphi) = (f, \varphi_{-\xi}). \quad (2.18)\]

For a shifted delta-function $\delta(x - \xi)$, we then define $\delta_{\xi}(x) = \delta(x - \xi)$ by

\[(\delta_{\xi}, \varphi) = (\delta, \varphi_{-\xi}) = \varphi(\xi) \quad \text{for all } \varphi \in \mathcal{D}. \quad (2.19)\]

We can also define a multiplication of $f \in \mathcal{D}'$ by an integrable function $g$, through the relation

\[(gf, \varphi) := (f, g\varphi). \quad (2.20)\]

Note that $gf$ again is a distribution. It follows that the product of a distribution and a regular distribution is defined. This is not necessarily the case for the product of two non-regular distributions.
Example 4.4 If we multiply the delta function by \( g(x) = x \), we obtain
\[
(g\delta, \varphi) = (\delta, g\varphi).
\]
Hence the test functions \( g\varphi \) are all 0 at \( x = 0 \). Consequently \( (g\delta, \varphi) = 0 \), or
\[
x\delta(x) \equiv 0.
\]
□

For an ordinary differentiable function \( f \) we find by partial integration
\[
(f', \varphi) = \left[ f(x)\varphi(x) \right]_{-\infty}^{\infty} - (f, \varphi') = -(f, \varphi'),
\]
(2.21)
since \( \varphi \) has compact support. Therefore, we define the derivative of a distribution \( f \) in general by
\[
(f', \varphi) := -(f, \varphi') \quad \text{for all} \quad \varphi \in \mathcal{D}.
\]
(2.22)
Clearly, \( f' \) is a distribution again. Note that \( f \) is, in this sense, arbitrarily many times differentiable.

An important application is the following. Let \( H(x) \) be defined by
\[
H(x) = \begin{cases} 
0, & \text{for } x < 0, \\
\frac{1}{2}, & \text{for } x = 0, \\
1, & \text{for } x > 0,
\end{cases}
\]
(2.23)
which is the so-called Heaviside function. It generates the Heaviside distribution if we write for any \( \varphi \in \mathcal{D} \)
\[
(H, \varphi) = \int_{-\infty}^{\infty} H(x)\varphi(x) \, dx = \int_{0}^{\infty} \varphi(x) \, dx.
\]
(2.24a)
From (2.22) we obtain
\[
(H', \varphi) = -(H, \varphi') = - \int_{0}^{\infty} \varphi'(x) \, dx = \varphi(0).
\]
(2.24b)
Hence we may identify the distribution \( H'(x) \) with \( \delta(x) \).

As will turn out in later chapters it is useful to have representations of the delta function as a convergent sequence of regular distributions. First, we introduce

Definition 4.5. Let \( \{f_i\}_{i \geq 0} \) be a sequence of distributions. Then this sequence converges to \( f \), denoted by \( f_i \to f \in \mathcal{D} \), if
\[
\lim_{i \to \infty} (f_i, \varphi) = (f, \varphi) \quad \text{for all} \quad \varphi \in \mathcal{D}.
\]
(2.25a)
Let \( I \) be some interval in \( \mathbb{R} \). If \( \{f_{\lambda}\}_{\lambda \in I} \) is a family of regular distributions, continuously parameterized by a parameter \( \lambda \), then we say: \( f_{\lambda} \to f \) as \( \lambda \to \lambda_0 \in I \), if
\[
\lim_{\lambda \to \lambda_0} (f_{\lambda}, \varphi) = (f, \varphi) \quad \text{for all} \quad \varphi \in \mathcal{D}.
\]
(2.25b)
2. DISTRIBUTIONS IN ONE VARIABLE

With these definitions we have the following

**Theorem 4.6.** Let \( f_\lambda \) be a locally integrable function with

(i) \( f_\lambda(x) \geq 0, \quad x \in \mathbb{R}, \quad \lambda > 0, \)

(ii) \( \int_{-\infty}^{\infty} f_\lambda(x) \, dx = 1, \quad \lambda > 0, \)

(iii) \( \lim_{\lambda \to 0} \int_{a}^{b} f_\lambda(x) \, dx = 1 \quad \text{for any} \quad a < 0 < b. \)

Then \( \{ f_\lambda \}_{\lambda \in (0, \infty)} \) is a family of regular distributions, with

\[ f_\lambda(x) \to \delta(x) \quad \text{for} \quad \lambda \to 0. \quad (2.26) \]

**Proof.** Consider for any \( \varphi \in \mathcal{D} \) the difference \( I(\lambda) := (f_\lambda, \varphi) - \varphi(0) \). From the properties of \( f_\lambda \) it follows that \( I(\lambda) = \int_{-\infty}^{\infty} f_\lambda(x)(\varphi(x) - \varphi(0)) \, dx \). We split the integration interval into three parts: \((-\infty, -a), (-a, a)\) and \((a, \infty)\) for some \( a > 0 \). Apparently \( \varphi \) is bounded on \( \mathbb{R} \), say, \( |\varphi(x)| < M \) for \( x \in \mathbb{R} \). Hence \( |\varphi(x) - \varphi(0)| \leq 2M \). Furthermore, \( \varphi \) is continuous, so for a given value of \( \varepsilon \), choose \( a \) such that \( |\varphi(x) - \varphi(0)| \leq \frac{1}{2}\varepsilon, \quad x \in (-a, a) \).

Finally, let \( \lambda_0 \) be such that \( \int_{-\infty}^{a} f_\lambda(x) \, dx > 1 - (4M)^{-1}\varepsilon \) for \( 0 < \lambda < \lambda_0 \). Then \( |I(\lambda)| \leq |\int_{-\infty}^{-a}() \, dx + \int_{a}^{\infty}() \, dx| + |\int_{-a}^{a}() \, dx| \leq 2M(4M)^{-1}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \). This shows (2.26).

A sequence \( f_\lambda \) as introduced in Theorem 4.6 is sometimes referred to as a delta sequence. Such a delta sequence has both theoretical and practical significance; the latter in a situation where one, e.g., needs numerical approximations.

**Example 4.7**

(i) The “top hat” sequence (see Fig. 4.1)

\[ f_\lambda(x) := \begin{cases} \frac{1}{2} \lambda^{-1}, & \text{for} \ x \in (-\lambda, \lambda), \\ 0 & \text{for} \ |x| > \lambda, \end{cases} \]

satisfies all requirements of Theorem 4.6.

![Figure 4.1. A top hat delta sequence](image)
(ii) Another well-known delta sequence is given by (see Fig. 4.2)

\[ f_{\lambda}(x) := \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{x^2}{2\lambda}\right), \quad x \in \mathbb{R}, \lambda > 0, \]

which is the probability density function for a normal distribution (in the ordinary statistical sense!) with variance \( \lambda \). The requirements from Theorem 4.6 can be verified by introducing the transformation \( y^2 = x^2/2\lambda \) and the error function

\[ \text{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-y^2} \, dy, \]

(cf. [6]), which has the property \( \text{erf}(\infty) = 1 \). This sequence has been important in probability theory, but also in parabolic problems, see Chapter 10. Finally, note that this \( f_{\lambda} \in C^\infty(\mathbb{R}) \).

\[ \text{Figure 4.2. An exponential delta sequence (\( \lambda = 0.01, 0.04, 0.09, 0.25, 0.49, 1.0 \)).} \]

3 Distributions in more variables

For \( f \) being a multivariate real function \( \mathbb{R}^n \to \mathbb{R} \), depending on, say, \( x = (x_1, x_2, \ldots, x_n)^T \) we can give straightforward generalisations of the foregoing concepts. The test space \( D \) now consists of functions in \( C^\infty(\mathbb{R}^n) \), defined by

\[ C_0^\infty(\mathbb{R}^n) := \{ \varphi \in C^\infty(\mathbb{R}^n) \mid \varphi \text{ has compact support} \}, \quad (3.27) \]

equipped with the following notion of convergence.

**Definition 4.8.** A sequence \( \{\varphi_i\}_{i \geq 0} \), with \( \varphi_i \in C_0^\infty(\mathbb{R}^n) \), is called convergent to 0 if

(i) there is a closed and bounded set \( \Omega \subset \mathbb{R}^n \), such that for all \( i \) \( \text{supp}(\varphi_i) \subset \Omega \),

(ii) \( \lim_{i \to \infty} \varphi_i(x) = 0 \) uniformly in \( \mathbb{R}^n \), and likewise all partial derivatives.
4. STRONG AND WEAK SOLUTIONS

A distribution is a continuous linear functional on $\mathcal{D}$. If $f(x)$ is locally integrable, then we can identify $f$ with its distribution, where
\[ (f, \varphi) := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x)\varphi(x) \, dx_1 \cdots dx_n. \] (3.28)

In particular, we can find the delta function $\delta(x)$, defined by
\[ (\delta, \varphi) = \varphi(0), \quad \text{for all } \varphi \in \mathcal{D}. \] (3.29)

Differentiation is defined by the relation
\[ \left( \frac{\partial f}{\partial x_i}, \varphi \right) = -\left( f, \frac{\partial \varphi}{\partial x_i} \right), \quad \text{for all } \varphi \in \mathcal{D}. \] (3.30)

Note that this “derivative” is again a distribution. Theorem 4.6 can be generalised straightforwardly.

**Example 4.9** We can give simple generalisations of the top hat and normal distribution representations of the multivariate delta function as follows
\[ f_{\lambda}(x) = \begin{cases} (2\lambda)^{-n} & \text{for } \|x\|_\infty \leq \lambda, \\ 0 & \text{for } \|x\|_\infty > \lambda, \end{cases} \] (3.31)
\[ f_{\lambda}(x) = \left( \frac{1}{\sqrt{2\pi\lambda}} \right)^n \exp \left( -\frac{\|x\|_2^2}{2\lambda} \right), \] (3.32)

where $x \in \mathbb{R}^n$, $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ and $\|x\|_2^2 = x_1^2 + \ldots + x_n^2$. □

If we have two distributions of a single variable we can form a direct product to obtain a multivariate distribution. Let $f_1$ and $f_2$ be distributions, then
\[ (f_1(x_1) f_2(x_2), \varphi(x_1, x_2)) := (f_1(x_1), (f_2(x_2), \varphi(x_1, x_2))), \] (3.33)

with the inner products defined in an obvious way. We have retained the arguments in $f_1$, $f_2$ and $\varphi$ for clarity.

**Example 4.10**
\[ \delta(x) = \delta(x_1) \delta(x_2) \ldots \delta(x_n). \]

4 Strong and weak solutions

In order to facilitate the following discussion we shall denote a general linear second order PDE in $d$ spatial variables, $x_1, \ldots, x_d$ say, as
\[ \mathcal{L}[u] = f, \] (4.34a)
where
\[ \mathcal{L}[u] := \sum_{i,j=1}^{d} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_j \frac{\partial u}{\partial x_i} + cu. \] (4.34b)

Let \( x := (x_1, x_2, \ldots, x_d)^T \in \Omega \subset \mathbb{R}^d \). If the operator \( \mathcal{L} \) is really second order, i.e. \( a_{ij} \neq 0 \) for at least one pair \((i, j)\), then we call \( u \) a strong solution of (4.34) if \( u \in C^2(\Omega) \). If \( \mathcal{L} \) is first order, all \( a_{ij} = 0 \), then \( u \) is a strong solution if \( u \in C^1(\Omega) \).

Assuming the coefficients in (4.34b) to be \( C^\infty \), we may look for solutions in distributional sense. We call \( u \) a generalised or weak solution if (4.34) is satisfied in distributional sense, i.e. if
\[ \langle \mathcal{L}[u], \varphi \rangle = \langle f, \varphi \rangle, \quad \text{for all } \varphi \in \mathcal{D}. \] (4.35)

In other words, \( \mathcal{L}[u] - f \) is identical to the zero-distribution. We may restrict \( \mathcal{D} \) to test functions with support within \( \Omega \). Then \( u \) is a weak solution in \( \Omega \). A strong solution is also a weak solution, but not the other way around.

We next introduce the concept of adjoint operator, defined in general by
\[ \langle \mathcal{L}[u], v \rangle := \langle u, \mathcal{L}^*[v] \rangle. \] (4.36)

If \( v \) is a test function, then using (3.30) we find in particular
\[ \mathcal{L}^*[v] := \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}v) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(b_i v) + cv. \] (4.37)

For the operator \( \mathcal{L} \) and its adjoint \( \mathcal{L}^* \), the following property holds.

**Property 4.11.** For \( u, v \in C^2(\mathbb{R}^d) \) we have
\[ v \mathcal{L}[u] - u \mathcal{L}^*[v] = \nabla \cdot w, \] (4.38a)
where the vector function \( w := (w_1, \ldots, w_d)^T \) is given by
\[ w_i = \sum_{j=1}^{d} \left( a_{ij} \frac{\partial u}{\partial x_j} - u \frac{\partial}{\partial x_j}(a_{ij}v) \right) + b_i uv, \quad i = 1, \ldots d. \] (4.38b)

**Proof.** The proof is trivial by direct substitution of (4.38b) into (4.38a). \( \square \)

Next, we have

**Corollary 4.12.** If \( u \) is a distribution and \( \varphi \in \mathcal{D} \), then
\[ \int_\Omega (\varphi \mathcal{L}[u] - u \mathcal{L}^*[\varphi]) \, dV = \int_{\partial \Omega} \mathbf{w} \cdot \mathbf{n} \, dS, \] (4.39)
where $\mathbf{n}$ is the unit outward normal of the boundary $\partial \Omega$ of $\Omega$.

**Proof.** Apply the divergence theorem (J.12) to (4.38a). 

**Example 4.13** Consider the Cauchy problem

$$\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = H(x), \quad x \in \mathbb{R}.$$

We see that the PDE has characteristics $x + t = c, \ c \in \mathbb{R}$. Hence we have a discontinuity along the line $x + t = 0$. So the solution appears to be

$$u(x, t) = H(x + t),$$

which clearly is not a classical solution. Let $\varphi$ be a test function and consider (cf. 3.30)

$$I := - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (u_t - u_x) \varphi \, dx \, dt = \int_{0}^{\infty} \int_{-\infty}^{\infty} u(x, t) \varphi_t \, dx \, dt.$$

We split the domain into the regions $R_1 = \{x + t > 0\}$ where $u = 1$ and $R_2 = \{x + t < 0\}$ where $u = 0$. Then

$$I = \int_{R_1} u(\varphi_t - \varphi_x) \, dx \, dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\varphi_t - \varphi_x) \, dx \, dt =$$

$$\int_{-\infty}^{0} \int_{-\infty}^{\infty} \varphi_t \, dx \, dt + \int_{0}^{\infty} \int_{-\infty}^{\infty} \varphi_t \, dx \, dt - \int_{0}^{\infty} \int_{-\infty}^{\infty} \varphi_x \, dx \, dt =$$

$$\int_{-\infty}^{0} -\varphi(x, -x) \, dx + \int_{0}^{\infty} -\varphi(x, 0) \, dx + \int_{0}^{\infty} \varphi(-t, t) \, dt = 0$$

because $\varphi$ vanishes along the borders of its domain, i.e. $\varphi(x, 0) = 0$. Hence $I = 0$ and so $u$ is a weak solution. 

The example above shows that the weak solution concept nicely captures “shock structures”. We shall encounter this in Chapter 12, where the “jump” condition in $u$ will be met again as a Rankine-Hugoniot condition.

We finally show that a weak solution with sufficient smoothness is always a strong solution.

**Theorem 4.14.** Let $u \in C^2(\Omega)$ be a weak solution of (4.34a) (or at least $C^1(\Omega)$ if $a_{ij} \equiv 0$). Then $u$ is a strong solution on $\Omega$.

**Proof.** From (4.35) and (4.39) we find

$$\int_{\Omega} \mathcal{L}[u] \varphi \, dV = \int_{\Omega} u \mathcal{L}^* \varphi \, dV = \int_{\Omega} f \varphi \, dV,$$

17:22 7 Oct 2003 64 version: 26-09-2003
since \( w \equiv 0 \) on \( \partial \Omega \). Hence
\[
\int_\Omega (\mathcal{L}[u] - f) \psi \, dV = 0.
\]
Since \( \mathcal{L}[u] - f \in C^0(\Omega) \) we conclude that \( \mathcal{L}[u] = f \). \( \square \)

In the sequel of this book we often encounter systems of PDEs of the form
\[
\frac{\partial u}{\partial t} = \mathcal{L}[u], \quad x \in \mathbb{R}, \ t > 0,
\]
(4.40)
where \( \mathcal{L} \) is a vector-valued differential operator containing only spatial derivatives, which is working on the components of the vector function \( u \in \mathbb{R}^m \). Analogously to (4.35) we can define a weak solution of (4.40) by the requirement
\[
\int_0^\infty \int_{\mathbb{R}^d} \left( \frac{\partial u}{\partial t} - \mathcal{L}[u] \right) \cdot \varphi \, dx \, dt = 0, \quad \text{for all } \varphi \in \mathcal{D}^m.
\]
(4.41)
In this definition, \( \varphi \) is a vector-valued test function with components in \( \mathcal{D} \). We will apply this definition in Chapter 12 to systems of hyperbolic equations.

5 Fundamental solutions

We now have the tools to generalise the findings of Section 1. The “solution” \( w \) found there will now be recognized as a weak solution.

Consider equation (4.34a). For a general linear differential operator \( \mathcal{L} \) and for any \( \xi \), we call \( w(x; \xi) \) a fundamental solution, if
\[
\mathcal{L}[w](x; \xi) = \delta(x - \xi), \quad x \in \mathbb{R}^d.
\]
(5.42)

**Example 4.15** A fundamental solution of the Laplace equation in \( \mathbb{R}^2 \), satisfying
\[
\nabla^2 w = \delta(x - \xi),
\]
is found to be given by
\[
w(x; \xi) := \frac{1}{2\pi} \ln(\|x - \xi\|_2).
\]
This is not evident, because when considered as an ordinary function, \( \nabla^2 w \) is identically zero, except at \( x = \xi \) where it does not exist. As a result, any integral \( \int_{\mathbb{R}^2} \varphi \nabla^2 w \, dx \) is zero, rather than equal to \( \psi(\xi) \). However, when we interpret \( \nabla^2 w \) as a distribution, the gradient is defined by Eq. (3.30), and we have
\[
\int_{\mathbb{R}^2} \varphi \nabla^2 w \, dx = -\int_{\mathbb{R}^2} \nabla \varphi \cdot \nabla w \, dx = -\int_0^{2\pi} \int_0^{\infty} \frac{\partial \varphi}{\partial r} \frac{\partial w}{\partial r} \, dr \, d\theta = \frac{1}{2\pi} \int_0^{2\pi} \varphi(\xi) \, d\theta = \varphi(\xi),
\]
with \( x - \xi \) written in polar co-ordinates \( r \) and \( \theta \). \( \square \)
5. FUNDAMENTAL SOLUTIONS

Now consider a general linear inhomogeneous problem, with linear boundary value operator $B$, where we would like the solution $u$ on a domain $\Omega$ say

\begin{align*}
L[u] &= f, \quad x \in \Omega, \quad (5.43a) \\
B[u] &= \beta, \quad x \in \partial \Omega. \quad (5.43b)
\end{align*}

Then we can construct $u$ as follows. First we seek a particular solution, $u_p(x)$ say, such that

\begin{align*}
L[u_p] &= f, \quad x \in \Omega, \quad (5.44a) \\
B[u_p] &= 0, \quad x \in \partial \Omega. \quad (5.44b)
\end{align*}

In order to find $u_p$ we employ special weak solutions derived from a fundamental solution by adding a suitable homogeneous (strong) solution of (5.44) (i.e. with $f = 0$) giving rise to so called Green’s functions $G(x; \xi)$

\begin{align*}
L[G](x; \xi) &= \delta(x - \xi), \quad x \in \Omega, \quad (5.45a) \\
B[G](x; \xi) &= 0, \quad x \in \partial \Omega. \quad (5.45b)
\end{align*}

(Note the difference between a fundamental solution and the Green’s function.) Since we can write $f(x)$ formally as a superposition of delta-functions as

\begin{equation}
 f(x) = \int_{\Omega} \delta(x - \xi) f(\xi) \, dV_\xi, \quad (5.46)
\end{equation}

where the integration is w.r.t. the $\xi$-coordinates, we can construct $u_p$ by convolution as

\begin{equation}
 u_p(x) = \int_{\Omega} G(x; \xi) f(\xi) \, d\xi. \quad (5.47)
\end{equation}

The complete solution of (5.43) is now found by adding a (homogenous) solution $u_h$ of

\begin{align*}
L[u_h] &= 0, \quad x \in \Omega \quad (5.48a) \\
B[u_h] &= \beta, \quad x \in \partial \Omega. \quad (5.48b)
\end{align*}

Such adding of a suitable homogenous solution to a particular solution (e.g. in order to make the boundary data fit) is called superposition.

Note that $u_h$, if it exists, need not be unique. This then gives rise to Fredholm’s alternative: either $u_h$ is unique on $\Omega$ or there exist infinitely many solutions for (5.48). The general solution of (5.43) is given by

\begin{equation}
 u(x) = \int_{\Omega} G(x; \xi) f(\xi) \, dV_\xi + u_h(x), \quad (5.49)
\end{equation}

which is an analogue of (1.3).
We remark that the fundamental solution \( w(x; \xi) \) can be found by translation of the argument from \( w(x; 0) \) if \( \mathcal{L} \) has constant coefficients, i.e.

\[
\w(x; \xi) = w(x - \xi; 0). \tag{5.50}
\]

We can generalise this concept of fundamental solution by taking \( t \) as an additional independent variable. So consider

\[
\mathcal{L}_1[u] + \mathcal{L}_2[u] = f, \tag{5.51}
\]

where \( \mathcal{L}_1 \) denotes a differential operator with respect to \( t \) and \( \mathcal{L}_2 \) a differential operator with respect to \( x \). For each \((\xi, \tau)\) a fundamental solution \( w(x, t; \xi, \tau) \) is then satisfying

\[
\mathcal{L}_1[w](x, t; \xi, \tau) + \mathcal{L}_2[w](x, t; \xi, \tau) = \delta(x - \xi) \delta(t - \tau), \quad x \in \mathbb{R}^d, \ t > 0, \tag{5.52a}
\]

\[
w(x, t; \xi, \tau) = 0, \quad x \in \mathbb{R}^d, \ t < \tau. \tag{5.52b}
\]

Note that this formulation takes care of the evolutionary character of the problem (causality principle) with time intervals \((\tau, \infty)\). We shall see later (see Chapter 10) that solutions of those problems can be found using such fundamental solutions. An alternative way is making use of Duhamel integrals, see next section.

## 6 Initial (boundary) value problems; Duhamel integrals

Most problems we are dealing with in this book have both a space and time dependence. In the linear case the equations are in fact either one of the following type

\[
\frac{\partial u}{\partial t} = \mathcal{L}_1[u] + f(x, t), \quad x \in \mathbb{R}^d, \ t > 0, \tag{6.53}
\]

\[
\frac{\partial^2 u}{\partial t^2} = \mathcal{L}_2[u] + f(x, t), \quad x \in \mathbb{R}^d, \ t > 0. \tag{6.54}
\]

Here \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are homogeneous (first or second order) differential operators in \( x \). In the chapters that will follow we shall investigate the solutions of the various types of PDE. Generally speaking, we need to specify a condition on \( \partial \Omega \), (a part of) the boundary of the spatial domain \( \Omega \), say

\[
\mathcal{B}[u] = \beta(t), \quad x \in \partial \Omega, \ t > 0. \tag{6.55}
\]

Moreover we need initial conditions at \( t = 0 \). For (6.53) they have the form

\[
u(x, 0) = a(x), \quad x \in \Omega. \tag{6.56}
\]

For (6.54) we have, besides (6.56) also

\[
\frac{\partial u}{\partial t}(x, 0) = b(x), \quad x \in \Omega. \tag{6.57}
\]
Now consider the following family of problems defined on \( \mathbb{R}^d \times (\tau, \infty) \), associated to (6.53), (6.55), (6.56),

\[
\frac{\partial w}{\partial t}(x, t; \tau) = \mathcal{L}_1[w](x, t; \tau), \quad x \in \mathbb{R}^d, \ t > \tau, \tag{6.58a}
\]

\[
\mathcal{B}[w](x, t; \tau) = 0, \quad x \in \partial \Omega, \ t > \tau, \tag{6.58b}
\]

\[
w(x, \tau; \tau) = f(x, \tau), \quad x \in \Omega. \tag{6.58c}
\]

These initial boundary value problems together build the actual solution; see the following theorem.

**Theorem 4.16 (Duhamel integral I).** Assume that the initial boundary value problem (6.53), (6.55), (6.56) with \( \beta \equiv 0, a \equiv 0 \) has a unique solution \( u \). Then it is given by the (so-called) Duhamel integral

\[
u(x, t) = \int_0^t w(x, t; \tau) \, d\tau, \tag{6.59}
\]

where \( w(\cdot, \cdot; \tau) \) is a solution of (6.58).

**Proof.** If we differentiate the right hand side in (6.59) with respect to \( t \) we get

\[
\frac{\partial}{\partial t} \int_0^t w(x, t; \tau) \, d\tau = w(x, t; t) + \int_0^t \frac{\partial w}{\partial t}(x, t; \tau) \, d\tau = f(x, t) + \int_0^t \mathcal{L}_1[w](x, t; \tau) \, d\tau = f(x, t) + \mathcal{L}_1 \left[ \int_0^t w(x, t; \tau) \, d\tau \right]
\]

Hence \( \int_0^t w(x, t; \tau) \, d\tau \) satisfies (6.53). Moreover it is easy to see that (6.55), with \( \beta \equiv 0 \), and (6.56), with \( a = 0 \), are satisfied, so that \( \int_0^t w(x, t; \tau) \, d\tau \) may be identified with the solution \( u(x, t) \).

**Example 4.17** Consider the problem

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x), \quad x \in \mathbb{R}, \ t > 0
\]

\[
u(x, 0) = 0, \quad x \in \mathbb{R}.
\]

It can be checked that the solution of this IVP is given by (6.59), with

\[
w(x, t; \tau) := \frac{1}{2\sqrt{\pi(t - \tau)}} \int_{-\infty}^{\infty} \exp \left( -\frac{(x - \xi)^2}{4(t - \tau)} \right) f(\xi) \, d\xi.
\]

\( \square \)

Similarly, consider a problem associated to (6.54), (6.55)-(6.57)

\[
\frac{\partial^2 w}{\partial t^2}(x, t; \tau) = \mathcal{L}_2[w](x, t; \tau), \quad x \in \mathbb{R}^d, \quad t > \tau, \quad (6.60a)
\]

\[B[w](x, t; \tau) = 0, \quad x \in \partial\Omega, \quad t > \tau, \quad (6.60b)\]

\[w(x, t; \tau) = 0, \quad x \in \mathbb{R}^d, \quad (6.60c)\]

\[\frac{\partial w}{\partial t}(x, t; \tau) = f(x, \tau), \quad x \in \mathbb{R}^d. \quad (6.60d)\]

Then we find

**Theorem 4.18 (Duhamel integral II).** Assume that the initial boundary value problem (6.54), (6.55), (6.56) with \(a = 0\) and (6.57) with \(b = 0\) has a unique solution. Then \(u\) is given by

\[u(x, t) = \int_0^t w(x, t; \tau) \, d\tau, \quad (6.61)\]

where \(w(x, t; \tau)\) is a solution of (6.60).

**Proof.** If \(w\) is a solution of (6.60) we find from differentiating the right hand side in (6.61) once

\[\frac{\partial}{\partial t} \int_0^t w(x, t; \tau) \, d\tau = \int_0^t \frac{\partial w}{\partial t}(x, t; \tau) \, d\tau. \quad \]

Differentiating this relation once more then gives

\[\frac{\partial^2}{\partial t^2} \int_0^t w(x, t; \tau) \, d\tau = \frac{\partial w}{\partial t}(x, t; t) + \int_0^t \frac{\partial^2 w}{\partial t^2}(x, t; \tau) \, d\tau = f(x, t) + \mathcal{L}_2\left[\int_0^t \mathcal{W}(x, t, \tau) \, d\tau \right], \quad \]

which shows that (6.61) is the solution of (6.54)-(6.57) with homogeneous conditions indeed.

If \(a \neq 0\) in (6.56), we find instead of (6.61)

\[u(x, t) = w(x, t; 0) a(x) + \int_0^t w(x, t; \tau) \, d\tau. \quad (6.62a)\]

Moreover, if \(b \neq 0\) in (6.57), we obtain

\[u(x, t) = v(x, t) a(x) + w(x, t; 0) b(x) + \int_0^t w(x, t; \tau) \, d\tau. \quad (6.62b)\]
where \( w \) is a solution of (6.60) and \( v \) a solution of

\[
\frac{\partial^2 v}{\partial t^2}(x, t) = \mathcal{L}[v](x, t), \quad x \in \Omega, \; t > 0, \tag{6.63a}
\]
\[
\mathcal{B}[v](x, t) = 0, \quad x \in \partial \Omega, \; t > 0, \tag{6.63b}
\]
\[
v(x, 0) = 1, \quad x \in \Omega, \tag{6.63c}
\]
\[
\frac{\partial v}{\partial t}(x, 0) = 0, \quad x \in \Omega. \tag{6.63d}
\]

The Duhamel integral is in fact nothing but superposition of fundamental solutions (with respect to time), see

**Theorem 4.19.** Let \( \mathcal{D}' \) be the set of distributions in the parameter \( t \).

(i) Then \( w \) satisfying (6.58), with \( f(x, t) \equiv 1 \), is a fundamental solution of

\[
\frac{\partial w}{\partial t}(x, t; \tau) = \mathcal{L}_1[w](x, t; \tau) + \delta(\tau), \tag{6.64a}
\]

(ii) \( w, \) satisfying (6.60), with \( f(x, t) \equiv 1 \), is a fundamental solution of

\[
\frac{\partial^2 w}{\partial t^2}(x, t; \tau) = \mathcal{L}_2[w](x, t; \tau) + \delta(\tau). \tag{6.64b}
\]

Here we have defined \( w(x, t) \equiv 0 \) for \( t < \tau \).

**Proof.**

(i) From (6.58) we obtain

\[
\int_0^\infty \left( \frac{\partial w}{\partial t} - \mathcal{L}_1[w] \right) \varphi \, dt = -\int_0^\infty \left( \frac{\partial \varphi}{\partial t} w + \varphi \mathcal{L}_1[w] \right) \, dt =
\]

\[
-\int_\tau^\infty \left( \frac{\partial \varphi}{\partial t} w + \varphi \mathcal{L}_1[w] \right) \, dt = \varphi(\tau) + \int_\tau^\infty \left( \frac{\partial w}{\partial t} - \mathcal{L}_1[w] \right) \varphi \, dt = \varphi(\tau).
\]

(ii) Likewise, from (6.60) we deduce

\[
\int_0^\infty \left( \frac{\partial w}{\partial t} - \mathcal{L}_2[w] \right) \varphi \, dt = \int_0^\infty \left( \frac{\partial^2 \varphi}{\partial t^2} w - \varphi \mathcal{L}_2[w] \right) \, dt =
\]

\[
\int_\tau^\infty \left( \frac{\partial^2 \varphi}{\partial t^2} w - \varphi \mathcal{L}_2[w] \right) \, dt = \varphi(\tau) + \int_\tau^\infty \left( \frac{\partial^2 w}{\partial t^2} - \mathcal{L}_2[w] \right) \varphi \, dt = \varphi(\tau).
\]
Example 4.20 Consider the same problem as in Example 4.17. As one can easily see a fundamental solution, i.e. a solution of
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \delta(x - \xi)\delta(t - \tau), \quad x \in \mathbb{R}, \ t > 0,
\]
\[u(x, t) = 0, \quad x \in \mathbb{R}, \ t < \tau,
\]
is given by
\[w(x, t; \xi, \tau) := \frac{1}{2\sqrt{\pi(t - \tau)}} \exp\left(-\frac{(x - \xi)^2}{4(t - \tau)}\right).
\]
Hence \(u(x, t; \tau) = \int_{-\infty}^{\infty} w(x, \xi, t, \tau) d\xi\). □

As a final application of the Duhamel integral consider initial boundary value problems on a semi-infinite domain, e.g. \(\Omega = [0, \infty)\), which have homogeneous source term and initial condition(s) (i.e. \(a \equiv 0\) in (6.56) and/or \(b \equiv 0\) in (6.57)). Let the BC (6.55) be given by
\[u(0, t) = \beta(t).
\]
Assume that the function \(w(x, t; \tau)\) can actually be written as
\[\bar{w}(x, t - \tau) := w(x, t; \tau).
\]
Let \(\bar{w}\) satisfy (6.53) or (6.54) with \(f \equiv 0\), (6.56) with \(a \equiv 0\) and (6.65)
\[
\frac{\partial \bar{w}}{\partial t} = \mathcal{L}_{i}[\bar{w}], \quad x \in \mathbb{R}, \ t > \tau, \quad (i = 1, 2),
\]
\[\bar{w}(0, t) = 1, \quad t > 0,
\]
\[\bar{w}(x, 0) = 0, \quad x \in \mathbb{R}.
\]
Then we have the following representation

Property 4.21. The solution of (6.53) or (6.54) with \(f \equiv 0\), (6.64) with \(a \equiv 0\) and (6.56) is given by
\[u(x, t) = \frac{\partial}{\partial t} \int_{0}^{t} \bar{w}(x, t - \tau) \beta(\tau) \, d\tau.
\]

Proof. Denote the right hand side in (6.68) by \(q(x, t)\), then
\[q(x, t) = \bar{w}(x, t)\beta(0) + \int_{0}^{t} \bar{w}(x, t - \tau) \beta'(\tau) \, d\tau.
\]
and using all properties of $\bar{w}$, we find
\[
\frac{\partial q}{\partial t}(x, t) = \frac{\partial}{\partial t} \bar{w}(x, t - \tau) \beta(0) + \int_0^t \frac{\partial}{\partial t} \bar{w}(x, t - \tau) \beta'(\tau) \, d\tau,
\]
\[
\frac{\partial^2 q}{\partial t^2}(x, t) = \frac{\partial^2}{\partial t^2} \bar{w}(x, t - \tau) \beta(0) + \int_0^t \frac{\partial^2}{\partial t^2} \bar{w}(x, t - \tau) \beta'(\tau) \, d\tau,
\]
\[
\mathcal{L}_i[q](x, t) = \mathcal{L}_i[\bar{w}](x, t) \beta(0) + \int_0^t \mathcal{L}_i[\bar{w}](x, t - \tau)) \beta'(\tau) \, d\tau, \quad (i = 1, 2).
\]
This shows that $q$ satisfies the PDE (6.53). Furthermore we have
\[
q(0, t) = \beta(0) + \int_0^t \beta'(\tau) \, d\tau = \beta(t),
\]
so that it also satisfies the boundary condition. \[\square\]

7 Discussion

- Distributions are a rather fundamental concept and a bit of an outsider in this book as far as rigour is concerned, cf. [89]. From a mathematical point of view they open entire new vistas in which concepts like differentiability can be treated for larger classes of problems, like in theory for pseudo-differential operators, see e.g. [112]. They will turn out to be essential, though, when it comes to hyperbolic equations. Here one cannot solve actual problems without taking recourse to weak solutions.

- Fundamental solutions have a value in their own right as will turn out e.g. in Chapters 8 and 10. Another interesting application is their use in the Boundary Element Method (BEM), where a boundary value problem first is transformed to an Integral Equation by using Green’s theorem after which the latter is discretised. For boundary value problems on infinite domains, as well as for certain other problems this method can be very attractive, see e.g. [18], [41].

Exercises

4.1. Consider the set of functions
\[
f(x; t) = \frac{t}{\pi(x^2 + t^2)}, \quad x \in \mathbb{R}, \ t > 0.
\]
Show that $\lim_{t \to 0} f(x; t) = \delta(x)$. 

17:22 7 Oct 2003 72 version: 26-09-2003
4.2. Let the following set of (locally) integrable functions \( f(x; t) \), for \( x \in \mathbb{R}, t > 0 \) be given, which have the properties

\[
\begin{align*}
&f(x; t) \geq 0, \\
&\int_{-\infty}^{\infty} f(x; t) \, dx = 1, \quad t > 0, \\
&\lim_{t \downarrow 0} \int_{a}^{b} f(x; t) \, dx = 1, \quad \text{for } a < 0 < b.
\end{align*}
\]

Show that \( \lim_{t \downarrow 0} f(x; t) = \delta(x) \).

4.3. Show that a (fundamental) solution of the problem \( \nabla^2 w(x; \xi) = \delta(\xi), \quad x \in \mathbb{R}^3 \) is given by \( w(x; \xi) = -\frac{1}{4\pi \|x - \xi\|^2} \).

4.4. Determine the Green’s function of the problem

\[
\frac{du}{dx} = 0, \quad x \in (\alpha, \beta), \\
u(\alpha) = u_\alpha, \quad u(\beta) = u_\beta.
\]

4.5. Given the form of a fundamental solution as found in exercise 3, determine a Greens function for the problem

\[
\nabla^2 u = 0, \quad x \in \Omega := S_{0,1}, \\
u = g(x), \quad x \in \partial \Omega.
\]

where \( S_{0,1} \) denotes the unit sphere of radius 1 centred at the origin. Hint: use an appropriate mirror point.

4.6. Consider the problem

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in \mathbb{R}, \quad t > 0, \\
u(x, 0) = f(x), \quad x \in \mathbb{R}.
\]

Use a Duhamel integral to show that the solution is given by

\[
\frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} \exp \left( -\frac{(x - \xi)^2}{4t} \right) f(\xi) \, d\xi.
\]

4.7. Under the same assumptions as in Property 4.21 show that the following holds

\[
u(x, t) = \int_{0}^{t} \beta(t) \frac{\partial}{\partial \tau} \bar{w}(x, t - \tau) \, d\tau.
\]
Chapter 6

The equations of continuum mechanics and electromagnetics

Important areas of applications pertaining to the methods described in this book are found in continuum physics. They are based on an almost axiomatic footing of conservation laws, completed with problem dependent constitutive relations. Some care is needed when the well-established laws for mass, momentum, and energy conservation are reformulated for a continuum. They will therefore be given in detail. The equations for compressible viscous flow and for linear elastic deformations will be written out in full. In view of their importance the equations for electromagnetic fields will also be derived, but without taking full account of possible coupling between electromagnetic forces and stresses in the material.

1 Introduction

The major areas of application of the problems and methods considered in this book are found in the physics of continua, which encompasses the theories of fluid dynamics, deformation of elastic media and electromagnetic phenomena. Therefore, we will summarise the respective theories here. In the context of this book this is necessarily very brief and concise. We will, for example, not consider the combined effect of electromagnetic and inertial forces. The reader is advised to consult the extensive existing literature for this and other details and for further background information.

Continuum physics considers the deformation and motion of matter under internal and external forces (inertia, stresses, gravity, or electromagnetic fields) at a macroscopic level, disregarding the molecular structure other than by its integrated effects. The prevailing equations are based on the postulates that mass, momentum and energy are conserved. Therefore these equations are called conservation equations. They are universal and do not contain the properties of the material in question; the number of unknowns is larger than the number of equations. Therefore, they are not sufficient to determine the problem, and we need in addition so-called constitutive relations. These relations represent the properties of the material considered, and their choice is part of the modelling process (see Chapter 7).
We will start with the equations for fluid flow and elastic deformation, leaving aside any electromagnetic effects. Then we will present the equations for electromagnetic fields, without dilating upon any mechanical coupling (except the production of heat).

2 Eulerian and Lagrangian coordinates

Consider the deformation and motion of material continuously distributed in some physical domain. We adopt the continuum hypothesis to consider the material as being made up of a coherent collection of 'particles', each consisting of sufficiently many, but not too many, molecules in order to allow us to speak pointwise of velocity, pressure, temperature etc. Thus, we study the matter at a macroscopic level and do not consider explicitly the interaction between the individual molecules. The particles interact with each other via contact forces (stresses) that depend on the material considered.

There are two approaches to describe a deformation or motion. In the Lagrangian description we follow particles that move with the continuum. In the alternative Eulerian description we maintain a fixed position and consider particles that pass this position. Let the position \( x \) of a particle at time \( t \) be written as

\[
x = \varphi(x^*, t),
\]

where \( x^* \) is a reference position, for which we take the position of this particle at an initial time \( t_0 \), i.e. \( x^* = \varphi(x^*, t_0) \). Thus a particle can be specified by its Lagrangian coordinates \( (x^*, t) \), when following the motion, or by its Eulerian coordinates \( (x, t) \). A generic variable \( f \) can be expressed in terms of the Lagrangian or the Eulerian coordinates, and we write \( f = f(x, t) = f^*(x^*, t) \). The velocity \( v^*(x^*, t) \) of a particle is the time rate of change of the position of this particle, expressed in Lagrangian coordinates, i.e.

\[
v^* := \frac{\partial \varphi}{\partial t}(x^*, t).
\]

The time rate of change of a variable \( f^*(x^*, t) \) may be expressed in Eulerian coordinates as the material or convective derivative of \( f \) and is given by

\[
\frac{\partial f^*}{\partial t}(x^*, t) = \frac{df}{dt}(\varphi(x^*, t), t) = \frac{\partial f}{\partial t}(x, t) + v(x, t) \cdot \nabla f(x, t).
\]

The first term on the right hand side of (2.3) is the local time derivative and the second term is the convective derivative, which is in fact the directional derivative of \( f \) in the direction of velocity \( v \).

The displacement vector is defined by

\[
\mathbf{u} = x - x^* = \varphi(x^*, t) - x^*.
\]

In the theory of small elastic deformations we need the linear deformation tensor or linear strain tensor

\[
\varepsilon := \frac{1}{2} \nabla \mathbf{u} + \frac{1}{2} (\nabla \mathbf{u})^T.
\]
Under the assumption of the deformations being small, the gradient $\nabla u$ may be either interpreted as to $x$ or $x^*$. In the theory of Newtonian viscous fluid flow (Section 8.2) we need the deformation velocity tensor or the rate of deformation tensor

$$D := \frac{1}{2} \nabla v + \frac{1}{2} (\nabla v)^T.$$ 

(Note that the definition of tensor $\nabla v$ is ambiguous. Both $\left( \frac{\partial v_i}{\partial x_j} \right)$ and $\left( \frac{\partial v_j}{\partial x_i} \right)$ occur. Due to the symmetry of $D$ this ambiguity is not important here. The same is true for $E$.)

**Example 6.1** A piece of rubber is stretched uniformly in all directions, such that its local coordinate system $x$ is related to its rest position $x^*$ by $x = (1 + \alpha)x^*$. The displacement vector is then

$$u = x - x^* = \alpha x^* = \frac{\alpha}{1+\alpha} x.$$  

The linear deformation tensor is thus for small $\alpha$ given by $E = \alpha I$. □

**Example 6.2** A velocity field directed along the $x$-axis with linear shear in $y$-direction, is given by $v = \alpha y e_x$. The deformation velocity tensor is then

$$D = \frac{1}{2} \alpha \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$ □

### 3 The transport theorem

A rigorous derivation of the conservation laws is based on the transport theorem, which we will derive first. The theorem is most relevant for a moving fluid, although it remains equally valid for a deforming solid.

Let $q(x, t)$ be a quantity per unit volume of the material (a fluid, say). Consider a control volume $\Omega(t)$ moving with the flow. $\Omega$ is called a material volume and its surface $\partial \Omega$ is called a material surface. Define

$$F(t) := \int_{\Omega(t)} q(x, t) \, dV.$$  

For example, if $q(x, t) = \rho(x, t)$ the mass density of the fluid, then $F(t)$ is the total mass of fluid contained in the control volume. The time rate of change of $F(t)$ is given by

$$\frac{dF(t)}{dt} = \int_{\Omega(t)} \left( \frac{\partial q}{\partial t}(x, t) + \nabla \cdot (q v)(x, t) \right) \, dV.$$  

Equation (3.2) is called the transport theorem (see also equation J.19).
4. CONSERVATION EQUATIONS

A concise proof is as follows. Consider $F(t+h) - F(t)$ asymptotically for small $h$. We write symbolically $\Omega(t + h) = \Omega(t) + d\Omega$. Then we have

$$F(t + h) - F(t) \approx \int_{\Omega(t) + d\Omega} q(x, t + h) \frac{\partial q}{\partial t}(x, t) \, dV - \int_{\Omega(t)} q(x, t) \, dV$$

$$\approx \int_{d\Omega} q(x, t) \, dV + h \int_{\Omega(t)} \frac{\partial q}{\partial t}(x, t) \, dV.$$

Apply locally near the surface $\partial\Omega$ of $\Omega$ an orthogonal coordinate system $x = \sigma + \lambda n$, where $\lambda = 0$ denotes the surface $\partial\Omega$ and $n$ is the unit outward normal of $\partial\Omega$. As the volume $\Omega$ is a material volume, it moves with the fluid velocity $v$, and therefore the surface $\partial\Omega$ has moved a distance $hv$ between time $t$ and $t + h$. So we have

$$\int_{d\Omega} q(x, t) \, dV = \int_{\partial\Omega} \int_0^{h(v \cdot n)} q(\sigma + \lambda n, t) \, d\lambda \, dS \approx \int_{\partial\Omega} q(\sigma) h(v \cdot n) \, dS = h \int_{\Omega} \nabla \cdot (qv) \, dV$$

where we used Gauss's divergence theorem for the last step. After dividing by $h$ and taking the limit $h \to 0$ we obtain (3.2).

Example 6.3 A flow field $v$, that keeps the content of any convected volume $\Omega(t)$ constant, is called incompressible. It satisfies

$$\frac{d}{dt} \int_{\Omega(t)} 1 \, dV = \int_{\Omega(t)} \nabla \cdot v \, dV = 0,$$

and therefore is divergence free, i.e. $\nabla \cdot v = 0$.

4 Conservation equations

Consider the deformation and motion of matter in some domain $\Omega(t)$, moving with the material (i.e. $\Omega$ is a material volume). From physics we know that the matter is subject to some very strict limitations, called conservation laws. These laws postulate that without source certain properties like mass, momentum and energy remain unchanged. Any such property $P$ can be described by a density $E$ (the amount of $P$ per unit volume) and an associated flux $F$ (the amount of $P$ that flows per unit time through a unit material surface normal to $F$) such that the change of $P$ of a given material volume $\Omega$ must be exactly equal to the sum of the net influx through the volume’s surface $\partial\Omega$ and any possible production from a source distribution $Q$:

$$\frac{d}{dt} \int_{\Omega} E \, dV = - \int_{\partial\Omega} F \cdot n \, dS + \int_{\Omega} Q \, dV,$$

where $n$ denotes the unit outward normal of $\partial\Omega$. As $\Omega$ moves with the flow with velocity $v$, this is according to the transport theorem and Gauss’s divergence theorem equivalent to

$$\int_{\Omega} \left( \frac{\partial E}{\partial t} + \nabla \cdot (Ev + F) - Q \right) \, dV = 0.$$.  

17:22 7 Oct 2003 78 version: 26-09-2003
CHAPTER 6. CONTINUUM MECHANICS AND ELECTROMAGNETICS

If this is true for any material volume in our region of interest, the integrand itself must vanish, so we have

\[ \frac{\partial E}{\partial t} + \nabla \cdot (Ev + F) = Q. \] (4.3)

It is clear that while \( F \) denotes the flux through a material surface, \( Ev + F \) is the flux through a fixed surface. Relation (4.3) is the general form of a conservation law for a continuum. In the following we will derive specific versions of this general format.

5 Conservation of mass

As there exist in the present context no mass sources, the mass of any material volume \( \Omega(t) \) is constant, so

\[ \frac{d}{dt} \int_{\Omega(t)} \rho \, dV = 0. \] (5.1)

Applying the transport theorem to (5.1) with \( q(x, t) = \rho \) gives

\[ \int_{\Omega(t)} \left( \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) \right) \, dV = 0. \] (5.2)

Since this conservation law holds for any \( \Omega(t) \), the differential form of the mass conservation law

\[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0, \] (5.3)

must be satisfied. Equation (5.3) is called the continuity equation, written in conservative form. By using the material derivative (2.3) we can rewrite it into convective form

\[ \frac{d\rho}{dt} + \rho \nabla \cdot v = 0. \] (5.4)

6 Conservation of momentum

The equations of motion of a continuum describe conservation of linear momentum and angular momentum. First, the law of conservation of linear momentum reads

\[ \frac{d}{dt} \int_{\Omega(t)} \rho v \, dV = \int_{\partial\Omega(t)} t(n) \, dS + \int_{\Omega(t)} f \, dV, \] (6.1)

with \( t(n) \) the stress vector, i.e. the internal or contact force field per unit area, acting on the boundary \( \partial\Omega \) of the material volume \( \Omega \) and \( f \) the specific (i.e. per unit mass) external or volume force field, acting on the material contained by \( \Omega \). In (6.1) we have explicitly written \( t(n) \) to denote the dependence of the stress vector on the outward unit normal \( n \) on \( \partial\Omega \). The conservation law in (6.1) states that the rate of change of momentum of the material contained in \( \Omega \), due to the movement of \( \Omega \) with velocity \( v \), is equal to the sum of contact forces and volume forces acting on the material.
Secondly, the law of conservation of angular momentum is expressed by
\[
\frac{d}{dt} \int_{\Omega(t)} \mathbf{x} \times \rho \mathbf{v} \, dV = \int_{\partial \Omega(t)} \mathbf{x} \times \mathbf{t}(\mathbf{n}) \, dS + \int_{\Omega(t)} \mathbf{x} \times \rho \mathbf{f} \, dV,
\]  
(6.2)
meaning that the rate of change of angular momentum of the material in \(\Omega(t)\), when the control volume moves with the continuum, is equal to the sum of the moment of the contact forces and the moment of the volume forces acting on it.

In the following, the conservation law of linear momentum will be used to develop the equations of motion for a continuum. From the conservation law of angular momentum a certain symmetry in the stress vector will be derived.

Consider the \(i\)-th \((i = 1, 2, 3)\) component of the conservation law of linear momentum (6.1). Application to its left-hand side of transport theorem (3.2) with \(q(x,t) = \rho(x,t)v_i(x,t)\) yields
\[
\frac{d}{dt} \int_{\Omega(t)} \rho v_i \, dV = \int_{\Omega(t)} \left( \frac{\partial}{\partial t} (\rho v_i) + \nabla \cdot (\rho \mathbf{v} v_i) \right) \, dV.
\]  
(6.3)
The stress vector \(\mathbf{t}(\mathbf{n})\) acting on a surface with normal \(\mathbf{n}\) is completely determined by the stress vectors \(\mathbf{t}(\mathbf{e}_j)\) for the unit vectors \(\mathbf{e}_j\) \((j = 1, 2, 3)\). The stress tensor \(\mathbf{T}\) is defined by
\[
\mathbf{T} = (\mathbf{T}_{ij}) = (\mathbf{t}(\mathbf{e}_1), \mathbf{t}(\mathbf{e}_2), \mathbf{t}(\mathbf{e}_3)) \quad \text{with} \quad \mathbf{T}_{ij} := t_i(e_j),
\]  
(6.4)
i.e. \(\mathbf{T}_{ij}\) is the \(i\)-th component of the stress vector \(\mathbf{t}\) acting on a surface with unit normal \(\mathbf{e}_j\). (Note that there is no uniformity in the nomenclature. Some authors define the stress tensor as \(\mathbf{T}^T\). Since \(\mathbf{T}\) will in general be symmetric, this is usually of no concern.) Applying the principle of local equilibrium, it follows that \([7, 23, 33]\)
\[
t_i(\mathbf{n}) = \sum_{j=1}^{3} \mathbf{T}_{ij} \mathbf{n}_j, \quad i = 1, 2, 3,
\]  
(6.5a)
or
\[
t(\mathbf{n}) = \mathbf{T} \mathbf{n}.
\]  
(6.5b)
The surface integral in (6.1) can now be replaced by a volume integral, using Gauss’ theorem (J.12), and we find
\[
\oint_{\partial \Omega(t)} \mathbf{t}(\mathbf{n}) \, dS = \int_{\partial \Omega(t)} \mathbf{T} \mathbf{n} \, dS = \int_{\Omega(t)} \nabla \cdot \mathbf{T}^T \, dV,
\]  
(6.6)
where \(\nabla \cdot \mathbf{T}^T\) is defined by
\[
\nabla \cdot \mathbf{T}^T := \begin{pmatrix} \nabla \cdot \mathbf{T}_{1*} \\ \nabla \cdot \mathbf{T}_{2*} \\ \nabla \cdot \mathbf{T}_{3*} \end{pmatrix}
\]  
(6.7)
and \(\mathbf{T}_{is}\) denotes the \(i\)-th row of \(\mathbf{T}\). Combining (6.1), (6.3) and (6.6), the \(i\)-th component of the conservation law of momentum reads
\[
\int_{\Omega(t)} \left( \frac{\partial}{\partial t} (\rho v_i) + \nabla \cdot (\rho \mathbf{v} v_i) \right) \, dV = \int_{\Omega(t)} \nabla \cdot \mathbf{T}_{is} \, dV + \int_{\Omega(t)} \rho f_i \, dV.
\]  
(6.8)
CHAPTER 6. CONTINUUM MECHANICS AND ELECTROMAGNETICS

This becomes in differential, conservative, form

$$
\frac{\partial}{\partial t}(\rho v_i) + \nabla \cdot (\rho vv_i) = \nabla \cdot T_i + \rho f_i, \quad i = 1, 2, 3,
$$

(6.9a)
or in vectorial notation

$$
\frac{\partial}{\partial t}(\rho v) + \nabla \cdot (\rho vv^T) = \nabla \cdot T^T + \rho f.
$$

(6.9b)

The divergence of the dyadic product $\rho vv^T$ is defined analogously to (6.7).

From the conservation of angular momentum (6.2) it transpires (see below) that $T = T^T$, in other words, the stress tensor $T$ is symmetric (for non-polar media). Therefore, in equation (6.9b) we can safely change $\nabla \cdot T^T$ into $\nabla \cdot T$. If, in addition, we use the continuity equation (5.3) we finally get the equation in its best known, convective, form

$$
\rho \frac{d}{dt} v = \nabla \cdot T + \rho f.
$$

(6.10)

This is known as Cauchy’s equation of motion [23].

A brief proof for the symmetry of $T$ is as follows [7]. By employing the transport theorem (3.2) and the continuity equation (5.3), we can derive for the left hand side of (6.2)

$$
\int_{\Omega(t)} x \times \rho v \, dV = \int_{\Omega(t)} \rho \frac{d}{dt} (x \times v) \, dV = \int_{\Omega(t)} \rho x \times \frac{d}{dt} v \, dV.
$$

(6.11)

The second integral in (6.2) can be rewritten by means of Gauss’ theorem (J.12) into

$$
\int_{\partial \Omega(t)} x \times t(n) \, dS = \int_{\partial \Omega(t)} (x \times T) \cdot n \, dS = \int_{\Omega(t)} (x \times (\nabla \cdot T^T + t^*) \, dV,
$$

(6.12)

where $t^*$ denotes a vector, related to the anti-symmetric part of $T$, given by

$$
t^* := \begin{pmatrix}
T_{32} - T_{23} \\
T_{13} - T_{31} \\
T_{21} - T_{12}
\end{pmatrix}.
$$

(6.13)

By combining (6.2) with (6.11) and (6.12) we get

$$
\int_{\Omega(t)} x \times (\rho \frac{d}{dt} v - \nabla \cdot T^T - \rho f) \, dV = \int_{\Omega(t)} t^* \, dV,
$$

(6.14)

Together with (5.3) and (6.9b) this implies that $t^*$ vanishes, i.e., $T$ is symmetric.

7 Conservation of energy

Finally, we have the law of conservation of energy. The rate of change of kinetic energy ($\frac{1}{2} \rho |v|^2$) and internal energy ($\rho e$) should equal the mechanical power of the stresses
8. Constitutive Relations and Thermodynamic Relations

The constitutive relations and volume forces $(\rho f \cdot v)$ acting on the material, plus the heat supplied by internal sources $(\rho r)$ and exchanged across the border $(-q \cdot n)$. This is given by

$$\frac{d}{dt} \int_{\Omega(t)} \rho E \, dV = \oint_{\partial \Omega(t)} (t(n) \cdot v) \, dS + \int_{\Omega(t)} \rho f \cdot v \, dV - \oint_{\partial \Omega(t)} q \cdot n \, dS + \int_{\Omega(t)} \rho r \, dV,$$

(7.1)

with $q$ the heat flux vector and $r$ the specific heat supply, for example by a distribution of radioactive sources or electric (Joule) heating (see equation 9.11). The specific energy $E$ is defined by

$$E := e + \frac{1}{2}|v|^2,$$

(7.2)

where $e$ the specific internal energy (a thermodynamic property) of the material and $\frac{1}{2}|v|^2$ the specific kinetic energy of the continuum. The term $-q \cdot n$ is the amount of energy per unit area and per unit time, which is transmitted through $\partial \Omega(t)$ to the fluid in material volume $\Omega(t)$.

By using the transport theorem (3.2) and Gauss’ theorem (J.12) we can convert this equation (7.1) into the following volume integral

$$\int_{\Omega(t)} \left( \frac{\partial}{\partial t} (\rho E) + \nabla \cdot (\rho v E) \right) \, dV = \int_{\Omega(t)} \left( \nabla \cdot (T v) + \rho f \cdot v - \nabla \cdot q + \rho r \right) \, dV,$$

(7.3)

which holds for any volume that moves with the material. Therefore, we have the energy equation in differential form

$$\frac{\partial}{\partial t} (\rho E) + \nabla \cdot (\rho v E) = \nabla \cdot (T v) + \rho f \cdot v - \nabla \cdot q + \rho r.$$

(7.4)

This may be further simplified by taking the inner product between the momentum equations (6.9b) and the velocity $v$, and subtracting the result from (7.4). This yields, using the symmetry of $T$, the equation in the following conservative form

$$\frac{\partial}{\partial t} (\rho e) + \nabla \cdot (\rho ve) = T : \nabla v - \nabla \cdot q + \rho r,$$

(7.5)

where the double inner product $T : \nabla v$ is given by (see equation L.4)

$$T : \nabla v := \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} \frac{\partial v_i}{\partial x_j}.$$  

(7.6)

If we use the equation of mass conservation (5.1) we obtain the equation in convective form

$$\rho \frac{d}{dt} e = T : \nabla v - \nabla \cdot q + \rho r.$$

(7.7)

8 Constitutive relations and thermodynamic relations

The conservation laws for deformable mediums are the continuity equation (5.3), the momentum equations (6.9b) and the energy equation (7.5). These 5 equations are less than
the 14 unknowns. Therefore, the mathematical description is not complete without some closure relations, describing properties of the matter and certain instantaneous (both in time and space) interactions between material parameters. These relations are called the constitutive equations and thermodynamic relations. Especially the constitutive equations are not as basic as the conservation laws, and their form will depend on the model adopted. In fact, this part of the model, the constitutive equations, may be called the physical model.

The conservation laws have to be completed with models for the stress tensor $\mathbf{T}$, the heat flux vector $\mathbf{q}$, thermodynamic relations, internal heat sources $r$ and the volume force field $\mathbf{f}$. Apart from possible apparent forces in an accelerating coordinate system (called Coriolis forces), gravity is usually the only external force field working on the matter and therefore $\mathbf{f} = g \mathbf{e}_g$ with $g$ the acceleration of gravity and $\mathbf{e}_g$ the unit vector in the direction of gravitation. Except when indicated otherwise, we will further assume no internal heat sources, so $r = 0$.

It should be noted that these constitutive relations between $\mathbf{T}$, $\mathbf{D}$ and $\mathbf{E}$ are based in the physics, and are therefore subject to certain compatibility conditions. The material may be isotropic [23], i.e. has no preferred directions, and the relation must be independent of orientation. The material may be homogeneous, in which case the relation is independent of position. A most important condition is their compatibility with the principle of objectivity. This means that the relation should be equivalent for any observer, i.e. in any frame of reference [23]. This may be formalized as follows. The transformation that connects two frames of reference by the relations

$$\mathbf{x}' = \mathbf{x}_0(t) + \mathbf{Q}(t) \mathbf{x}, \quad t' = t - t_0$$  \hspace{1cm} \text{(8.1a)}

where $\mathbf{Q}$ is a rotation, i.e. an orthogonal tensor with $\det \mathbf{Q} = 1$, is called an observer transformation. If such an observer transformation identifies a scalar field $\phi$, vector field $\mathbf{u}$ and a tensor field $\mathbf{T}$ to corresponding fields $\phi'$, $\mathbf{u}'$ and $\mathbf{T}'$ in the following way

$$\phi'(\mathbf{x}', t') = \phi(\mathbf{x}, t), \quad \mathbf{u}'(\mathbf{x}', t') = \mathbf{Q}(t) \mathbf{u}(\mathbf{x}, t), \quad \mathbf{T}'(\mathbf{x}', t') = \mathbf{Q}(t) \mathbf{T}(\mathbf{x}, t) \mathbf{Q}^T(t),$$ \hspace{1cm} \text{(8.1b)}

these fields are called objective. For a further discussion on this important restriction we refer to the literature.

In the remainder of this section we will consider some examples of physical models, defined by constitutive equations and, where relevant, supplemented by suitable thermodynamic relations.

A simple but important example is for the problem of heat conduction in rigid material. This assumption of rigidity implies the absence of any deformation or any response to external forcing. The stress tensor, therefore, plays no rôle. The resulting heat equation is similar to the equation that describes mass diffusion.

We will further consider two important types of material, characterized by their stress tensor, viz. viscous fluids and elastic material. A fluid at rest (i.e. in equilibrium) sustains normal stresses by compression, but cannot sustain any shear stress, and $\mathbf{T}$ is only dependent of the rate of deformation tensor $\mathbf{D}$ (equation 2.6). Elastic material, on the other hand, is characterized by a response to any deformation and for small deformations the stress tensor depends on the deformation tensor $\mathbf{E}$ (equation 2.5). If the material is purely elastic, the stress depends on $\mathbf{E}$ only, and vanishes when $\mathbf{E}$ vanishes (the material is in its undeformed reference state).
8. CONSTITUTIVE RELATIONS AND THERMODYNAMIC RELATIONS

8.1 Heat conduction and mass diffusion

If the material is rigid, allowing no motion or deformation, all conservation laws are trivially satisfied, except for the heat balance in the energy equation. If the internal energy \( e \) is a function of the temperature \( T \) only, we can define \( C = \frac{\partial e}{\partial T} \), the heat capacity or specific heat of the material. Furthermore, if the heat flux satisfies Fourier's law of heat conduction

\[
q = -\kappa(T)\nabla T,
\]

with \( \kappa \) the coefficient of heat conduction, equation (7.7) reduces to

\[
\rho C \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + \rho r. \tag{8.3}
\]

It is worthwhile to mention here that diffusion of heat and diffusion of molecules of a solute in a liquid are essentially similar processes. If the concentration of the solute is \( c \) and its flux \( j \), a widely used constitutive law that relates \( j \) to \( c \) in the absence of a temperature gradient is given by [68]

\[
j = -D(c)\nabla c, \tag{8.4}
\]

which is called Fick's law, the analogue of Fourier’s law. From the equation of mass conservation for the solute

\[
\frac{\partial c}{\partial t} + \nabla \cdot j = Q,
\]

where \( Q \) is a source of solute, we obtain the diffusion equation

\[
\frac{\partial c}{\partial t} = \nabla \cdot (D \nabla c) + Q. \tag{8.5}
\]

8.2 Newtonian viscous fluid

For fluids it is useful to split the stress tensor into a part depending on the thermodynamic pressure \( p \), representing the stationary normal components, and a viscous part \( \tau \), related to the velocity gradients.

\[
T = -pI + \tau, \tag{8.6}
\]

where \( \tau \) is called the viscous stress tensor. Let us recall the just derived conservation equations

- mass: \( \frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \),
- momentum: \( \frac{\partial}{\partial t} (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \mathbf{v}^T) = -\nabla p + \nabla \cdot \tau + \rho \mathbf{f} \),
- energy: \( \frac{\partial}{\partial t} (\rho e) + \nabla \cdot (\rho \mathbf{v} e) = -\nabla \cdot q + \nabla \cdot (-\rho \mathbf{v} + \tau \mathbf{v}). \)

Depending on the application, it is often convenient to introduce the specific enthalpy \( h := e + p/\rho \), or entropy \( s \) and absolute temperature \( T \) via the fundamental law of thermodynamics for a reversible process, i.e.

\[
Td\sigma = de + pd\rho^{-1} = dh - \rho^{-1}d\rho. \tag{8.8}
\]
With the convective derivative (2.3), and noting that \( \tau : \nabla v = \nabla \cdot (\tau v) - v \cdot (\nabla \cdot \tau) \) since \( \tau \) is symmetric, the above conservation laws may be rewritten into their convective form

\[
\frac{d}{dt} \rho = -\rho \nabla \cdot v,
\]
\[
\rho \frac{d}{dt} v = -\nabla p + \nabla \cdot \tau + \rho f,
\]
\[
\rho \frac{d}{dt} e = -\nabla \cdot q - p \nabla \cdot v + \tau : \nabla v,
\]
\[
\rho \frac{d}{dt} h = \frac{d}{dt} p - \nabla \cdot q + \tau : \nabla v,
\]
\[
\rho T \frac{d}{dt} s = -\nabla \cdot q + \tau : \nabla v.
\]

Equations (8.9c, 8.9d, 8.9e) are three equivalent forms of the energy equation. For an ideal fluid \( e \) and \( h \) depend on \( T \) only and we may introduce \( de = C_V dT \) and \( dh = C_P dT \) such that we get

\[
\text{energy 1: } \rho C_V \frac{d}{dt} T = -\nabla \cdot q - p \nabla \cdot v + \tau : \nabla v,
\]
\[
\text{energy 2: } \rho C_P \frac{d}{dt} T = \frac{d}{dt} p - \nabla \cdot q + \tau : \nabla v.
\]

For an incompressible fluid we have \( T ds = C_V dT \) resulting in \( \rho C_V \frac{d}{dt} T = -\nabla \cdot q + \tau : \nabla v. \)

\( C_V = T \left( \frac{\partial e}{\partial T} \right)_P \) is the heat capacity or specific heat at constant volume. \( C_P = T \left( \frac{\partial h}{\partial T} \right)_V \) is the heat capacity or specific heat at constant pressure [68]. In general (for an ideal gas) they may depend on temperature. For a perfect gas they are constants. (There is no uniformity in nomenclature. Some authors use perfect gas for what we define as ideal gas.) For an incompressible fluid like a liquid \( C_V \) is practically equal to \( C_P \). A notable exception is matter during phase transition (fusion, vaporization), where all the added heat (enthalpy) is used for the solid-liquid or liquid-vapor transition, rather than an increase of temperature. The amount of heat involved with the phase transition is usually called latent heat.

We will consider here an ideal, heat conducting and viscous fluid, which is described by the following relations.

\[
\text{Ideal gas relation: } p = \rho RT, \quad \text{(8.10a)}
\]
\[
\text{Fourier’s heat flux model: } q = -\kappa \nabla T, \quad \text{(8.10b)}
\]
\[
\text{Newton’s viscous stress tensor: } \tau = 2 \mu \mathbf{D} + \lambda (\nabla \cdot \mathbf{v}) \mathbf{I}, \quad \text{(8.10c)}
\]

where \( \lambda \) and \( \mu \) are viscosity coefficients and \( \mathbf{D} \) the deformation velocity tensor, defined in equation (2.6), while \( R \) is the specific gas constant. For air \( R = 286.73 \text{ J/kg K} \). For an ideal gas we have further the relationship \( \mathcal{R} = C_P - C_V \). Under these assumptions, the equations (8.9a) and (8.9b), usually supplemented by an energy equation (8.9f) or (8.9g), are called the Navier-Stokes equations. (There is no uniformity in the nomenclature. Sometimes this name is only given to the momentum equation (8.9b), and sometimes only to the equations for incompressible flow).

For an isentropic process we have

\[
C_P dT - \rho^{-1} dp = C_V dT - p \rho^{-2} d\rho = 0, \quad \text{so } \frac{dp}{d\rho} = \frac{\rho}{\gamma}, \quad \text{(8.11)}
\]
where $\gamma = C_P/C_V$ is the specific-heat ratio ($= 1.4$ for air). From the definition of the speed of sound $c^2 := (\partial p/\partial \rho)_s$, we have

$$c = (\gamma p/\rho)^{1/2} \quad \text{or} \quad c = (\gamma RT)^{1/2}. \quad (8.12)$$

It is instructive to introduce (see equation (L.2)) the deviatoric deformation velocity tensor $D' := D - \frac{1}{3}(\nabla \cdot v)I$, such that

$$T = (-p + (\lambda + \frac{2}{3}\mu)(\nabla \cdot v))I + 2\mu D'. \quad (8.13)$$

The first part of the stress represents the fluid’s resistance against dynamic compression, the second part against shear. It shows that the mechanical pressure $p_m := -\frac{1}{3} \text{tr}(T)$ is not equivalent to the thermodynamic pressure $p$. From continuity equation (5.3) the difference is found to be proportional to relative changes of density

$$p - p_m = (\lambda + \frac{2}{3}\mu)(\nabla \cdot v) = (\lambda + \frac{2}{3}\mu)\frac{1}{\rho} \frac{d}{dt} \rho$$

and $\lambda + \frac{2}{3}\mu$ is called the coefficient of bulk viscosity (or expansion viscosity, second viscosity). If, according to Stokes’ hypothesis, the fluid is in local thermodynamic equilibrium and both pressures are the same, this coefficient vanishes (see [7, 17, 33]). The coefficient $\mu$ is sometimes called the coefficient of dynamic viscosity, in contrast to the ratio $\nu = \mu/\rho$ which is called the coefficient of kinematic viscosity. It should be noted that the viscosity coefficients in general depend on the temperature.

Important simplifications are obtained (see Chapter 7) if we may neglect viscosity (the fluid is called a gas) or if the fluid is incompressible (the fluid is called a liquid). In the latter case the energy equation is decoupled from the mass and momentum equations, and may be solved separately.

Finally we note that at a free surface $S$ of a fluid the so-called surface tension produces a pressure jump across $S$, which is proportional to the sum of the principal curvatures of the surface. The factor of proportionality $\sigma$ (say) is usually called “surface tension”, but this is really a force per length. If $n$ denotes a unit vector field, (outward) normal to $S$, then the pressure jump $[p]_S := p_{\text{inside}} - p_{\text{outside}}$ is given by [36]

$$[p]_S = \sigma \nabla \cdot n \quad \text{at} \quad S. \quad (8.14)$$

It can be proved that any smooth $n$ yields the same $\nabla \cdot n$ at $S$. The so-called contact angle between the fluid surface and a solid surface (for example, of the container of the fluid), is—in equilibrium— a material property that does not depend on the shape of the fluid.

8.3 Linear elastic and viscoelastic deformations

It can be shown that for deformations, small enough to allow linearization, of homogeneous isotropic elastic material the stress tensor satisfies Hooke’s law, viz. the linear relation

$$T = \lambda \text{tr}(\varepsilon)I + 2\mu \varepsilon, \quad (8.15a)$$
or, written out component wise,

\[ t_{ij} = \lambda \delta_{ij} \left( e_{11} + e_{22} + e_{33} \right) + 2\mu e_{ij} \]  \hspace{1cm} (8.15b)

where we followed the tradition to write here the Cartesian components of \( \mathbf{T} \) as \( t_{ij} \) and of \( \mathbf{E} \) as \( e_{ij} \). The material parameters \( \lambda \) and \( \mu \) are called Lamé coefficients. The material is called linear elastic.

From the observation that with small deformations density changes of the material are negligible, we may consider \( \rho \) constant and we can ignore the continuity equation (5.3). Similarly, in the absence of heat sources the energy equation (7.7) may usually be decoupled from the elastic deformation problem. Another important simplification implied by the assumption of small displacements is the fact that the acceleration in equation of motion (6.10) simplifies to a double time-derivative of displacement \( \mathbf{u} \). By eliminating \( e_{ij} = \frac{1}{2} \left( \frac{\partial}{\partial x_i} u_j + \frac{\partial}{\partial x_j} u_i \right) \) we may finally obtain Navier’s equations

\[ \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} + \rho \mathbf{f}. \]  \hspace{1cm} (8.16)

If a piece of elastic material is uniformly stretched in one direction, we have, say, \( t_{11} > 0 \), while the other \( t_{ij} = 0 \). This does not imply that only \( e_{11} > 0 \). The material is usually contracted in the transverse directions. From the inverted relation (8.15), \( 2\mu \mathbf{E} = \mathbf{T} - \frac{\lambda}{3\lambda + 2\mu} \text{tr}(\mathbf{T}) \mathbf{I} \), we find that

\[ e_{11} = \frac{\lambda + \mu}{\mu(3\lambda + 2\mu)} t_{11}, \quad e_{22} = e_{33} = -\frac{\lambda}{2\mu(3\lambda + 2\mu)} t_{11}, \quad e_{ij} = 0 \quad (i \neq j). \]

From the ratios \( t_{11}/e_{11} \) and \( -e_{22}/e_{11} \) two constants naturally appear:

\[ E := \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} \]  \hspace{1cm} (Young’s modulus),  \hspace{1cm} \[ \nu := \frac{\mu}{2(\lambda + \mu)} \]  \hspace{1cm} (Poisson’s ratio).

\( E \) is positive, \( \nu \) is less than 0.5 and usually positive. The inverted equation (8.15) is thus

\[ E \mathbf{E} = (1 + \nu) \mathbf{T} - \nu \text{tr}(\mathbf{T}) \mathbf{I}. \]

Since \( \int_\Omega \nabla \cdot \mathbf{u} \, dV = \int_\partial \Omega \mathbf{u} \cdot \mathbf{n} \, dS \) describes (linearised) the increase of a volume \( \Omega \) by a displacement field \( \mathbf{u} \), the material is incompressible (no volume changes) if \( \nabla \cdot \mathbf{u} = \text{tr}(\mathbf{E}) = 0 \). As \( \text{tr}(\mathbf{T}) = \frac{E}{E_0} \text{tr}(\mathbf{E}) \) is to be finite, while \( E \) is finite, we have necessarily \( \nu = 0.5 \) for incompressible elastic material.

Viscoelastic material is characterised by a stress-strain relation that depends on the history of the deformation and allows a certain amount of dissipation. The simplest and most common models are

- **Kelvin-Voigt model:**

\[ \mathbf{T} = \lambda (\text{tr}(\mathbf{E}) + \theta_1 \text{tr}(\dot{\mathbf{E}})) \mathbf{I} + 2\mu (\mathbf{E} + \theta_2 \dot{\mathbf{E}}), \]  \hspace{1cm} (8.17a)

where the dot denotes a derivative with respect to time. \( \lambda \) and \( \mu \) are equivalent Lamé’s coefficients, while \( \theta_1 \) and \( \theta_2 \) are time parameters. The factor \( 3\lambda \theta_1 + 2\mu \theta_2 \) of \( \text{tr}(\dot{\mathbf{E}}) \) is usually small enough to be neglected. Since there exists an equilibrium with constant stress, this material is sometimes called a viscoelastic solid.
9. MAXWELL’S EQUATIONS

- Maxwell model:

\[
E \dot{\varepsilon} = (1 + \nu) (\dot{T} + \gamma_1 T) - \nu (\text{tr}(\dot{T}) + \gamma_2 \text{tr}(T)) I
\]  

(8.17b)

where \( E \) and \( \nu \) are equivalents of Young’s modulus and Poisson’s ratio, while \( \gamma_1 \) and \( \gamma_2 \) are time parameters. As a constant stress produces a constant flow rate, this material is sometimes called a viscoelastic fluid.

Both models may be considered as special cases of further generalisations.

Example 6.4 In the case of pure shear stress in the 1,2-direction (i.e. \( t_{12} \) is constant), we have with Kelvin-Voigt material the equation

\[
t_{12} = 2\mu (e_{12} + \theta_2 \dot{e}_{12}),
\]

yielding an exponential decay to a limiting deformation gradient

\[
e_{12}(t) = \frac{C}{1+\theta} e^{t/\theta}.
\]

In the case of pure shear in the 1,2-direction with a fixed velocity gradient (i.e. \( \dot{e}_{12} \) is constant), we have with Maxwell material the equation

\[
E \dot{e}_{12} = (1 + \nu) (t_{12} + \gamma_1 t_{12}),
\]

yielding an exponential decay to a limiting shear stress

\[
t_{12}(t) = \frac{E}{1+\nu} e^{t/\gamma_1}.
\]

□

9 Maxwell’s equations

The above discussion aimed at a derivation from first principles of the Navier-Stokes equations, describing the motion of fluids, and the equations of linear elastic deformations. Although this relates to the main area of application considered in this book, we cannot leave unmentioned another monument in applied mathematics, the equations of Maxwell for electromagnetic fields.

Electric charge is described by a charge density \( Q \) and a current density \( J \), corresponding to charges in motion. These charges and currents produce electromagnetic fields, described by: (i) the electric field intensity \( E \), that applies a force \( qE \) to a point charge \( q \), and (ii) the magnetic-flux density or magnetic induction \( B \), that applies a torque \( m \times B \) to a magnetic dipole with magnetic moment \( m \). Further, we introduce the derived fields (iii) \( D \), the electric displacement, and (iv) \( H \), the magnetic field intensity.

For these fields we have the following equations.

Coulomb’s law. The net effect of a charge distribution in a fixed volume \( \Omega \) is equivalent to the total flux of electric displacement \( D \) out through \( \Omega \)’s surface \( \partial \Omega \) (with \( n \) the outward unit normal),

\[
\int_{\Omega} Q \, dV = \oint_{\partial \Omega} D \cdot n \, dS.
\]  

(9.1)

Ampère-Maxwell’s law. A current or a changing electric displacement causes a magnetic field as follows. The work done by a magnetic field \( H \) along a closed contour \( C \) is equal to the flux of the total current \( J + \frac{\partial}{\partial t} D \) through the enclosed surface \( S \)

\[
\oint_{C} H \cdot dl = \int_{S} \left( J + \frac{\partial}{\partial t} D \right) \cdot n \, dS.
\]  

(9.2)
Faraday-Henry’s law. A changing magnetic field causes an electric field as follows. The electromotive force induced by $E$ around a circuit $C$ is equal but opposite to the rate of change of the magnetic flux $B$ through the enclosed surface $S$

$$\oint_C E \cdot d\ell = -\frac{d}{dt} \oiint_S B \cdot n \, dS.$$  \hspace{1cm} (9.3)

Absence of free magnetic poles. This yields a vanishing total flux of magnetic induction $B$ through the surface of a volume $\Omega$

$$\oiint_{\partial \Omega} B \cdot n \, dS = 0.$$  \hspace{1cm} (9.4)

Maxwell’s equations. As these equations are valid for any volume $\Omega$ or contour $C$, the identities are valid locally and may be reformulated into the differential equations

$$\nabla \cdot D = Q, \quad \nabla \times H = J + \frac{\partial}{\partial t} D,$$

$$\nabla \times E + \frac{\partial}{\partial t} B = 0, \quad \nabla \cdot B = 0,$$  \hspace{1cm} (9.5)

known as Maxwell’s equations.

By taking the divergence of Ampére-Maxwell’s law (in differential form) we obtain the law of conservation of electric charge,

$$\frac{\partial}{\partial t} Q + \nabla \cdot J = 0.$$  \hspace{1cm} (9.6)

In integral form this says that the variation of the total charge of a fixed volume $\Omega$ (in the absence of sources) is equal to the net flux of the current into $\Omega$’s surface $\partial \Omega$,

$$\frac{d}{dt} \int_\Omega Q \, dV = -\oiint_{\partial \Omega} (J \cdot n) \, dS.$$  \hspace{1cm} (9.7)

Taking the divergence of Faraday-Henry’s law yields that $\nabla \cdot B$ is stationary (and therefore equals zero if it starts that way). So Maxwell’s equations consist of only 7 independent relations for 16 unknowns. Therefore, we need additional relations to describe the field uniquely.

9.1 Constitutive relations

The fields $D, H$ and $J$ are related to $E$ and $B$ via constitutive equations that depend on the problem considered. For example, in vacuum and isotropic diamagnetic and paramagnetic media we have the simple linear relationship

$$B = \mu H,$$  \hspace{1cm} (9.8)

where $\mu$ is the magnetic permeability and is denoted by $\mu_0$ for vacuum. Its numerical value is $\mu_0 = 1.2566371 \cdot 10^{-6} \text{H/m}$. In vacuum and isotropic dielectric media we have the linear relation

$$D = \epsilon E,$$  \hspace{1cm} (9.9)
9. Maxwell’s Equations

where $\epsilon$ is the electric permittivity and is denoted by $\epsilon_0$ for vacuum. Its numerical value is $\epsilon_0 = 8.854185 \cdot 10^{-12}$ F/m. Note that $\mu_0 \epsilon_0 c^2 = 1$, where $c = 2.99792458 \cdot 10^8$ m/s is the speed of light in vacuum, and $\mu_0 = 4\pi \cdot 10^{-7}$ H/m.

The relation between the current and the electric field, the generalized Ohm’s law, is for a wide range of conditions linear and given by

$$J = \sigma E,$$  \hspace{1cm} (9.10)

where $\sigma$ is the conductivity of the medium.

**Example 6.5** A stationary point charge has a charge distribution $Q(x) = q \delta(x)$. As the field is stationary, we have $\nabla \times E = 0$. Hence, $E$ is conservative and may be written as the gradient of a potential $E = -\nabla \phi$. If the field is in vacuum, we have $D = \epsilon_0 E$, and so $\phi$ satisfies Poisson’s equation $\nabla^2 \phi = -q \delta(x)$. In infinite space this has solution $\phi = \frac{q}{4\pi \epsilon_0 r}$, where $r = |x|$.

9.2 Energy conservation and Poynting’s theorem

Energy is dissipated in conducting media (otherwise $J = 0$) by Joule heating, given (per unit volume) by the power density $J \cdot E$. With Ohm’s law (9.10) this simplifies to $J \cdot E = \sigma |E|^2$.

This is illustrated by the following energy conservation law. If we define the rate of change of energy density $\frac{\partial u}{\partial t} := E \cdot \frac{\partial}{\partial t} D + B \cdot \frac{\partial}{\partial t} H$ and the energy flux vector $S := E \times H$ (also called the Poynting vector), Maxwell’s equations may be recast into the identity

$$\frac{\partial u}{\partial t} + \nabla \cdot S = -J \cdot E,$$ \hspace{1cm} (9.11)

known as Poynting’s theorem. It shows that the rate of change of electromagnetic energy within a certain volume plus the energy flowing out through the boundaries per unit time is equal to minus the work done by the field inside the volume.

9.3 Electromagnetic waves and Lorentz’s force

Disturbances of an electromagnetic field propagate like waves. For linear material, where $B = \mu H$ and $D = \epsilon E$, Maxwell’s equations can be recast into a set of wave equations as follows.

As $B$ is solenoidal (divergence free), it can be written as the curl of a vector potential $A$. Since $\nabla \times \nabla = 0$, this vector potential is defined up to a scalar potential $\alpha$, so we write

$$B = \nabla \times (A + \nabla \alpha).$$

From $\nabla \times (E + \frac{\partial}{\partial t} A + \frac{\partial}{\partial t} \nabla \alpha) = 0$ it follows that there is a scalar potential $\psi$, such that

$$E = -\nabla \psi - \frac{\partial}{\partial t} A$$
CHAPTER 6. CONTINUUM MECHANICS AND ELECTROMAGNETICS

because we can absorb $\frac{\partial}{\partial t} \alpha$ into $\psi$. This yields

$$\nabla^2 \psi + \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -\frac{1}{\varepsilon} Q.$$  

With the vector identity $\nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, we have

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left( \nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \psi}{\partial t} \right) = -\mu J,$$

where we introduced the speed of light $c$ in the medium considered, with $\mu \varepsilon = c^{-2}$. If we choose $\psi$ and $\mathbf{A}$ such that $\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \psi}{\partial t} = 0$, the **Lorenz gauge condition**, we have finally the set of wave equations

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu J, \tag{9.12a}$$

$$\nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = -\frac{1}{\varepsilon} Q. \tag{9.12b}$$

Note that the freedom in $\alpha$ provides us with a fairly large class of possible potentials $\psi$.

Other gauge conditions are possible, for example the **Coulomb gauge condition**, where $\nabla \cdot \mathbf{A} = 0$, leading to a Poisson, rather than wave, equation for $\psi$

$$\nabla^2 \psi = -\frac{1}{\varepsilon} Q.$$  

We finally remark that the **Lorentz force** $q (E + v \times B)$ describes the force acting on a point charge $q$, moving with velocity $v$ in the presence of an electromagnetic field. The path of the particle may be determined by recalling Newton’s equations, describing the change of momentum due to an external force.

**Example 6.6** In free space, in the absence of charge $Q$ or current $J$, in a medium satisfying the linear relations (9.8) and (9.9), we have the time-harmonic plane-wave solution (see Chapter 3) given by the real part of

$$E = E_0 e^{ik \cdot x - i\omega t}, \quad H = H_0 e^{ik \cdot x - i\omega t},$$

where $k = k \kappa$ is the wave vector, and $E_0 = E_0 \epsilon$ and $H_0 = H_0 \mu$ are the vectorial amplitudes. Unit vectors $\kappa, \epsilon$ and $h$ denote the direction of propagation, and the polarisation of $E$ and $H$ respectively. They form an orthonormal triple with $k = \epsilon \times h$, $\epsilon = h \times k$, and $h = k \times \epsilon$. The modulus of the wave vector is $k = \omega/c$ where $c = (\mu / \varepsilon)^{1/2}$, while the moduli of the vectorial amplitudes satisfy $E_0 = Z H_0$ where $Z = (\mu / \varepsilon)^{1/2}$ is the impedance of the medium.

10 Discussion

- As myriads of applications of mathematics are found in physics, in particular in continuum physics, it is useful to have the pertaining equations concisely summarized. It is, however, very rare that the equations are used in the very comprehensive way as presented. Usually, the problem is much more limited, and it is wise to simplify the equations first before an attempt is made to tackle them mathematically. This highly non-trivial step is called “modelling”, and the next chapter will be devoted to it.

17:22 7 Oct 2003 91 version: 26-09-2003
Exercises

• The section on conservation laws is not only of interest purely physically. They form
the basis of many numerical methods, that are based on the integral formulations of
the conservation laws, and are known as Finite Volume Methods.

Exercises

6.1. Verify that the equations (8.9), with relations (8.10), for inviscid non-conducting
fluids simplify to the Euler equations

\[ \frac{\partial}{\partial t} \rho = -\rho \nabla \cdot v, \quad \rho \frac{\partial}{\partial t} v = -\nabla p + \rho f, \quad \frac{\partial}{\partial t} s = 0. \]

6.2. Verify that the equations (8.9), with relations (8.10) for incompressible fluids (\( \rho = \rho_0 \)) of constant viscosity simplify to

\[ \nabla \cdot v = 0, \quad \rho_0 \frac{\partial}{\partial t} v = -\nabla p + \mu \nabla^2 v + \rho_0 f, \quad \rho_0 T \frac{\partial}{\partial t} s = -\nabla \cdot q + \tau : \nabla v. \]

6.3. (a) Derive for a perfect gas the following relation between entropy, pressure and
density

\[ s - s_0 = C_V \ln \rho - C_P \ln \rho. \]

(b) Show that a homentropic (\( s \) is constant) perfect gas is a polytropic gas (\( p/\rho^n \) is
constant).

(c) Show that for a homentropic perfect gas flow the following relation holds

\[ \frac{1}{\rho} \nabla \rho = \frac{C_P}{C_V} \nabla \rho = \frac{\gamma - 1}{\gamma} \nabla \rho. \]

(d) Prove Kelvin’s Theorem of the conservation of circulation: in an inviscid homentropic
flow field any circulation

\[ \Gamma := \oint_{\partial A} v \cdot d\ell, \]

where \( \partial A \) is a closed material contour (the boundary of a material surface \( A \),
i.e. moving with the flow) is a constant, while being convected with the flow.

6.4. Show, by using equation (8.16), that \( \nabla^2 (\nabla \cdot u) = 0 \) and \( \nabla^4 u = 0 \) if \( f = \frac{\partial^2 u}{\partial t^2} = 0. \)

6.5. Verify the identity

\[ \frac{\partial^2 e_{ij}}{\partial x_i \partial x_j} + \frac{\partial^2 e_{kl}}{\partial x_k \partial x_l} = \frac{\partial^2 e_{ik}}{\partial x_i \partial x_k} + \frac{\partial^2 e_{jl}}{\partial x_j \partial x_l}, \]

and show that this yields 6 compatibility relations for the components of \( \mathbf{E} \).

6.6. Show that the harmonic plane wave \( u = U e^{ik \cdot x - i\omega t} \) with frequency \( \omega \), wave vector \( k \), amplitude \( U \) and polarization vector \( v \), that satisfies Navier’s equation (8.16) with \( f = 0 \) has the dispersion relation

\[ \rho \omega^2 = (\lambda + 2\mu)(k \cdot k) \]

6.7. Consider two electromagnetic media satisfying relations (9.8) and (9.9) with constants $\epsilon_1, \mu_1$ and $\epsilon_2, \mu_2$. The media are separated from each other by a plane interface with unit normal vector $\kappa$ directed from medium 1 to 2. The interface is given by $\kappa \cdot \mathbf{x} = 0$. In medium 1 a plane wave propagates in direction $\kappa$ (see Example 6.6) and reflects and transmits at the interface. In medium 1 the wave is given by

$$E = E_0 (e^{i \mathbf{k}_1 \cdot \mathbf{x}} + R e^{-i \mathbf{k}_1 \cdot \mathbf{x}}) e^{-i \omega t}, \quad H = Z_1^{-1} E_0 (e^{i \mathbf{k}_1 \cdot \mathbf{x}} - R e^{-i \mathbf{k}_1 \cdot \mathbf{x}}) e^{-i \omega t},$$

while in medium 2 it is given by

$$E = E_0 T e^{i \mathbf{k}_2 \cdot \mathbf{x}} e^{-i \omega t}, \quad H = Z_2^{-1} E_0 T e^{i \mathbf{k}_2 \cdot \mathbf{x}} e^{-i \omega t}.$$

Given that at an interface the normal components of $D$ and $B$, and the tangential components of $E$ and $H$ are continuous, determine the reflection and transmission coefficients $R$ and $T$. 
Chapter 7
The art of modelling

This chapter describes the ideas and principles of modelling a real life problem. It is necessarily a bit contemplating in nature. In Section 1 we discuss how modelling may be defined. There are various ways to model real life situations. We then discuss the kinds of models that can be distinguished in Section 2. They all relate to the (thrifty) way a real world problem description is translated to a mathematical formulation and the usefulness of the conclusions drawn from the latter for the actual problem at hand. In order to decide which part of the resulting equations really matters, it is necessary to make these dimensionless. To do this one has to scale the parameters and variables first, which is treated in Section 3. We show that by Buckingham’s theorem there is always a subset of relevant (redefined) variables in which the problem can be formulated. As a result one can often indicate large or small coefficients for some of the contributions in the equation. If they are small enough we can neglect them, at least in most of the domain. This kind of treating the equations of mass, momentum, and energy for fluid flow is discussed to some extent in Section 4. It is shown which simpler equations (models) may result from them.

1 Introduction

Mathematics has, historically, its major sources of inspiration in applications. It is just the unexpected question from practice that forces one to go off the beaten track. Also it is usually easier to portray properties of a mathematical abstraction with a concrete example at hand. Therefore, it is safe to say that most mathematics is applied, applicable or emerges from applications.

Before mathematics can be applied to a real problem, the problem must be described mathematically. We need a mathematical representation of its primitive elements and their relations, and the problem must be formulated in equations and formulas, to render it amenable to formal manipulation and to clarify the inherent structure. This is called mathematical modelling. An informal definition could be:
1. INTRODUCTION

Describing a real-world problem in a mathematical way by what is called a model, such that it becomes possible to deploy mathematical tools for its solution. The model should be based on first principles and elementary relations and it should be accurate enough, such that it has reasonable claims to predict both quantitative and qualitative aspects of the original problem. The accuracy of the description should be limited, in order to make the model not unnecessarily complex.

This is evidently a very loose definition. Apart from the question what is meant with: a problem being described in a mathematical way, there is the confusing paradox that we only know the precision of our model, if we can compare it with a better model, but this better model is exactly what we try to avoid as it is usually unnecessarily complex! In general we do not know a problem and its accompanying model well enough to be absolutely sure that the sought description is both consistent, complete and sufficiently accurate for the purpose, and not too formidable for any treatment. A model is, therefore, to a certain extent a vague concept. Nevertheless, modelling plays a key rôle in applied mathematics, since mathematics cannot be applied to any real world problem without the intermediate steps of modelling. Therefore, a more structured approach is necessary, which is the aim of the present chapter.

Some people define modelling as the process of translating a real-world problem into mathematical terms. We will not do so, as this definition is too wide to include the subtle aspects of “limited precision” (to be discussed in a minute). Therefore we will introduce the word mathematizing, defined as the process of translating a real-world problem into mathematical terms. It is a translation in the sense that we translate from the inaccurate, verbose “everyday” language to the language of mathematics. For example, the geometrical presence and evolution of objects in space and time may be described parametrically in a suitable coordinate system. Any properties or fields that are expected to play a rôle may be formulated by functions in time and space, explicitly or implicitly, for example as a differential equation.

Mathematizing is an elementary but not trivial step. In fact, it forms probably the single most important step in the progress of science. It requires the distinction, naming, and exact specification of the essential relevant elementary objects and their interrelations, where mathematics acts as a language in which the problem is described. If theory is available for the mathematical problem obtained this way, the problem considered may be subjected to the strict logic of mathematics, and reasoning in this language will transcend over the limited and inaccurate ordinary language. Mathematizing is therefore, apart from providing the link between the mathematical world and the real world, also important for science in general.

A very important point to note is the fact that such a mathematized formulation is always at some level simplified. The earth can be modelled by a point or a sphere in astronomical applications, or by an infinite half-space or modelled not all in problems of human scale. Based on the level of simplification, sophistication or accuracy, we can associate an inherent hierarchy to the set of possible descriptions. A model may be too crude, but also it may be too refined. It is too crude if it just doesn’t describe the problem considered, or if the numbers it produces are not accurate enough to be acceptable. It is
too refined if it includes irrelevant effects that make the problem untreatable, or make the model so complicated that important relations or trends remain hidden.

The ultimate goal for mathematizing a problem is a deeper understanding and a more profound analysis and solution of the problem. Usually, a more refined problem translation is more accurate but also more complicated and more difficult – if not impossible! – to analyse and solve than a simpler one. Therefore, not every mathematical translation is a good one. We will call a good mathematical translation a model or mathematical model if it is lean or thrifty in the sense, that it describes our problem quantitatively or qualitatively in a suitable or required accuracy with a minimal number of essentially different parameters and variables. (We say “essentially different”, in view of a reduction that is always possible by writing the problem in dimensionless form. See Buckingham’s Theorem 7.10 below.) Again, this definition is rather subjective, as it greatly depends on the context of the problem considered and our knowledge and resources. So there will rarely be one “best” model. At the same time, it shows that modelling, even if relying significantly on intuition, is part of the mathematical analysis.

2 Models

We will distinguish the following three classes of models.

- Systematic models.

Other possible names are asymptotic models or reducing models. The starting point here is to use available complete models, which are adequate, but over-complete so that effects are included which are irrelevant, uninteresting, or negligibly small, and thereby making the mathematical problem unnecessarily complex. By using available additional information (order of magnitude of the parameters) assumptions can be made which minimize in a systematic way the over-complete model into a good model by taking a parameter that is already large or small to its asymptotic limit: small parameters are taken zero, large parameters become infinite, an almost symmetry becomes a full symmetry.

Examples of systematic models are found in particular in the well-established fields of continuum physics, considered in Chapter 6. An ordinary flow is usually described by a model which is reduced from the full, i.e. compressible and viscous, Navier-Stokes equations. This will be elaborated in detail in Section 4.

Example 7.1 Consider the following simple convection-diffusion problem. For a temperature field $T$ and given velocity field $v$ in space $x$ and time $t$ and thermal diffusion coefficient $\alpha$ we have the assumed “over-complete” model

$$\frac{\partial T}{\partial t} + v \cdot \nabla T = \alpha \nabla^2 T,$$

which is, however, difficult to solve. If we have reasons to assume that the diffusion term may be ignored, we obtain the reduced problem

$$\frac{\partial T}{\partial t} + v \cdot \nabla T = 0,$$

(\#)
2. MODELS

which is far more attractive than the full problem, as it may be solved exactly. Along the streamlines \( x = \xi(t) \) given by

\[
\frac{d\xi}{dt} = v,
\]
equation (*) simplifies to

\[
\frac{d}{dt}T(\xi(t), t) = 0
\]
with solution \( T = T(\xi(t), t) = \text{constant} \). □

**Example 7.2** The dynamics of an idealised frictionless undriven pendulum consisting of a point mass suspended by a weightless cord of length \( L \) may be described by the difficult nonlinear pendulum equation

\[
\frac{d^2\theta}{dt^2} = -\frac{g}{L} \sin \theta, \quad \theta(0) = \theta_0, \quad \frac{d\theta}{dt}(0) = 0,
\]
where \( g \) is the gravitational acceleration and \( \theta = \theta(t) \) is the angle at time \( t \) of the cord with the vertical. Since \( \frac{d\theta}{dt}(0) = 0 \) and the pendulum is undriven, \( \theta(t) \) will never be larger than \( \theta_0 \). So if \( |\theta_0| \ll 1 \), we may assume that \( \theta \) is always much smaller than 1, and we may approximate, at least for some time, the nonlinear term \( \sin \theta \) by \( \theta \). This yields the much simpler model

\[
\frac{d^2\theta}{dt^2} = -\frac{g}{L} \theta, \quad \theta(0) = \theta_0, \quad \frac{d\theta}{dt}(0) = 0,
\]
which has solution \( \theta(t) = \theta_0 \cos(\omega t) \), with \( \omega = \sqrt{g/L} \). □

• **Constructing models**

Other possible names are building block models or lumped-parameter models. Here we build our problem description step by step from low to high, from simple to more complex, by adding effects and elements lumped together in building blocks, until the required accuracy or adequacy is obtained.

**Example 7.3** The following problem of air release by a simple air pump may be an example of a building block model. Consider a pump of cross section \( S \) and length \( a(t) \), which depends on the piston position. Initially, \( a(0) = L \). Under pressure, the enclosed volume of air \( Sa(t) \) leaves the pump through a small hole, forming a jet of cross section \( S_j \) and (mean) velocity \( v_j \). From time \( t = 0 \), a spring pushes against the piston with a force \( F = \lambda a(t) \). Assuming any inertia effects of the piston to be much smaller than the inertia of the flow, the piston force is balanced by a pressure increase from atmospheric pressure \( p_\infty \) outside to the value \( p_0 \) inside the pump. So \( F = S(p_0 - p_\infty) \).

\[
\begin{align*}
S \quad \text{[]} \quad L \quad \text{[]} \quad F \quad \text{[]} \quad v_j \quad \text{[]} \quad S_j
\end{align*}
\]
Assuming incompressible air of density $\rho_0$ and a practically vanishing velocity inside the pump, conservation of mass (see Eq. (6.5.1)) tells us that \{the variation inside\} = \{what goes out\}. In formula this is
\[
\frac{d}{dt}(\rho_0 S a(t)) = -\rho_0 v_j S_j
\]
leading to
\[
v_j = -\frac{S}{S_j} \frac{da}{dt}.
\]
From Bernoulli’s law (to be introduced below; see Eq. (4.10)), relating pressure $p$ and velocity $v$ by $p + \frac{1}{2}\rho_0 v^2 = \text{constant}$, and noting that the pressure inside the jet is equal to the atmospheric pressure $p_\infty$ (the jet cannot support a pressure difference), we can deduce that
\[
p_0 = p_j + \frac{1}{2}\rho_0 v_j^2 = p_\infty + \frac{1}{2}\rho_0 v_j^2
\]
resulting into
\[
\frac{1}{2}\rho_0 v_j^2 = \frac{\lambda a}{S}.
\]
Together we have the model
\[
\frac{da}{dt} = -K \sqrt{a}, \quad \text{with } K = \frac{S_j}{S} \sqrt{\frac{2\lambda}{\rho_0 S}},
\]
which is easily solved by
\[
a(t) = L \left(1 - \frac{1}{2}\frac{K t}{\sqrt{L}}\right)^2 \quad \text{along } 0 \leq t \leq \frac{2\sqrt{L}}{K}.
\]

**Example 7.4** A brilliant example of a constructive model is the Bernoulli-Euler model of elastic deformation of slender bars, in which case the bar is described by a flexible line of vanishing cross section. The essentials of the theory was developed long before the general results described by the equations (6.8.15, 6.8.16) were available. In principle, the equations for the line should be implied by the general three-dimensional theory by utilizing the slenderness of the bar in the limit to zero. This, however, is not straightforward. Therefore, the classical derivation is still important [38, 109].

For simplicity we will restrict the analysis to the case of deformation and motion in the vertical plane. Torsion and friction with any surroundings are neglected. The two-dimensional equations of motion are written as a differential equation for the position vector.

The line is described by the position vector $x(s, t)$ as a function of curve length $s$ and time $t$, with natural local coordinate $s$ such that $|x'| = 1$, where $|\cdot| = \sqrt{\cdot}$ and $|\cdot| = \frac{1}{s} [\cdot]$ (see for example [75]). Introduce the right-hand orthogonal basis $\{t, n, b\}$, consisting of the tangential unit vector $t = x'$, the principal normal unit vector $n$, and binormal unit vector $b$, such that $b = t \times n$, $n = b \times t$, $t = n \times b$.

The curvature vector is $k = t' = x''$, with curvature $|k| = |\kappa|$ defined such that $k = \kappa n$. The torsion or second curvature vector is $b' = -\tau n$, with torsion $\tau$. Note that $n' = -\kappa t + \tau b$. 

2. MODELS

Introduce a bar element of length $ds$, loaded by an external line load $q$ and internal forces $F$ and moments $M$ at the both ends. The basic equations are derived from the equilibrium of the dynamic forces, equilibrium of the moments, and from the constitutive equations as follows [69].

For a beam there is a moment around $b$ (bending) and around $t$ (torsion) so $M = M_b b + M_t t$. Torsion will be assumed to be zero, and $M_b$ is given by the following Bernoulli hypothesis. See Figure 7.1.

![Figure 7.1. Sketch of bar element. Side view (i) and cross section (ii).](image)

Consider a small bar of length $\ell$ and cross section $A$, bent over an angle $\psi$. From $R \psi = \ell$ and $(R + \xi) \psi = \ell + d \ell$ it follows that $d \ell / \ell = \xi / R = \xi \kappa$. The residual force $df$ at a cross sectional slice $dA$, that causes the bar to bend, is with Hooke’s law given by $df = E\kappa \xi dA$. The moment applied by $df$ is then

$$M_b = \int_A \xi df = E\kappa \int_A \xi^2 dA = EI\kappa,$$

where bending stiffness $EI$ is the product of Young’s modulus $E$ and the second moment of cross sectional area $I$.

Since the force $F$ is the only cause of the deformation, $F$ lies in the plane of tangent and principal normal, so $F = Tt + Sn$, where $T$ is called the normal force and $S$ the shearing force. The dynamic force equilibrium $dF + q \, ds = m_0 \ddot{x} \, ds$ (where $m_0$ is the mass per unit length) and the moment equilibrium $dM + d \times F = 0$ yield

$$F' + q = m_0 \ddot{x}, \quad M' + t \times F = 0.$$

From the vector identity $t \times (M' + t \times F) = t \times M' + T t - F = 0$ we obtain $(t \times M')' + (T t)' + q = m_0 \ddot{x}$. With (*) and $T = M_T = 0$, we have $t \times M' = \left[ t \times (M_b b) \right]' - \left[ t \times (M_T t) \right]' = 0$.
CHAPTER 7. THE ART OF MODELLING

Canonical models.
Another possible name is characteristic models or quintessential models. Here an existing model is further reduced to describe only the essence of a certain aspect of the problem considered. These models are particularly important if the mathematical analysis of a model from one of the other categories is lacking available theory. The development of such theory is usually hindered by too much irrelevant details. These models are useful for the understanding, but usually far away from the original full problem setting and therefore not suitable for direct industrial application.

Example 7.5 The Navier-Stokes equations for incompressible viscous flow (see Chapter 7), given in terms of a velocity and pressure field

\[ \frac{\partial}{\partial t} \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{v}, \quad \nabla \cdot \mathbf{v} = 0, \]

are in general very complex. Especially the coupling between the nonlinear and viscous terms yielding instabilities and turbulence is complicated and difficult to analyse. Therefore, Burgers proposed to consider the following very simplified version
3. NON-DIMENSIONALISATION AND SCALING

of it, where the pressure gradient has been neglected, and only behaviour in one dimension is taken into account. This equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}$$

is called Burgers’ equation [127]. A great deal of insight is obtained by the remarkable transformation

$$u = -2\nu \frac{1}{\psi} \frac{\partial \psi}{\partial x}$$

found independently by Cole (1951) and Hopf (1950), by which the nonlinear equation is reduced to a linear equation, related to the heat equation

$$\frac{\partial \psi}{\partial t} - \nu \frac{\partial^2 \psi}{\partial x^2} = C(t)\psi$$

where $C(t)$ is an arbitrary function of $t$. This equation is well understood and allows many exact solutions. \[\square\]

Note that an asymptotic model may start as a building-block model, which is only found at a later stage to be too comprehensive. Similarly, a canonical model may reduce from an asymptotic model if the latter appears to contain a particular, not yet understood effect, which should be investigated in isolation before any progress with the original model can be made.

3 Non-dimensionalisation and scaling

Modelling means that one has to decide which effects are relevant and should be included, and which are irrelevant and can be ignored. More in general, we may expect a hierarchy in relevance, from most dominant, via less relevant and locally irrelevant to absolutely unimportant effects or contributions. Relevant and irrelevant are rather vague qualifications. To make this operational we will relate them to small and large terms in our mathematical description (equations, etc.).

3.1 General concepts

Small and large have no absolute meaning, as long as we have not defined our “measuring stick”. To illustrate this we may imagine the following science fiction scenery. Suppose we are lost in outer space, with all planets, stars, and galaxies so far away that they are only seen as sizeless spots on our retina. Then, a rock drifts slowly into our field of vision. As long as we are not close enough for a stereoscopic view with both our eyes, we are not able to compare its size or distance with anything we know. There is no way to estimate if it is big and far away, or small and nearby. Only the rock itself is our scale of reference. A similar experience is found when we look into a microscope of unknown magnification. An object, visible but not recognizable, may be as big as an amoeba, or as small as a virus or a molecule. Re-interpreting the famous saying of Protagoras, Man is the measure of all things, nothing we observe is small or large, fast or slow, in any absolute sense. It is only by comparison that these qualifications have a meaning.
The next question is: what do we use for comparing. We can use an absolute or universal measuring stick, like a meter or a kilogram, to archive the observations and be able to reproduce them exactly again. However, we use a natural scale, like typical sizes in the problem itself, if we want to classify the type of phenomena.

The following concepts are important in this respect.

When we model, we need to understand the problem in advance to a certain degree, such that we are able to formulate the relevant physical laws and relations. Therefore, in modelling the natural scaling is the appropriate one to use. We introduce for all our dependent and independent variables typical values, taken from the problem in question. For example, a length $L$ for the independent spatial variable $x$, and a velocity $V$ for the dependent variable $v$, and thus an intrinsic time $L/V$ for time coordinate $t$. We refer to this as inherent scaling.

When more than one problem parameters in the same units is available, for example a length $L$ and width $D$, or a time $L/V$ and an inverse frequency $\omega^{-1}$, it is inevitable that if one is selected for the scaling, the combination with the other gives us a new parameter, like $D/L$ or $\omega L/V$, which is now independent of the units (meters, seconds) and is therefore called a dimensionless parameter. Incidentally, this meaning of the word “dimension” has nothing to do with the mathematical meaning of the number of independent basis vectors in a vector space. Dimensionless parameters are very important for a systematic classification of types of problems. They measure the relative importance of certain effects in an absolute way.

If the model is a proper one, reflecting the intrinsic relations between the variables, it should not depend on the arbitrary use of meters or inches, etc. The equations should be dimensionally homogeneous in the following sense. Let a physical quantity $q_0$ be described by $n$ quantities $q_1, q_2, \ldots, q_n$. Each quantity $q$ has a dimensional unit or dimension $[q]$, derived from a set of $r$ independent base units $d_i$, e.g. the SI base units [m, kg, s, A, K, mol, cd] [111]. Let $q_0$ be given by a relation

$$q_0 = f(q_1, q_2, \ldots, q_n).$$

This relation should be equivalent for all choices of sets of independent base units. Terms that are added, like $q_1 + q_2$ should have the same dimensions, i.e. $[q_1] = [q_2]$. The dimension of a product should be the product of the dimensions, i.e. if $q_0 = q_1 q_2$ then $[q_0] = [q_1][q_2]$. Terms that occur as the argument of a dimensionless function, like sin or exp, should have dimension 1, i.e. be dimensionless. So if $\sin(q_1 q_2)$ occurs, then $[q_1][q_2] = 1$. It follows that any dimension of $q_j$ may be written as a product over $d_i$, so

$$[q_j] = d_1^{\mu_{1j}} d_2^{\mu_{2j}} \ldots d_r^{\mu_{rj}} = \prod_{i=1}^{r} d_i^{\mu_{ij}}.$$

Write each quantity as $q_j = u_j[q_j]$. Then the relation

$$u_0 \prod_{i=1}^{r} d_i^{\mu_{io}} = f \left( u_1 \prod_{i=1}^{r} d_i^{\mu_{1i}}, \ldots, u_n \prod_{i=1}^{r} d_i^{\mu_{ni}} \right)$$

17:22 7 Oct 2003 103 version: 26-09-2003
3. NON-DIMENSIONALISATION AND SCALING

should be homogeneous in any \( d_i \), i.e. any transformation \( d_i \mapsto \lambda d_i \) should produce on the right-hand side \( \lambda^{\mu_0} f(q_1, q_2, \ldots, q_n) \). We refer to this as the Principle of Dimensional Homogeneity.

**Example 7.6** Consider the following model of a quantity \( x \) satisfying the equation
\[
ax^2 + bx + c = 0.
\]
Assume that \( x \) denotes a length, with units in meters, denoted by \([x] = m\), and \( c \) is a velocity with units in meters per second, or \([c] = m/s\). If the equation is dimensionally homogeneous with \([ax^2] = [bx] = [c]\), the units of the other parameters \( a \) and \( b \) cannot be else but \([a] = [c]/[x^2] = 1/\text{ms}\) and \([b] = [c]/[x] = 1/s\). Therefore, we can scale time and length on several combinations to obtain a reduced problem as follows.

- If \( b, c \neq 0 \):
  \[
  x := \frac{c}{b} X, \quad a := \frac{b^2}{c} \alpha, \quad \alpha X^2 + X + 1 = 0;
  \]
- if \( ac > 0 \):
  \[
  x := \sqrt{\frac{a}{c}} X, \quad b := \sqrt{ac} \beta, \quad X^2 + \beta X + 1 = 0;
  \]
- if \( a \neq 0 \):
  \[
  x := \frac{b}{a} X, \quad c := \frac{b^2}{a} \gamma, \quad X^2 + X + \gamma = 0.
  \]

The constants \( \alpha, \beta \) and \( \gamma \) are dimensionless constants, parameterizing the respective reduced problem. It should be noted that any of these scalings are equivalent (no information is lost), but they are not equally useful. The preferred reduction is the one in which \( x \) is scaled on a value typically occurring in the situation considered, and \( X \) is henceforth of order unity. So a careful inspection of the range of numerical values of \( x \) and the parameters \( a, b, c \) is essential. Only then the dimensionless parameter \( \alpha, \beta \) or \( \gamma \) can tell us more about the behaviour of \( X \). \( \square \)

**Example 7.7** Consider an object \( \Omega \) of typical size \( L \) that has initially a temperature distribution \( T(x, 0) = T_0(x) \). The temperature \( T \) satisfies the following heat diffusion equation with thermal diffusion constant \( \alpha \).

![Figure 7.2. A temperature distribution](image)

\[
\begin{align*}
\frac{\partial T}{\partial t} &= \alpha \nabla^2 T, \quad x \in \Omega, \quad t > 0, \\
T(x, t) &= 0, \quad x \in \partial\Omega, \quad t > 0, \\
T(x, 0) &= T_0(x), \quad x \in \Omega.
\end{align*}
\]

The edges of the object are kept at a constant temperature \( T(\partial\Omega, t) = 0 \) (Fig. 7.2). Note that the steady state solution is \( T(x, t) \equiv 0 \). So any gradient of \( T \) is always coupled to a variation...
in time. We scale \( x \) on \( L \), the only length scale in the problem. As the problem is linear, it is not really necessary to scale \( T \), but we could use the mean, or maximum value of \( T_0 \). There is no explicit time scale, \( t_0 \) say, in the problem, for example from an external source. If we leave it unspecified for the moment, and write \( x = L \xi \) and \( t = t_0 \tau \), then we obtain

\[
\frac{1}{t_0} \frac{\partial T}{\partial \tau} = \frac{\alpha}{L^2} \nabla^2 \xi T.
\]

As is also clear from the equation, the only parameter with the dimension of time is the number \( L^2/\alpha \). Therefore, as long as no steady state is achieved, the balance between decay and diffusion implies that the typical decay time (the half-life, say) is given, in order of magnitude, by this number. It is thus the natural time to scale on and so we have \( t_0 = L^2/\alpha \).

\[\square\]

**Example 7.8** A piece of metal \( \Omega \) of size \( L \) is heated, from an initial state \( T(x, t) \equiv 0 \), to a temperature distribution \( T \) by applying an electric field with potential \( \psi \) and typical voltage \( V \) (Fig. 7.3). This heat source, amounting to the energy dissipation of the electric field (see Section 6.9.2), is given by the inhomogeneous term \( \sigma |\nabla \psi|^2 \) in the following inhomogeneous heat equation

\[
\rho C \frac{\partial T}{\partial t} = \kappa \nabla^2 T + \sigma |\nabla \psi|^2, \quad x \in \Omega, \quad t > 0,
\]

\[
T(x, t) = 0, \quad x \in \partial\Omega, \quad t > 0,
\]

\[
T(x, 0) = 0, \quad x \in \Omega.
\]

The edges are kept at \( T = 0 \). If we introduce the formal scaling \( T = T_0 u, \ t = t_0 \tau, \ x = L \xi, \ \) and \( \psi = V \Psi \), then we get

\[
\frac{\rho C T_0}{t_0} \frac{\partial u}{\partial \tau} = \frac{\kappa T_0}{L^2} \nabla^2 \xi u + \frac{\sigma V^2}{L^2} |\nabla \Psi|^2.
\]

Assuming a balance between the storage (1-st) and dissipation (2-nd) term during the initial phase of the process (although details may vary with the applied field \( \psi \)), it follows that the generated heat is dissipated through the metal with a typical decay time of \( O(\rho CL^2/\kappa) \), which is therefore a natural choice for the scaling time \( t_0 \).

Assuming a balance between the dissipation and source (3-d) term in the stationary state, it follows that the final temperature of the stationary state is typically \( O(\sigma V^2/\kappa) \), which is therefore a suitable choice for \( T_0 \), the temperature for scaling.

Note that the boundary conditions are rather important. If the edges were thermally isolated, we would, at least initially, have no temperature gradients scaling on \( L \). Only the storage term would balance the source term, and there would be no other temperature to scale on than \( \sigma V^2 t_0/\rho CL^2 \). In other words, the temperature would rise approximately linearly in time.

See for an extensive description Example 7.17.  

\[\square\]
3. NON-DIMENSIONALISATION AND SCALING

3.2 Dimensional Analysis

In any description of reality, the variables and parameters have physical dimensions and are therefore dimensionally related. For example, any relation or physical law is only universally valid if the expression or expressions are dimensionally homogeneous, and independent of the physical dimensions used. This fact alone implies that the number of essentially independent, i.e. dimensionless, groups of variables and parameters is less than the number of variables and parameters.

**Definition 7.9.** Dimensionless groups consisting of a combination of only problem parameters are called dimensionless numbers. Groups consisting of a combination of parameters and several variables are called similarity variables.

As the number of dimensionless groups tells us something about the complexity of the model involved (and for a good model this should not be too high; see the above sections 1,2), we are interested to determine how many dimensionless groups are at most possible.

A fundamental result quantifying this is known as Buckingham’s Π-theorem ($\Pi$ stands for the products).

**Theorem 7.10 (Buckingham’s Π-theorem).** If a physical quantity is described by $n$ variables and parameters in $r$ dimensions, the number of essentially different problem parameters is at most $n - r$. The number of possible dimensionless groups, including the primary quantity, is at most $n - r + 1$.

**Proof.** Evidently, $n \geq r$. Suppose we have a physical variable $q_0$, described by (together) $n$ variables and parameters: $q_1, \ldots, q_n$, in $r$ independent physical dimensions (e.g. the SI units): $d_1, \ldots, d_r$. All $d_i$ occur at least once. So there exists a relation like

$$q_0 = f(q_1, q_2, \ldots, q_n).$$

From the Principle of Dimensional Homogeneity it follows that $f$ can be written as

$$q_0 = q_1^{\gamma_1} q_2^{\gamma_2} \ldots q_n^{\gamma_n} \Phi(R_1, \ldots, R_m),$$

where $\Phi$ depends on $m$ dimensionless groups of $q_1, \ldots, q_n$ of the form $R_k = q_1^{\alpha_{1k}} \ldots q_n^{\alpha_{nk}}$. $m$ is the number to be determined.

Let $[q]$ denote the physical dimension, or unit of scale, of $q$. Since each group is dimensionless, we have

$$[R_k] = [q_1^{\alpha_{1k}} q_2^{\alpha_{2k}} \ldots q_n^{\alpha_{nk}}] = [q_1^{\alpha_{1k}}][q_2^{\alpha_{2k}}] \ldots [q_n^{\alpha_{nk}}] = 1$$

for $k = 1 \ldots m$, while the dimension $[q_j]$ of each $q_j$ may be written as

$$[q_j] = d_1^{\mu_{1j}} d_2^{\mu_{2j}} \ldots d_r^{\mu_{rj}} = \prod_{i=1}^r d_i^{\mu_{ij}}.$$
So we have

\[
[R_k] = \left[ \prod_{j=1}^{n} q_j^{a_{jk}} \right] = \prod_{j=1}^{n} \prod_{i=1}^{r} d_i^{\alpha_{ij}} = \prod_{i=1}^{r} \sum_{j=1}^{n} \mu_{ij} = 1.
\]

This is only possible if any of the exponents of \( d_i \) is zero. In other words, \( m \), the number of possible groups, is the number of non-trivial solutions of

\[
\sum_{j=1}^{n} \mu_{ij} \xi_j = 0 \quad \text{for} \quad i = 1, 2, \ldots, r,
\]

or in matrix notation

\[
\begin{pmatrix}
\mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\
\mu_{21} & \ddots & & \\
\vdots & & \ddots & \\
\mu_{r1} & \cdots & \cdots & \mu_{rn}
\end{pmatrix}
\begin{pmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_n
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}.
\]

Hence, we find to have at least \( n - r \) non-trivial solutions, as the number of solutions is equal to the dimension \( n \) of the solution vector minus the rank of the matrix which is at most \( r \), the number of equations. However, the base units \( d_1, \ldots, d_r \) are by assumption independent and occur at least once, so all \( r \) equations are independent and the rank of the matrix is exactly \( r \), and we have \( n - r \) non-trivial solutions.

On the other hand, as long as we have not solved the problem in detail we may not be certain that all \( R_1, \ldots, R_{n-r} \) are indeed necessary to describe the problem. Therefore, the number of dimensionless groups is at most \( n - r \).

In the same way we find for \( q_0 \)

\[
[q_0] = d_1^{\mu_{10}} d_2^{\mu_{20}} \cdots d_r^{\mu_{r0}} = \prod_{i=1}^{r} d_i^{\mu_{i0}} = \prod_{i=1}^{r} \sum_{j=1}^{n} \mu_{ij} \gamma_j,
\]

that the possible \( \gamma_1, \ldots, \gamma_n \) satisfy

\[
\begin{pmatrix}
\mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\
\mu_{21} & \ddots & & \\
\vdots & & \ddots & \\
\mu_{r1} & \cdots & \cdots & \mu_{rn}
\end{pmatrix}
\begin{pmatrix}
\gamma_1 \\
\gamma_2 \\
\vdots \\
\gamma_n
\end{pmatrix}
= \begin{pmatrix}
\mu_{10} \\
\mu_{20} \\
\vdots \\
\mu_{r0}
\end{pmatrix}.
\]

and thus have in general (at most) \( n - r + 1 \) solutions: one solution of the inhomogeneous problem and (at most) \( n - r \) solutions of the homogeneous problem.

\begin{example}
Consider the drag \( D \) - the reaction force due to the surrounding flow - of a sphere of radius \( a \) moving with velocity \( V \) in a viscous fluid with viscosity \( \mu \) and density \( \rho \).
\end{example}
We assume, as our model, that the drag $D$ is only dependent on $\rho$, $V$, $\mu$ and $a$. (This is true for relatively low velocities, an infinite medium and a relatively large sphere).

Now we verify the dimensions of the parameters $[D] = \text{kg m/s}^2$, $[\rho] = \text{kg/m}^3$, $[V] = \text{m/s}$, $[\mu] = \text{kg/m s}$, $[a] = \text{m}$. Presented in the form of a table, with at each entry the corresponding exponent of the base units kg, m and s, this is

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$V$</th>
<th>$\mu$</th>
<th>$a$</th>
<th>$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>kg</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>m</td>
<td>-3</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>s</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
</tbody>
</table>

We conclude that we can express $D$ in a functional relationship with at most $4 - 3 = 1$ dimensionless constant. The best known form is the Reynolds number: $Re = \frac{\rho VL}{\mu}$. The drag is usually either scaled on the pressure difference $\frac{1}{2}\rho V^2 a^2$ between front and back of the sphere, or on the viscous friction $aV\mu$ due to wall shear stress. This leads to the following functional relations

$$D = \frac{1}{2}\rho V^2 a^2 F(Re) = aV\mu G(Re), \quad \text{where } Re = \frac{\rho VL}{\mu}.$$ 

The first one is the proper scaling for nearly inviscid flow ($Re$ large), and the second one for very viscous flow ($Re$ small).

**Example 7.12** A famous example, originally due to G.I. Taylor [16, 15], is the analysis of the shock front propagation of a very intense (e.g. nuclear bomb) explosion. From physical considerations the radius of the shock wave front $R$ depends, during the early stages of the explosion when the pressure inside the shock wave is much higher than outside, only on the time interval $t$, since the explosion, the initial energy $E$, and the initial air density $\rho_0$. Since $[R] = \text{m}$, $[t] = \text{s}$, $E = \text{kg m}^2/\text{s}^2$, and $\rho_0 = \text{kg/m}^3$, we have only $3 - 3 = 0$ dimensionless groups. In other words, we can express $R$ as

$$R = \text{constant} \left( \frac{E}{\rho_0} \right)^{1/5} t^{2/5}.$$ 

The full solution to the appropriate gas dynamical problem showed that the constant has a value close to unity.

**Example 7.13** The resonance frequency $\omega$ of a freely suspended membrane (no resonance cavity) is determined by the air density $\rho_a$, sound speed $c_a$, and the membrane tension $T$, density $\sigma$ and diameter $a$. As $[\rho_a] = \text{kg/m}^3$, $[c_a] = \text{m/s}$, $[T] = \text{kg/s}^2$, $[\sigma] = \text{kg/m}^2$ and $[a] = \text{m}$, we have $5 - 3 = 2$ dimensionless numbers determining $\omega$. A possible choice is

$$\omega = \frac{c_M}{a} F(\frac{c_M}{c_a}, \frac{\rho_a a}{\sigma})$$

where for convenience we introduced $c_M = (T/\sigma)^{1/2}$, the propagation speed of transversal waves in the membrane in the absence of air loading.

**Example 7.14** The height $h$ of a drop of liquid at rest on a horizontal surface with the effect of gravity being balanced by surface tension is a function of liquid density $\rho$, volume $L^3$, acceleration of gravity $g$, surface tension $\gamma$ and contact angle $\alpha$. As $[h] = \text{m}$, $[\rho] = \text{kg/m}^3$,
[L] = m, [g] = m/s², [γ] = kg/s², and [α] = 1, we have 5 − 3 = 2 dimensionless number. Possible choices are (α is already dimensionless)

\[
h = L F(Bo, α) = \left( \frac{γ}{ρg} \right) \frac{1}{2} G(Bo, α), \quad \text{where} \quad Bo = \frac{ρgL^2}{γ}.
\]

Bo is known as the Bond number. The first form is suitable when Bo is small (high relative surface tension). The drop becomes spherical and \( h \) is comparable with \( L \). The second form is the proper scaling when Bo is large (low relative surface tension). The drop will be flat as a pancake and \( h \) is comparable with \( \sqrt{γ/ρg} \) [92]. □

3.3 Similarity solutions.

If the problem contains no other length scale than the spatial variable \( x \) itself and no other time scale than the time variable \( t \) itself, dimensionless groups can only occur by combinations of \( x \) and \( t \). Thus, dimensional analysis leads naturally to similarity solutions. This is easiest explained by an example.

**Example 7.15** Consider the following heat conduction problem. A very long thermally isolated bar, initially at uniform temperature zero, is heated at one end by a constant flux. There is no source at the other end. The bar is modelled as semi-infinite (Fig. 7.4), with a crosswise constant temperature distribution, while the temperature \( T \) is described by the following one-dimensional equation for heat conduction, with constant heat diffusion coefficient \( α \),

\[
\frac{∂T}{∂t} = α \frac{∂^2T}{∂x^2}, \quad x \in [0, ∞), \quad t > 0,
\]

with initial and boundary conditions

\[
\frac{∂T}{∂x} (0, t) = -Q_0, \quad t > 0,
\]

\[
T(x, 0) = 0, \quad x \in [0, ∞),
\]

\[
0 ≤ T(x, t) < ∞ \quad x → ∞, \quad t > 0.
\]

If we try to scale, we find that we modelled any explicit length, time or temperature scale out of our problem. So we can only make dimensionless on the available implicit scales:

- As there is no length scale in \( x \) or \( t \), the intrinsic length scale can only be \( \sqrt{αt} \).
- The only temperature in the problem is \( Q_0x \) or \( Q_0\sqrt{αt} \).

Therefore, we assume:

\[
T(x, t) = Q_0xg(η),
\]

where the similarity variable \( η \) is given by

\[
η = \frac{x}{\sqrt{4αt}}.
\]
It follows that $g$ satisfies the following reduced ordinary differential equation

$$\frac{1}{2} \eta g'' + (1 + \eta^2) g' = 0,$$

with boundary conditions

$$\lim_{\eta \to \infty} g(\eta) = 0, \quad \lim_{\eta \to 0} \left( g(\eta) + \eta g'(\eta) \right) = -1.$$

This has the solution

$$g(\eta) = \frac{1}{\eta \sqrt{\pi}} \exp(-\eta^2) - \text{erfc}(\eta),$$

where $\text{erfc}(x) := \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-\xi^2} \, d\xi$. Hence we have

$$T(x, t) = Q_0 \left[ \sqrt{\frac{4 \alpha t}{\pi}} \exp(-\eta^2) - x \text{erfc}(\eta) \right].$$

Note that there exists no stationary solution!

**Example 7.16** In the convection problem

$$\frac{\partial u}{\partial t} + U_0 \frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = H(x), \quad x \in \mathbb{R},$$

there is the length scale given by $U_0 t$ and a length scale, say $L$, hidden in the initial profile $H(x)$, as $x$ cannot occur on its own. The dimensions of $u$ and $H$ are the same, say $H_0$, and we write $H(x) = H_0 h(x/L)$. We scale $x = \xi, t = \frac{U_0}{L} \tau, u = H_0 v$, and $v(\xi, 0) = h(\xi)$, to get

$$\frac{\partial v}{\partial \tau} + \frac{\partial v}{\partial \xi} = 0,$$

with solution $v(\xi, \tau) = h(\xi - \tau)$.

**Example 7.17** Consider the edge singularity of the time-dependent temperature field generated in a homogeneous and isotropic conductor by an electric field (see Eqs. (6.9.5)). The electric current density $J$ and the electric field $E$ satisfy Ohm’s law (6.9.10) $J = \sigma E$, where $\sigma$ is the electric conductivity, i.e. the inverse of the specific electric resistance. For an effectively stationary current flow the conservation of electric charge (6.9.6) leads to a vanishing divergence of the electric current density, $\nabla \cdot J = 0$. The electric field $E$ satisfies $\nabla \times E = 0$, and therefore has a potential $\psi$, with $E = -\nabla \psi$, satisfying $\nabla \cdot (\sigma \nabla \psi) = 0$. The electric conductivity $\sigma$ is a material parameter which is quite strongly dependent on temperature. Nevertheless, to make progress we will assume a constant $\sigma$, independent of $T$. This, then, leads to the Laplace equation for $\psi$

$$\nabla^2 \psi = 0.$$

The heat dissipated as a result of the work done by the field per unit time and volume (Ohmic heating) is given by Joule’s law $J \cdot E$ (see Section 6.9.2), and leads to the heat-source distribution

$$\sigma |\nabla \psi|^2.$$
Since energy is conserved, the net rate of heat conduction and the rate of increase of internal energy are balanced by the heat source (equation (6.8.3) with \( \rho \) the above heat source), which yields the equation for temperature \( T \)

\[
\rho C \frac{\partial T}{\partial t} = \kappa \nabla^2 T + \sigma |\nabla \psi|^2. 
\]

The thermal conductivity \( \kappa \), the density \( \rho \) and the specific heat of the material \( C \) are mildly dependent on temperature, but we assume these parameters constant.

![Figure 7.5. A wedge shaped conductor, heated by an electric field](image)

Since we are interested in the rôle of the edge only, the conductor is modelled, in cylindrical \((r, \phi)\)-coordinates, as an electrically and thermally isolated infinite wedge-shaped two-dimensional region (without any geometrical length scale; Fig. 7.5) \( 0 \leq \phi \leq \nu \) with an electric field with potential

\[
\psi(x, y) = \left(\frac{\nu}{\pi}\right) A r^{\pi/\nu} \cos(\phi \pi/\nu),
\]

while the temperature distribution \( T \) due to the heat generated by this source is then given by

\[
\rho C \frac{\partial T}{\partial t} = \kappa \nabla^2 T + \sigma A^2 r^{2\pi/\nu - 2}
\]

with boundary and initial conditions

\[
\frac{\partial T}{\partial \phi} = 0 \quad \text{at} \quad \phi = 0, \quad \phi = \nu, \quad T(x, y, 0) = 0.
\]

Since there are no other (point) sources at \( r = 0 \), we have the additional condition of a finite field at the origin: \( 0 \leq T(0, 0, t) < \infty \). Boundary conditions and the symmetric source imply that \( T \) is a function of \( t \) and \( r \) only, so that equation (*) reduces to

\[
\rho C \frac{\partial T}{\partial t} = \kappa \left( \frac{\partial^2 T}{\partial r^2} + \frac{1}{r} \frac{\partial T}{\partial r} \right) + \sigma A^2 r^{2\pi/\nu - 2}.
\]

Owing to the homogeneous initial and boundary conditions, the infinite geometry, and the fact that the source is a monomial in \( r \), homogeneous of the order \( 2\pi/\nu - 2 \), there is no length scale in the problem other than \( \sqrt{\frac{\kappa t}{\rho C r^2}} \), while the temperature \( T \) can only scale on \( \frac{A^2}{r^{2\pi/\nu}} \). This indicates that a similarity solution is possible of the following form

\[
T(r, t) = \frac{\sigma}{4\kappa} A^2 r^{2\pi/\nu} h(X), \quad X = \frac{\rho C r^2}{4\kappa t},
\]

where \( X \) is a similarity variable, reducing equation (†) to

\[
X^2 h'' + X \left( 1 + \frac{\pi}{\nu} + X \right) h' + \frac{\pi^2}{\nu^2} h = -1,
\]

with boundary conditions, corresponding to the behaviour near \( r = 0 \) and \( t = 0 \),

\[
0 \leq X^{\pi/\nu} h(X) < \infty \quad \text{if} \quad X \downarrow 0, \quad h(X) \to 0 \quad \text{if} \quad X \to \infty.
\]

17:22 7 Oct 2003 111 version: 26-09-2003
This equation may be recognized as related to the confluent hypergeometric equation in \(-X\). It has the solution, with the required behaviour in \(X = 0\), given by

\[ h(X) = \text{constant} \cdot X^{-\pi/\nu} M(-\frac{\pi}{\nu}; 1; -X) - \frac{2}{\nu} \]

where \(M(a; b; z)\) is Kummer’s function or the regular confluent hypergeometric function \([6, Ch.13]\). From the asymptotic expansion of \(M(-\pi/\nu; 1; -X)\) and the condition for \(X \to \infty\), the unknown constant is found to be \((\nu/\pi)/\Gamma(1+\pi/\nu)\).

Putting everything together, we have the solution

\[
T(r, t) = \frac{\sigma v^2 A^2}{4\pi^2 \kappa} \Gamma\left(1 + \frac{\pi}{\nu}\right) \left(\frac{\rho Cr^2}{4\kappa t}\right)^{-\pi/\nu} M\left(-\frac{\pi}{\nu}; 1; -\frac{\rho Cr^2}{4\kappa t}\right) \left.\right] - \frac{1}{\nu}.
\]

At the edge we have \(T(0, t) \sim r^{\pi/\nu}\). This shows, together with the radial temperature distribution given in Fig. 7.6, a marked difference in behaviour between outward \((\nu < \pi)\) and inward \((\nu > \pi)\) pointed wedges. For the first category, the temperature at the corner rises smoothly and so slowly that it always remains behind the temperatures for larger \(r\). For the other category it is just the other way around. The corner temperature rises abruptly and so quickly that the values for larger \(r\) are always lower.

**Example 7.18** Since the range of our human audible sensitivity is incredibly large \((10^4 \text{ in energy})\), the loudest and quietest levels are practically infinitely far away. Therefore, we have no reference or scaling level to compare with, other than the sound itself we are hearing. As a result, variations in sound loudness \(dL\) are perceived proportional to relative variations of the physical sound intensity \(dI/I\) and thus \(L\) varies logarithmically in \(I\). As the intensity \((\text{time-averaged energy flux}) I\) is, for a single tone, proportional to the mean squared acoustic pressure \(p_{\text{rms}}^2\), we have for suitable constants \(K\) and \(L_0\) the relation \(L = K \log(p_{\text{rms}}) + L_0\).

When

\[
L = 2 \log_{10}(p_{\text{rms}}/p_0)
\]

for a reference value \(p_0 = 2 \cdot 10^{-5} \text{ Pascal}\), we call \(L\) the Sound Pressure Level in Bels. The usual unit is one tenth of it, the decibel.

4 Scaling and reduction of the Navier-Stokes equations

One of the most fruitful usages of scaling is the hierarchy it provides in comprehensive and rich models. In most applications, such over-complete models are not truly adapted...
CHAPTER 7. THE ART OF MODELLING

...to the problem, and therefore not a true, "lean" model as we discussed above. From a suitable, inherent scaling the order of magnitude of the various contributions or effects can be estimated, leading to a hierarchy on which we can base an asymptotic modelling.

We will give here a (maybe the) most important example, the scaling and reduction of the compressible Navier-Stokes equations. We recall the mass, momentum and energy conservation equations (6.8.9) with constitutive relations (6.8.10) as derived in Chapter 6.

4.1 Scaling, non-dimensionalisation.

Assume that we have the following typical scales for velocity, pressure, density and temperature.

\[ v \text{ with } v_0, \Delta v; \quad p \text{ with } p_0, \Delta p; \quad \rho \text{ with } \rho_0, \Delta \rho; \quad T \text{ with } \Delta T; \quad x \text{ with } L; \quad t \text{ with } f^{-1} \]

where a subscript "0" refers to the primary variable, and \( \Delta \) denotes a typical difference. The typical time is written as the inverse of a frequency \( f \). Note that sometimes we have more candidates for a scaling parameter. For example, the typical frequency may be enforced by an external source or instability, or may be inherited from the intrinsic hydrodynamic frequency \( v_0/L \), or the inverse diffusion time \( \kappa/\rho_0 C_P L^2 \).

The scaled equations (6.8.9a), (6.8.9b), and (6.8.9g) are now, symbolically, given by

mass:

\[ f \Delta \rho, \frac{v_0 \Delta \rho}{L} = \frac{\rho_0 \Delta v}{L}, \frac{\Delta T}{L} \]

momentum:

\[ f \rho_0 \Delta v, \frac{\rho_0 v_0 \Delta v}{L} = \frac{\Delta p}{L}, \frac{\mu \Delta v}{L^2} \]

energy:

\[ f \rho_0 C_P \Delta T, \frac{\rho_0 C_P v_0 \Delta T}{L} = f \Delta \rho, \frac{v_0 \Delta \rho}{L}, \frac{\kappa \Delta T}{L^2}, \frac{\mu (\Delta v)^2}{L^2} \]

From gas laws we know that the typical sound speed squared \( c_0^2 \) scales with both \( p_0/\rho_0 \) and \( \Delta p/\Delta \rho \) (not exactly, only in order of magnitude) so we can take \( p_0 = \rho_0 c_0^2 \) and \( \Delta \rho = \Delta p/c_0^2 \). Furthermore, we will take here the common situation where \( \Delta v = O(v_0) \), and \( \kappa \) and \( \mu \) are constants. (This may not be always the case!) With a rescaling such that all coefficients become dimensionless, we have

\[ \frac{f L \Delta p}{v_0 \rho_0 c_0^2} = 1, \quad \frac{\Delta p}{\rho_0 v_0^2}, \frac{\mu}{\rho_0 v_0 L}, \quad \frac{f L}{v_0}, 1 = \frac{\Delta p}{\rho_0 v_0^2}, \frac{\mu}{\rho_0 v_0 L}, \quad \frac{f L \Delta p}{\rho_0 C_P v_0 \Delta T}, \frac{\Delta p}{\rho_0 C_P \Delta T}, \frac{\kappa \rho_0 C_P v_0 L}{\rho_0 C_P L \Delta T}. \]

From these table we can collect by inspection all the potentially relevant dimensionless numbers that may occur in problems described by these Navier-Stokes equations.

4. SCALING AND REDUCTION OF THE NAVIER-STOKES EQUATIONS

4.2 Some dimensionless groups with their common names

In a natural way a large number of dimensionless numbers or groups, characterizing a flow, have appeared. Each represent a certain balance between two or more effects. If the typical values used for the scaling are well-chosen, these numbers already provide an enormous amount of information about the physical problem considered. A more complete list, derived also from other fields, is given in Appendix, Section M.

Reynolds: \( Re = \frac{\rho_0 v_0 L}{\mu} \), Prandtl: \( Pr = \frac{C_p \mu}{\kappa} \), Mach: \( M = \frac{v_0}{c_0} \),

Strouhal: \( Sr = \frac{f L}{v_0} \), Fourier: \( Fo = \frac{\kappa f^{-1}}{\rho_0 C_p L^2} \), Euler: \( Eu = \frac{\Delta p}{\rho_0 v_0^2} \),

Helmholtz: \( He = \frac{f L}{c_0} \), Eckert: \( Ec = \frac{v_0^2}{C_p \Delta T} \), Pécelt: \( Pe = \frac{\rho_0 C_p v_0 L}{\kappa} \).

These dimensionless numbers have the following physical interpretation.

- The Reynolds number \( Re \) describes how important viscous forces are compared to inertial forces, and tells us if either viscosity or inertia may be neglected.
- The Prandtl number \( Pr \), describing the relative importance of viscous against heat diffusion, depends only on the material, and is for most gases and fluids of order 1.
- The Mach number \( M \), comparing the occurring velocities with the speed of sound, tells us whether the stationary velocity is so high that compressibility effects should be taken into account.
- The Strouhal number \( Sr \) compares an externally enforced frequency \( f \) with the hydrodynamically induced frequency \( v_0/L \).
- The Fourier number \( Fo \) compares a time scale \( f^{-1} \) with the typical time necessary for the diffusion of heat along a distance \( L \).
- The Euler number \( Eu \) compares the available pressure difference with the typical pressure difference that can be expected from hydrodynamical effects alone.
- The Helmholtz number \( He \) compares the typical wave length of sound with the size of a scattering object or a source, which tells us a lot about the effectivity of the scatterer or source.
- The Eckert number \( Ec \) compares the kinetic energy of the flow with available differences in enthalpy.
- The Pécelt number \( Pe \) compares forced convection of heat with heat conduction.

Note that a dimensionless number is not always best described by a balance between two effects. In that case the problem at hand may be better described by another selection of numbers. This is not difficult, as many dimensionless numbers are related. For example, \( Pe = Pr Re \), \( He = Sr M \), and \( Sr Fo Pr Re = 1 \).
CHAPTER 7. THE ART OF MODELLING

The scaled equations are now, in dimensionless variables,

\[ Eu M^2 \left( Sr \frac{\partial \rho}{\partial t} + v \cdot \nabla \rho \right) = -\rho \nabla \cdot v \]  
\[ Sr \rho \frac{\partial v}{\partial t} + \rho v \cdot \nabla v = -Eu \nabla \rho + \frac{1}{Re} \nabla \cdot \tau \]  
\[ Sr \frac{\partial T}{\partial t} + \rho v \cdot \nabla T = Eu Ec \left( Sr \frac{\partial p}{\partial t} + v \cdot \nabla p \right) + \frac{1}{Pe} \nabla^2 T + \frac{Ec}{Re} \tau : \nabla v \]

4.3 Asymptotic reductions of the Navier-Stokes equations

When \( M, Re, Sr, Ec \), etc. become small or large, we may derive from equations (4.1) various reduced models by assuming the respective terms to vanish or dominate completely. Other terms are sometimes best taken equal to one, in particular if one of the involved scales is inherent (i.e. a result) rather than externally enforced.

A Fourier number \( Fo = 1 \) corresponds to a time scale \( \rho_0 C_P L^2/\kappa \), determined by heat diffusion. A Strouhal number \( Sr = 1 \) corresponds to a frequency \( f = v_0/L \), which is connected to hydrodynamical convection processes, for example the shedding of vortices from a blunt trailing edge. For a cylinder of diameter \( D \) experiments indeed show that the shedding frequency is practically equal to \( f_S = 2v_0/D \). The Euler number \( Eu = 1 \) corresponds to a present pressure gradient comparable to hydrodynamic pressure variations, \( \rho_0 v_0^2 \), which is the usual situation in incompressible, inviscid (separated) flow around non-streamlined bodies. We have to select \( Eu Re = \frac{\Delta p L}{\rho_0 v_0 \mu} = 1 \) if the pressure gradient is mainly balanced by viscous forces, for example in very slow or very viscous flow. The combination \( Eu M^2 = \frac{\Delta p}{\rho_0 c_0^2} = 1 \) corresponds to a pressure scaled on pressure variations due to compressible effects, which is common for acoustic problems (pressure-density coupling while \( M \) is small) or high speed problems (\( v_0 \) comparable with \( c_0 \)). We select \( Eu Ec = \frac{\Delta p}{C_P \Delta T} = 1 \) if enthalpy changes are mainly coupled to pressure variations and the flow is at least for the larger part isentropic (cf. Eq. (6.8.8)).

In the following we will derive some important examples of asymptotic models, systematically deduced from the Navier-Stokes equations.

• **Incompressible flow.** When the Mach number \( M \) tends to zero, while we take \( Eu = Sr = 1 \), we have incompressible flow described by

\[ \nabla \cdot v = 0, \quad \rho \left( \frac{\partial v}{\partial t} + v \cdot \nabla v \right) = -\nabla p + \frac{1}{Re} \nabla \cdot \tau. \]

Very often an incompressible flow will have a uniform constant density, but this is not necessary. Note that the energy equation does not disappear, but is decoupled from the other equations if viscosity and density may be taken independent of the temperature. The pressure does not play a rôle any more thermodynamically, as only its gradient occurs as a reaction force. If the fluid is Newtonian and \( \rho \) and \( \mu \) are constant (they may be taken equal to 1), we obtain for (4.2) the form

\[ \nabla \cdot v = 0, \quad \frac{\partial v}{\partial t} + v \cdot \nabla v = -\nabla p + \frac{1}{Re} \nabla^2 v. \]
4. SCALING AND REDUCTION OF THE NAVIER-STOKES EQUATIONS

The equation for mass conservation can be solved by introducing a stream function \( \Psi \), defined by \( \mathbf{v} = \nabla \times \Psi \). This is particularly useful in 2D flow when \( \Psi = (0, 0, \psi)^T \) and \( \mathbf{v} = (\psi_y, -\psi_x, 0)^T \). By taking the curl of the momentum equation we remove the dependence on pressure and obtain (use (J.8-J.11)) the following equation for the transport of vorticity \( \omega := \nabla \times \mathbf{v} \)

\[
\frac{\partial \omega}{\partial t} + \mathbf{v} \cdot \nabla \omega = \omega \cdot \nabla \mathbf{v} + \frac{1}{Re} \nabla^2 \omega. \tag{4.4}
\]

Note that in 2D \( \omega \cdot \nabla \mathbf{v} = 0 \), so (4.4) becomes a convection-diffusion equation in \( \omega \).

**Inviscid compressible flow.** If the Reynolds number tends to infinity, usually also the Peclet number does, because \( Pe = Pr Re \) and the Prandtl number is for most fluids and gases of order 1. If we further take \( Eu = Sr = M = Ec = 1 \), we obtain a compressible inviscid flow described by

\[
\frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho = -\rho \nabla \cdot \mathbf{v}, \tag{4.5a}
\]

\[
\rho \left( \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla p, \tag{4.5b}
\]

\[
\rho \left( \frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T \right) = \frac{\partial p}{\partial t} + \mathbf{v} \cdot \nabla p. \tag{4.5c}
\]

In terms of entropy, the last equation is equivalent to \( \frac{\partial}{\partial t} s = 0 \), so the flow is isentropic everywhere where the assumptions hold.

**Stokes flow.** If the velocities of a viscous flow are so low that \( Re \) tends to zero, while the time scales are all determined by the flow itself (no external forcing) such that \( Sr \) remains finite, we need to scale the pressure gradients on the inverse Reynolds number in order to have flow at all, i.e. \( Eu \) tends to infinity such that \( Eu Re = 1 \). If, in addition, the velocities remain so small compared to the sound speed, i.e. \( M \ll 1 \), that \( Eu M^2 \) tends to zero, we obtain the very viscous incompressible, or Stokes flow, given by

\[
\nabla \cdot \mathbf{v} = 0, \quad -\nabla p + \nabla \cdot \mathbf{\tau} = 0. \tag{4.6}
\]

Again, it should be noted that the energy equation is not negligible, but only decoupled from the other equations (provided the viscosity is not temperature dependent).

**Sound waves.** Consider small pressure-density perturbations in an atmosphere of uniform mean pressure. Assume that the frequencies are relatively high and the typical velocities are small but large enough to neglect viscosity, such that \( M \to 0 \) while we choose \( Eu = Sr = M^{-1} \) and \( Ec = M \), and \( Re \) and \( Pe \) remain finite or large. We then retain

\[
\frac{\partial \rho}{\partial t} = -\rho \nabla \cdot \mathbf{v}, \quad \frac{\partial \mathbf{v}}{\partial t} = -\nabla p, \quad \frac{\partial T}{\partial t} = \frac{\partial p}{\partial t}. \]

Written in terms of entropy \( s \), the last equation is equivalent to \( \frac{\partial}{\partial t} s = 0 \). This means that pressure and density perturbations are coupled isentropically by \( \frac{\partial}{\partial t} p = c^2 \frac{\partial}{\partial t} \rho \). Noting that \( \rho c^2 \) is proportional to the mean pressure which is constant, we can now eliminate \( \rho \) and \( \mathbf{v} \).
to obtain finally the wave equation
\[
\frac{\partial^2 p}{\partial t^2} = \nabla \cdot (c^2 \nabla p). \tag{4.7}
\]
If \( \rho \) and \( c \) are constant we have the usual wave equation with constant coefficients.

- **Convection-diffusion.** In a given flow field with \( Sr = 1, Pe \) finite, and \( Eu Ec \) and \( Ec / Re \) negligible, we get for the temperature the convection-diffusion problem
\[
\rho \left( \frac{\partial T}{\partial t} + v \cdot \nabla T \right) = \frac{1}{Pe} \nabla^2 T. \tag{4.8}
\]
We end with two important reductions, not immediately obtainable from small parameter considerations.

- **Potential flow.** If the flow is irrotational, i.e. the vorticity vector \( \omega = \nabla \times v = 0 \), a scalar velocity potential \( \phi \) may be introduced with
\[
v = \nabla \phi.
\]
For example, in inviscid homentropic flow, any vorticity is convected with the flow (see Kelvin’s Theorem, Exercise 6.3d), and if the flow starts irrotational it stays that way. In incompressible flow this potential is independent of pressure (except indirectly via boundary conditions) and satisfies Laplace’s equation
\[
\nabla^2 \phi = 0. \tag{4.9}
\]

- **Bernoulli’s law.** In stationary, incompressible and inviscid flow of constant density \( \rho_0 \) the momentum equation
\[
\rho_0 \left( \frac{\partial v}{\partial t} + v \cdot \nabla v \right) = -\nabla p
\]
can be integrated along a streamline, leading to an equation, known as Bernoulli’s law, that describes conservation of mechanical energy density. By using the vector identity
\[
v \cdot \nabla v = \frac{1}{2} \nabla |v|^2 + \omega \times v
\]
and noting that \( (\omega \times v) \cdot d\ell = 0 \), we can integrate along a stream line
\[
\int \left[ \frac{1}{2} \rho_0 |v|^2 + \rho_0 \omega \times v + \nabla p \right] \cdot d\ell = \int \nabla \left[ \frac{1}{2} \rho_0 |v|^2 + p \right] \cdot d\ell = 0,
\]
to get the famous and useful identity
\[
\frac{1}{2} \rho_0 |v|^2 + p = \text{constant}. \tag{4.10}
\]
Bernoulli’s equation in irrotational flow is valid everywhere rather than only along a streamline. By introducing a potential, we can generalise to unsteady flow with gravity in \( z \)-direction as follows
\[
\frac{\partial \phi}{\partial t} + \frac{1}{2} \nabla |\phi|^2 + \frac{p}{\rho_0} + gz = C(t), \tag{4.11}
\]
where $C$ is an arbitrary function of time. A typical example is water. Bernoulli’s equation may be generalised to include compressibility if the flow is barotropic, i.e. the pressure is a function of density alone. For example, if we have an irrotational inviscid homentropic perfect gas flow, we obtain with the result of Exercise 6.3c

$$\frac{\partial \phi}{\partial t} + \frac{1}{2} |\nabla \phi|^2 + \frac{\gamma}{\gamma - 1} \frac{p}{\rho} = C(t), \quad (4.12)$$

where $\gamma$ is the specific-heat ratio and $C$ is an arbitrary function of time.

5 Discussion

• The mathematical solution of a real world problem starts with the modelling phase, where the problem is described in a mathematical representation of its primitive elements and their relations.

• As the solution is not served by unnecessary complexity, we are interested in an adequate mathematical description with the lowest number of essentially independent parameters and variables. We will call this a model.

• A model may be constructed ad-hoc, by collecting as much as possible pieces of information and combining this into what we call a building block model. Usually, this is the first type of model if a new area of research is explored. Once the foundations are laid, it is not necessary anymore to start from scratch, but we can use the complete and comprehensive descriptions that have become available. By removing any unnecessary details these usually overcomplete models can be simplified to the model required. This is what we call a systematic or asymptotic model.

• A very important aspect in the modelling is the identification of a hierarchy of importance, to distinguish between the important, the less important, and the unimportant effects. From this hierarchy we may decide which aspects can be included and which can be neglected in the model. An important tool in this respect is the notion of scaling and dimensionless numbers. If this is done properly, it is possible to characterize a problem without any calculation.

Exercises

7.1. Simplify, by suitable scaling of the variables, the Korteweg-de Vries equation

$$\frac{\partial u}{\partial t} + e(1 + \frac{3}{2}u) \frac{\partial u}{\partial x} + \frac{1}{2} e h^2 \frac{\partial^2 u}{\partial x^2} = 0$$

and the linearized Boussinesq equation

$$\frac{\partial^2 u}{\partial t^2} - e^2 \frac{\partial^2 u}{\partial x^2} - \frac{1}{2} h^2 \frac{\partial^4 u}{\partial x^4 \partial t^2} = 0,$$
such that the coefficients become just equal to 1.

7.2. Non-dimensionalize the telegraph equation (*) in Example 1.4

\[
\frac{\partial^2 u}{\partial t^2} + (a + b) \frac{\partial u}{\partial t} + abu - c^2 \frac{\partial^2 u}{\partial x^2} = 0,
\]

such that the resulting problem depends on one dimensionless number only.

7.3. The drag \( D \) of a moving ship, due to viscous effects and wave generation, depends on its length \( L \), velocity \( V \), water viscosity \( \mu \), gravity acceleration \( g \) and water density \( \rho \). The dimensional units are \([D]=\text{kg m/s}^2\), \([L]=\text{m}\), \([V]=\text{m/s}\), \([\mu]=\text{kg/ms}\), \([g]=\text{m/s}^2\), and \([\rho]=\text{kg/m}^3\). By how many dimensionless groups is the problem completely described? Give an example of such a set of dimensionless groups (these are not uniquely defined).

7.4. Analyse Example 7.17 in terms of Buckingham’s theorem. Verify the dimensional groups, including their number. Note that \([\sigma]=\text{A}^2\text{s}^3/\text{m}^3\text{kg}\), \([\rho C]=\text{kg/ms}^2\text{K}\), \([\psi]=\text{kg m}^2/\text{s}^3\text{A}\), \([\kappa]=\text{kg m/s}^3\text{K}\).

7.5. Reconsider Example 7.1. Make the problem complete by adding boundary and initial conditions. Make the problem dimensionless by scaling on the inherent time and length scales. Determine conditions, in terms of a dimensionless number, for which the diffusion term can be neglected. Note that the order of the differential equation is then reduced from 2 to 1. What are the consequences for the boundary and/or initial conditions?

7.6. Consider equation (†) of example 7.4 to describe a stationary suspended flexible bar of length \( L \). Make the problem complete by adding suitable boundary conditions at \( x=0 \) and \( x=D \) where \( 1-D/L \) is positive and not small. Note the two integration constants \( H_0 \) and \( V_0 \), so we need four conditions. Make the problem dimensionless by scaling lengths on \( L \) and forces on \( QL \). Under what condition, in terms of a dimensionless parameter, can we neglect bending stiffness (i.e. the term multiplying \( \phi_{ss} \))? The result describes a cable with vanishing bending stiffness, or catenary. What are the consequences for the boundary conditions? Solve this equation.

7.7. Material of concentration \( c \) is diffused from a container located at \(|x|\leq a\) through a membrane at \(|x|=a\) into the outer medium \(|x|>a\). In the container and in the medium the diffusion is described by

\[
\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left( D \frac{\partial c}{\partial x} \right).
\]

The diffusion coefficient in the container is \( D=D_1 \), in the outer medium \( D=D_2 \). Initially, \( c=0 \) for \(|x|>a\), and \( c=c_0 \) in \(|x|\leq a\). At the interfaces \( x=\pm a \) we have \( c \) and \( Dc_x \) to be continuous. Conservation of mass requires that \( \int_{-\infty}^{\infty} c(x,t) \, dx \) is constant.

Describe the problem in dimensionless variables. Approximate the solution for \( t \) large. Hint: use the fact that eventually the majority of the material is diffused to the outer medium; then derive a similarity solution. Note the symmetry in \( x \).

7.8. We are interested to know at what distance \( D \) a boat of height \( H \) is still visible above the horizon. Criticize the following (incorrect) model.
The height of a person is negligible compared to the earth radius $R$. So our field of visibility is just in the tangent plane of the earth at the position of the observer. A boat is visible in this plane if $\cos(D/R) \geq \frac{R}{R+H}$. Since $D/R$ and $H/R$ are small this is equivalent to $D^2 \leq 2RH$.

7.9. A simple model for the temperature $T$ in the ground, at time $t$ and depth $z$, is

$$\rho C \frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2}, \quad -\infty < z \leq 0$$

where $T(0,t) = T_0(t)$ is the given air temperature at surface $z = 0$. $\rho$ denotes the density, $C$ the specific heat and $\kappa$ the thermal conductivity of the soil. This is for dry sand: $\rho = 1600$ kg/m$^3$, $C = 800$ J/kg K, $\kappa = 0.3$ W/m K, and for saturated sand: $\rho = 2000$ kg/m$^3$, $C = 1480$ J/kg K, $\kappa = 2.2$ W/m K. Estimate, by scaling, the typical penetration depth of the yearly temperature variations (steady state). Compare this with an exact solution if $T_0$ varies harmonically like $T_0(t) = A + B \cos(\omega t)$ where $\omega = \frac{2\pi}{1 \text{ year}}$, and $T$ is in steady state. If the temperature for $z \to -\infty$ is below $0^\circ$ C, we call this soil permafrost.
Chapter 8
The analysis of elliptic equations

In this chapter we discuss analytical methods for elliptic equations. We define in Section 1 several boundary value problems for elliptic equations and investigate uniqueness of the solution. The concepts of eigenvalues and eigenfunctions of an elliptic boundary value problem is introduced in Section 2. An important analytical solution method for linear elliptic equations is separation of variables and this is discussed in Section 3. In Section 4, we introduce the so-called fundamental solution of the Poisson equation, which is a solution of the Poisson equation with a Dirac delta function as source term. If furthermore, the fundamental solution satisfies certain homogeneous boundary conditions, it is called a Green’s function. Next, in Section 5, we derive integral representations for the solution of elliptic boundary value problem using these Green’s functions. A qualitative description of the solution of the Poisson equation is based on the maximum principle and is presented in Section 6. As an application of elliptic equations, we study in Section 7 the Stokes equation, we describe creeping flow. In particular, we compute the (slow) flow around a sphere.

1 The Laplace operator

Elliptic differential equations occur typically in problems which describe stationary situations, i.e. the time has no explicit rôle. The simplest and most well known elliptic equation is the Laplace equation, defined on a domain \( \Omega \subset \mathbb{R}^d \) (\( d = 1, 2, 3 \)) say,

\[
L[u] := \nabla^2 u = 0, \quad x \in \Omega.
\]

(1.1)

In the inhomogeneous case we have the Poisson equation

\[
L[u] := \nabla^2 u = f(x), \quad x \in \Omega.
\]

(1.2)

A further type, often encountered, is the Helmholtz equation, which is actually related to the eigenvalue problem of (1.1)

\[
L[u] := \nabla^2 u - \lambda u = 0, \quad x \in \Omega, \quad \lambda \in \mathbb{R}.
\]

(1.3)
1. THE LAPLACE OPERATOR

The Laplace operator $\nabla^2$ also occurs quite often in time dependent problems, like the heat equation or the wave equation. This provides for additional interest to investigate problems like (1.1), (1.2) and (1.3).

1.1 Problem types

The Laplace operator typically occurs in situations where the flux $f$ of a variable is proportional to its gradient. As in Example 1.1 we may consider a concentration $c$, which causes a flow to areas with lower concentration, i.e.

$$f = -D \nabla c,$$

with $D > 0$ the diffusion coefficient. If we apply the Gauss’ divergence theorem (see Appendix J.12) to an arbitrary volume $W \subset \Omega$ we find

$$-\int_W \nabla \cdot (D \nabla c) \, dV = -\oint_{\partial W} (D \nabla c) \cdot n \, dS = \oint_{\partial W} f \cdot n \, dS.$$

Without sources or sinks, the net flux through $\partial W$ should be 0. If furthermore the diffusion coefficient $D$ is constant, we obtain equation (1.1). Any solution in $C^2(\Omega)$ of (1.1) is called a harmonic function.

In order to define a solution more precisely, we have to specify a boundary condition. Three common cases are distinguished for $x \in \partial \Omega$:

$$u(x) = a(x), \quad \text{(Dirichlet)}$$

$$\frac{\partial u}{\partial n}(x) = b(x), \quad \text{(Neumann)}$$

$$\alpha u(x) + \beta \frac{\partial u}{\partial n}(x) = c(x), \quad \alpha, \beta \neq 0, \quad \text{(Robin)}$$

with $n$ the outward unit normal on $\partial \Omega$. $\frac{\partial}{\partial n} u$ denotes the normal derivative, i.e. $\frac{\partial}{\partial n} u := n \cdot \nabla u$.

Condition (1.6a) is called a Dirichlet boundary condition, (1.6b) is called a Neumann boundary condition, and finally, (1.6c) is called a Robin boundary condition. We can easily establish uniqueness of a solution of (1.1) and either one of the two boundary conditions (1.6a) or (1.6c). It is instructive to illustrate these problems for the one-dimensional case, where the Poisson equation reduces to an ordinary differential equation.

Example 8.1 Consider the two-point Dirichlet boundary value problem

$$L[u] := \frac{d^2u}{dx^2} = f(x), \quad x \in (0, \pi),$$

$$u(0) = u(\pi) = 0$$

where $f$ is piecewise smooth on $(0, \pi)$. We now use Fourier sine series to find the solution in a formal way. We take the Ansatz

$$u(x) = \sum_{k=1}^{\infty} a_k \sin(kx),$$
which has the advantage that $u$ satisfies the boundary conditions identically if the series converges uniformly and the found solution is continuous at the end points (Appendix C). This is to be verified afterwards. We expand $f(x)$ also in a Fourier sine series,

$$f(x) = \sum_{k=1}^{\infty} f_k \sin(kx),$$

where

$$f_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(kx) \, dx, \quad k = 1, 2, \ldots$$

and $f_k \to 0$ for $k \to \infty$. We then find

$$\sum_{k=1}^{\infty} \left(k^2 a_k + f_k\right) \sin(kx) = 0.$$

From uniqueness of the Fourier coefficients of the null function it follows that $k^2 a_k = -f_k$, yielding the solution

$$u(x) = -\sum_{k=1}^{\infty} \frac{f_k}{k^2} \sin(kx), \quad x \in (0, \pi),$$

which is indeed uniformly convergent. \hfill \Box

**Example 8.2** Consider the two-point Neumann boundary value problem

$$\mathcal{L}[u] := \frac{d^2 u}{dx^2} = f(x), \quad x \in (0, \pi),$$

$$\frac{du}{dx}(0) = \frac{du}{dx}(1) = 0,$$

where $f$ and $f'$ are both piecewise smooth on $(0, \pi)$. Because of the form of the boundary conditions, it seems advantageous to expand the solution $u(x)$ in a Fourier cosine series, i.e.

$$u(x) = b_0 + \sum_{k=1}^{\infty} b_k \cos(kx).$$

If $u'$ converges uniformly (to be verified afterwards), the boundary conditions are automatically satisfied. Likewise we have

$$f(x) = f_0 + \sum_{k=1}^{\infty} f_k \cos(kx),$$

with coefficients $f_k$ given by

$$f_0 = \frac{1}{\pi} \int_0^{\pi} f(x) \, dx, \quad f_k = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(kx) \, dx, \quad k = 1, 2, \ldots$$

Note that $f_k = O(k^{-1})$ for $k \to \infty$ (see Corollary 3.7). Substituting these expansions into the differential equation, we find the relation

$$\sum_{k=0}^{\infty} \left(k^2 b_k + f_k\right) \cos(kx) = 0.$$
so that \( k^2 b_k + f_k = 0 \). In particular \( f_0 = 0 \), implying that \( f(x) \) should satisfy the consistency condition

\[
\int_0^\pi f(x) \, dx = 0,
\]

stating that the average value of \( f(x) \) over \((0, \pi)\) vanishes. As a consequence \( b_0 \) is undetermined and the solution reads:

\[
u(x) = b_0 - \sum_{k=1}^{\infty} \frac{f_k}{k^2} \cos(kx),
\]

with \( b_0 \) arbitrary. Thus \( u(x) \) is determined up to an additive constant. Since \( f_k/k = O(k^{-2}) \), it is readily verified that \( u' \) converges uniformly and satisfies the boundary conditions. \( \square \)

Example 8.3 Consider the two-point Helmholtz boundary value problem

\[
\mathcal{L}[u] := \frac{d^2 u}{dx^2} - \lambda u = 0, \quad x \in (0, \pi),
\]

\[
u(0) = u(\pi) = 0.
\]

If we solve this linear ordinary differential equation in a formal way, we look for solutions of the form \( u(x) = e^{\mu x} \). Substituting this solution into the differential equation, we find that \( \mu \) should satisfy the characteristic equation

\[
\mu^2 - \lambda = 0.
\]

So as general solution we obtain

\[
u(x) = A e^{\sqrt{\lambda} x} + B e^{-\sqrt{\lambda} x}, \quad A, B \in \mathbb{R}.
\]

Using the boundary conditions, we find

\[
A + B = 0, \quad A e^{\sqrt{\lambda} \pi} + B e^{-\sqrt{\lambda} \pi} = 0.
\]

For arbitrary \( \lambda \) this system has only the trivial solution \( A = B = 0 \). Nontrivial solutions (the eigenvalue problem) exist if its determinant is zero, i.e.

\[
\begin{vmatrix}
1 & 1 \\
e^{\sqrt{\lambda} \pi} & e^{-\sqrt{\lambda} \pi}
\end{vmatrix} = e^{-\sqrt{\lambda} \pi} - e^{\sqrt{\lambda} \pi} = 0.
\]

This relation implies that \( e^{\sqrt{\lambda} \pi} = 1 \), which has the solutions \( 2\sqrt{\lambda} \pi = k 2\pi i \) \((k = 0, 1, 2, \ldots)\). Apparently, the only possible values of \( \lambda \) that allow solutions are given by

\[
\lambda = \lambda_k = -k^2, \quad k = 0, 1, 2, \ldots.
\]

Choosing \( A = -B = -\frac{1}{2} i \), we find the corresponding solutions

\[
u(x) = u_k(x) = \sin(kx), \quad k = 0, 1, 2, \ldots.
\]

So we have either no, or infinitely many solutions. Note that \( k = 0 \) corresponds to the trivial solution \( u_0(x) \equiv 0 \) and should therefore be discarded. \( \square \)
1.2 Uniqueness

One can simply investigate uniqueness of the Laplace equation (1.1), for a solution satisfying either one of the boundary conditions (1.6). This is done in the next theorem.

**Theorem 8.4.** A harmonic function, satisfying the Dirichlet boundary condition (1.6a) is unique. A harmonic function satisfying the Neumann boundary condition (1.6b) is unique but an additive constant. If \( \text{sign}(\alpha) = \text{sign}(\beta) \) then a harmonic function satisfying the Robin boundary condition (1.6c) is unique.

**Proof.** First consider the boundary condition (1.6a) with \( a(x) \equiv 0 \). Using the first identity of Green (J.15) we obtain

\[
\int_{\Omega} \left( u \nabla^2 u + \nabla u \cdot \nabla u \right) \, dV = \oint_{\partial \Omega} u \frac{\partial u}{\partial n} \, dS,
\]

so that

\[
\int_{\Omega} |\nabla u|^2 \, dV = 0,
\]

whence \( u \) is constant in \( \Omega \). Because of continuity we conclude that \( u(x) \equiv 0 \). If we now would have two harmonic solutions, \( u_1(x) \) and \( u_2(x) \) say, both satisfying (1.6a), then obviously \( u_1(x) - u_2(x) \equiv 0 \) in \( \Omega \), which implies uniqueness. For the boundary condition (1.6b) with \( b(x) \equiv 0 \) we derive in a similar manner (*) for a harmonic function \( u(x) \). This trivially implies \( u(x) \) to be constant in \( \Omega \). If \( u_1(x) \) and \( u_2(x) \) are two harmonic functions satisfying the boundary condition (1.6b) then \( u_1(x) - u_2(x) \) can be identified with \( u(x) \) as before. Hence a solution is unique apart from an additive constant. Finally, for a harmonic function \( u(x) \) satisfying the boundary condition (1.6c) we obtain

\[
\int_{\Omega} |\nabla u|^2 \, dV = - \oint_{\partial \Omega} \frac{\alpha}{\beta} u^2 \, dS \leq 0.
\]

Obviously this can only be true if \( u(x) \equiv 0 \). Uniqueness then follows in the same fashion as for the first case.

2 Eigenvalues and eigenfunctions

Eigenvalue problems play an important rôle, either directly, e.g. the Helmholtz equation or indirectly, when determining the character of the partial differential equation. In this section we mainly aim at the latter aspect. We first consider the one-dimensional case.
2. EIGENVALUES AND EIGENFUNCTIONS

2.1 The 1-D eigenvalue problem

The simplest eigenvalue problem reads

\[ \frac{d^2u}{dx^2} = \lambda u, \quad 0 < x < a, \quad (2.1a) \]
\[ u(0) = u(a) = 0. \quad (2.1b) \]

Note that equation (2.1a) is just the one-dimensional Helmholtz equation. The parameter \( \lambda \) is called an eigenvalue, and its value is determined by the condition that (2.1) should have a nontrivial solution. The general solution of equation (2.1a) can be written as,

\[ u(x) = A e^{\sqrt{\lambda}x} + B e^{-\sqrt{\lambda}x}, \quad (2.2) \]

where \( A \) and \( B \) follow from the boundary conditions (2.1b). Applying the boundary conditions to the solution (2.2), we find the equations

\[ A + B = 0, \quad A e^{\sqrt{\lambda}a} + B e^{-\sqrt{\lambda}a} = 0. \quad (2.3) \]

This homogeneous system has only a nontrivial solution, if its determinant is zero, i.e.

\[ \begin{vmatrix} 1 & 1 \\ e^{\sqrt{\lambda}a} & e^{-\sqrt{\lambda}a} \end{vmatrix} = e^{-\sqrt{\lambda}a} - e^{\sqrt{\lambda}a} = 0. \quad (2.4) \]

From (2.4), we conclude that \( e^{2\sqrt{\lambda}a} = 1 = e^{2k\pi i} \), and the solutions to this equation read

\[ \lambda = \lambda_k = -\left(\frac{k\pi}{a}\right)^2, \quad k = 0, 1, 2, \ldots. \quad (2.5) \]

Taking into account that \( B = -A \), we find for the corresponding eigenfunctions

\[ u(x) = u_k(x) = \sin\left(\frac{k\pi x}{a}\right), \quad k = 0, 1, 2, \ldots. \quad (2.6) \]

Any multiple of \( u_k(x) \) is an eigenfunction as well. Note that \( u_0(x) \equiv 0 \) and therefore \( \lambda_0 \) and \( u_0(x) \) should be discarded as non-trivial solutions.

As in Chapter 3 we can define an inner product

\[ (f, g) := \int_0^a f(x)g(x) \, dx. \quad (2.7) \]

It is simple to see then that \( \langle u_k, u_l \rangle = 0 \) for \( k \neq l \). In other words, the eigenfunctions \( u_k \) are orthogonal w.r.t. the inner product (2.7). We can put this in a more formal setting by considering the operator

\[ \mathcal{L}[u] := \frac{d^2}{dx^2}u. \quad (2.8) \]

If we apply partial integration twice, we find for a suitable “test function” \( v = v(x) \) with \( v(0) = v(a) = 0 \) the relation

\[ (\mathcal{L}[u], v) = \int_0^a \frac{d^2u}{dx^2}(x)v(x) \, dx = \int_0^a u(x)\frac{d^2v}{dx^2}(x) \, dx = (u, \mathcal{L}[v]). \quad (2.9) \]
CHAPTER 8. THE ANALYSIS OF ELLIPTIC EQUATIONS

This is a special property of this operator. More specifically, we define the adjoint operator $A^*$ of an operator $A$ by (cf. Chapter 4)

$$(u, A[v]) = (A^*[u], v).$$

Then we see from (2.9) that $L$ defined in (2.8) has a symmetry property i.e.

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

(2.11)

We call this $L$ therefore self-adjoint. More generally, if we have the operator $L$ defined by

$$L[u] := \frac{d}{dx}(p(x) \frac{du}{dx}) + q(x)u(x),$$

(2.12)

then $L$ is self-adjoint. We leave it to an exercise to see that an operator with an explicit first derivative term is not self-adjoint. There is a remarkable similarity with symmetric matrices. Indeed, symmetric matrices have an orthogonal system of eigenvectors, which correspond to real eigenvalues. This similarity is often exploited in numerical approaches, which try to preserve the self-adjointness of (2.12) in a discrete form, see also Chapter ??.

We can also derive the following (cf. (2.9)). Let $\lambda$ be an eigenvalue of (2.12) and $u$ a corresponding eigenfunction. Then

$$\lambda(u, u) = (L[u], u) = -\int_0^a p(x) (\frac{du}{dx})^2 dx + \int_0^a q(x)u^2 dx.$$  

(2.13)

Hence $\lambda$ is real, which is in agreement with the eigenvalues of (2.1) and is in line with what we know for symmetric matrices. Moreover, we see that $\lambda < 0$ for $p > 0$, and $q \leq 0$. With these requirements on $p$ and $q$ we have found the analogue of what we call definite negative symmetric matrices (a matrix $C$ is negative definite if $z^T C z < 0$ for any $z > 0$).

Remark: Sometimes one rather prefers to use the operator $-L$, in order to have strictly positive eigenvalues.

Example 8.5 Consider the operator $L$, with

$$L[u] := \frac{d^2u}{dx^2} + qu, \quad q \in \mathbb{R},$$

and let $u(0) = u(a) = 0$. If we try and find the eigenvalues as we did for (2.1), we obtain

$$\lambda_k = -\left(\frac{k \pi}{a}\right)^2 + q, \quad k = 1, 2, \ldots .$$

Clearly all eigenvalues are negative if $q < (\pi/a)^2$, which is slightly more relaxed than the requirement $q \leq 0$ as we used before.

Example 8.6 If we have purely Neumann boundary conditions, we get eigensolutions different from (2.5) and (2.6). Thus consider the boundary value problem

$$\frac{d^2u}{dx^2} = \lambda u, \quad 0 < x < a,$$

$$\frac{du}{dx}(0) = \frac{du}{dx}(a) = 0.$$
2. EIGENVALUES AND EIGENFUNCTIONS

Analogously to the derivation of the eigensolution (2.5) and (2.6), we find a general solution of the form (2.2), where the coefficients $A$ and $B$ have to satisfy the equations

$$A - B = 0, \quad A e^{\sqrt{\lambda} a} - B e^{-\sqrt{\lambda} a} = 0.$$ 

We again find the eigenvalues

$$\lambda_k = -\left(\frac{k \pi}{a}\right)^2, \quad k = 0, 1, 2, \ldots,$$

but now corresponding to the eigenfunctions read

$$u_k(x) = \cos \left(\frac{k \pi x}{a}\right), \quad k = 0, 1, 2, \ldots.$$ 

Note that $u_0(x) \equiv 1$ corresponds to a constant which can be added to the solution of the Neumann problem.

There is a close relationship between eigenfunctions and Fourier series (cf. Chapter 3). If we would expand a function $f$ with $f(0) = f(a) = 0$ on $[0, a]$, then we obtain the Fourier sine series

$$f(x) = \sum_{k=1}^{\infty} f_k \sin \left(\frac{k \pi x}{a}\right), \quad (2.14)$$

so $f(x)$ is actually expanded in terms of eigenfunctions of (2.1). This property holds more generally for eigenfunctions and has a host of consequences. An appropriate setting for this is variational calculus, which is outside the scope of this text, however.

2.2 Eigenvalue problems in more dimensions

The analysis of eigenvalue problems can be extended to higher dimensions. The starting point is a self-adjoint generalization of the Helmholtz equation defined on a domain $\Omega$ with homogeneous Dirichlet boundary condition, i.e.

$$\mathcal{L}[u] := \nabla^2 u + qu = \lambda u, \quad x \in \Omega, \quad (2.15a)$$

$$u(x) = 0, \quad x \in \partial \Omega. \quad (2.15b)$$

We first introduce an inner product on $\Omega$. Let $u$ and $v$ be defined on $\Omega$ and satisfy the homogeneous Dirichlet boundary condition as in (2.15b), then

$$(u, v) := \int_{\Omega} u v \, dV. \quad (2.16)$$

Our first goal is to show that $\mathcal{L}$ is self-adjoint. We have, using the second identity of Green,

$$(\mathcal{L}[u], v) = \int_{\Omega} (\nabla^2 u + qu) v \, dV = \int_{\Omega} u (\nabla^2 v + q v) \, dV = (u, \mathcal{L}[v]) = (\mathcal{L}^*[u], v). \quad (2.17)$$
CHAPTER 8. THE ANALYSIS OF ELLIPTIC EQUATIONS

Note that we have used that \( v = u = 0 \) on \( \partial \Omega \). If we now let \( u \) be an eigensolution corresponding to \( \lambda \) (presupposing its existence), then we find by Green’s first identity that

\[
\lambda (u, u) = (L[u], u) = \int_{\Omega} (\nabla^2 u + qu) u \, dV = -\int_{\Omega} (|\nabla u|^2 - qu^2) \, dV. \tag{2.18}
\]

Hence we see that \( \lambda \in \mathbb{R} \). Moreover, if \( q < 0 \), then \( \lambda < 0 \).

Now let \( u_k \) and \( u_l \) be eigenfunctions corresponding to the eigenvalues \( \lambda_k \) and \( \lambda_l \), respectively, with \( \lambda_k \neq \lambda_l \), then

\[
u_l u_k \Delta u_k - u_k \Delta u_l = u_l \nabla^2 u_k - u_k \nabla^2 u_l = (\lambda_k - \lambda_l) u_k u_l. \tag{2.19}\]

After integration we obtain from (2.19)

\[
\int_{\Omega} (u_l \nabla^2 u_k - u_k \nabla^2 u_l) \, dV = (\lambda_k - \lambda_l) \int_{\Omega} u_k u_l \, dV. \tag{2.20}
\]

For the left hand side in (2.20), we can apply Green’s second identity, which then results in

\[
\oint_{\partial \Omega} \left( u_l \frac{\partial u_k}{\partial n} - u_k \frac{\partial u_l}{\partial n} \right) \, dS = (\lambda_k - \lambda_l) \int_{\Omega} u_k u_l \, dV. \tag{2.21}
\]

Clearly the integral on the left in (2.21) is zero, from which we derive

\[
(u_k, u_l) = \int_{\Omega} u_k u_l \, dV = 0, \tag{2.22}
\]

i.e. the eigensolutions form an orthogonal set.

For a general domain \( \Omega \), one has to use numerical methods to find (approximation of) eigenvalues. If we have a rectangular domain, say \( \Omega = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq a, 0 \leq y \leq b\} \), we can simply obtain the eigensolutions from what we found for the one-dimensional case when \( q(x) = 0 \). It is left as an exercise to show that

\[
\lambda_{k,l} = -\pi^2 \left( \frac{k^2}{a^2} + \frac{l^2}{b^2} \right), \tag{2.23a}
\]

\[
u_{k,l}(x, y) = \sin \left( \frac{k\pi x}{a} \right) \sin \left( \frac{l\pi y}{b} \right), \quad (k, l = 1, 2, \ldots) \tag{2.23b}
\]

are eigenvalues and eigenfunctions, respectively. We conclude this section with a general property which we give without proof (see e.g. [30]).

**Property 8.7.** Let the area of a 2-dimensional domain be \( A \). Then the eigenvalues of the Laplace equation can be ordered such that for the \( m \)-th eigenvalue, \( \lambda_m \) say, we have

\[
\lambda_m \sim -\frac{4\pi m}{A}. \tag{2.24}
\]
Example 8.8 Consider a Dirichlet problem on the unit square \( \Omega := \{(x, y) \in \mathbb{R}^2 \mid 0 < x < 1, 0 < y < 1\} \). From (2.23a) we see that \( \lambda_{k,l} = -\pi^2 (k^2 + l^2) \). Let \( m(\lambda) \) be the number of eigenvalues, still larger than a fixed negative \( \lambda \), or equivalently, the number of index pairs \((k, l)\) satisfying \( k^2 + l^2 < |\lambda|/\pi^2 \). We can relabel the eigenvalues \( \{\lambda_{k,l}\} \) as \( \{\mu_m\} \) by letting \( \lambda_{k,l} \) gradually decrease starting from \( \lambda_{1,1} \). In Figure 8.1 we have indicated the set of index pairs \( \{(k, l)\} \) which lie in the first quadrant of a circle with radius \( \frac{1}{\pi} \sqrt{|\lambda|} \). The area of this quarter circle is equal to \( |\lambda|/4\pi \). Hence,

\[
m(\lambda) \sim \frac{|\lambda|}{4\pi}.
\]

Reverting the argument, we obtain \( \mu_m \sim -4\pi m \), in agreement with (2.24).

![Figure 8.1](image.png)

**Figure 8.1.** The eigenvalues contained in first quadrant of the circle \( k^2 + l^2 < |\lambda|/\pi \).

3 Separation of variables

The method of separation of variables is useful for linear problems with constant coefficients and homogeneous boundary conditions. In this section we will apply the method to the two-dimensional Laplace equation (1.1). It is based on the assumption that the solution \( u = u(x, y) \) can be written as a product of a function \( v \), say, depending solely on \( x \) and a function \( w \), say, depending solely on \( y \), i.e.

\[
u(x, y) = v(x)w(y).
\]

If we substitute this in (1.1) we obtain

\[
\frac{d^2 v}{dx^2}(x)w(y) = -v(x)\frac{d^2 w}{dy^2}(y).
\]
CHAPTER 8. THE ANALYSIS OF ELLIPTIC EQUATIONS

Dividing by \(v(x)w(y)\) we can write this as

\[
\frac{1}{v(x)} \frac{d^2v}{dx^2}(x) = -\frac{1}{w(y)} \frac{d^2w}{dy^2}(y). \tag{3.2}
\]

Since the left hand side in (3.2) is a function of \(x\) only and the right hand side of \(y\) only, they must be independent of both, i.e. constant. Let us denote this constant by \(\lambda\) then we apparently have the two eigenvalue problems

\[
\frac{d^2v}{dx^2}(x) = \lambda v(x), \tag{3.3a}
\]
\[
\frac{d^2w}{dy^2}(y) = -\lambda w(y). \tag{3.3b}
\]

The constant \(\lambda\) is called the separation constant. Let us now assume that \(\Omega\) is the unit square, i.e. \(0 < x, y < 1\), and that e.g. the following BC are given

\[
u(0, y) = u(1, y) = 0, \quad u(x, 0) = u(x, 1) = x(1 - x). \tag{3.4}\]

From the homogeneous boundary conditions we conclude that \(v(0) = v(1) = 0\) and so (3.3a) is a genuine eigenvalue problem. We then find that the eigenvalues \(\lambda_k\) are given by

\[
\lambda_k = -k^2\pi^2, \quad k = 1, 2, \ldots, \tag{3.5a}
\]

and the corresponding eigenfunctions by

\[
v_k(x) = \sin(k\pi x), \quad k = 1, 2, \ldots; \tag{3.5b}
\]

cf. Section 2.1. The functions \(w(y) = w_k(y)\) corresponding to \(\lambda_k\) can be determined from the ODE (3.3b) and we find

\[
w_k(y) = \alpha_k e^{k\pi y} + \beta_k e^{-k\pi y}, \quad k = 1, 2, \ldots, \tag{3.6}
\]

for some \(\alpha_k, \beta_k\). In order to determine the desired solution \(u(x, y)\), we apply the superposition principle, i.e. we assume that

\[
u(x, t) = \sum_{k=1}^{\infty} v_k(x)w_k(y) = \sum_{k=1}^{\infty} (\alpha_k e^{k\pi y} + \beta_k e^{-k\pi y}) \sin(k\pi x). \tag{3.7}
\]

This superposition is possible, since the PDE is linear. Note that (3.7) is in fact a Fourier-sine series in the \(x\)-variable, so that the homogeneous boundary conditions are automatically satisfied if the series converges uniformly. This is to be verified afterwards. The coefficients \(\alpha_k, \beta_k\) follow from the remaining boundary conditions. First we have to find a Fourier-sine series for the function \(x(x - 1)\). We obtain

\[
x(x - 1) = \sum_{k=1}^{\infty} \gamma_k \sin(k\pi x), \tag{3.8a}
\]
3. SEPARATION OF VARIABLES

where

\[
\gamma_k = \begin{cases} 
0, & \text{for } k \text{ even,} \\
-\frac{8}{(k\pi)^2}, & \text{for } k \text{ odd.}
\end{cases}
\]  

(3.8b)

The coefficients now follow from comparing (3.8a) and (3.7) for \(x = 0\) and \(x = 1\). This gives the set of equations

\[
\begin{align*}
\alpha_k + \beta_k &= \gamma_k, \\
\alpha_k e^{k\pi} + \beta e^{-k\pi} &= \gamma_k.
\end{align*}
\]

(3.9a, 3.9b)

With some straightforward arithmetic one finds from (3.9) the coefficients resulting in the final (indeed uniformly converging) series solution

\[
u(x, y) = \sum_{k=1}^{\infty} \gamma_k \sin(k\pi x) \frac{\cosh(k\pi (y - \frac{1}{2}))}{\cosh(\frac{3}{2}k\pi)}.
\]  

(3.10)

The method of separation of variables apparently requires knowledge of the eigenvalues and eigenfunction of the separate systems. This means that the boundary conditions should allow the separation to result in homogeneous boundary conditions for either one of the subproblems. The method is not restricted to cartesian coordinates. Indeed, consider the PDE

\[
\nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} = 0,
\]  

(3.11)

defined on a finite wedge \(\{(r, \phi) : 0 < r < R, 0 < \phi < \Phi\}\), where \(u\) satisfies the BC

\[
u(r, 0) = u(r, \Phi) = 0, \quad u(R, \phi) = g(\phi),
\]  

(3.12)

for some \(g(\phi)\). Writing

\[
u(r, \phi) = v(r)w(\phi),
\]

we obtain from (3.11)

\[
\begin{align*}
\frac{r^2}{v} \frac{d^2 v}{dr^2} + \frac{r}{v} \frac{dv}{dr} &= -\frac{1}{w} \frac{d^2 w}{d\phi^2} = -\lambda.
\end{align*}
\]  

(3.13)

We now have the two eigenvalue problems

\[
\frac{d^2 w}{d\phi^2} = \lambda w, \quad (3.14a)
\]

\[
\frac{r^2}{v} \frac{d^2 v}{dr^2} + \frac{r}{v} \frac{dv}{dr} = -\lambda v. \quad (3.14b)
\]

From the homogeneous boundary conditions in (3.12) we see that \(w(0) = w(\Phi) = 0\) and consequently \(w\) is the solution of a genuine eigenproblem, with the typical eigenvalues

\[
\lambda_k = -\omega_k^2, \quad \omega_k := \frac{k\pi}{\Phi}, \quad (k = 1, 2, \ldots),
\]

(3.15)
and corresponding eigenfunctions

\[ u_k(\phi) = \sin(\omega_k \phi), \quad (k = 1, 2, \ldots). \]  

(3.16)

For each \( \lambda_k \) the equation for \( v \), can now be solved by simple substitution of this \( \lambda_k \) in (3.14b). Substituting a solution of the form \( v(r) = r^{\mu} \), we can easily see that \( v_k(r) \) has the form

\[ v(r) = \alpha_k r^{\omega_k} + \beta_k r^{-\omega_k}. \]  

(3.17)

By requiring \( v_k \) to be bounded we see that \( \beta_k = 0 \). The resulting solution of the boundary value problem (3.11), (3.12) is then given by

\[ u(r, \phi) = \sum_{k=1}^{\infty} \alpha_k r^{\omega_k} \sin(\omega_k \phi). \]  

(3.18)

Like in the previous case, if we have a Fourier-sine expansion of the function \( g(\phi) \) we can formally determine the coefficients \( \alpha_k \).

**Example 8.9** Denote by \( \gamma_k \) the coefficients of the Fourier-sine expansion of \( g(\phi) \), then we have

\[ \gamma_k = \frac{2}{\Phi} \int_{0}^{\Phi} g(\phi) \sin(\omega_k \phi) d\phi. \]

Hence from the boundary condition at \( r = R \) we obtain

\[ \alpha_k = \gamma_k R^{\omega_k}, \]

so the final solution reads

\[ u(r, \phi) = \sum_{k=1}^{\infty} \gamma_k \left( \frac{r}{R} \right)^{\omega_k} \sin(\omega_k \phi). \]

This result is interesting as it shows the smoothness of the solution on a wedge in the neighbourhood of the corner point. Indeed, if we consider e.g. \( \frac{\partial u}{\partial r} \) we obtain

\[ \frac{\partial u}{\partial r} = \frac{1}{R} \sum_{k=1}^{\infty} \gamma_k \left( \frac{r}{R} \right)^{\omega_k-1} \sin(\omega_k \phi). \]

We see that already for \( \Phi > \pi \) the first term \( (k = 1) \) is not bounded. The conclusion, more generally therefore is that corners in a domain imply less smoothness. In particular for re-entrant corners (i.e. those with angles larger than \( \pi \) this already holds true for the first derivative. This corner problem has, of course, consequences when solving a problem numerically.

4 Fundamental solutions

Before looking at general Poisson problems, as we shall do in Section 5, it is meaningful to investigate the fundamental solution \( w \) of the Poisson equation, i.e. the solution of the equation equation

\[ \nabla^2 w(x; \xi) = \delta(x - \xi). \]  

(4.1)
We look for symmetry solutions, depending on the distance \( r = \|x - \xi\|_2 \) only; i.e. circular or spherically symmetric solutions in the two-dimensional or three-dimensional space, respectively. This simplifies the problem to an ordinary differential equation. Indeed, let us denote this solution by \( \tilde{w}(r) \), then it satisfies the differential equation

\[
\nabla^2 \tilde{w} = \frac{d^2 \tilde{w}}{dr^2} + \frac{d - 1}{r} \frac{d \tilde{w}}{dr} = \delta(r), \quad (d = 2, 3).
\]  (4.2)

Clearly, (4.2) has the general solution, for \( r > 0 \)

\[
\tilde{w}(r) = \begin{cases} 
A \ln r + B & \text{if } d = 2, \\
A + B & \text{if } d = 3,
\end{cases}
\]  (4.3)

with \( A, B \in \mathbb{R} \). For \( d = 3 \) we may choose \( \tilde{w}(r) \to 0 \), for \( r \to \infty \), implying that \( B = 0 \).

We shall also choose \( B = 0 \) for \( d = 2 \). Therefore we are left to find \( A \), such that (4.2) is satisfied altogether. To this end we define a ball \( B(0; \rho) \) around \( 0 \) with radius \( \rho \) and denote by \( \partial B(0; \rho) \) its sphere. From Gauss’ theorem, we then formally obtain

\[
\int_{B(0; \rho)} \nabla^2 \tilde{w}(r) \, dV = \oint_{\partial B(0; \rho)} \frac{\partial \tilde{w}}{\partial n}(r) \, dS = \oint_{\partial B(0; \rho)} \frac{d \tilde{w}}{dr}(r) \, dS.
\]  (4.4)

Since \( \int_{B(0; \rho)} \nabla^2 \tilde{w}(r) \, dV = 1 \), we can determine \( A \). For \( d = 2 \) we find

\[
\oint_{\partial B(0; \rho)} \frac{d \tilde{w}}{dr}(r) \, dS = \oint_{\partial B(0; \rho)} \frac{A}{\rho} \, dS = 2\pi A,
\]

from which we conclude that \( A = 1/(2\pi) \). Likewise, for \( d = 3 \) we obtain

\[
\oint_{\partial B(0; \rho)} \frac{d \tilde{w}}{dr}(r) \, dS = \oint_{\partial B(0; \rho)} \frac{-A}{\rho^2} \, dS = -4\pi A,
\]

so that in this case \( A = -1/(4\pi) \). Using this in (4.3) we obtain for the fundamental solution

\[
w(x; \xi) = \begin{cases} 
\frac{1}{2\pi} \ln \|x - \xi\|_2 & \text{if } d = 2, \\
-\frac{1}{4\pi \|x - \xi\|_2} & \text{if } d = 3.
\end{cases}
\]  (4.5)

We leave it to an exercise to show that \( w(x; \xi) \) is the (weak) solution of (4.1).

**Example 8.10** Consider the scalar problem

\[
\mathcal{L}[u] := \frac{d^2 u}{dx^2} = 0.
\]

We like to find a fundamental solution \( w(x; \xi) \) satisfying

\[
\frac{d^2 w}{dx^2}(x; \xi) = \delta(x - \xi).
\]
Since the general solution of the ordinary differential equation for \( x \neq \xi \) can be written as \( A + Bx \), we may take as ansatz for \( w \)

\[
w(x; \xi) = \begin{cases} 
A_1 + B_1 x & \text{if } x < \xi, \\
A_2 + B_2 x & \text{if } x > \xi.
\end{cases}
\]

Using integration, we obtain for sufficiently small \( \rho \)

\[
\frac{dw}{dx}(\xi + \rho; \xi) - \frac{dw}{dx}(\xi - \rho; \xi) = \int_{\xi - \rho}^{\xi + \rho} \delta(x - \xi) \, dx = 1.
\]

Applying this condition to the solution above, we find

\[
B_2 - B_1 = 1.
\]

Since \( w \) apparently is continuous at \( x = \xi \), we also find

\[
A_1 + B_1 \xi = A_2 + B_2 \xi.
\]

Hence we obtain the fundamental solution

\[
w(x; \xi) = \begin{cases} 
A_1 + B_1 x & \text{if } x \leq \xi, \\
A_1 - \xi + (B_1 + 1)x & \text{if } x > \xi,
\end{cases}
\]

where \( A_1 \) and \( B_1 \) are arbitrary constants. \( \square \)

Of course the fundamental solution relates to the (homogeneous) operator only. In the next section we shall construct solutions of (4.1) which satisfy the homogeneous form of the boundary condition as well. It is constructive to illustrate this already here for the one-dimensional case; see next example.

**Example 8.11** Consider the fundamental solution derived in Example 8.5. We now like to solve the boundary value problem

\[
\mathcal{L}[w](x; \xi) := \frac{d^2 w}{dx^2}(x; \xi) = \delta(x - \xi), \quad x \in (0, 1),
\]

\[
w(0; \xi) = w(1; \xi) = 0.
\]

Applying the boundary conditions to the general form in Example 8.5 we find

\[
A_1 = 0, \quad A_1 - \xi + B_1 + 1 = 0.
\]

Therefore, the fundamental solution is given by

\[
w(x, \xi) = \begin{cases} 
x(\xi - 1) & \text{if } x \leq \xi, \\
(x - 1)\xi & \text{if } x > \xi.
\end{cases}
\]

\( \square \)
Consider the Poisson problem (1.2). We can formally write the right hand side \( f(x) \) as
\[
f(x) = \int_\Omega \delta(x - \xi) f(\xi) \, dV_\xi, \quad x \in \Omega,
\]
where the subscript \( \xi \) denotes integration with respect to \( \xi \). Using this representation of \( f(x) \), we can derive a particular solution \( u_p(x) \) in the following way. We multiply (4.1) with \( f(\xi) \) and integrate with respect to \( \xi \) over the domain \( \Omega \). This way we find
\[
u_p(x) = \int_\Omega w(x; \xi) f(\xi) \, dV_\xi.
\]
Inserting the fundamental solution (4.5) we obtain
\[
u_p(x) = \begin{cases} 
\frac{1}{2\pi} \int_\Omega \ln \|x - \xi\|_2 f(\xi) \, dV_\xi, & \text{if } d = 2, \\
-\frac{1}{4\pi} \int_\Omega \frac{1}{\|x - \xi\|_2^2} f(\xi) \, dV_\xi, & \text{if } d = 3.
\end{cases}
\]
In general \( u_p(x) \) does not satisfy the prescribed boundary condition (see (1.6)). However, it is clear that \( u(x) - u_p(x) \) satisfies the Laplace equation (1.1). Hence there exists a harmonic function \( u_h(x) \) such that
\[
u(x) = \nu_p(x) + u_h(x).
\]
This is called superposition. The harmonic function \( u_h(x) \) has to be determined such that \( u(x) \) satisfies the prescribed boundary condition. Suppose, we have a Dirichlet problem, then it is obvious that \( u_h(x) \) is the solution of the boundary value problem
\[
\nabla^2 u_h = 0, \quad x \in \Omega, \tag{5.5a}
\]
\[
u_h(x) = u(x) - u_p(x), \quad x \in \partial \Omega. \tag{5.5b}
\]
In a similar way, we can derive a solution for a Neumann or Robin boundary value problem.

**Example 8.12** Consider a two-dimensional Dirichlet problem on the half space \( \Omega := \{ x, y \in \mathbb{R}^2 \mid y > 0 \} \) (Fig. 8.2). The fundamental solution (4.5), and consequently also the particular solution (5.3), does not satisfy the homogeneous boundary condition on the line \( y = 0 \). In order to overcome this problem, we have to modify the fundamental solution. This can be done in the following way. Let \( \xi = (\xi, \eta) \) be an arbitrary point in \( \Omega \). Then we take a mirror point \( \xi^* := (\xi, -\eta) \) with respect to the line \( y = 0 \) and modify the fundamental solution as follows
\[
w(x; \xi) = \frac{1}{2\pi} \ln \|x - \xi\|_2 - \frac{1}{2\pi} \ln \|x - \xi^*\|_2 = \frac{1}{2\pi} \ln \left( \frac{\|x - \xi\|_2}{\|x - \xi^*\|_2} \right).
\]
It is obvious that this fundamental solution satisfies the homogeneous boundary condition \( w(x; \xi) = 0 \) on the line \( y = 0 \). Moreover, since \( \xi^* \notin \Omega \), \( w(x; \xi) \) is also a solution of (4.1). The solution of the Dirichlet problem is given by
\[
u(x) = \frac{1}{2\pi} \int_\Omega \ln \left( \frac{\|x - \xi\|_2}{\|x - \xi^*\|_2} \right) f(\xi) \, dV_\xi.
\]
Another way to use this superposition principle is to find a particular solution $u_p(x)$ satisfying the inhomogeneous equation however with homogeneous boundary conditions and a harmonic function $u_h(x)$ satisfying the Laplace equation with the appropriate boundary conditions. This leads to the introduction of the Green’s function $G(x; \xi)$. The Green’s function $G(x; \xi)$ for the Dirichlet problem (1.2) and (1.6a) is by definition the solution of the following boundary value problem

$$\nabla^2 G(x; \xi) = \delta(x - \xi), \quad x \in \Omega, \quad (5.6a)$$

$$G(x; \xi) = 0, \quad x \in \partial\Omega. \quad (5.6b)$$

Using superposition, we immediately see that the Green’s function $G(x; \xi)$ equals the fundamental solution $w(x; \xi)$ (see Section 4) apart from a harmonic function. Using theorem 8.4 we therefore conclude that it is unique. From the second identity of Green

$$\int_{\Omega} (u \nabla^2 v - v \nabla^2 u) \, dV = \oint_{\partial\Omega} \left( u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) \, dS, \quad (5.7)$$

we can find a representation of $u(x)$ as announced. Indeed, if we take $v = G(x; \xi)$, then we obtain

$$u(\xi) = \int_{\Omega} G(x; \xi) f(x) \, dV + \oint_{\partial\Omega} a(x) \frac{\partial G}{\partial n}(x; \xi) \, dS. \quad (5.8)$$

Also for Neumann problems, we can construct a solution by superposition. First we remark that the divergence theorem provides for a constraint, relating the source term $f(x)$ and the boundary function $b(x)$ (see (1.6b)). Indeed, we have

$$\int_{\Omega} f(x) \, dV = \int_{\Omega} \nabla^2 u \, dV = \oint_{\partial\Omega} \frac{\partial u}{\partial n} \, dS = \oint_{\partial\Omega} b(x) \, dS. \quad (5.9)$$

As a candidate for the Green’s function $G(x; \xi)$ we might take the function satisfying equation (4.1) and the homogeneous Neumann boundary condition. However, this choice
of $G(x; \xi)$ obviously does not satisfy the constraint (5.9). Another possibility is to define the Green’s function as the solution of the following boundary value problem

$$\nabla^2 G(x; \xi) = \delta(x - \xi) - \frac{1}{\kappa}, \quad \kappa := \int_{\Omega} dV, \quad x \in \Omega, \quad (5.10a)$$

$$\frac{\partial G}{\partial n}(x; \xi) = 0, \quad x \in \partial\Omega, \quad (5.10b)$$

where the right hand side of (5.10a) has been modified to enforce (5.9). We then obtain from (5.7) with $v = G(x; \xi)$

$$u(\xi) = \int_{\Omega} G(x; \xi) f(x) dV + \frac{1}{\kappa} \int_{\Omega} u(x) dV - \oint_{\partial\Omega} G(x; \xi) b(\xi) dS. \quad (5.11)$$

Firstly, note that for this Green’s function the constraint (5.9) is satisfied. Indeed, we have

$$\oint_{\partial\Omega} \frac{\partial G}{\partial n} (x; \xi) dS = \int_{\Omega} \nabla^2 G(x; \xi) dV = 1 - \frac{1}{\kappa} \int_{\Omega} dV = 0.$$

Secondly, we remark that the first integral in (5.11) represents the particular solution and the last one, the harmonic function. The term $\frac{1}{\kappa} \int_{\Omega} u(x) dV$ is just an additive constant.

Finally, for the boundary value problem with Robin boundary condition (1.6c), we define the Green’s function $G(x; \xi)$ as solution of

$$\nabla^2 G(x; \xi) = \delta(x - \xi), \quad x \in \Omega, \quad (5.12a)$$

$$\alpha G(x; \xi) + \beta \frac{\partial G}{\partial n}(x; \xi) = 0, \quad x \in \partial\Omega. \quad (5.12b)$$

Analogously to the previous derivations, we obtain the following integral representation of $u(\xi)$:

$$u(\xi) = \int_{\Omega} G(x; \xi) f(x) dV - \oint_{\partial\Omega} \left( \frac{\alpha}{\beta} u(x) + \frac{\partial u}{\partial n}(x) \right) G(x; \xi) dS$$

$$= \int_{\Omega} G(x; \xi) f(x) dV - \frac{1}{\beta} \oint_{\partial\Omega} G(x; \xi) c(x) dS, \quad (5.13)$$

provided $\beta \neq 0$.

A particular solution of (5.10a) can be found by subtracting the following solution $v_p(x)$ (which is independent of $\xi$) of the fundamental solution $w(x; \xi)$:

$$v_p(x) = \frac{1}{4\kappa} ||x||_2^2, \quad \text{if } d = 2, \quad (5.14a)$$

$$v_p(x) = \frac{1}{6\kappa} ||x||_2^3, \quad \text{if } d = 3. \quad (5.14b)$$

The Green’s function $G(x; \xi)$ then differs from $w(x; \xi) - v_p(x)$ by a harmonic function, which can be found formally from solving a (homogeneous) Neumann problem with known BC.
CHAPTER 8. THE ANALYSIS OF ELLIPTIC EQUATIONS

Example 8.13 Consider the following Dirichlet boundary value problem on the half space

$$\nabla^2 u = f(x), \quad x \in \Omega := \{(x, y) \in \mathbb{R}^2 | y > 0\},$$

$$u(x, 0) = a(x), \quad u(x, y) \to 0 \quad \text{for} \quad y \to \infty,$$

where the boundary function $a(x) \to 0$ as $|x| \to \infty$. From Example 8.12 we derive for the Green's function with $\xi = (\xi, \eta)$

$$G(x, y; \xi, \eta) = \frac{1}{2\pi} \ln \left( \frac{(x-\xi)^2 + (y-\eta)^2}{(x-\xi)^2 + (y+\eta)^2} \right)^{\frac{1}{2}}.$$

Hence on $\partial$ we find

$$\frac{\partial G}{\partial n}(x, 0; \xi, \eta) = \frac{\eta}{\pi (x-\xi)^2 + \eta^2}.$$

So, formally, the solution of the boundary value problem above is given by

$$u(\xi, \eta) = \frac{1}{2\pi} \int_0^\infty \int_{-\infty}^\infty \ln \left( \frac{(x-\xi)^2 + (y-\eta)^2}{(x-\xi)^2 + (y+\eta)^2} \right)^{\frac{1}{2}} f(x, y) \, dx \, dy$$

$$+ \frac{\eta}{\pi} \int_{-\infty}^\infty a(x) \frac{1}{(x-\xi)^2 + \eta^2} \, dx.$$

The last example shows the power and weakness of an analytical approach. On the one hand, by mere construction, one can show that a solution exists, and one can give estimates for it. On the other hand, the expressions one obtains, are often complicated and usually not directly solvable in closed form.

We would like to point at another important fact. The expressions we have derived in this section for the solutions can be seen as inverting the operator form for the original problem: that is both the equation and the boundary condition. In particular, the Green’s function can be interpreted as (“constrained”) inverse of the Laplace operator. In the next chapter we will investigate numerical methods for such problems, where this “inversion” will be met again, now in terms of matrices.

6 The maximum principle

The Laplace operator has a nice property that gives rise to a number of useful results. If the second derivative of a scalar function $u(x)$ is zero, it is linear. On a finite interval the absolute maximum is attained at either end of the interval. For a Poisson problem this is described in

Property 8.14 (Maximum principle). Let $\nabla^2 u(x) = 0$ for all $x \in \Omega$. Then $u(x)$ satisfies the inequalities:

$$m := \min_{\xi \in \partial\Omega} u(\xi) \leq u(x) \leq \max_{\xi \in \partial\Omega} u(\xi) =: M.$$

(6.1)
6. THE MAXIMUM PRINCIPLE

Proof. Define the function $v_\varepsilon(x) := u(x) + \varepsilon \|x\|^2_2$ for $\varepsilon > 0$. Clearly we have

$$\nabla^2 v_\varepsilon(x) = \nabla^2 u(x) + 2d\varepsilon > 0, \quad x \in \Omega, \quad (d = 2, 3).$$

(\star)

Suppose, $v_\varepsilon(x)$ has a maximum in the interior of $\Omega$, say at $x_0$, then

$$\nabla^2 v_\varepsilon(x_0) \leq 0.$$

Since this contradicts the inequality in (\star), we conclude that $v_\varepsilon$ can only attain its maximum at the boundary $\partial \Omega$. If we denote the latter by $M_\varepsilon$, we derive the required upper bound by letting $\varepsilon \downarrow 0$. The lower bound follows from a similar argument, now using $-u(x)$ and $v_\varepsilon(x) := -u(x) + \varepsilon \|x\|^2_2$ instead.

The theorem above is sometimes also referred to as the minimum-maximum principle. In particular the “maximum principle” can be extended to Poisson problems.

Property 8.15. Let $\nabla^2 u(x) \geq 0$ for all $x \in \Omega$. Then $u(x)$ attains its maximum at the boundary, i.e.

$$u(x) \leq \max_{x \in \partial \Omega} u(\xi).$$

(6.2)

Proof. It is easy to see that the arguments in the proof of property 8.14 still apply for this case.

The latter property is quite powerful in that it gives a possibility to compare solutions of two Poisson problems. The following property is stating this more precisely.

Property 8.16 (Comparison theorem). Consider the two Poisson equations $\nabla^2 u_1(x) = f_1(x)$ and $\nabla^2 u_2(x) = f_2(x)$ with $f_1(x) \geq f_2(x)$ for all $x \in \Omega$, then

$$u_1(x) \leq u_2(x) + \max_{x \in \partial \Omega} (u_1(x) - u_2(x)).$$

(6.3)

Proof. Since $\nabla^2(u_1(x) - u_2(x)) \geq 0$ the result follows directly from property 8.15.

Corollary 8.17. A Dirichlet problem has a unique solution, which depends continuously on the boundary data.

Proof. If there were two solutions then the difference is a harmonic function satisfying the homogeneous boundary condition. We can apply property 8.14 to conclude that this difference must be zero. The continuous dependence is a consequence of the comparison theorem.
7 The Stokes equations

The problems met in practice are often more complex than the ones discussed thus far. A typical example of these are the Navier-Stokes equations in fluid dynamics and their creeping flow simplification, the Stokes equations. Here we shall not dwell on their derivation nor their application. This is done in Chapter 7 and Chapter 8. The first complication we meet is that a flow in one dimension is not of interest and thus a nontrivial formulation requires a Laplacian, operating on a vector. Let $\mathbf{u}$ and $p$ be a velocity and pressure field. Then the Stokes equations for a domain $\Omega \subset \mathbb{R}^d (d = 2, 3)$ read:

$$\nabla^2 \mathbf{u} - \nabla p = 0, \quad (7.1a)$$
$$\nabla \cdot \mathbf{u} = 0. \quad (7.1b)$$

Note that $\nabla^2 \mathbf{u}$ is to be taken componentwise. From (7.1a), it is clear that $\nabla p$, rather than $p$, is a (dependent) variable, so we always need to specify $p$ somewhere in the domain. We shall moreover prescribe the following Dirichlet boundary condition

$$\mathbf{u} \cdot n = b(x), \quad x \in \partial \Omega. \quad (7.2)$$

First consider the case where $b(x) \equiv 0$. We now like to show that the solution of boundary value problem (7.1) and (7.2) is unique. We shall need a bit of vector calculus to do this. First we use the relation (see Appendix J)

$$\nabla \cdot (p \mathbf{u}) = p \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla p. \quad (7.3)$$

Applying the divergence theorem to this relation and taking into account (7.1b) and (7.2) we find

$$\int_{\Omega} \mathbf{u} \cdot \nabla p \, dV = \int_{\Omega} \nabla \cdot (p \mathbf{u}) \, dV = \oint_{\partial \Omega} (p \mathbf{u}) \cdot n \, dS = \oint_{\partial \Omega} p b(x) \, dS. \quad (7.4)$$

Since $b(x) \equiv 0$ we obtain

$$\int_{\Omega} \mathbf{u} \cdot \nabla p \, dV = 0, \quad (7.5)$$

stating that the velocity field and the pressure gradient are orthogonal. Furthermore, we have the following relation for the velocity field $\mathbf{u} := (u, v, w)^T$

$$\nabla \cdot (\mathbf{u} \cdot \nabla \mathbf{u}) = |\nabla \mathbf{u}|^2 + \mathbf{u} \cdot \nabla^2 \mathbf{u}, \quad (7.6)$$

where $\mathbf{u} \cdot \nabla \mathbf{u} := (u \cdot u_x, u \cdot u_y, u \cdot u_z)^T$ and $|\nabla \mathbf{u}|^2 := |\nabla u|^2 + |\nabla v|^2 + |\nabla w|^2$. Integrating this relation yields

$$\int_{\Omega} \nabla \cdot (\mathbf{u} \cdot \nabla \mathbf{u}) \, dV = \int_{\Omega} |\nabla \mathbf{u}|^2 \, dV + \int_{\Omega} \mathbf{u} \cdot \nabla^2 \mathbf{u} \, dV. \quad (7.7)$$

The divergence theorem now implies for the left hand side of (7.7)

$$\int_{\Omega} \nabla \cdot (\mathbf{u} \cdot \nabla \mathbf{u}) \, dV = \oint_{\partial \Omega} (\mathbf{u} \cdot \nabla \mathbf{u}) \cdot n \, dS = 0. \quad (7.8)$$

17:22 7 Oct 2003 141 version: 01-10-2003
The zero value follows from (7.2). Let us now take the inner product of $\nabla^2 u - \nabla p$ and $u$; this trivially gives
\[
\int_{\Omega} u \cdot \nabla^2 u \, dV - \int_{\Omega} u \cdot \nabla p \, dV = 0. \tag{7.9}
\]
The second term in (7.9) is zero on account of (7.5). So the first integral in (7.9) is zero as well. Combining (7.7), (7.8) and (7.9) we therefore conclude that
\[
\int_{\Omega} |\nabla u|^2 \, dV = 0. \tag{7.10}
\]
Clearly $|\nabla u|^2 = 0$ only if $\nabla u = 0$, so $u$ is constant. On account of the boundary condition (7.2), with $b(x) \equiv 0$ we thus conclude that $u(x) \equiv 0$. It is trivial now to see that this implies uniqueness of the general problem (7.1), (7.2).

The linearity of the problem allows us to construct fundamental solutions $u_i$, $p_i$, the so-called Stokeslets; cf. Section 4. They are named after G.G. Stokes but are really first discovered by H.A. Lorentz in 1896; see [76]. In two dimensions they are the solution of

\[
\begin{align*}
\nabla^2 u_i - \nabla p_i &= \delta(x - \xi) e_i, \quad (i = 1, 2), \tag{7.11a} \\
\n\nabla \cdot u_i &= 0. \tag{7.11b}
\end{align*}
\]

Here $e_1 = (1, 0)^T$ and $e_2 = (0, 1)^T$. We have for $u_i = (u_i^1, u_i^2)^T$

\[
\begin{align*}
 u_i^1(x; \xi) &= \frac{1}{4\pi} \left( \delta_{ij} \ln ||x - \xi||^2 - \frac{(x_j - \xi_j)(x_i - \xi_i)}{||x - \xi||^2} \right), \quad (i, j = 1, 2), \\
p_i(x; \xi) &= -\frac{x_i - \xi_i}{2\pi ||x - \xi||^2}. \tag{7.12b}
\end{align*}
\]

These Stokeslets can be used to derive an integral formulation for the Stokes equation; for more details see [?].

We conclude this section with an example of one of the few Stokes problems that can be solved analytically. It regards the flow past a sphere of radius 1. We assume that the velocity $u$ is equal to zero at the sphere and equal to 1 in the direction of the flow (say the $z$-direction) at infinity. The pressure $p$ approaches a limit value there. It will give us an opportunity to show some more vector calculus. To start with, we shall use the spherical coordinates $(r, \theta, \phi)$, and we obtain with $u = (u_r, u_\theta, u_\phi)^T$

\[
\nabla \cdot u = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 u_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta u_\theta) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} u_\phi = 0. \tag{7.13}
\]

We assume axial symmetry, implying that $\frac{\partial}{\partial \phi} = 0$ and $u_\phi = 0$. The construction of the solution now employs the notions of a stream function, $\psi = \psi(r, \theta)$ say. In particular, we require

\[
u_r = \frac{1}{r \sin \theta} \frac{\partial \psi}{\partial \theta}, \quad u_\theta = -\frac{1}{r \sin \theta} \frac{\partial \psi}{\partial r}. \tag{7.14}
\]

Substituting these relations into (7.13) (with $u_\phi = 0$) gives

\[
\frac{1}{r^2 \sin \theta} \frac{\partial^2 \psi}{\partial r \partial \theta} - \frac{1}{r^2 \sin \theta} \frac{\partial^2 \psi}{\partial \theta \partial r} = 0.
\]

17:22 7 Oct 2003 142 version: 01-10-2003
So we have $\nabla \cdot \mathbf{u} = 0$ indeed. The relations in (7.14) imply that the velocity can be written as

$$\mathbf{u} = \nabla \times \tilde{\psi}, \quad \tilde{\psi} := \frac{\psi}{r \sin \theta} \mathbf{e}_\phi.$$  

(7.15)

Next, it is convenient to introduce the vorticity vector $\omega = \nabla \times \mathbf{u}$. Obviously, $\tilde{\psi}$ and $\omega$ are related by

$$\nabla \times (\nabla \times \tilde{\psi}) = \omega.$$  

(7.16)

Applying twice the curl operator to the vector $\tilde{\psi}$ defined in (7.15) results in a vector in the same direction. More precisely, we have

$$\omega = -\frac{1}{r \sin \theta} \mathcal{D}[\psi] \mathbf{e}_\phi,$$  

(7.17)

where the differential operator $\mathcal{D}$ is defined by

$$\mathcal{D}[\psi] := \frac{\partial^2 \psi}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \theta^2} - \cot \theta \frac{\partial \psi}{\partial \theta}.$$  

(7.18)

Using the relation $\nabla^2 \mathbf{u} = \nabla (\nabla \cdot \mathbf{u}) - \nabla \times \omega$ (see Appendix J), we can rewrite (7.1a) as

$$\nabla \times \omega = -\nabla p.$$  

(7.19)

In order to eliminate the pressure gradient, we apply the curl operator to (7.19) and find

$$\nabla \times (\nabla \times \omega) = 0.$$  

(7.20)

Analogous to the derivation of (7.17) we also find

$$\nabla \times (\nabla \times \omega) = \frac{1}{r \sin \theta} \mathcal{D}^2[\psi] \mathbf{e}_\phi.$$  

(7.21)

Hence we conclude from (7.20) that

$$\mathcal{D}^2[\psi] = 0.$$  

(7.22)

At the sphere $\mathbf{u} = 0$, so we obtain as boundary conditions for $\psi$

$$\frac{\partial \psi}{\partial r} = \frac{\partial \psi}{\partial \theta} = 0, \quad \text{at} \quad r = 1,$$  

(7.23)

so apparently $\psi(r, \theta) = \text{Const}$ at $r = 1$ and therefore we may as well take $\psi(1, \theta) = 0$. As for the asymptotic boundary condition we see from Figure 8.3 that $u_r \sim \cos \theta$ and $u_\theta \sim -\sin \theta$. Using this in (7.14) gives $\frac{\partial \psi}{\partial \theta} \sim r^2 \sin \theta \cos \theta$, so

$$\psi(r, \theta) \sim \frac{1}{2} r^2 \sin^2 \theta, \quad \text{for} \quad r \to \infty;$$  

(7.24)

Note that an integration constant is immaterial.

Returning now to (7.22), we apply separation of variables, i.e.

$$\psi(r, \theta) = f(r) g(\theta).$$  

(7.25)
Actually, with some further (tedious) analysis it turns out that $g(\theta)$ may be taken equal to $\sin^2 \theta$. Straightforward computation then reveals

$$D[\psi] = \left( \frac{d^2 f}{dr^2} - \frac{2f}{r^2} \right) \sin^2 \theta.$$  \hspace{1cm} (7.26)

Once more applying the rule (7.26) we find

$$D^2[\psi] = \left( \frac{d^4 f}{dr^4} - \frac{4 d^2 f}{r^2 dr^2} + \frac{8 d f}{r^3 dr} - \frac{8 f}{r^4} \right) \sin^2 \theta = 0.$$  \hspace{1cm} (7.27)

We have finally arrived at an ordinary differential equation for which the general solution is given by

$$f(r) = \alpha r^4 + \beta r^2 + \gamma r + \delta \frac{1}{r}.$$  \hspace{1cm} (7.28)

Using the condition at infinity (7.24) we find $\alpha = 0$, $\beta = \frac{1}{4}$ and using the boundary condition at the sphere (7.23) we then find $\gamma = -\frac{3}{4}$, $\delta = \frac{1}{4}$. We thus find for the stream function

$$\psi(r, \theta) = \left( \frac{1}{2} r^2 - \frac{3}{4} r + \frac{1}{4r} \right) \sin^2 \theta.$$  \hspace{1cm} (7.29)

The actual sought flow field $u$ is then given by

$$u_r = \left( 1 - \frac{3}{2r} + \frac{1}{2r^3} \right) \cos \theta, \quad u_\theta = \left( -1 + \frac{3}{4r} + \frac{1}{4r^3} \right) \sin \theta,$$  \hspace{1cm} (7.30)

while for the pressure $p$ we obtain

$$p = -\frac{3 \cos \theta}{2r^2} + p_0.$$  \hspace{1cm} (7.31)

with $p_0$ the constant pressure at infinity.
CHAPTER 8. THE ANALYSIS OF ELLIPTIC EQUATIONS

8 Discussion

- Problems of elliptic type are probably the most common. Or at least, elliptic operators appear in a large number of problems in mathematical physics. As we have seen in Chapter 2 we encounter the Laplace operator in parabolic problems and hyperbolic problems as well. A hyperbolic problem like the wave equation reduces to an elliptic problem when we consider only time-harmonic solutions. When the equation is of parabolic type elliptic problems appear as steady-state problems, letting the time go to infinity. Thus we may consider situations like the temperature distribution in a room with heat sources and sinks or the concentration in a vessel with constant replenishing draining [27, 28]. The name potential problem comes from electrostatics. It refers to the potential from electrical charges, following from Maxwell’s equation. In fact Maxwell’s equations are a good example of a hyperbolic problem, [59]

- In fluid mechanics it arises as an equation for the velocity in an incompressible irrotational flow or the stream function in a two-dimensional incompressible flow, [17]. Typical mechanical problems involve the deflection of a thin membrane in two dimensions. For deflection of beams one encounters the so called biharmonic equation. This is still elliptic, but contains a \( \nabla^4 \) operator. There are many similarities with harmonic analysis, [34]

- In Chapter ?? we encounter a number of elliptic problems arising in a practical context. To start with, in Section ?? the groundwater table is modelled as a nonlinear boundary value problem in one dimension. This nonlinearity arises in the second derivative term, which contains a small coefficient moreover. It is shown when and how an asymptotic expansion can “solve” this problem. Section ?? deals with chemical reactions in small pellets. It is described as a the steady state of a reaction-diffusion equation of a spherically symmetric problem. Here too, solutions can be obtained in terms of an asymptotic expansion. An example of a viscous flow problem leading to the Stokes equation is the forming of glas products, see ???. The problem discussed here can be solved analytically using thin layer approximations. A final problem involving elliptic equations is the transfer of heat in a multilayered material. This topic is discussed in Section ??; here the Green’s function turns out to be a useful tool to obtain the solution.

Exercises

8.1. In this exercise we determine the spherically symmetric fundamental solution \( \tilde{w}(x) \) of the \( d \)-dimensional Poisson equation, i.e. \( w(x) = \tilde{w}(r) \) with \( r = \|x\|_2 \).

(a) Show that \( \tilde{w}(r) \) is a solution of the ordinary differential equation

\[
\frac{d^2 \tilde{w}}{dr^2} + \frac{d-1}{r} \frac{d \tilde{w}}{dr} = \delta(r).
\]
Exercises

(b) For \( d > 2 \) and the requirement that \( \lim_{r \to \infty} \tilde{w}(r) = 0 \), show that

\[
\tilde{w}(r) = \frac{A_d}{r^d},
\]

where the integration constants \( A_d \) satisfy

\[
A_d \int_{B(0,1)} \frac{d}{dr}(r^{2-d}) dS = 1.
\]

One can prove that the constants \( A_d \) can be found from

\[
A_d = \frac{1}{(2 - d) B_d},
\]

where \( B_d \) is the surface area of the \( d \)-dimensional unit sphere. NB. \( B_d = 2\pi^{d/2} \Gamma(d/2) \).

8.2. Let \( \nabla^2 u(x) = 0 \) for all \( x \in \Omega \) and \( u(x) > 0 \) for all \( x \in \partial \Omega \). Prove that \( u(x) > 0 \) for all \( x \in \Omega \).

8.3. Consider the Neumann problem

\[
\nabla^2 u = c, \quad x \in \Omega := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\},
\]

\[
\frac{\partial u}{\partial n} = 1, \quad x \in \partial \Omega,
\]

where \( c \) is a constant. Show that this boundary value problem has only solutions for \( c = 2 \); determine all these solutions.

8.4. Let \( \nabla^2 u_i(x) = 0 \) (\( i = 1, 2, \ldots, N \)) for all \( x \in \Omega \). Let \( u_{i-1}(x) \leq u_i(x) \leq u_{i+1}(x) \) (\( i = 2, 3, \ldots, N - 1 \)) for all \( x \in \partial \Omega \). Show that \( u_{i-1}(x) \leq u_i(x) \leq u_{i+1}(x) \) (\( i = 2, 3, \ldots, N - 1 \)) for all \( x \in \Omega \).

8.5. Let \( \nabla^2 u(x) = 0 \) for all \( x \in \Omega \). Split the boundary \( \partial \Omega \) into two simply connected parts, \( \partial \Omega_1 \) and \( \partial \Omega_2 \), such that \( \partial \Omega_1 \cup \partial \Omega_2 = \partial \Omega \). Let \( u(x) \) satisfy the boundary conditions

\[
u(x) = \alpha(x), \quad x \in \partial \Omega_1,
\]

\[
\frac{\partial u}{\partial n}(x) = \beta(x), \quad x \in \partial \Omega_2.
\]

Show that the solution \( u(x) \) is unique.

8.6. Consider the domain \( \Omega := \{(x, y) \in \mathbb{R}^2 \mid x > 0, \ y > 0\} \). Construct a Green’s function for the Dirichlet and Neumann problem using mirror points.

8.7. Let \( a, b \in \mathbb{R} \) with \( a < b \). Define the domain \( \Omega := \{(x, y) \in \mathbb{R}^2 \mid a < y < b\} \). Construct a Green’s function for the Dirichlet problem on \( \Omega \) using mirror points. Show that the Green’s function is symmetric, i.e.

\[
G(x; \xi) = G(\xi, x).
\]

8.8. Consider on the unit disc \( \Omega := \{(r, \phi) \in \mathbb{R}^2 \mid 0 \leq r < 1, -\pi \leq \phi < \pi\} \) the following Neumann problem

\[
\nabla^2 u = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} = 0,
\]

17:22 7 Oct 2003 146 version: 01-10-2003
CHAPTER 8. THE ANALYSIS OF ELLIPTIC EQUATIONS

satisfying the BC
\[
\frac{\partial u}{\partial n}(1, \phi) = g(\phi),
\]
where the function \( g(\phi) \) satisfies the constraint
\[
\int_{-\pi}^{\pi} g(\phi) \, d\phi = 0.
\]
Find a formal solution.

8.9. Prove the mean value theorem: If \( u(x) \) is harmonic in \( \Omega \subset \mathbb{R}^3 \) and \( B(x; \rho) \) denotes the sphere in \( \Omega \) with centre \( x \) and radius \( \rho \), then
\[
u(x) = \frac{1}{4\pi\rho^2} \int_{B(x;\rho)} u(y) \, dS_y,
\]
where the subscript \( y \) denotes that the integration is carried out in the \( y \)-variables.

8.10. Consider on the strip \( \Omega := \{ (x, y) \in \mathbb{R}^2 \mid 0 < x < 1, 0 < y \} \) the following Dirichlet problem
\[
\nabla^2 u = 0, \quad x \in \Omega
\]
\[
u(0, y) = u(1, y) = 0, \quad 0 \leq y < \infty
\]
\[
u(x, 0) = x(1-x), \quad 0 \leq x \leq 1
\]
\( u \) is bounded in \( \Omega \).

Show by separation of variables that the solution is given by
\[
u(x, y) = \frac{8}{\pi^2} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^3} \, e^{-2k+1}\pi y \sin(2k+1)\pi x.
\]

8.11. The Helmholtz equation may not allow for the maximum principle. Prove this for the one-dimensional case. So consider the equation
\[
\frac{d^2 u}{dx^2}(x) + \lambda u = 0, \quad x \in \Omega := (0, \pi),
\]
where \( \lambda > 1 \).

8.12. The maximum principle may also apply to nonlinear problems. For this, consider the boundary value problem
\[
\nabla^2 u(x) = \lambda \, e^{u(x)}, \quad x \in \Omega,
\]
\[
u(x) = 0, \quad x \in \partial \Omega.
\]
(a) Prove that \( u(x) > 0 \) for all \( x \in \Omega \) if \( \lambda > 0 \).
(b) Construct a comparison function to estimate \( u(x) \).

8.13. If we apply an affine coordinate transformation (corresponding to rigid body motions) the Laplace operator remains invariant. Prove this for \( x \in \mathbb{R}^2 \).
Exercises

8.14. Consider the Helmholtz equation of a unit square, i.e.
\[ \nabla^2 u(x) + ku(x) = 0, \quad x \in \Omega := \{(x, y) \in \mathbb{R}^2 \mid 0 < x, y < 1\}, \]
\[ u(x) = 0, \quad x \in \partial \Omega, \]
where \( k < 0 \). Compute the eigenvalues and eigenfunctions.

8.15. Show that the operator
\[ \mathcal{L}[u] := p(x) \frac{d^2 u}{dx^2} + q(x) \frac{du}{dx} \]
is not self-adjoint unless \( q(x) = \frac{d}{dx} p(x) \).

8.16. Consider the eigenvalue problem
\[ \mathcal{L}[u] := \frac{d^2 u}{dx^2} = \lambda u, \quad x \in (0, 1), \]
\[ u(0) = 0, \quad \frac{du}{dx}(1) = 0. \]
Determine the eigenvalues and eigenfunctions.

8.17. Consider the eigenvalue problem
\[ \mathcal{L}[u] := \frac{d^2 u}{dx^2} = \lambda u, \quad x \in (0, 1), \]
\[ u(0) - \frac{du}{dx}(0) = 0, \quad u(1) = 0. \]
Show that the eigenvalues are negative.

8.18. Consider the equation
\[ \mathcal{L}[u] := \nabla^2 u + qu = \lambda u, \quad x \in \Omega, \]
where \( q < 0 \). Show that the eigenvalues \( \lambda_k \leq 0 \) if we impose homogeneous Dirichlet or Neumann boundary conditions. Also show that \( \lambda_k \leq 0 \) if we impose the homogeneous Robin boundary condition (1.6c) with \( \text{sign}(\alpha) = \text{sign}(\beta) \).

8.19. Show that for the operator
\[ \mathcal{L}[u] := \nabla^2 u + a(x) \frac{\partial u}{\partial x} + b(x) \frac{\partial u}{\partial y} \]
the adjoint operator is given by
\[ \mathcal{L}^*[v] = \nabla^2 v - \frac{\partial}{\partial x}(a(x)v) - \frac{\partial}{\partial y}(b(x)v). \]
Chapter 10

Analysis of parabolic equations

The analytical theory of parabolic equations in this chapter starts (Section 1) with deriving solutions for Cauchy problems, i.e. pure initial-value problems. The dependence of (fundamental) solutions on the time and space variables turns out to respect certain symmetries (Section 2). We study so-called similarity solutions in Section 3, where we also derive formulae for the diffusion operator in other than cartesian coordinates. Next we investigate the rôle of boundary conditions in Section 4. First we analyse problems on finite (spatial) domains. A special class of problems is formed by PDE with moving boundaries, so-called Stefan problems (Section 5). The last section, 6, is devoted to steady-state solutions and travelling-wave solutions. A point of interesting point here is that stationary solutions constitute proper solutions of corresponding elliptic boundary value problems (which are the subject of Chapter ??).

1 Cauchy problems

Parabolic equations arise in a variety of applications, mainly associated with diffusive processes. In Section 1.1 we saw an example of (chemical) diffusion. Heat flow also has a diffusive character, often called conduction. This may be combined with convection if the medium is a fluid in motion. In mechanics of fluids internal friction, called viscosity, produces diffusion of momentum. In this section we shall consider the simplest form of a diffusive problem, viz. the linear heat equation on infinite domains.

1.1 The heat equation in one space dimension

In order to investigate solutions of parabolic equations we consider the following initial value problem for the simplest (non-dimensional) form of the so-called heat equation

\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in \mathbb{R}, \ t > 0, \]  
\[ u(x, 0) = v(x), \quad x \in \mathbb{R}, \]  

(1.1a)

(1.1b)
where $|v|$ and $|u|$ are integrable over $\mathbb{R}$. (Note that limiting conditions for $x \to \pm \infty$ are just to be expected if we portray the present IVP as the limit of an IBVP. See example 10.2.) We seek a solution of the form
\[ u(x, t) = p(x)q(t), \]
(1.2)
i.e., we separate the independent variables so that the functions $p$ and $q$ in the product depend on either one of the variables $x$ and $t$ exclusively. Upon substitution we find
\[ \frac{dq}{dt} = \frac{d^2 p}{dx^2} q, \]
or, assuming $p(x), q(t) \neq 0$,
\[ \frac{1}{q} \frac{dq}{dt} = \frac{1}{p} \frac{d^2 p}{dx^2}. \]
(1.3)
The left hand side of (1.3) depends solely on $t$ and the right hand side on $x$. This is only possible if both sides are equal to some constant $\lambda$, the separation constant. (As will appear below, it is sufficient to assume that $\lambda$ is real.) So we obtain the two eigenvalue problems
\[ \frac{d^2 p}{dx^2} = \lambda p, \]
(1.4a)
\[ \frac{dq}{dt} = \lambda q. \]
(1.4b)
From (1.4b) we conclude that only value $\lambda \leq 0$ are allowed for a stable solution $q(t)$. Indeed, if $\lambda > 0$, then (1.4a) would have solutions $p(x)$ which are exponentially increasing for either $x \to -\infty$ or $x \to \infty$. Also for $\lambda = 0$, equation (1.4a) has an unbounded solution, given by $p(x) = C_1 x + C_2$. Therefore, in order for the solution $u(x, t)$ of (1.1a) to be bounded we assume (with $\kappa$ real)
\[ \lambda = -\kappa^2 < 0. \]
(1.5)
A general solution of (1.4) is then given by
\[ p(x)q(t) = C e^{i\kappa x - \kappa^2 t}, \]
(1.6)
where $C$ is a constant. Here we recognize a planar wave from the dispersion relation for the heat equation (1.1a). Apparently we can view this planar wave as a Fourier mode. By Fourier analysis (Chapter 3), a general bounded solution of (1.1) can now be found by superposition over all possible $\kappa$. The suggested solution is now given by
\[ u(x, t) = \int_{-\infty}^{\infty} \hat{\nu}(\kappa) e^{i\kappa x - \kappa^2 t} \, d\kappa, \]
(1.7)
where, using the initial value $u(x, 0) = v(x)$,
\[ \hat{\nu}(\kappa) := \frac{1}{2\pi} \int_{-\infty}^{\infty} v(\xi) e^{-i\kappa \xi} \, d\xi. \]
(1.8)
CHAPTER 10. ANALYSIS OF PARABOLIC EQUATIONS

By substituting (1.8) in (1.7) and interchanging the order of integration, we obtain

\[ u(x, t) = \int_{-\infty}^{\infty} v(\xi) \psi(x - \xi, t) \, d\xi, \quad (1.9a) \]

where

\[ \psi(\xi, t) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi \zeta} e^{-\xi^2 t} \, d\kappa = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(x - i\xi / 2t)^2 t - \xi^2 / 4t} \, d\kappa \]

\[ = \frac{e^{-\xi^2 / 4t}}{2\pi} \int_{-\infty}^{\infty} e^{-\xi^2 / 4t} \, d\kappa = \frac{e^{-\xi^2 / 4t}}{2\pi \sqrt{t}} \int_{-\infty}^{\infty} e^{-s^2} \, ds = \frac{e^{-\xi^2 / 4t}}{2\sqrt{\pi t}}. \quad (1.9b) \]

We shifted the \( \kappa \)-contour by an amount of \( i \xi / 2t \) into the complex plane. Altogether we obtain the solution

\[ u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} v(\xi) \exp\left(-\frac{(x - \xi)^2}{4t}\right) \, d\xi. \quad (1.10a) \]

We remark that the derivation closely follows that of the Duhamel integral in Section 4.6. One can easily verify that this solution satisfies the equation (1.1a). By changing the coordinate of integration to \( \xi = x + 2\eta \sqrt{t} \) we find

\[ u(x, t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} v(x + 2\sqrt{t} \eta) e^{-\eta^2} \, d\eta. \quad (1.10b) \]

The initial condition is now readily verified to be

\[ u(x, 0) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} v(x) e^{-\eta^2} \, d\eta = v(x) \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\eta^2} \, d\eta = v(x). \]

In Section 1.2 we shall show that it is also unique.

Next, we show that \( \psi(\xi, t) \) is equivalent to a fundamental solution; see (4.5). By taking \( v(x) = \delta(x) \) in (1.10a) we immediately see that \( \psi \) satisfies the initial value problem

\[ \frac{\partial \psi}{\partial t} - \frac{\partial^2 \psi}{\partial \xi^2} = 0, \quad \xi \in \mathbb{R}, t > 0, \quad (1.11a) \]

\[ \psi(\xi, 0) = \delta(\xi), \quad \xi \in \mathbb{R}. \quad (1.11b) \]

A fundamental solution \( w(x, \xi, t, \tau) \) of (1.1) satisfies

\[ \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} = \delta(x - \xi) \delta(t - \tau), \quad x \in \mathbb{R}, t > \tau, \quad (1.12a) \]

\[ w(x, \xi, t, \tau) = 0, \quad t < \tau. \quad (1.12b) \]

**Property 10.1.** The function \( w(x, \xi, t, \tau) = \psi(x - \xi, t - \tau)H(t - \tau) \), where \( H \) is the Heaviside step function, is a fundamental solution of (1.1).
1. CAUCHY PROBLEMS

Proof. $H(t - \tau) \equiv 0$ for $t < \tau$. Furthermore, the difference between

$$w_t = \psi_t H + \psi \delta = \psi_t H + \delta(x - \xi)\delta(t - \tau),$$

$$w_{xx} = \psi_{\xi\xi} H$$

is just $\delta(x - \xi)\delta(t - \tau)$. \[\square\]

It follows that for $t > \tau$ we can identify $w(x, \xi, t, \tau)$ with $\psi(x - \xi, t - \tau)$. We remark that the $\psi$ found in (1.9b) apparently satisfies

$$\int_{-\infty}^{\infty} \psi(\xi, t) d\xi = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\xi^2/4t} d\xi = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\eta^2} d\eta = 1.$$ (1.13a)

Hence $\psi$ can be seen as a probability distribution. In fact it is the famous Gaussian or normal distribution. From the rightmost side of (1.13a) we see that

$$\text{erf}(x) := \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-\mu^2} d\mu,$$ (1.13b)

the so-called error function, which plays such an important rôle in statistics, is related to the problem here. The same is true for the complementary error function defined by

$$\text{erfc}(x) := \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-\mu^2} d\mu = 1 - \text{erf}(x).$$ (1.13c)

We will frequently encounter this function further on.

Example 10.2 If we do not restrict ourselves, in problems on infinite domains, to bounded solutions we may easily encounter solutions that grow rather than decay in time. This growth may be associated to possible sources at infinity. Consider the examples

$$u(x, 0) = e^x \quad \text{with} \quad u(x, t) = e^{x+t},$$

$$u(x, 0) = \frac{1}{2}x^2 \quad \text{with} \quad u(x, t) = \frac{1}{2}x^2 - t.$$ \[\square\]

Example 10.3 Rather than equation (1.1a) one may have

$$\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2},$$

where $k$ is a positive constant, called the (thermal) diffusivity. We may apply a change of variables, $s := tk$ or $z := x/\sqrt{t}$ to obtain either $u_s = u_{ss}$ or $u_t = u_{zz}$. \[\square\]
1.2 The heat equation in \(d\) space dimensions

Our formulation in Section 1.1 was chosen such that it lends itself easily to heat equations in more than one space variable. So consider

\[
\frac{\partial u}{\partial t} = \nabla^2 u, \quad x \in \mathbb{R}^d, \quad t > 0, \quad (1.14a)
\]
\[
u(x, 0) = v(x), \quad x \in \mathbb{R}^d. \quad (1.14b)
\]

where \(|v|\) and \(|u|\) are integrable over \(\mathbb{R}^d\). A fundamental solution \(w\) (cf. Section 4.5) then satisfies (cf. (1.12))

\[
\frac{\partial w}{\partial t} - \nabla^2 w = \delta(x - \xi)\delta(t - \tau), \quad x \in \mathbb{R}^d, \quad t \in \mathbb{R} \quad (1.15a)
\]
\[
w(x, \xi, t, \tau) = 0, \quad t < \tau. \quad (1.15b)
\]

By analogy of Property 10.1 we have

\[
w(x, \xi, t, \tau) = \left(\frac{1}{2\sqrt{\pi (t - \tau)}}\right)^d \exp\left(-\frac{\|x - \xi\|^2}{4(t - \tau)}\right) H(t - \tau). \quad (1.16)
\]

Superposition then gives (cf. Chapter 4)

\[
u(x, t) = \left(\frac{1}{2\sqrt{\pi t}}\right)^d \int_{\mathbb{R}^d} v(\xi) \exp\left(-\frac{\|x - \xi\|^2}{4t}\right) d\xi. \quad (1.17)
\]

This solution exists and is unique:

**Theorem 10.4.** Let \(v\) be continuous and integrable on \(\mathbb{R}^d\). Then the function given by (1.17) satisfies initial value problem (1.14) uniquely.

**Proof.** By virtue of the exponential and the fact that \(v\) is bounded, the defining integral of \(u\) is uniformly integrable for all \(x\) and \(t > 0\). So \(u\) and likewise \(\int w; \nu(x) d\xi\) and \(\int \nabla^2 w; \nu(x) d\xi\) exist, and we can interchange differentiation and integration. Since \(w_t = \nabla^2 w\), (1.14a) readily follows. To show that \(u\) also satisfies the initial condition, we change variables \(\xi = x + 2\eta \sqrt{t}\), to obtain

\[
u(x, t) = \left(\frac{1}{\sqrt{t}}\right)^d \int_{\mathbb{R}^d} v(x + 2\eta \sqrt{t}) e^{-\|\eta\|^2} d\eta.
\]

With identity (1.13a), this is clearly equal to \(v(x)\) for \(t = 0\).

To prove that it is unique we assume that \(u_1\) and \(u_2\) are two solutions. Then the difference \(w(x, t) := u_1(x, t) - u_2(x, t)\) satisfies the initial value problem with \(v(x, t) \equiv 0\). From Green’s first identity we have

\[
d \int_{\mathbb{R}^d} \nabla^2 w^2 dV = \int_{\mathbb{R}^d} w \frac{\partial w}{\partial t} dV = \int_{\mathbb{R}^d} w \nabla^2 w dV = -\int_{\mathbb{R}^d} |\nabla w|^2 dV < 0.
\]
Hence
\[ 0 \leq \int_{\mathbb{R}^d} w^2(x, t) \, dV \leq \int_{\mathbb{R}^d} w^2(x, 0) \, dV = 0 \]
so \( w \equiv 0. \)

1.3 Problems on half spaces

It is of interest to study the solution of problems which are defined on a semi-infinite domain with values prescribed at the boundary. Although such a setting can also be handled in higher dimensions, we restrict ourselves to the one-dimensional case, with the domain \([0, \infty)\). We are looking for the solution of
\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \geq 0, \quad t > 0, \] (1.18)
satisfying either
\[ u(x, 0) = v(x), \quad x \geq 0, \] (1.19a)
\[ u(0, t) = 0, \quad t > 0, \] (1.19b)
for some given initial value \( v(x) \), or
\[ u(x, 0) = 0, \quad x \geq 0, \] (1.20a)
\[ u(0, t) = \beta(t), \quad t > 0, \] (1.20b)
for some boundary value \( \beta(t) \). Clearly, the conditions (1.19) are different from but related to the one we encountered in (1.1b).

Formally we may try to use a fundamental solution approach by employing a continuation of the IC on the left interval \((-\infty, 0]\). Let \( \tilde{v}(x) \) be equal to \( v(x) \) for \( x \geq 0 \) and some continuous extension for \( x < 0 \), but such that the representation
\[ u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} \tilde{v}(\xi) \exp\left(-\frac{(x - \xi)^2}{4t}\right) \, d\xi, \] (1.21)
still satisfies boundary condition (1.19b). Simple substitution then reveals
\[ \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} \tilde{v}(\xi) \exp\left(-\frac{\xi^2}{4t}\right) \, d\xi = \frac{1}{2\sqrt{\pi t}} \int_{0}^{\infty} (\tilde{v}(\xi) + \tilde{v}(-\xi)) \exp\left(-\frac{\xi^2}{4t}\right) \, d\xi = 0, \] (1.22)
which is true for all \( t \) if
\[ \tilde{v}(\xi) = -\tilde{v}(-\xi), \quad \xi < 0. \] (1.23)
As a consequence we find the representation
\[ u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{0}^{\infty} v(\xi) \left[ \exp\left(-\frac{(x - \xi)^2}{4t}\right) - \exp\left(-\frac{(x + \xi)^2}{4t}\right) \right] \, d\xi. \] (1.24)
Example 10.5 An important application is the heat shock. Suppose an object with a certain given temperature is suddenly exposed to a heat source (or sink). If this has a temperature different from the object, this means that the temperature profile initially exhibits a discontinuity. The most simple model is given by assuming that the initial temperature is constant, i.e. 

\[ v(x) = v_0 > 0. \]

(See Example 7.15 for the related problem of a given heat flux.) The solution of initial boundary value problem (1.18)-(1.19) is then given by

\[
 u(x, t) = \frac{v_0}{2\sqrt{\pi t}} \left[ \int_{-\infty}^{x-\xi/\sqrt{4t}} \exp\left( -\frac{(x-\xi)^2}{4t} \right) \, d\xi - \int_{x+\xi/\sqrt{4t}}^{\infty} \exp\left( -\frac{(x+\xi)^2}{4t} \right) \, d\xi \right].
\]

Substituting \( z := -\frac{x-\xi}{2\sqrt{t}} \) in the first integral and \( z := \frac{x+\xi}{2\sqrt{t}} \) in the second one gives

\[
 u(x, t) = \frac{v_0}{\sqrt{\pi t}} \left[ \int_{-\infty/-\xi/2\sqrt{t}}^{x} e^{-z^2} \, dz - \int_{x/2\sqrt{t}}^{\infty} e^{-z^2} \, dz \right] = v_0 \text{erf} \left( \frac{x}{2\sqrt{t}} \right).
\]

In Figure 10.1 we have displayed various stages of the solution for \( v_0 = 1. \) □

Example 10.6 We can also apply the analysis to a two-sided heat shock. Indeed, consider an object that is suddenly heated in a finite region, \([-1, 1]\) say. A very simple model is then given by the piecewise constant initial condition

\[
 v(x) = \begin{cases} 
 0 & \text{if } -\infty < x < -1 \\
 v_0 & \text{if } -1 \leq x \leq 1 \\
 0 & \text{if } 1 < x < \infty 
\end{cases}.
\]
2. THE HEAT EQUATION WITH SPATIAL SYMMETRIES

where \( v_0 \neq 0 \), yielding the solution

\[
u(x, t) = \frac{v_0}{2\sqrt{\pi t}} \int_{-1}^{1} \exp\left(-\frac{(x - \xi)^2}{4t}\right) d\xi,
\]
or, upon substituting \( z := (x - \xi)/2\sqrt{t} \),

\[
u(x, t) = \frac{v_0}{\sqrt{\pi}} \int_{(x-1)/2\sqrt{t}}^{(x+1)/2\sqrt{t}} e^{-z^2} dz = \frac{v_0}{2} \left( \text{erf}\left(\frac{x+1}{2\sqrt{t}}\right) - \text{erf}\left(\frac{x-1}{2\sqrt{t}}\right) \right),
\]

which can also be found by employing the solution of the previous example. \( \square \)

The other situation, viz. initial boundary value problem (1.18), (1.20) can be solved using the solution of the IBVP (1.18), (1.19). First consider the problem where \( \beta(t) \equiv 1 \). It is simple to see that

\[
u(x, t) = 1 - \text{erf}\left(\frac{x}{2\sqrt{t}}\right) = \text{erfc}\left(\frac{x}{2\sqrt{t}}\right), \tag{1.25}\]
satisfies (1.18), (1.20). We can now invoke the Duhamel integral (see Theorem 4.16) to obtain the solution. Note that we can define solutions \( w(x, t - \tau) \) by

\[
w(x, t - \tau) := \text{erfc}\left(\frac{x}{2\sqrt{t - \tau}}\right). \tag{1.26}\]

Hence we obtain as formal solution of (1.18), (1.20) from (4.6.68)

\[
u(x, t) = \frac{\partial}{\partial t} \int_0^t \beta(\tau) \text{erfc}\left(\frac{x}{2\sqrt{t - \tau}}\right) d\tau. \tag{1.27}\]

We can work this out to get

\[
u(x, t) = \frac{x}{2\sqrt{\pi}} \int_0^t \frac{\beta(\tau)}{(t - \tau)^{3/2}} \exp\left(-\frac{x^2}{4(t - \tau)}\right) d\tau. \tag{1.28}\]

One may check that this solution satisfies the initial and boundary conditions.

Other than for the simplest forms of \( \beta(\tau) \) this formula does not provide an explicit answer. Nevertheless it may be very useful for order of magnitude estimates, or the analysis of trends or asymptotic behaviour. On the other hand, if we are interested in actual numbers for the general case, we have to evaluate the integral numerically. Although this is not a major problem (the apparent square root singularity in \( \tau = t \) is completely cancelled by the exponential), we have to compare the effort with other, more direct, numerical methods for solving the problem.

2 The heat equation with spatial symmetries

In section 1.2 we formulated the solution of the \( d \)-dimensional heat equation

\[
\frac{\partial u}{\partial t} = \nabla^2 u, \quad x \in \mathbb{R}^d, \ t > 0, \tag{2.29}\]

17:22 7 Oct 2003 156 version: 26-09-2003
for a given initial value for \( u \). If boundary and initial conditions are cylindrically or spherically symmetric, so is the solution, and it is beneficial to rewrite the Laplacian \( \nabla^2 \) in the corresponding coordinates (see Appendix K) and suppress any of the non-radial derivatives. With circular (in 2-D) or cylindrical (in 3-D) symmetry, with \( r^2 = x^2 + y^2 \), we have

\[
\frac{\partial u}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right), \quad \text{or} \quad \frac{\partial u}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{\partial^2 u}{\partial z^2}.
\] (2.30a)

With spherical symmetry, with \( r^2 = x^2 + y^2 + z^2 \), we have

\[
\frac{\partial u}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial u}{\partial r} \right).
\] (2.30b)

In general we have thus for a radially symmetric field in \( d = 1, 2, 3 \)-dimensional space

\[
\frac{\partial u}{\partial t} = \frac{1}{r^{d-1}} \frac{\partial}{\partial r} \left( r^{d-1} \frac{\partial u}{\partial r} \right) \quad \text{for} \quad r = |x|, \ x \in \mathbb{R}^d, \ t > 0.
\] (2.30c)

Evidently, these PDEs are defined on the half space \( r \geq 0 \). If \( u \) is smooth at \( r = 0 \), the radial symmetry induces the boundary condition

\[
\frac{\partial u}{\partial r} = 0.
\] (2.31)

Finally we remark that if for a circular or cylindrically symmetric problem the \( r \)-domain consists of values \( r \geq R_0 > 0 \), for some \( R_0 \), we can use the transformation

\[
v(\rho, t) := u(r, t), \quad \rho := \ln r,
\] (2.32a)

resulting for (2.30a) in

\[
\frac{\partial v}{\partial t} = e^{-2\rho} \frac{\partial^2 v}{\partial \rho^2}, \quad \text{or} \quad \frac{\partial v}{\partial t} = e^{-2\rho} \frac{\partial^2 v}{\partial \rho^2} + \frac{\partial^2 v}{\partial z^2}.
\] (2.32b)

A 3-D spherically symmetric problem can be transformed, for \( r \geq R_0 \), for some \( R_0 > 0 \), by

\[
v(r, t) := u(r, t)r,
\] (2.33a)

to a form equivalent to the 1-D standard heat equation

\[
\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial r^2}.
\] (2.33b)

It shows that any of the foregoing 1-D solutions correspond immediately to a similar 3-D solution.

3 Similarity solutions

In Section (7.3), we found by dimensional arguments that in the absence of explicit length and time scales other than \( x \) and \( t \) themselves dimensionless groups can only occur by
mutual combinations of $x$ and $t$. In the present problems related to the heat equation, the prevailing combination appeared to be $x/\sqrt{t}$. In the same way the dependent variable (the temperature, say) will scale on a given temperature if this is available, but will have to depend on a combination of $x$ or $t$ and the flux, if no temperature but only a flux is given. These observations give rise to so-called similarity solutions (see also Section 2.5).

As these symmetry properties should be present in the problem independently of the physics, it seems useful to look for similarity solutions, as soon as the spatial and time coordinates have an infinite extension. Assume the problem is radially symmetric, so $u = u(r, t)$ where $d = 1, 2, \text{ or } 3$, and $r$ denotes the Euclidean distance to the origin. Consider therefore again

$$
\frac{\partial u}{\partial t} = \frac{1}{r^{d-1}} \frac{\partial}{\partial r} \left( r^{d-1} \frac{\partial u}{\partial r} \right) \quad \text{for } r = |x|, \; x \in \mathbb{R}^d, \; t > 0.
$$

We will look for solutions of the form

$$
u(r, t) = r^m f(\eta), \quad \eta = \frac{r}{t^n}.
$$

As $r = t^n \eta$, there is no need to include any factor of the type $t^p$. Differentiation with respect to $t$ and $r$ yields

\begin{align*}
\frac{\partial u}{\partial t} &= -nm^{m-2} \eta^3 t^{2n-1} f'(\eta), \\
\frac{\partial u}{\partial r} &= r^{m-1} (mf(\eta) + \eta f'(\eta)), \\
\frac{\partial^2 u}{\partial r^2} &= r^{m-2} (m(m-1)f(\eta) + 2mnf'(\eta)) + \eta^2 f''(\eta)
\end{align*}

(3.35)

(the prime denotes differentiation with respect to $\eta$). After substitution in (2.30c) we find

$$
\eta^2 f'' + \eta(2m + d - 1) f' + m(m + d - 2) f + n\eta^3 t^{2n-1} f' = 0
$$

which has only proper solutions if $n = \frac{1}{2}$. So we finally have the equation

$$
\eta^2 f'' + \eta\left(\frac{1}{2} \eta^2 + 2m + d - 1\right) f' + m(m + d - 2) f = 0.
$$

(3.37)

A solution of this equation for arbitrary $m$ and $d$ can be given in terms of hypergeometric or similar functions (see [6]), but this is too general to be of interest. It is more expedient to look for specific solutions once $d$ and $m$ are known. The value of $m$ has to follow from the available boundary conditions, and is sometimes immediately clear from dimensional arguments. A typical case may be found in example 7.15.

**Example 10.7** An important example of a similarity solution is the field of a steady point source. This is most efficiently found from the fundamental solution given in (1.16). This gives us instantly the field of a stroke of heat from the point source $\delta(x)\delta(t)$, i.e.

$$
u(x, t) = \left(\frac{1}{2\sqrt{\pi t}}\right)^d \exp\left(-\frac{r^2}{4t}\right).
$$

By integration of this solution with respect to time, we find the field of a steady point source $\delta(x)H(t)$ at the origin, switched on at $t = 0$, thus satisfying

$$\frac{\partial U}{\partial t} - \nabla^2 U = \delta(x)H(t).$$

In one dimension this is

$$U_1(x, t) = \int_0^t \frac{e^{-r^2/4t}}{\sqrt{4\pi t}} \, dr = \sqrt{\frac{t}{\pi}} \exp\left(-\frac{x^2}{4t}\right) - \frac{1}{2} \left| x \right| \text{erfc}\left(\frac{|x|}{\sqrt{4t}}\right),$$

which we already encountered in example 7.15. In two dimensions we have

$$U_2(r, t) = \int_0^t \frac{e^{-r^2/4t}}{4\pi t} \, dr = \frac{1}{4\pi} E_1\left(\frac{r^2}{4t}\right),$$

where $E_1(z)$ is known as the exponential integral [6]. In three dimensions we obtain

$$U_3(r, t) = \int_0^t \frac{e^{-r^2/4t}}{(4\pi t)^{3/2}} \, dr = \frac{1}{4\pi r} \text{erfc}\left(\frac{r}{\sqrt{4t}}\right).$$

It may be verified that indeed the respective fluxes out of the source are given by

$$\lim_{h \downarrow 0} \left[-\frac{\partial U}{\partial x}\right]_{x = h} = \lim_{h \downarrow 0} \left[\frac{1}{2} \text{sign}(x) \text{erf}(\frac{|x|}{\sqrt{4t}})\right]_{x = h} = 1,$$

$$\lim_{h \downarrow 0} \left[-2\pi \frac{\partial U}{\partial r}\right]_{r = h} = \lim_{h \downarrow 0} \left[\exp\left(-\frac{r^2}{4t}\right)\right]_{r = h} = 1,$$

$$\lim_{h \downarrow 0} \left[-4\pi r \frac{\partial U}{\partial r}\right]_{r = h} = \lim_{h \downarrow 0} \left[\frac{r}{4\pi t} \exp\left(-\frac{r^2}{4t}\right) + \text{erfc}\left(\frac{r}{\sqrt{4t}}\right)\right]_{r = h} = 1.$$

An important point to observe here is the fact that the field of a point source of constant output sometimes (if $d = 3$) converges to a stationary state, but not always (if $d = 1, 2$). It depends on $d$, the number of spatial dimensions of the problem. For large $t$ we find asymptotically [6]

$$U_1(x, t) = \sqrt{\frac{t}{\pi}} - \frac{1}{2} \left| x \right| + O(t^{-1/2}),$$

$$U_2(r, t) = \frac{1}{4\pi} \left[\ln(t) - \ln(4r^2) - \gamma\right] + O(t^{-1}),$$

$$U_3(r, t) = \frac{1}{4\pi r} + O(t^{-1/2}),$$

where $\gamma = 0.5772\ldots$. Apparently, only the capacity of three-dimensional space is big enough to absorb all the heat of a stationary source!

4 Initial boundary value problems

When the problem is defined on a domain with boundaries we need to provide boundary conditions (BC). Sometimes these boundaries are part of the evolution process; we shall address the latter in the next section. In order to fix our thoughts we think of $u$ as a temperature here. Consider, as before, the PDE with initial value

$$\frac{\partial u}{\partial t} = \nabla \cdot (k \nabla u) = k \nabla^2 u, \quad x \in \Omega \subset \mathbb{R}^d, \ t > 0 \quad (4.38a)$$

$$u(x, 0) = v(x), \quad (4.38b)$$
where \( k \) denotes the thermal diffusion coefficient, while \( -k \nabla u \) is the heat flux vector. The heat flux across a surface with unit normal vector \( n \) is thus given by \( -k \frac{\partial u}{\partial n} \). The first form of (4.38) is more general and is valid if \( k \) depends on \( x \) or \( u \) (or both). We assume that \( \Omega \) is a compact simply connected domain, with boundary \( \partial \Omega \). Apart from a prescribed initial value, the following boundary conditions are usually considered

(i) If the temperature at the boundary is prescribed, we call this a BC of Dirichlet type
\[
u(x, t) = g(x, t), \quad x \in \partial \Omega, \ t > 0. \tag{4.39}\]

(ii) If the heat flux across the boundary is prescribed, we have a BC of Neumann type
\[
-k \frac{\partial u}{\partial n}(x, t) = g(x, t), \quad x \in \partial \Omega, \ t > 0. \tag{4.40}
\]

Here \( \frac{\partial u}{\partial n} := n \cdot \nabla u \), where \( n \) is the unit outward normal vector on \( \partial \Omega \).

(iii) A linear combination of Dirichlet and Neumann type BC is called a Robin or Newton BC
\[
-k \frac{\partial u}{\partial n}(x, t) = s(u(x, t) - g(x, t)), \quad x \in \partial \Omega, \ t > 0, \tag{4.41}
\]

with the heat transfer coefficient \( s > 0 \). Put in physical terms, one might say that the heat flux is proportional to the difference between the temperature of the medium and the ambient.

(iv) If the loss of heat is caused by radiation, we have a radiative BC
\[
-k \frac{\partial u}{\partial n}(x, t) = \sigma (u(x, t)^4 - u_\infty^4), \quad x \in \partial \Omega, \ t > 0, \tag{4.42}
\]

where \( u_\infty \) is a specified ambient temperature of the surrounding medium, and parameter \( \sigma \) quantifies the emissivity properties of the surface.

We remark that BCs without external forcing, i.e. where \( g(x, t) \equiv 0 \), are called homogeneous.

**Example 10.8** Consider the initial boundary value problem
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in (0, 1), \ t > 0,
\]
\[
u(x, 0) = v(x), \quad x \in (0, 1)
\]
\[
u(0, t) = u(1, t) = 0, \quad t > 0.
\]

Clearly the BC are of Dirichlet type. Assume that \( v \) and \( v' \) are piecewise smooth, such that the coefficients of the Fourier-sine expansion
\[
v(x) = \sum_{k=1}^{\infty} v_k \sin(k\pi x)
\]
CHAPTER 10. ANALYSIS OF PARABOLIC EQUATIONS

converge according to $v_k = O(1/k)$ as $k \to \infty$ (Corollary 3.7). Consider the following formal series

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

Due to the exponential the coefficients behave as $v_k e^{-k^2 \pi^2 t} = O(k^{-n})$ as $k \to \infty$ for any $n$ if $t > 0$, so the series and all of its derivatives converge uniformly for $t > 0$ (Appendix C). As a result $u$ is continuous while differentiation and summation may be exchanged. It is then easily verified that $u$ satisfies the differential equation, initial profile $v$ and boundary values $u(0, t) = u(1, t) = 0$. Note in particular that in the heat equation no discontinuity in any derivative can be sustained: (practically) any initial profile $v$ yields immediately, for any $t > 0$, an infinitely differentiable temperature distribution $u$. □

Example 10.9 Consider the initial boundary value problem

$$\begin{align*}
\frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2}, \quad x \in (0, 1), \ t > 0, \\
u(x, 0) &= v(x), \quad x \in (0, 1) \\
\frac{\partial u}{\partial x}(0, t) &= \frac{\partial u}{\partial x}(1, t) = 0, \quad t > 0.
\end{align*}$$

This is a BC of Neumann type. Analogous to the previous example, we assume that $v$ and $v'$ are piecewise smooth and expand $v$ into the Fourier-cosine series

$$v(x) = \sum_{k=0}^{\infty} v_k \cos(k \pi x),$$

of which the coefficients converge like $v_k = O(1/k)$ as $k \to \infty$. The formal solution

$$u(x, t) = \sum_{k=0}^{\infty} v_k e^{-k^2 \pi^2 t} \cos(k \pi x)$$

and any of its derivatives converge uniformly for $t > 0$ to a continuous function. So it satisfies the equation and initial condition. In particular it satisfies the boundary conditions. □

Uniqueness of solutions can be shown by integral arguments.

**Theorem 10.10.** The solution of the linear initial boundary value problem (4.38), i.e. $k$ is independent of $u$, with any linear boundary condition (4.39), (4.40) or (4.41), is unique.

**Proof.** Let $u_1$ and $u_2$ be two solutions, then $w(x, t) := u_1(x, t) - u_2(x, t)$ satisfies either initial boundary value problem with $g(x, t) \equiv 0$ and $v(x, t) \equiv 0$. From a variation of Green’s first identity we have (for a fixed $\Omega$)

$$\begin{align*}
\frac{d}{dt} \int_{\Omega} w^2 \, dV &= \int_{\Omega} w \frac{\partial w}{\partial t} \, dV = \int_{\Omega} w \nabla \cdot (k \nabla w) \, dV = \int_{\Omega} \nabla \cdot (k w \nabla w) - k |\nabla w|^2 \, dV \\
&= -\int_{\Gamma} k |\nabla w|^2 \, dS + \int_{\Omega} k w \frac{\partial w}{\partial n} \, dS = \begin{cases} -\int_{\Omega} k |\nabla w|^2 \, dV \\
-\int_{\Omega} k |\nabla w|^2 \, dV - \int_{\partial \Omega} s w^2 \, dS \end{cases} < 0.
\end{align*}$$

5. MOVING BOUNDARIES; STEFAN PROBLEMS

(The first option corresponds to a Dirichlet or Neumann BC, the other to a Robin BC.) Hence

$$0 \leq \int_{\Omega} w^2(x, t) \, dV \leq \int_{\Omega} w^2(x, 0) \, dV = 0$$

so $w \equiv 0$. \qed

5 Moving boundaries; Stefan problems

Often diffusion processes involve a moving boundary. Typical examples are contact problems between two materials like two fluids or solid material and a dissolvent, between two phases of the same material like water and ice, and between wet and dry material. In these problems the boundary is determined by some additional kinematic constraint. In Chapter ?? we shall consider one such problem in more detail. Here we will restrict ourselves to a class of problems referred to as Stefan problems [133]. One may typically think of ice of $0^\circ$C that is melting at a surface, the front, denoted by $S$. Apparently $S = S(t)$. In 1-D, its most simple form, the (dimensionless) temperature $u(x, t)$ of the water satisfies the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < S(t), \ t > 0. \quad (5.43)$$

Before $t = 0$, the temperature is everywhere $u(x, t) = 0$, but at $t = 0$ the temperature at one end, $x = 0$, is suddenly increased to unity

$$u(0, t) = 1, \quad t \geq 0. \quad (5.44a)$$

This enforces the input of energy because the temperature elsewhere is lower. At $x = S(t)$, the interface with the ice, we have a continuous temperature, so

$$u(S, t) = 0, \quad t > 0. \quad (5.44b)$$

As the position of the interface is unknown, we need an additional condition. This extra condition is found from the physics of the phase transition. The amount of specific latent heat, released during the water-to-ice phase transition, is again to be added when the ice melts. As the front travels with a certain speed, the produced heat travels with the same speed, and is therefore equal to the heat flux. As we assumed that the ice is of constant temperature, i.e. $u_{ic} \equiv 0$, the flux into the ice vanishes, and we get (non-dimensionally)

$$-\frac{\partial u}{\partial x}(S(t), t) = \alpha \frac{dS}{dt}, \quad S(0) = 0. \quad (5.45)$$

The constant $\alpha$ in (5.45) is called the Stefan constant. Let us try and solve problem (5.43-5.45) by introducing, following (3.34-3.37), the similarity solution

$$u(x, t) = x^m f(\eta), \quad \eta := \frac{x}{2\sqrt{t}}. \quad (5.46)$$

From (5.44a) it follows that $m = 0$, while from (5.44b) we can write for some constant $\gamma$

$$S = 2\gamma \sqrt{t}.$$
Equation (5.43) and boundary conditions turn into
\[ f'' + 2\eta f' = 0, \quad f(0) = 1, \quad f(\gamma) = 0, \quad f'(\gamma) = -2\alpha \gamma, \] (5.46)
along \( 0 \leq \eta \leq \gamma \), with solution
\[ f(\eta) = 1 - \alpha \gamma e^{\frac{\eta^2}{4\sqrt{\pi}}} \text{erf}(\eta), \] (5.47)
while the unknown \( \gamma \) is given by the real positive root of
\[ \alpha \gamma \sqrt{\pi} e^{\gamma^2} \text{erf}(\gamma) = 1. \] (5.48)
Altogether we have thus
\[ u(x, t) = 1 - \frac{\text{erf}(x/2\sqrt{t})}{\text{erf}(\gamma)}. \] (5.49)

We may solve (5.48) numerically to obtain \( \gamma \) as a function of the Stefan parameter \( \alpha \). See for example figure 10.2. For small \( \gamma \) we can approximate \( \gamma \sqrt{\pi} e^{\gamma^2} \text{erf}(\gamma) \) by \( 2\gamma^2 + \ldots \), so for large \( \alpha \) we have \( \gamma \simeq 1/\sqrt{2\alpha} \). For large \( \gamma \) we have a dominating exponential that produces for small \( \alpha \) a \( \gamma \simeq (-\frac{1}{2} \ln \alpha)^{\frac{1}{2}} \).

Figure 10.2. Similarity coordinate \( \gamma \) of interface as a function of Stefan problem parameter \( \alpha \).

Example 10.11 This single-phase problem, where the temperature varies only in the water, is easily generalised to a double-phase problem, where the temperature varies both in ice and water. In this case we need the more general Stefan condition
\[ \rho L \frac{dS}{dt} = \left[ -k_w \frac{\partial}{\partial x} T_w + k_i \frac{\partial}{\partial x} T_i \right]_{x=S(t)}, \] (5.50)
(now dimensionally) where \( L \) denotes the amount of specific latent heat, released during the water-to-ice phase transition.
Consider a Stefan problem of melting ice in the semi-infinite domain, \( x \in [0, \infty) \). Let the phase change be at \( x = S \). Let the temperature of the ice at \( t = 0 \) be given by \( T_0 < 0 \). It
6. LONG-TIME BEHAVIOUR OF SOLUTIONS

follows that the temperature for \( x \to \infty \) also equals \( T_0 \). At time \( t = 0 \) the temperature at \( x = 0 \) is suddenly increased to \( T_1 > 0 \). We have

\[
\begin{align*}
\frac{\partial T_{wa}}{\partial t} &= k_{wa} \frac{\partial^2 T_{wa}}{\partial x^2}, & 0 < x < S(t), \ t > 0, & T_{wa}(0, t) = T_1, \\
\frac{\partial T_{ic}}{\partial t} &= k_{ic} \frac{\partial^2 T_{ic}}{\partial x^2}, & S(t) < x < \infty, \ t > 0, & T_{ic}(x, 0) = T_{ic}(\infty, t) = T_0, \\
\alpha S &= \left[ -k_{wa} \frac{\partial}{\partial x} T_{wa} + k_{ic} \frac{\partial}{\partial x} T_{ic} \right]_{x=S(t)}, & T_{ic}(x, t) = T_{wa}(x, t) = 0 & x=S(t).
\end{align*}
\]

Inspired by the previous solution we find solutions of the form

\[
S = 2\gamma \sqrt{t}, \quad T_{wa} = T_1 \left( 1 - \frac{\text{erf}(x/2\sqrt{k_{wa}t})}{\text{erf}(\gamma_{wa})} \right), \quad T_{ic} = T_0 \left( 1 - \frac{\text{erfc}(\gamma_{ic})}{\text{erfc}(\gamma_{ic})} \right)
\]

where \( \gamma_{wa} = \gamma/\sqrt{k_{wa}} \) and \( \gamma_{ic} = \gamma/\sqrt{k_{ic}} \). The constant \( \gamma \) is to be determined from the equation

\[
\alpha = T_1 \exp(-\gamma_{wa}^2) \gamma_{wa} \text{erf}(\gamma_{wa}) + T_0 \exp(-\gamma_{ic}^2) \gamma_{ic} \text{erfc}(\gamma_{ic}).
\]

From a physical point of view it is more sensible to use the enthalpy, rather than the temperature as dependent variable for the melting ice problem, as we have both “sensible” temperature and latent heat. If the enthalpy, being the sum of both is given as a function of the temperature. This then closes the equations. Let us denote this enthalpy by \( H \), then we have for given \( H = H(T) \)

\[
\begin{align*}
\frac{\partial H}{\partial t} &= \frac{\partial^2 T}{\partial x^2}, & 0 < x < S(t), \ t > 0, & (5.51a) \\
T(0, t) - \beta \frac{\partial T}{\partial x}(0, t) &= 1, & (5.51b) \\
T(x, 0) &= 0. & (5.51c)
\end{align*}
\]

To start with we know the temperature at \( t = 0 \) say, and so we know the initial enthalpy as well. Two typical graphs of \( H(T) \) are given in Figure 10.3. One shows a simple discontinuous \( H \), with a discontinuity at \( T = 0 \) between \( H_- \) and \( H_+ \), whereas the other one shows a \( H \) of a material with a “mushy region”, i.e. where the phase change is more gradually taking place (as happens in melting of alloys).

6 Long-time behaviour of solutions

The solution of parabolic problems are typically diffusive and smoothing steep gradients. As a result, any initial value tends to be “forgotten”, and is therefore sometimes not as important as the long-time behaviour. We will consider some occasions where this is the case.

6.1 Linear initial boundary value problem

Parabolic equations are of special interest to study phenomena which tend to a *steady state* situation or *equilibrium*. In fact, when considering the PDE

\[
\frac{\partial u}{\partial t} = \mathcal{L}[u],
\]

(6.52)

where \( \mathcal{L} \) is a second order differential operator, it is natural to ask for solutions for which

\[
\frac{\partial u}{\partial t} \to 0,
\]

i.e. \( u \) will ultimately satisfy the *stationary elliptic equation*

\[
\mathcal{L}[u] = 0.
\]

This shows that there is an intimate relation between some parabolic and elliptic PDEs. Indeed, one often even tries to solve a difficult elliptic boundary value problem numerically by embedding it in a corresponding parabolic problem like (6.52) and solve the latter by implicit time stepping. This so-called *false transient* approach is the basis of quite a few numerical methods. As we saw in example 10.7 this will not always work and the problems should be analyzed carefully. At least the elliptic problem should have a solution, and the solution of the associated parabolic problem should converge for large time to this solution.

We shall illustrate the long-time behaviour of solutions of parabolic problems by a simple 1-D case, to start with linear ones in this subsection. Consider for \( U(x) \) the steady elliptic problem

\[
\mathcal{L}[U] = a \frac{\partial^2 U}{\partial x^2} + b \frac{\partial U}{\partial x} + c U = 0, \quad x \in (0, 1),
\]

(6.53a)

\[
U(0) = p, \quad U(1) = q.
\]

(6.53b)
where $a > 0$, $b$, $c$, $p$ and $q$ are constants. Suppose that we try to approach this solution by the auxiliary parabolic problem

\[
\frac{\partial y}{\partial t} = a \frac{\partial^2 y}{\partial x^2} + b \frac{\partial y}{\partial x} + cy, \quad x \in (0, 1), \ t > 0, \quad (6.54a)
\]

\[
y(x, 0) = v(x) + U(x), \quad x \in (0, 1), \quad (6.54b)
\]

\[
y(0, t) = p, \quad y(1, t) = q, \quad t > 0, \quad (6.54c)
\]

where $v$ is an arbitrary, in general unknown, but reasonably smooth, initial error to the finally sought solution. We are interested to know under which conditions solution $y$ asymptotically tends to $U$. In other words, when does the difference $u = y - U$, satisfying

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + cu, \quad x \in (0, 1), \ t > 0, \quad (6.55a)
\]

\[
u(x, 0) = v(x), \quad x \in (0, 1), \quad (6.55b)
\]

\[
u(0, t) = u(1, t) = 0, \quad t > 0, \quad (6.55c)
\]

tends to zero for $t \to \infty$.

After trying a transformation of the type $u(x, t) = e^{at-\beta x} \tilde{u}(x, \gamma t)$ we find that the transformation $u(x, t) = \exp[\frac{c}{\alpha^2}t - \frac{\beta}{\alpha^2}x] \tilde{u}(x, at)$ reduces the problem to the one in Example 10.8. So we deduce a Fourier series solution of (6.55) of the following form

\[
u(x) e^{\frac{\beta}{\alpha^2}t} = \sum_{k=1}^{\infty} A_k e^{-\pi^2k^2at} \sin(k\pi x), \quad (6.56)
\]

where the coefficients $A_k$ are found from the Fourier sine series expansion

\[
v(x) e^{\frac{\beta}{\alpha^2}t} \sum_{k=1}^{\infty} A_k \sin(k\pi x), \quad (6.57)
\]

in order to satisfy condition (6.55b). Although in general (6.57) is not uniformly convergent, the coefficients will at least decay like $A_k = O(k^{-1})$ as $k \to \infty$. So, with the help of the exponential, series (6.56) will converge uniformly for any $t > 0$, and indeed $u(0, t) = u(1, t) = 0$ (Appendix C).

We conclude that $u(x, t) \to 0$ for $t \to \infty$ if the least attenuated mode tends to zero, i.e. if $c < \frac{\beta^2}{\alpha^2} + \pi^2a$. In this case solution $y(x, t)$ will approach the stationary solution $U(x)$.

The corresponding problem with boundary conditions of Neumann type is more involved. Consider the initial boundary value problem

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + cu, \quad x \in (0, 1), \ t > 0, \quad (6.58a)
\]

\[
u(x, 0) = v(x), \quad x \in (0, 1), \quad (6.58b)
\]

\[
u(0, t) = \frac{\partial u}{\partial x}(1, t) = 0, \quad t > 0. \quad (6.58c)
\]
After a similar transformation as before, and separation of variables we can construct a Fourier-type series expansion

$$ u(x, t) = A_0 e^{ct} + e^{(c - \frac{b^2}{4a} t) x} \sum_{k=1}^{\infty} A_k e^{-\pi^2 k^2 at} f_k(x), $$

where $\{f_k(x)\}$ form the orthonormal set

$$ f_k(x) = \left(\frac{2}{4a^2 \pi^2 k^2 + b^2}\right)^{\frac{1}{2}} \left(2a \pi k \cos k \pi x + b \sin k \pi x\right). $$

Except for $A_0$ (note: the solution is not unique), the coefficients $A_k$ are determined by

$$ A_k = \int_0^1 (v(x) - A_0) e^{\frac{b^2}{4a} x} f_k(x) \, dx. $$

We conclude that the restrictions on $c$ for a decaying $u$ are slightly more stringent than in the previous case. Unless we are able to make sure that $A_0 = 0$, we must have $c < 0$ for $u$ to tend to zero for large $t$.

**Example 10.12** Let $\Omega := [0, 1] \times [0, 1]$. Consider the following initial boundary value problem

$$ \begin{align*}
\frac{\partial u}{\partial t} &= \nabla^2 u + 1, \quad x \in \Omega, \quad t > 0 \\
u(x, y, t) &= 0, \quad x \in \partial \Omega, \quad t > 0, \\
u(x, y, 0) &= 0, \quad x \in \Omega.
\end{align*} $$

We shall use the Duhamel integral (see Theorem 4.16). To this end we have to find a solution of

$$ \begin{align*}
\frac{\partial w}{\partial t}(x, t; \tau) &= \nabla^2 w(x, t; \tau), \quad x \in \Omega, \quad t > \tau, \\
w(x, t; \tau) &= 0, \quad x \in \partial \Omega, \quad t > \tau, \\
w(x, \tau; \tau) &= 1, \quad x \in \Omega.
\end{align*} $$

Because of the boundary conditions it convenient to have Fourier basis modes of the form $\sin(j \pi x) \sin(k \pi y)$ only. We find

$$ 1 = 16 \sum_{j,k \geq 0} \frac{\sin \alpha_j x \sin \alpha_k y}{\alpha_j \alpha_k}, \quad \text{where} \quad \alpha_j := (2j + 1)\pi. $$

Using e.g. the dispersion relation we find the (uniformly converging) expansion

$$ w(x, y, t; \tau) = 16 \sum_{j,k \geq 0} \frac{\sin \alpha_j x \sin \alpha_k y}{\alpha_j \alpha_k} e^{-\alpha_j^2 + \alpha_k^2 (t - \tau)}. $$

Hence

$$ u(x, y, t) = 16 \sum_{j,k \geq 0} \frac{\sin \alpha_j x \sin \alpha_k y}{\alpha_j \alpha_k} \int_0^t e^{-\alpha_j^2 + \alpha_k^2 (t - \tau)} \, d\tau. $$
6. LONG-TIME BEHAVIOUR OF SOLUTIONS

If \( t \to \infty \) we find
\[
\begin{align*}
  u(x, y, t) & \to 16 \sum_{j,k \geq 0} \frac{\sin(\alpha_j x) \sin(\alpha_k y)}{\alpha_j \alpha_k (\alpha_j^2 + \alpha_k^2)} =: v(x, y).
\end{align*}
\]

By direct substitution and noting the uniform convergence we may verify immediately that this
is the solution of the Poisson problem
\[
\begin{align*}
  \nabla^2 v &= 1, \quad x \in \Omega, \\
  v(x, y) &= 0, \quad x \in \partial \Omega.
\end{align*}
\]

6.2 Equilibrium and travelling-wave solutions for nonlinear problems

Quite often diffusion equations arise in the analysis of chemical reactions, population dy-
namics or the modelling of epidemics. Typically, the equation includes a non-linear source
or reaction term and takes on the form
\[
\frac{\partial u}{\partial t} = \nabla \cdot (D \nabla u) + r(u). \tag{6.62}
\]

The reaction term \( r \) arises from the application. For example, \( u \) may be a temperature and \( r \)
the energy release from the chemical reaction; or \( u \) may be a concentration of some species
and \( r \) a source or sink. A common model for \( r \) is the polynomial
\[
r(u) = \gamma (u - \alpha)(\beta - u) \tag{6.63}
\]
where \( \alpha < \beta \) and \( \gamma > 0 \). If we further assume that \( D \) is a positive constant, we may rescale
this equation by the transformation
\[
\tilde{u} := \frac{u - \alpha}{\beta - \alpha}, \quad \tilde{x} := \left( \frac{\gamma}{D} (\beta - \alpha) \right)^{1/2} x, \quad \tilde{t} := \gamma (\beta - \alpha) t \tag{6.64}
\]
(where we will omit the tilde henceforth) into the standard form
\[
\frac{\partial u}{\partial t} = \nabla^2 u + u(1 - u). \tag{6.65}
\]

This is known as Fisher's equation. It was introduced originally to model the spread of a
gene in a population [135].

For demonstrating the basic ideas, we restrict our discussion of (6.62) to the 1-D case
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + r(u). \tag{6.66}
\]

Clearly, wherever \( r(u) \) is positive it represents a source, and a sink where it is negative.
Any zero of \( r \), say \( u = u_0 \), is evidently also a solution of the equation if the boundaries
are compatible with this solution, for example if the walls are isolated such that no heat or
matter is lost. These stationary solutions are so-called equilibrium solutions. It is not clear
in advance if any such solution will be attained for \( t \to \infty \) if we start in its neighbourhood. This will be seen to be critically dependent on the sign of \( r'(u_0) \). Let us therefore study the IBVP – with isolated boundaries and arbitrary initial condition – for a perturbation \( \delta e(x, t) \) of \( u = u_0 \), given by

\[
\frac{\partial u}{\partial t} := u_0 + \delta e(x, t),
\]

(6.67)

where \( \delta \) is small and \( r'_0 := r'(u_0) \neq 0 \). After linearisation of \( r(u) = \delta r'_0 e + \ldots \) (where we neglect any term smaller than \( O(\delta) \)), we have the following problem for \( e(x, t) \)

\[
\frac{\partial e}{\partial t} = \frac{\partial^2 e}{\partial x^2} + r'_0 e, \quad x \in (0, 1),
\]

(6.68a)

\[
e(x, 0) = v(x), \quad x \in (0, 1),
\]

(6.68b)

\[
\frac{\partial e}{\partial x}(0, t) = \frac{\partial e}{\partial x}(1, t) = 0, \quad t > 0.
\]

(6.68c)

This has the same form as the problem we studied in (6.58). We can immediately infer that \( u = u_0 \) is an asymptotically stable equilibrium, i.e. \( e \to 0 \) as \( t \to \infty \), if \( r'_0 < 0 \). Similarly, it is asymptotically unstable if \( r'_0 > 0 \). In particular, for the Fisher equation (6.65) we obtain that \( u(x, t) \equiv 1 \) is stable, while \( u(x, t) \equiv 0 \) is unstable.

It should be noted that other equilibrium solutions than the zeros of \( r(u) \) are in principle also possible. They are, however, in general more difficult to analyse.

**Example 10.13** The stability of solutions is an important issue in many reaction-diffusion equations. A typical question is whether the solution becomes unbounded in finite time, so-called “blow up”. We shall give a simple example of such a problem. Consider the PDE with IC and BC

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u^3, \quad x \in (0, 1), \quad t > 0,
\]

\[
u(x, 0) = 1, \quad x \in (0, 1),
\]

\[
\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(1, t) = 0, \quad t > 0.
\]

One easily verifies that a solution exists which is independent of \( x \) and hence given by

\[
u(x, t) = \frac{1}{\sqrt{1-2t}}.
\]

It has the remarkable property that \( u \to \infty \) for \( t \to \frac{1}{2} \). \[\square\]

Reaction-diffusion equations allow for further analysis by looking for travelling-wave solutions (see Section 2.5). These may be considered as similarity solutions of a particular type (Section 7.3.3, Example 7.16). As an example we consider Fisher’s equation

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1 - u), \quad x \in (-\infty, \infty), \quad t > 0.
\]

(6.69)

We now try to find similarity solutions of the following form

\[
U(\xi) := U(x/c - t) = u(x, t).
\]
6. LONG-TIME BEHAVIOUR OF SOLUTIONS

Figure 10.4. Phase portrait of Fisher’s travelling-wave problem (6.71) with $c = 2.25$. Arrows indicate positive $\xi$-direction. Note the trajectory that connects saddle point $(1, 0)^T$ with stable node $(0, 0)^T$.

By substituting this in (6.69) we obtain the autonomous ordinary differential equation

$$U'' + c^2 U' + c^2 U(1 - U) = 0, \quad \xi \in (-\infty, \infty),$$

(6.70)

where the prime denotes differentiation with respect to $\xi$. The physically interesting solutions are those that remain finite for $\xi \to \pm \infty$. Note that $U$ depends on $c^2$ so the behaviour for left and right running waves is the same.

To facilitate the analysis it is useful to rewrite (6.70) as a first order system, i.e. with

$$\begin{align*}
y_1 &= U, \\
y_2 &= U'.
\end{align*}$$

$$y_1' = y_2, \quad y_2' = -c^2 y_2 - c^2 y_1 (1 - y_1).$$

(6.71)

Insight into the behaviour of possible solutions is obtained by using the phase plane (see e.g. [77]), i.e. graphs (trajectories) of $y_2$ as a function of $y_1$; see Figure 10.4. Of particular importance are the stationary points $(0, 0)^T$ and $(1, 0)^T$. The Jacobi matrix of the system (6.71), linearised around $(0, 0)^T$, is given by

$$\begin{pmatrix} 0 & 1 \\ -c^2 & -c^2 \end{pmatrix}.$$  

The corresponding eigenvalues

$$\lambda_{1,2} = -\frac{1}{2}c^2 \pm \frac{1}{2}\sqrt{c^4 - 4c^2}$$

are negative if $|c| \geq 2$, and complex with a negative real part otherwise. Hence the point $(0, 0)^T$ is a stable spiral point for $|c| < 2$ and a stable node for $|c| \geq 2$. At the other stationary point $(1, 0)^T$ one finds the Jacobi matrix

$$\begin{pmatrix} 0 & 1 \\ c^2 & -c^2 \end{pmatrix}.$$
CHAPTER 10. ANALYSIS OF PARABOLIC EQUATIONS

with the real eigenvalues
\[ \lambda_{1,2} = -\frac{1}{2}c^2 \pm \frac{1}{2} \sqrt{c^4 + 4c^2}, \]
one of which is always negative and the other always positive. So this point is an unstable saddle point (see Figure 10.4). Note that both conclusions for \((0, 0)^T\) and \((1, 0)^T\) are entirely in agreement with the results of (6.68).

Most of the trajectories are physically not interesting, because they tend to infinity for either \(\xi \to \infty\) or \(-\infty\). For each \(c\) there is one, however, that starts at \((1, 0)^T\) and ends at \((0, 0)^T\). This trajectory (see Figure 10.4) corresponds to a solution we are looking for. This solution may be understood as a transition of the system from the unstable stationary state \(u(x, t) = 0\) to the stable stationary state \(u(x, t) = 1\).

For models where \(u\) cannot be negative we have to assume additionally that \((0, 0)^T\) is no spiral point, i.e. \(|c| \geq 2\). In Figure 10.5 we have depicted the right-running wave for the value \(c = 2.25\) and normalized by \(U(0) = \frac{1}{2}\). As the activity of the wave is exponentially localized around \(x - ct = 0\) (cf. Equation 6.59) this travelling-wave solution is sometimes called a solitary wave or soliton.

Figure 10.5. The solitary wave of Fisher’s travelling-wave problem (6.71) with \(c = 2.25\) and \(U(0) = \frac{1}{2}\).

Exercises

10.1. Consider the following initial boundary value problem

\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - cu + \sin \pi x + \frac{1}{8} \sin 3\pi x, \quad x \in (0, 1), \quad t > 0, \]
\[ u(x, 0) = 0, \quad x \in (0, 1), \]
\[ u(0, t) = u(1, t) = 0, \quad t > 0. \]
Exercises

Find its solution $u(x, t)$.

10.2. Consider the initial boundary value problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in (0, 1), \quad t > 0,$$

$$u(x, 0) = v(x), \quad x \in (0, 1),$$

$$u(0, t) = u(1, t) = 0, \quad t > 0.$$

(a) Find the “Green’s” function $w(x, t; \tau)$ such that

$$u(x, t) = \int_0^t w(x, t; \tau) v(\tau) \, d\tau.$$

(b) Determine $u(x, t)$ for $v(x) = \sin^2(\pi x)$.

10.3. Consider the problem of Exercise 2, now with BC

$$\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(1, t) = 0, \quad t > 0.$$

(a) Compute the solution $u(x, t)$.

(b) Determine $\lim_{t \to \infty} u(x, t)$ both from the expression under (a) and by direct computation.

10.4. Solve the initial boundary value problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - cu + \frac{1}{2}(1 - \cos \pi x), \quad x \in (0, 1), \quad t > 0,$$

$$u(x, 0) = 2(1 - \cos \pi x), \quad x \in (0, 1),$$

$$\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(1, t) = 0, \quad t > 0.$$

10.5. Solve the initial boundary value problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in (-1, 1), \quad t > 0,$$

$$u(x, 0) = 1, \quad x \in (-1, 1),$$

$$u(-1, t) = u(1, t) = 0, \quad t > 0.$$

10.6. A sphere with radius $R$ has a constant initial temperature $u_0$. At $t = 0$ the sphere is cooled at $u = 0$.

(a) Show that the temperature $u(r, t)$ satisfies the initial boundary value problem

$$\frac{\partial u}{\partial t} = k \left( \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} \right), \quad r \in (0, R), \quad t > 0,$$

$$u(r, 0) = u_0, \quad r \in (0, R),$$

$$\frac{\partial u}{\partial r}(0, t) = u(R, t) = 0, \quad t > 0.$$
(b) Determine \( u(r, t) \).

10.7. Use Green’s functions to solve the following problem, defined for \( 0 < x < L \) (for some \( L \)) and \( t > 0 \)
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + g(x, t), \quad x \in (0, L), \ t > 0,
\]
\[
u(x, 0) = v(x), \quad x \in (0, L),
\]
\[
u(0, t) = 0, \ \frac{\partial u}{\partial x}(L, t) + u(L, t) = 0, \quad t > 0.
\]

10.8. Consider the following initial boundary value problem for \( x > 0, \ t > 0 \)
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x > 0, \ t > 0,
\]
\[
u(x, 0) = 0, \quad x > 0,
\]
\[
u(0, t) = t^\alpha, \quad t > 0,
\]
for some parameter \( \alpha > 0 \). Find a similarity solution \( \hat{u}(\hat{t}) \) of the form
\[
u(x, t) = t^\alpha \hat{u}(\hat{t}).
\]

10.9. Find the solution of the following initial boundary value problem with periodic boundary conditions
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in (-1, 1), \ t > 0,
\]
\[
u(x, 0) = v(x), \quad x \in (-1, 1),
\]
\[
u(-1, t) = u(1, t), \ \frac{\partial u}{\partial x}(-1, t) = \frac{\partial u}{\partial x}(1, t) \quad t > 0.
\]

10.10. Find a similarity solution, for \( x > 0, \ t > 0 \) of
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 1, \quad x > 0, \ t > 0,
\]
\[
u(x, 0) = 0, \quad x > 0,
\]
\[
u(0, t) = 0, \quad t > 0.
\]

10.11. Let \( 0 < a < \frac{1}{2} \).

(a) Find an ODE for the travelling-wave solution \( \hat{u}(\xi) := \hat{u}(x - bt) \) of
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - u(1 - u)(a - u).
\]
(b) Draw a phase plane for \( \hat{u} \) and investigate when a solution exists subject to the conditions
\[
\hat{u}(\xi) \rightarrow 1 \quad \text{for} \quad \xi \rightarrow \infty, \quad \hat{u}(\xi) \rightarrow a \quad \text{for} \quad \xi \rightarrow -\infty.
\]
10.12. Use separation of variables to compute the solution of
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < 1, \ t > 0,
\]
\[u(x, 0) = \sin(\pi x) \cos(\pi x), \quad 0 < x < 1,
\]
\[u(0, t) = u(1, t) = 0, \quad t > 0.\]

10.13. Use separation of variables to compute the solution of
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < 1, \ t > 0,
\]
\[u(x, 0) = \sin(\pi x) \cos(\pi x), \quad 0 < x < 1,
\]
\[u(0, t) = 0, \quad \frac{\partial u}{\partial x}(1, t) = 0, \quad t > 0.\]

10.14. Consider the initial boundary value problem
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \alpha u, \quad 0 < x < 1, \ t > 0,
\]
\[u(x, 0) = v(x), \quad 0 < x < 1,
\]
\[u(0, t) = u(1, t) = 0, \quad t > 0.\]

(a) Determine the formal solution.

(b) Show that this solution is stable for \(\alpha < \pi^2\).

10.15. Consider the problem
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \alpha u, \quad 0 < x < 1, \ t > 0,
\]
\[u(x, 0) = 0, \quad 0 < x < 1,
\]
\[u(0, t) = u(1, t) = 0, \quad t > 0.\]

(a) Determine the stationary solution, i.e. the solution for \(t \to \infty\).

(b) Use the result in (a) to compute the solution.
Chapter 12

Analysis of hyperbolic equations

This chapter is devoted to hyperbolic problems. We mainly restrict ourselves to equations with one space variable only. In Section 1 we start with first order scalar equations and describe solutions in terms of characteristics. Difficulties are encountered for nonlinear equations, for which characteristics may intersect or fan out. We introduce the corresponding solutions, viz. shock waves and rarefaction waves. It turns out that a classical approach, with smooth solutions, is not suited to treat these phenomena. For that reason, we introduce in Section 2 so-called weak solutions. As an important example, we investigate the solution of the Riemann problem, which is an initial value problem with piecewise constant initial condition. Extension of the previous to systems is discussed in the next two sections. First, in Section 3, we introduce the definition of hyperbolicity. A system is hyperbolic if its Jacobi matrix has real eigenvalues and linearly independent eigenvectors. This then means that the system can be diagonalised. Subsequently, we can apply the scalar theory to each of the resulting equations. Furthermore, for a linear system of two unknowns we give a concise description of the method of characteristics. Next, in Section 4, we consider weak solutions of the Riemann problem. In particular, we introduce the elementary solutions, viz. shock waves, rarefaction waves and contact discontinuities. In Section 5 we apply this theory to the shallow water equations. In Section 6 the wave equation is treated. We derive the solution in one space dimension first and study next higher dimensional problems. The last section deals with choosing the proper boundary conditions. Since the characteristics determine the propagation directions of the solution, these determine what boundary conditions can be prescribed.

1 First-order scalar equations

In Chapter 2 we introduced first-order scalar equations. In this section we shall have a closer look at these equations, to start with the linear case. After that we deal with the nonlinear case in more detail. A crucial element in the construction of solutions will turn out to be the behaviour of the characteristics.
1. FIRST-ORDER SCALAR EQUATIONS

1.1 Semilinear equations

Consider the first-order scalar equation, cf. (2.1.1),

\[ \frac{\partial u}{\partial t} + b \frac{\partial u}{\partial x} = c, \quad (a, b \neq 0). \]  \hspace{1cm} (1.1)

Equation (1.1) is a scalar *hyperbolic* equation. In this subsection we allow the coefficients \(a, b\) and \(c\) to depend on \(x\) and \(t\) but not on \(u\), so that we have a linear equation in \(u\). We showed in Chapter 2 that the partial differential equation (1.1) leads to the following set of ordinary differential equations

\[ \frac{dx}{dt} = \frac{b}{a}, \]  \hspace{1cm} (1.2a)

\[ \frac{du}{dt} = \frac{c}{a}. \]  \hspace{1cm} (1.2b)

Equation (1.2a) defines the location of the (base) characteristics \(C\) in the \((x, t)\)-plane and (1.2b) the solution \(u\), as a function of \(t\), along these base characteristics. In the following we will omit the adjective base, and use the term characteristics to denote solution curves of (1.2a) in the \((x, t)\)-plane. In order to find solutions of (1.1), it will be sufficient to study the ordinary differential equations (1.2). Let us first consider the homogeneous problem, i.e. \(c(x, t) = 0\). If \(u\) is given at a single point on some characteristic \(C\), it is completely determined along \(C\) by virtue of (1.2b). Consequently, we may not prescribe a solution \(u(x, t)\) on a characteristic unless it satisfies (1.2b). The solution would then exist on this single characteristic only. Thus, consider a curve \(J\), not being a characteristic that intersects each characteristic at most once. So for some interval \(I \subset \mathbb{R}\) let

\[ J := \{(x(\sigma), t(\sigma)) \mid \sigma \in I\}, \]  \hspace{1cm} (1.3)

then we may prescribe \(u\) on \(J\), say

\[ u(x(\sigma), t(\sigma)) = v(\sigma), \quad \sigma \in I, \]  \hspace{1cm} (1.4)

where \(v\) is some given function. We can now compute \(u\) from the differential equations (1.2) and the condition (1.4).

Next, we assume that \(a\) and \(b\) are constant. Special cases are when \(J\) coincides with either the \(x\)-axis or \(t\)-axis. In the first case we have

\[ J = \{(x(\sigma), 0)^T \mid \sigma \in \mathbb{R}\}, \]

and we may identify \(\sigma\) with \(x\). The characteristic intersecting \(J\) at a point \((x_0, 0)^T\), say, is given by

\[ x - \frac{b}{a} t = x_0. \]

The solution along this characteristic reads

\[ u(x, t) = v(x_0) = v\left(x - \frac{b}{a} t\right). \]  \hspace{1cm} (1.5)
In a similar vein, if we would have

$$\mathcal{J} = \{ (0, t(\sigma))^T \mid \sigma \geq 0 \},$$

we may identify $\sigma$ with $t$. At the characteristic intersecting $\mathcal{J}$ at a point $(0, t_0)^T$, we then have

$$t - \frac{a}{b}x = t_0.$$  

Consequently, for the solution along this characteristic we find

$$u(x, t) = v(t_0) = v(t - \frac{b}{a}x).$$  

(1.6)

Note that $b/a$ can be interpreted as a velocity, implying that the solution in (1.5) is just the initial solution $v(x)$ propagated over a distance $bt/a$. A similar interpretation holds for (1.6).

**Example 12.1** Consider the partial differential equation

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

The characteristics of this equation are given by $x - t = C$ ($C \in \mathbb{R}$). If we prescribe $u(x, t) = v(t)$ on $C = \{(x, t) \mid x = t\}$, i.e. the characteristic through the origin, we apparently need $v(t)$ to be constant. The solution then only exists on $C$. Prescribing $u(x, t)$ along any line $\mathcal{J}$ that is not parallel to this curve $C$ will be meaningful then. Suppose we take for $\mathcal{J}$ the $x$-axis and choose

$$u(x, 0) = \sin x.$$  

Clearly we then have

$$u(x, t) = \sin(x - t),$$  

implying that $u(x, t)$ is constant along the characteristics indeed. \(\square\)

If $a$ and $b$ are not constant and $c \neq 0$, we can still find an integral representation of the solution along characteristics from (1.2b). Suppose a characteristic $C$ intersects the 'initial' line $\mathcal{J}$ at the point $(x_0, t_0)^T = (x(s_0), t(s_0))^T$. We can express the value $u(x, t)$, for $(x, t)^T \in C$, in terms of the 'initial' value $u(x(s_0), t(s_0))$ as follows:

$$u(x, t) = u(x(s_0), t(s_0)) + \int_{t_0}^{t} \frac{c(x(\tau), \tau)}{a(x(\tau), \tau)} \, d\tau. \quad (1.7)$$

In particular, if $u$ is given along the $x$-axis, i.e.

$$u(x, 0) = v(x),$$  

and if the coefficients $a$ and $b$ are constant, we find from (1.7)

$$u(x, t) = v(x_0) + \frac{1}{a} \int_{0}^{t} c(x(\tau), \tau) \, d\tau$$  

$$= v(x - \frac{b}{a}t) + \frac{1}{a} \int_{0}^{t} c(x - \frac{b}{a}(t - \tau), \tau) \, d\tau,$$  

(1.8)

where we have used that $x - \frac{b}{a}t = x_0 = x(\tau) - \frac{b}{a}\tau$ along characteristics. Also for more general $c$, i.e. $c = c(x, t, u)$, in which case the partial differential equation is sometimes called *semilinear*, it is fairly simple to find a solution; see the next example.
Example 12.2 Consider the initial value problem

\[ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = -u, \]
\[ u(x, 0) = \sin x. \]

From (1.2) we conclude that

\[ u(x(t), t) = Ce^{-t} \quad (C \in \mathbb{R}) \]

along the characteristic with \( x(t) - t = x_0 \). Applying the initial condition we obtain

\[ u(x_0, 0) = \sin x_0 = C. \]

So the solution is given by

\[ u(x, t) = e^{-t} \sin(x - t). \]

\[ \Box \]

If \( a \) and \( b \) are not constant the characteristics are not straight lines in general. Yet, the preceding outline for the construction of a solution is still valid, as is shown by the following example.

Example 12.3 Consider the initial value problem

\[ \frac{\partial u}{\partial t} + x \frac{\partial u}{\partial x} = 0, \]
\[ u(x, 0) = \sin x. \]

From (1.2) we see that the characteristic through the 'initial' point \((x_0, 0)\) is given by \( x(t) = x_0 e^t \), along which \( u(x(t), t) = C \). From the initial condition we find \( u(x_0, 0) = \sin x_0 = C \) and consequently, the solution is given by

\[ u(x, t) = \sin(x e^{-t}). \]

\[ \Box \]

It is important to realise that the 'initial condition' function \( v \) (the function of initial values) does not need to be smooth, or even continuous. Indeed, it is just a representation of a collection of initial values for the solution defined on the various characteristics. If \( v \) has a discontinuity, then the solution is not smooth either. In fact one may wonder whether a solution can still satisfy (1.1). We shall see that it can, in a so-called weak sense, in Section 2.

1.2 Quasilinear equations

If the coefficients \( a, b \) and \( c \) depend on \( u \) as well, we call the partial differential equation quasilinear; cf. example 1.1.2. Since we have assumed that \( a \neq 0 \), it is not restrictive to take \( a(x, t, u) \equiv 1 \). So consider the equation

\[ \frac{\partial u}{\partial t} + b(x, t, u) \frac{\partial u}{\partial x} = c(x, t, u). \quad (1.9) \]

Moreover, if we assume that \( b = b(u) \), i.e. \( b \) does not explicitly depend on \( x \) or \( t \), we can define a flux function \( f(u) \) as follows:

\[ f(u) := \int_{u_0}^u b(v) \, dv, \quad (1.10) \]
where $u_0$ is some reference value for $u$. We rather rewrite equation \(1.9\) in so-called conservation form as follows:

$$
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = c(x, t, u). \tag{1.11}
$$

Equation (1.11) has the form of a conservation equation. It is also known as a transport equation.

Below we consider the case $c(x, t, u) \equiv 0$. In order to determine a solution of equation (1.11), or equivalently (1.9), we proceed like before. So the governing equations for the solution along a characteristic $C$ read

$$
\frac{dx}{dt} = b(u) = f'(u), \tag{1.12a}
$$

$$
\frac{du}{dt} = 0. \tag{1.12b}
$$

From (1.12a) we see that the location of characteristics depends on $u$, in contrast to the linear case where we could compute the characteristics independently from the solution $u$. We can, in principle, compute $u$ from these equations, if $u$ is given on some 'initial' curve $\mathcal{J}$ intersecting the characteristics at most once. Let $\mathcal{J}$ be the $x$-axis, so

$$u(x, 0) = v(x), \tag{1.13}$$

for some given function $v$. From (1.12b) we conclude that $u$, and therefore also $b(u)$, is constant along a characteristic. Integration of the differential equations in (1.12) is then trivial, and we find the following solution

$$x - x_0 = b(v(x_0)) t, \tag{1.14a}$$

$$u(x, t) = v(x_0) = v(x - b(v(x_0)) t), \tag{1.14b}$$

which holds on the characteristic through the point $(x_0, 0)^T$; see also example 2.2.2. Equation (1.14a) implicitly defines $x_0$ as a function of $x$ and $t$, i.e. $x_0 = x_0(x, t)$. By substituting the latter relation into (1.14b), we can find the solution $u(x, t)$.

A well known example of (1.9) is the (inviscid) Burgers’ equation. (In Chapter 12 and ?? we simply write Burgers’ equation to denote the inviscid Burgers’ equation.)

**Example 12.4** Consider the Burgers’ equation

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

Clearly $b(u) = u$ and we have $u(x, t) = v(x - v(x_0) t)$ along the characteristics given by $x - x_0 = v(x_0) t$. $\square$

As before, we note that the initial condition is propagated along characteristics, now with a speed that depends on the location. This induces a problem, certainly from a mathematical point of view.
1. FIRST-ORDER SCALAR EQUATIONS

Figure 12.1. Characteristics of the Burgers’ equation that intersect (left) or fan out (right).

Example 12.5 Consider the Burgers’ equation again. Let the initial condition \( v(x) \) be piecewise constant, say

\[
v(x) := \begin{cases} 
\alpha & \text{if } x < 0, \\
\beta & \text{if } x \geq 0.
\end{cases}
\]

The characteristic through a point \( (x_0, 0)^T \) is given by

\[x = x_0 + v(x_0)t, \quad x_0 \in \mathbb{R},\]

and has a slope \( \frac{dx}{dt} = 1/v(x_0) \). Therefore, we can encounter two typical situations, viz. \( \alpha > \beta \) and \( \alpha \leq \beta \). In the first case, the slope of characteristics emanating from the negative \( x \)-axis is smaller than the slope of characteristics emanating from the positive \( x \)-axis; see figure 12.1. This would imply a multivalued solution where characteristics intersect. In the second case the characteristics on the left have a larger slope than the characteristics on the right, leading to a wedge-shaped region in the \((x, t)\)-plane where the solution is not defined. \( \square \)

Apparently, the nonlinearity is causing the problems met in example 12.5. We emphasize that it is not necessarily a consequence of discontinuous initial data. To see this, we shall analyse the influence of the initial condition in more detail. For simplicity we restrict ourselves to the Burgers’ equation. Consider as an example the initial condition \( v(x) = \sin \pi x \) \((0 \leq x \leq 1)\) shown in figure 12.2; \( v(x) \) is monotonically increasing on \((0, \frac{1}{2})\) and monotonically decreasing on \((\frac{1}{2}, 1)\). From (1.14) we conclude that the initial condition \( v(x) \) is propagated along characteristics with velocity \( b(v(x)) = v(x) \). This means that the characteristics, emanating from \((0, \frac{1}{2})\), fan out and consequently the initial solution on this interval expands. On the other hand, characteristics originating from \((\frac{1}{2}, 1)\) are approaching each other, leading to a compression of the initial solution on this interval. This means that the left part of the solution overtakes the right part, leading to an increasingly steeper profile as shown in figure 12.2. The solution will eventually break down when \( \frac{dx}{dt} \) tends to infinity at some point \((x^*, t^*)^T\), say, where a discontinuity starts. We can compute \( t^* \) as follows. Consider the characteristic through \((x_0, 0)^T \) with \( x_0 \in (\frac{1}{2}, 1) \) where the initial solution is monotonically decreasing. The location of this characteristic is given by
Figure 12.2. The solution of (1.12) at \( t = 0, 0.2, t^* = 1/\pi \) and 0.8 for the initial condition \( v(x) = \sin \pi x \).

the relation (1.14a), from which we can deduce

\[
(1 + v'(x_0)t)\frac{\partial x_0}{\partial x} = 1. \tag{1.15}
\]

Furthermore, the solution \( u \) along this characteristic is implicitly given by (1.14b). From this relation and (1.15) we can compute \( \frac{\partial u}{\partial x} \) to find

\[
\frac{\partial u}{\partial x} = v'(x_0)\frac{\partial x_0}{\partial x} = v'(x_0)(1 + v'(x_0)t)^{-1}. \tag{1.16}
\]

Obviously \( \frac{\partial u}{\partial x} \to -\infty \) when \( 1 + v'(x_0)t^* = 0 \) for some \( t^* > 0 \). Note that this condition also implies that \( x_0 \) cannot be determined anymore from relation (1.14a). The time \( t^* \) when a discontinuity first emerges is thus given by

\[
t^* = -1/ \min_{0 \leq x \leq 1} v'(x), \tag{1.17}
\]

and only holds when \( v'(x) < 0 \) somewhere.

If we would compute the solution from (1.14) beyond the critical time \( t^* \) we would obtain the physically incorrect triple valued function shown in figure 12.2. To determine the
Proof. We have the problem
\[
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \varepsilon \frac{\partial^2 u}{\partial x^2},
\]
(1.18)
including the so-called viscous term \( \varepsilon \frac{\partial^2 u}{\partial x^2} \) with \( 0 < \varepsilon \ll 1 \), instead of (1.11). Equation (1.11) is an appropriate model only if \( \varepsilon \) is small and the solution is smooth. In this case the viscous term \( \varepsilon \frac{\partial^2 u}{\partial x^2} \) is negligible. However, when a discontinuity starts to develop, equation (1.11) looses its validity and we must return to (1.18). In the vicinity of the emerging discontinuity the term \( \varepsilon \frac{\partial^2 u}{\partial x^2} \) becomes gradually larger, thus balancing the left hand side in (1.18) and preventing break-down of the solution. For decreasing \( \varepsilon \), the solution becomes gradually steeper. In fact, one can prove that the \textit{vanishing viscosity solution}, for which \( \varepsilon \to 0 \), is the discontinuous solution discussed above [72, 61].

1.3 Nonlinear equations

Surprisingly, it appears that a generalisation is possible of the above results to the general, fully nonlinear case. Under relatively mild conditions of smoothness (derivatives should exist, two characteristics should not pass through the same point) and consistency of boundary conditions (no more than one boundary condition prescribed on the same characteristic) any first order scalar partial differential equation may be rewritten into characteristic form, i.e. as a system of ordinary differential equations.

\textbf{Theorem 12.6 (Characteristic form of 1st order PDE).} The \( n \)-th dimensional first-order nonlinear scalar partial differential equation in \( u = u(x) \), given by

\[
F(x, u, q) = 0, \quad q = \nabla u,
\]
where \( F \) is smooth and with consistent boundary values, may be recast into the following system of ordinary differential equations

\[
\frac{d\chi}{ds} = \frac{\partial F}{\partial q}, \quad \frac{du}{ds} = q \cdot \frac{\partial F}{\partial q}, \quad \frac{dq}{ds} = -q \frac{\partial F}{\partial u} - \frac{\partial F}{\partial x}
\]

(\( \frac{\partial F}{\partial q} \) denotes the gradient with respect to \( q \); similar for \( \frac{\partial F}{\partial x} \)), where the curve \( x = \chi(s) \), with parameter \( s \), is called a characteristic. Since \( s \) is only an auxiliary variable, other equivalent forms exist. They are easily constructed by varying the defining equation \( F = 0 \).

\textbf{Proof.} We have the problem \( F(x_1, \ldots, x_n, u, q_1, \ldots, q_n) = 0 \), where \( q_i = \frac{\partial u}{\partial x_i} \). We look for characteristics, given by \( x_i = \chi_i(s) \), along which the partial differential equation can be written as an ordinary differential equation. If we compare the expressions

\[
\frac{dF}{dx_i} = \sum_{j=1}^{n} \frac{\partial q_j}{\partial x_i} \frac{\partial F}{\partial q_j} + \frac{\partial u}{\partial x_i} \frac{\partial F}{\partial u} + \frac{\partial F}{\partial x_i} = 0
\]

\[
\frac{dq_i}{ds} = \sum_{j=1}^{n} \frac{\partial q_j}{\partial x_j} \frac{d\chi_j}{ds} = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j \partial x_i} \frac{d\chi_j}{ds} = \sum_{j=1}^{n} \frac{\partial q_j}{\partial x_i} \frac{d\chi_j}{ds},
\]
we observe that if we take
\[ \frac{d\chi_j}{ds} = \frac{\partial F}{\partial q_j} \]
we obtain
\[ \frac{dq_i}{ds} = \sum_{j=1}^{n} \frac{\partial q_j}{\partial x_i} \frac{d\chi_j}{ds} = \sum_{j=1}^{n} q_j \frac{\partial F}{\partial q_j} \]
and the problem is indeed rewritten in characteristic form.

Example 12.7 A famous example is the eikonal equation
\[ F(x, u, q) = q \cdot q - a^2(x, y) = 0, \quad q = \nabla u, \]
which becomes along the characteristics \( x = \chi(s) \)
\[ \frac{d\chi}{ds} = 2q, \quad \frac{du}{ds} = 2q \cdot q, \quad \frac{dq}{ds} = 2a \nabla a. \]
A much neater form, with the same characteristics but a different parametrization, is found if we take
\[ F(x, u, q) = \frac{1}{2} \left( \frac{q \cdot q}{a^2(x, y)} - 1 \right) = 0, \quad q = \nabla u. \]
Then we have
\[ \frac{d\chi}{ds} = \frac{q}{a^2}, \quad \frac{du}{ds} = \frac{q \cdot q}{a^2} = 1, \quad \frac{dq}{ds} = \frac{1}{a} \nabla a, \]
with \( u = s + u(0) \). The parametrization corresponds now with the level surfaces of \( u \).

2 Weak formulation of first-order scalar equations

As we saw in the previous section, a solution of (1.11) which is computed from the ODE system (1.12) is not necessarily continuous, let alone differentiable. This means that we have to reconsider our concept of a solution of equation (1.11). For this we need distribution theory; cf. Chapter 4.

2.1 Weak solutions

Let us start by observing that hyperbolic conservation equations are often derived in integral form, rather than as a differential equation. As an example, think of gas flowing in a tube of constant cross section. Let \( x \) denote the coordinate along the tube, \( \rho(x, t) \) and \( v(x, t) \) the mass density and flow velocity, respectively, at position \( x \) and time \( t \). Then conservation of mass in an arbitrary segment \( (x_1, x_2) \) is given by the relation
\[ \frac{d}{dt} \int_{s_1}^{s_2} \rho(x, t) \, dx = (\rho v)(x_1, t) - (\rho v)(x_2, t), \quad (2.1) \]
2. WEAK FORMULATION OF FIRST-ORDER SCALAR EQUATIONS

stating that the increase of mass in \((x_1, x_2)\) is balanced by the net influx of mass. If we replace \(\rho\) by a variable \(u\) and the mass flux \(\rho v\) by a generic flux \(f(u)\), equation (2.1) generalises to:

\[
\frac{d}{dt} \int_{x_1}^{x_2} u(x, t) \, dx = f(u(x_1, t)) - f(u(x_2, t)).
\]

(2.2)

By integrating this equation over an arbitrary time interval \([t_1, t_2]\) we find

\[
\int_{x_1}^{x_2} (u(x, t_2) - u(x, t_1)) \, dx = \int_{t_1}^{t_2} \left( f(u(x_1, t)) - f(u(x_2, t)) \right) \, dt.
\]

(2.3)

If \(u\) and \(f(u)\) are continuously differentiable, this equation is equivalent to

\[
\int_{t_1}^{t_2} \int_{x_1}^{x_2} \left( \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} \right) \, dx \, dt = 0.
\]

(2.4)

Since this equation should hold for arbitrary \(x_1, x_2, t_1\) and \(t_2\) the integrand has to be zero necessarily, i.e.

\[
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0.
\]

(2.5)

A function \(u\) is called a weak solution of (2.5), if it satisfies (2.3) for arbitrary \(x_1, x_2, t_1\) and \(t_2\). Note that a solution of (2.5) is always a solution of (2.3); the converse need not to be true.

Since the verification of (2.3) for arbitrary \(x_1, x_2, t_1\) and \(t_2\) is rather cumbersome, we prefer another definition of weak solution, which is based on distribution theory; see Chapter 4. Here, we define the space of test functions \(D\) as follows

\[
D = C^1_0(\mathbb{R} \times [0, \infty)) := \{ \varphi \in (C^1(\mathbb{R} \times [0, \infty)) \mid \varphi\ \text{has compact support for any} \ t \}. \quad (2.6)
\]

The basic idea is then to multiply equation (2.5) by such a test function \(\varphi(x, t)\), integrate over \(\mathbb{R} \times [0, \infty)\) and subsequently apply partial integration. Using the fact that \(\varphi(x, t)\) vanishes for \(|x| + t \to \infty\) we obtain

\[
\int_0^\infty \int_{-\infty}^{\infty} \left( u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) \, dx \, dt = -\int_{-\infty}^{\infty} u(x, 0) \varphi(x, 0) \, dx.
\]

(2.7)

This relation then gives rise to the following definition.

**Definition 12.8.** A function \(u(x, t)\) is called a weak solution of conservation law (2.5) if (2.7) holds for all test functions \(\varphi \in C^1_0(\mathbb{R} \times [0, \infty))\).

Obviously, when \(u(x, t)\) satisfies (2.5) it is a weak solution. The converse is only true when \(u(x, t)\) is continuously differentiable. In the following, when we speak of a solution of (2.5), we mean a weak solution in the sense of this definition.

Note that relation (2.7) allows for discontinuous solutions. However, not every discontinuous function can be a solution of (2.5) as is shown in the following theorem.
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

Theorem 12.9. Let \( u \) be a piecewise smooth solution of (2.5) that has a discontinuity across a curve \( \mathcal{E} : x = x(t) \). Then \( u \) satisfies the condition

\[
[f(u)]^+ = s[u]^+,
\]

(2.8)

with \([v]^+ := v(x(t) + , t) - v(x(t) -, t) (v = u, f(u))\) the jump of \( v \) across \( \mathcal{E} \) and \( s \) the speed of \( \mathcal{E} \).

**Proof.** Assume that \( \mathcal{E} \) separates a domain \( \Omega = \text{supp}(\varphi) \) in a left part \( \Omega_\ell \) and a right part \( \Omega_r \); see figure 12.3. The solution is smooth both in \( \Omega_\ell \) and in \( \Omega_r \). Since (2.5) holds in \( \Omega_\ell \), we have

\[
\int_{\Omega_\ell} \left( \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} \right) \varphi \, dx \, dt = 0
\]

for every test function \( \varphi \in \mathcal{D} \). We can rewrite this equation as follows:

\[
\int_{\Omega_\ell} \left( \frac{\partial (u \varphi)}{\partial t} + \frac{\partial (f(u) \varphi)}{\partial x} \right) \, dx \, dt = \int_{\Omega_\ell} \left( u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) \, dx \, dt.
\]

If we apply the two-dimensional Gauss’ theorem to the integral on the left hand side, we find

\[
- \int_{\partial \Omega_\ell} \varphi (u \, dx - f \, dt) = \int_{\Omega_\ell} \left( u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) \, dx \, dt,
\]

with \( \partial \Omega_\ell \) the boundary of \( \Omega_\ell \). Next, using that \( \varphi(x, t) = 0 \) for \((x, t) \in \partial \Omega \cap \{t > 0\} \), with \( \partial \Omega \) the boundary of \( \Omega \), we obtain

\[
- \int_{a}^{b} u(x, 0) \varphi(x, 0) \, dx - \int_{E} \varphi(u_\ell \, dx - f(u_\ell) \, dt) = \int_{\Omega_\ell} \left( u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) \, dx \, dt, \quad (*)
\]

with \( x_n \) the intersection of \( \mathcal{E} \) with the \( x \)-axis and \( u_\ell := u(x(t) -, t) \), i.e. the limit value of \( u \) just left of the discontinuity. Carrying out a similar procedure for \( \Omega_r \), we find

\[
- \int_{c}^{d} u(x, 0) \varphi(x, 0) \, dx + \int_{E} \varphi(u_r \, dx - f(u_r) \, dt) = \int_{\Omega_r} \left( u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) \, dx \, dt, \quad (**)
\]

with \( u_r := u(x(t) + , t) \) the limit value of \( u \) from the right of \( \mathcal{E} \). The integral over \( \mathcal{E} \) in (***) is evaluated in the same direction as in (*), see figure 12.3, and therefore has a + sign in front. Adding (*) and (**) we obtain

\[
- \int_{a}^{b} u(x, 0) \varphi(x, 0) \, dx + \int_{E} \varphi([u]^+ \, dx - [f(u)]^+ \, dt) = \int_{\Omega} \left( u \frac{\partial \varphi}{\partial t} + f(u) \frac{\partial \varphi}{\partial x} \right) \, dx \, dt.
\]

Combining this relation with (2.7) and taking into account that \( \varphi(x, t) \) vanishes outside \( \Omega \), we find

\[
\int_{E} \varphi([u]^+ \, dx - [f(u)]^+ \, dt = 0.
\]

This relation holds for arbitrary test functions \( \varphi \in \mathcal{D} \), so that \([u]^+ \, dx - [f(u)]^+ \, dt = 0\). Finally, since \( s = \frac{dx}{dt} \), this implies relation (2.8). \( \Box \)

2. WEAK FORMULATION OF FIRST-ORDER SCALAR EQUATIONS

The propagation speed $s$ of a discontinuity is thus given by

$$f(u_r) - f(u_\ell) = s(u_r - u_\ell), \quad (2.9)$$

with $u_r := u(x(t)+t)$ and $u_\ell := u(x(t)-t)$ the limit values of $u(x,t)$ just right and left of the discontinuity, respectively. Relation (2.9) is called the Rankine-Hugoniot jump condition. Inserting (1.10) into (2.9), we find the following alternative expression for $s$:

$$s = \frac{1}{u_r - u_\ell} \int_{u_\ell}^{u_r} b(v) \, dv, \quad (2.10)$$

i.e. $s$ is the average advection velocity $b(v)$ over the interval $\text{int}(u_\ell, u_r)$.

**Example 12.10** Consider the Burgers’ equation, written in conservation form,

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{1}{2} u^2 \right) = 0,$$

subject to the following piecewise constant initial condition,

$$u(x, 0) = \begin{cases} 
\alpha & \text{if } x < 0, \\
\beta & \text{if } x > 0.
\end{cases}$$

Let $\alpha > \beta$, so that we have a discontinuity. We apparently have for the speed $s$ of the discontinuity

$$s = \frac{dx}{dt} = \frac{\beta^2 - \alpha^2}{2\beta - \alpha} = \frac{1}{2}(\beta + \alpha).$$

The discontinuity is thus a straight line with a directional coefficient being the average of those of the characteristics to the left and the right, respectively. □

It is important to note that the weak solution depends on the formulation of the conservation equation as the next example clearly shows.
Example 12.11 Consider again the Burgers’ equation from the previous example, subject to the same initial condition. If we multiply this equation by $u$, we can easily derive the following conservation equation for $w := u^2$:

$$\frac{\partial w}{\partial t} + \frac{\partial}{\partial x} (\frac{2}{3} w^{3/2}) = 0. \quad (*)$$

If we apply the Rankine-Hugoniot jump condition (2.9) to $(*)$, we find for the propagation speed $s$ of a discontinuity

$$s = \frac{2}{3} \left( \frac{w_{L}^{3/2} - w_{R}^{3/2}}{w_{L} - w_{R}} \right) = \frac{2}{3} \frac{\beta^3 - \alpha^3}{\beta^2 - \alpha^2} = \frac{2}{3} \frac{\beta^2 + \beta \alpha + \alpha^2}{\beta + \alpha}.$$

This is generally not equal to the propagation speed $\frac{1}{2} (\beta + \alpha)$ found in example 12.10. So, although the Burgers’ equation and equation $(*)$ are equivalent for smooth solutions, they have different weak solutions.

What we learn from this example is that we cannot manipulate the strong formulation of a problem when dealing with discontinuous solutions. In fact, the Rankine-Hugoniot jump condition is the reformulation of the correct physical conservation law across a discontinuity. Stated otherwise, the Rankine-Hugoniot jump condition is an extra condition which should be imposed for discontinuous solutions along with the corresponding partial differential equation.

It is instructive to look back again now to the problem depicted in figure 12.2. The breaking at $t = t^*$ starts a shock which then propagates at speed $s$ given by (2.9). From the shock speed we can determine the location of the shock. Now consider the initial profile in figure 12.2 which is zero outside $(0, 1)$ Then if $(a, b)$ is a sufficiently large interval and the flux is proportional to $u$, we have from the integral form of the conservation equation (2.5)

$$\frac{d}{dt} \int_{a}^{b} u(x, t) \, dx = f(u(a, t)) - f(u(b, t)) = 0. \quad (2.11)$$

Hence $\int_{a}^{b} u(x, t) \, dx$ is constant. At a time point $t > t^*$ a classical solution would have a positive and a negative integral part as shown in figure 12.4. Of course, the shape as depicted does not make sense practically as we would have a multivalued solution, yet we may formally do the integration. The total effect would be the same as when we would have integrated up to the point $x_s$ (starting from $a$); see figure 12.4. Conservation means that our weak solution should also be conserved, and hence we may identify the point $x_s$ with the propagated breaking point on the shock line. Actually one may revert the argument and determine $x_s$ graphically: choose the point $x_s$ such that area $A_L$ equals area $A_R$.

2.2 The Riemann problem

The initial value problem for a conservation equation with piecewise constant initial condition and having one single discontinuity, is known as the Riemann problem. The canonical
2. WEAK FORMULATION OF FIRST-ORDER SCALAR EQUATIONS

Form of the Riemann problem for (2.5) reads

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad x \in \mathbb{R}, \ t > 0,$$  \hspace{1cm} (2.12a)

$$u(x, 0) = \begin{cases} u_\ell & \text{if } x < 0, \\ u_r & \text{if } x > 0, \end{cases}$$  \hspace{1cm} (2.12b)

Note that if \( u(x, t) \) is a solution of (2.12) then also \( u(\alpha x, \alpha t) \) for arbitrary \( \alpha > 0 \), implying that the solution of (2.12) is a similarity solution of the form \( u(x, t) = \tilde{u}(x/t) \); cf. Section 2.5. In Section 1 we showed that solutions of an initial value problem like (2.12) are propagated along characteristics with velocity \( b(u) = f'(u) \). Based on this observation we can distinguish two cases, viz. \( b(u_\ell) > b(u_r) \) and \( b(u_\ell) < b(u_r) \), to be discussed below separately.

**Case 1.** \( b(u_\ell) > b(u_r) \)

The characteristics emanating from the negative \( x \)-axis have a slope smaller than the slope of the characteristics coming from the positive \( x \)-axis. As a consequence, characteristics intersect, which would lead to multivalued solutions. Instead we have a discontinuous solution. We can easily verify by substitution into (2.7) that the solution of the Riemann problem (2.12) is indeed given by

$$u(x, t) = \begin{cases} u_\ell & \text{if } x < st, \\ u_r & \text{if } x > st, \end{cases}$$  \hspace{1cm} (2.13)

where \( s \) is defined in (2.9). The solution in (2.13) represents a discontinuity, traveling with speed \( s \) and is called a shock wave; \( s \) is called the shock speed. A typical shock wave and the corresponding characteristics is shown in figure 12.5. Note that the characteristics move into the shock for increasing \( t \).

**Case 2.** \( b(u_\ell) < b(u_r) \).
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

Figure 12.5. Shock wave and corresponding characteristics.

Figure 12.6. Rarefaction wave and corresponding characteristics.

In this case the characteristics emanating from the negative $x$-axis have a larger slope than those emanating from the positive $x$-axis. So we have separating characteristics and we do not expect a discontinuous solution. The solution of (2.12) is now given by

$$u(x, t) = \begin{cases} 
    u_\ell & \text{if } x < b(u_\ell)t, \\
    w(x/t) & \text{if } b(u_\ell)t < x < b(u_r)t, \\
    u_r & \text{if } x > b(u_r)t,
\end{cases} \quad (2.14)$$

where $w(\eta)$ is the solution of the relation $b(w(\eta)) = \eta$. This solution is called a rarefaction wave. It is continuous, despite the fact that the initial condition is discontinuous, and consists of the constant states $u(x, t) = u_\ell$ and $u(x, t) = u_r$, connected by the intermediate solution $u(x, t) = w(x/t)$. The latter solution is a similarity solution of (2.5) as we can easily verify by direct substitution. The constant states are an immediate consequence of (1.12). A typical rarefaction wave and the corresponding characteristics are depicted in figure 12.6.

Example 12.12 Consider once more the Burgers' equation, cf. example 12.10. Since the convection velocity $b(u) = u$, we have for the solution either a shock wave when $u_\ell > u_r$ or a...
2. WEAK FORMULATION OF FIRST-ORDER SCALAR EQUATIONS

Figure 12.7. Characteristics corresponding to an expansion shock of the Burgers’ equation.

rarefaction wave when \( u_\ell < u_r \). The shock wave solution reads

\[
u(x, t) = \begin{cases} 
  u_\ell & \text{if } x < st, \\
  u_r & \text{if } x > st,
\end{cases}
\]

with shock speed \( s = \frac{1}{2}(u_\ell + u_r) \). The rarefaction wave is given by

\[
u(x, t) = \begin{cases} 
  u_\ell & \text{if } x < u_\ell t, \\
  x/t & \text{if } u_\ell t < x < u_r t, \\
  u_r & \text{if } x > u_r t.
\end{cases}
\]

We can show that (2.13) is a solution of the Riemann problem (2.12) by substitution into (2.7), regardless whether \( b(u_\ell) > b(u_r) \) or \( b(u_\ell) < b(u_r) \). In the latter case this solution is called an expansion shock. In figure 12.7 we have sketched the corresponding characteristics. Note that the characteristics move out of the shock for increasing \( t \). This poses a problem, since for \( b(u_\ell) < b(u_r) \) we have at least two solutions of the Riemann problem (2.12). The expansion shock is physically not correct and should be discarded. One of the reasons is that this solution is not stable under small perturbation in the initial data as is demonstrated in the next example.

Example 12.13 Consider Burgers’ equation subject to the following, piecewise linear, initial condition

\[
u(x, 0) = \begin{cases} 
  0 & \text{if } x < 0, \\
  x/\delta & \text{if } 0 < x < \delta, \\
  1 & \text{if } x > \delta,
\end{cases}
\]

with \( 0 < \delta \ll 1 \). Note that for \( \delta \to 0 \) we obtain the standard Riemann problem. The characteristics in this initial value problem fan out and we can compute its solution from
(1.14). This way we find
\[
\begin{align*}
u(x, t) &= \begin{cases} 
0 & \text{if } x < 0, \\
x/(t + \delta) & \text{if } 0 < x < t + \delta, \\
1 & \text{if } x > t + \delta.
\end{cases}
\end{align*}
\]
Clearly, for \( \delta \to 0 \) this solution changes into the rarefaction wave (2.14) which is completely different from an expansion shock. \( \square \)

The next example concerns the modelling of traffic flow; cf. example 1.1.2.

**Example 12.14** A simple model for traffic flow on a highway is given by the following conservation equation:
\[
\frac{\partial n}{\partial t} + \frac{\partial f(n)}{\partial x} = 0, \quad f(n) := u_m n(1 - n/n_m),
\]
where \( x \) is the coordinate along the highway, \( n(x, t) \) is the density of cars, \( n_m \) the maximum density and \( u_m \) the maximum speed of vehicles. Obviously \( n_m, u_m > 0 \). Consider the corresponding Riemann problem with initial condition
\[
n(x, 0) = \begin{cases} 
n_\ell & \text{if } x < 0, \\
n_r & \text{if } x > 0,\end{cases}
\]
where \( 0 \leq n_\ell, n_r \leq n_m \). We can easily verify that the advection velocity \( b(n) \) is given by
\[
b(n) = u_m \left(1 - \frac{2n}{n_m}\right),
\]
which is a monotonically decreasing function of \( n \). This implies that a shock wave occurs when \( n_\ell < n_r \). The shock wave solution is given by
\[
n(x, t) = \begin{cases} 
n_\ell & \text{if } x < st, \\
n_r & \text{if } x > st,\end{cases}
\]
with shock speed \( s = u_m \left(1 - (n_\ell + n_r)/n_m\right) \). Note that the shock speed can be either positive or negative, depending on the values of \( n_\ell \) and \( n_r \). This corresponds e.g. with the situation that cars approach a red traffic light. On the other hand, when \( n_\ell > n_r \), we have a rarefaction wave given by
\[
n(x, t) = \begin{cases} 
n_\ell & \text{if } x < b(n_\ell)t, \\
\frac{1}{2}n_m \left(1 - \frac{x}{u_m t}\right) & \text{if } b(n_\ell)t < x < b(n_r)t, \\
n_r & \text{if } x > b(n_r)t.\end{cases}
\]
This solution describes e.g. the situation that cars speed up after the traffic light has turned green. \( \square \)

We have seen that not every discontinuous solution of (2.12) is physically correct. Therefore, we like to have a simple criterion to determine whether a discontinuous solution is admissible. In fact, the physically relevant solution is the solution of equation (1.18) for \( \varepsilon \to 0 \). One can show that this vanishing viscosity solution for the Burgers’ equation
2. WEAK FORMULATION OF FIRST-ORDER SCALAR EQUATIONS

reduces to a shock wave when \( u_\ell > u_r \) and to a rarefaction wave when \( u_\ell < u_r \); see e.g. [61]. However, the computation of the solution of (1.18) is often very tedious and not very practical to work with. A simple criterion is suggested by the requirement that characteristics move into a shock for increasing \( t \), as shown in figure 12.5. This then gives rise to the following definition.

**Definition 12.15 (Lax’ entropy condition).** A discontinuous solution of (2.5), that has a convex flux function, satisfies the entropy condition, if

\[
b(u_\ell) > s > b(u_r),
\]

with \( s \) the propagation speed of the discontinuity, given by (2.9).

One can show that this so-called entropy solution is the unique, physically correct solution, cf. [104]. When this condition is generalized to the Euler equations for compressible gas flow, one can prove that the entropy of the flow increases across the discontinuity, in agreement with the second law of thermodynamics, and therefore (2.15) is referred to as the entropy condition.

**Example 12.16** The flux function for the Burgers’ equation is convex, and therefore the entropy condition (2.15) simply reduces to \( u_\ell > u_r \). On the other hand, for the traffic flow problem we have a concave flux function, leading to the criterion \( n_\ell < n_r \). □

Integrating the ODE system (1.12) along the characteristic that goes into the shock from the left gives a relation between \( u_\ell \) and the initial data. Likewise, the characteristic going into the shock from the right provides a relation for \( u_r \). Together with the jump condition (2.9), these relations suffice to compute the three unknowns \( u_\ell, u_r \) and \( s \).

A more general definition, which is also applicable when \( f(u) \) is neither convex nor concave is the following; cf. [82].

**Definition 12.17 (Oleinik’s entropy condition).** A weak solution of (2.5) is the entropy solution if all discontinuities, which propagate at speed \( s \) given by (2.9), satisfy

\[
\frac{f(u) - f(u_\ell)}{u - u_\ell} \geq s \geq \frac{f(u) - f(u_r)}{u - u_r},
\]

for all \( u \) between \( u_\ell \) and \( u_r \).

We have seen that for convex or concave flux functions the Riemann problem (2.12) has either a shock or a rarefaction wave as solution. For general flux functions, the entropy solution might involve both as demonstrated by the next example.

**Example 12.18** A model equation for two-phase flow is the Buckley-Leverett equation, given by (2.5) with flux function

\[
f(u) := \frac{u^2}{u^2 + a(1-u)^2},
\]

\( \star \)
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

see figure 12.8. Here $u$ typically represents the saturation of water, so that $0 \leq u \leq 1$ and $a$ is a constant. It is applied in oil reservoir simulation; for more details see e.g. [51, 72]. Consider the Riemann problem for $(\ast)$ with initial data

$$u(x, 0) = \begin{cases} 1 & \text{if } x < 0, \\ 0 & \text{if } x > 0. \end{cases}$$

The characteristic velocity $b(u)$ is given by

$$b(u) = \frac{2au(1 - u)}{(u^2 + a(1 - u)^2)^2}.$$  

Note that $b(1) = b(0) = 0$, so that all characteristics emanate vertically from the $x$-axis. For all intermediate values $u$ ($0 < u < 1$) we have $b(u) > 0$. As a consequence, the two constant states $u_r = 0$ and $u_l = 1$ have to be connected by a shock followed by a rarefaction wave. The solution is thus given by

$$u(x, t) = \begin{cases} 1 & \text{if } x < 0, \\ w(x/t) & \text{if } 0 < x < st, \\ 0 & \text{if } x > st, \end{cases}$$

where $w = w(\eta)$ is a similarity solution satisfying $b(w(\eta)) = \eta$ and where $s$ is the shock speed; see figure 12.8. Suppose the shock connects the right state $u_r = 0$ with an intermediate state $u_s$ ($0 < u_s < 1$), then according to (2.9) we have

$$s = f(u_s)/u_s.$$  

To determine the value of $u_s$, we have to invoke the entropy condition (2.16), which in this case reduces to

$$\frac{f(u) - f(u_s)}{u - u_s} \geq s \geq \frac{f(u)}{u}. \quad (\ast\ast)$$

It will turn out that $u_s$ satisfies the relation

$$f(u_s)/u_s = f'(u_s),$$

which means that the straight line through $(0, 0)$ and $(u_s, f(u_s))$ is tangent to $y = f(u)$ in $u_s$. Moreover, the shock moving with speed $s = f'(u_s)$ is then parallel to the characteristic just left of it. Suppose, the shock were connected to a state $u^* < u_s$, then the shock speed $f(u^*)/u^*$ would be smaller than $f'(u_s)$ leading to a triple valued function. On the other hand, if the shock were connected to a state $u^* > u_s$, then the entropy condition $(\ast\ast)$ would be violated. □

3 First order systems

In this section we shall deal with first order systems of hyperbolic equations, to start with linear systems in Section 3.1 followed by nonlinear ones in Section 3.2. Finally, in Section 3.3, we briefly describe the method of characteristics for a system of two equations, which gives a nice illustration of the role of characteristics for hyperbolic equations.
3. **FIRST ORDER SYSTEMS**

3.1 Linear systems

Consider the first order system of equations

\[
A \frac{\partial u}{\partial t} + B \frac{\partial u}{\partial x} = c, \tag{3.1}
\]

with \( A \) and \( B \) constant \( m \times m \) matrices and \( c \) an \( m \)-vector, possibly depending on \( x, t \) and \( u \). In Section 2.2 we considered the case where \( A \) was nonsingular. We can easily generalise this procedure for singular matrices. For this we have to consider the **generalised eigenvalue problem**: Find left eigenvectors \( t^T \) and corresponding eigenvalues \( \lambda \) such that

\[
\lambda t^T A = t^T B. \tag{3.2}
\]

Obviously, equation (3.2) only holds for nonzero \( t^T \) if

\[
\det(\lambda A - B) = 0. \tag{3.3}
\]

The linear combination

\[
\lambda A - B; \quad \lambda \in \mathbb{C}, \tag{3.4}
\]

is called a *matrix pencil* [39]. This matrix pencil is called *singular* if \( \det(\lambda A - B) = 0 \) for all \( \lambda \in \mathbb{C} \), else it is called *regular*. We now assume the matrix pencil (3.4) to be regular. In this case there exist \( m \) eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_m \) such that (3.3) holds. Furthermore, we assume that these eigenvalues are real and that the matrix is not defect, i.e. there exist a complete set of (left) eigenvectors \( t_1^T, t_2^T, \ldots, t_m^T \). We can take together all (generalised) eigenvalue problems, giving the relation

\[
TB = \Lambda \Lambda^T, \tag{3.5}
\]

where the matrices \( A \) and \( T \) are defined by

\[
\Lambda := \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m), \quad T := \begin{pmatrix} t_1^T \\ t_2^T \\ \vdots \\ t_m^T \end{pmatrix}, \tag{3.6}
\]

**Figure 12.8.** Flux function and solution of the Buckley-Leverett equation.
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

Note that $T$ is nonsingular since its rows $t^T_k$ are linearly independent.

Left multiplication of (3.1) by $T$ results in

$$TA \frac{\partial u}{\partial t} + \Lambda TA \frac{\partial u}{\partial x} = Tc.$$  \hspace{1cm} (3.7)

In analogy of what we did in Section 2.2, we define the characteristic variable $\tilde{u}$ by

$$\tilde{u} := TAu,$$  \hspace{1cm} (3.8)

and since $T$ is a constant matrix, we obtain the decoupled system, cf. (2.2.13a),

$$\frac{\partial \tilde{u}}{\partial t} + \Lambda \frac{\partial \tilde{u}}{\partial x} = \tilde{c} := Tc.$$  \hspace{1cm} (3.9a)

Written componentwise, we have

$$\frac{\partial \tilde{u}_k}{\partial t} + \lambda_k \frac{\partial \tilde{u}_k}{\partial x} = \tilde{c}_k \quad (k = 1, 2, \ldots, m).$$  \hspace{1cm} (3.9b)

When this decoupling is possible, we call the system (3.1) hyperbolic. We thus have the following formal definition.

**Definition 12.19.** The system (3.1) is called hyperbolic, if there exists a real diagonal matrix $\Lambda$ and a nonsingular matrix $T$ such that (3.5) holds.

**Example 12.20** Consider tidal waves travelling along a straight canal of uniform depth $h$. Let $x$ denote the coordinate along the canal. For small amplitude waves, the water elevation $\eta(x, t)$ above the still water level satisfies the standard wave equation \[85\]

$$\frac{\partial^2 \eta}{\partial t^2} = a^2 \frac{\partial^2 \eta}{\partial x^2},$$

with $a := \sqrt{gh}$ and $g$ the gravitational acceleration. Introducing the auxiliary variables

$$u_1 := \frac{1}{a} \frac{\partial \eta}{\partial t}, \quad u_2 := \frac{\partial \eta}{\partial x}, \quad u := \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

we can e.g. reformulate the wave equation for $\eta$ as the $2 \times 2$ linear system

$$\frac{\partial u}{\partial t} + B \frac{\partial u}{\partial x} = 0,$$  \hspace{1cm} (*)

with the coefficient matrix $B$ given by

$$B := - \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}.$$  

The eigenvalues $\lambda_1, \lambda_2$ and corresponding eigenvectors $t^T_1, t^T_2$ are given by:

$$\lambda_1 = -a < 0, \quad \lambda_2 = a > 0, \quad t^T_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad t^T_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

and consequently, the linear system (*) is hyperbolic. \qed

3. FIRST ORDER SYSTEMS

Since there is no bias with respect to \( A \) or \( B \) we rather may consider pairs \((\lambda, \mu)\) of solutions of

\[
\det(\lambda A - \mu B) = 0.
\]  

(3.10)

Note that for each pair \((\lambda, \mu)\) satisfying (3.10) any nonzero multiple of it also satisfies this equation. Analogously to the previous, the matrix pencil

\[
\lambda A - \mu B, \quad \lambda, \mu \in \mathbb{C}
\]  

(3.11)

is called singular if (3.10) holds for all \( \lambda, \mu \in \mathbb{C} \), else it is called regular. We shall again assume that the matrix pencil is regular. In this case there exist \( m \) eigenvalues pairs \((\lambda_1, \mu_1), (\lambda_2, \mu_2), \ldots, (\lambda_m, \mu_m)\), apart from a multiplicative constant. Note that for each nontrivial eigenvalue pair \((\lambda_k, \mu_k)\), at least one of them is nonzero. The appropriate generalisation of (3.5) then reads: There exist nonsingular matrices \( T, S \) and real diagonal matrices \( \Lambda_A, \Lambda_B \) such that

\[
T AS^{-1} = \Lambda_A, \quad T B S^{-1} = \Lambda_B,
\]  

(3.12)

i.e. the matrices \( A \) and \( B \) can be diagonalised simultaneously. In this case equation (3.1) is called hyperbolic; note that both \( A \) and \( B \) may be singular.

Premultiplying (3.1) by \( T \) and substituting

\[
\tilde{u} := Su,
\]  

(3.13)

yields the decoupled system

\[
\Lambda_A \frac{\partial \tilde{u}}{\partial t} + \Lambda_B \frac{\partial \tilde{u}}{\partial x} = \tilde{c} := Tc.
\]  

(3.14)

If a diagonal element of \( \Lambda_A \), say \( \lambda_{A,k} \), is zero, then for the corresponding characteristic \( C_k \), we have

\[
\frac{dt}{dx} = 0,
\]

i.e. \( C_k \) is parallel to the \( x \)-axis. This implies that the information is propagating with infinite velocity along this characteristic. We shall explicitly exclude such cases here and in the sequel. As a consequence we may assume \( A \) to be nonsingular.

3.2 Quasilinear systems

We now consider the more general case where the matrices \( A \) and \( B \) may depend on \( x, t \) and \( u \):

\[
A(x, t, u) \frac{\partial u}{\partial t} + B(x, t, u) \frac{\partial u}{\partial x} = c(x, t, u).
\]  

(3.15)

Assuming now that \( A \) is nonsingular (cf. what we said in the previous section), we may as well take it equal to \( I \), i.e. we consider the system

\[
\frac{\partial u}{\partial t} + B(x, t, u) \frac{\partial u}{\partial x} = c(x, t, u).
\]  

(3.16)
As a straightforward generalisation of definition 12.19, we can define hyperbolicity of the system (3.16).

**Definition 12.21.** The system (3.16) is called hyperbolic in \((x, t, u)\), if there exists a real diagonal matrix \(\Lambda(x, t, u)\) and a nonsingular matrix \(T(x, t, u)\) such that

\[
T(x, t, u)B(x, t, u) = \Lambda(x, t, u)T(x, t, u).
\]  

(3.17)

In this definition, \(\Lambda(x, t, u) = \text{diag} (\lambda_1(x, t, u), \lambda_2(x, t, u), \ldots, \lambda_m(x, t, u))\) and \(T(x, t, u)\) is the matrix of corresponding left eigenvectors; cf. (3.6). Note that hyperbolicity of (3.16) depends on \(x, t\) and \(u\). In the following we will suppress this dependency and simply write \(\Lambda\) instead of \(\Lambda(x, t, u)\); etc.

Left multiplication of (3.16) by the matrix \(T\) gives

\[
T \frac{\partial u}{\partial t} + \Lambda T \frac{\partial u}{\partial x} = T c := \tilde{c}.
\]  

(3.18)

Next, we introduce the *characteristic variables* \(\tilde{u}\) by the relation

\[
d\tilde{u} := T du,
\]  

(3.19)

where \(d\) is an arbitrary differential. Equation (3.19) is a *Pfaffian differential equation* and expresses the differentials of \(\tilde{u}\) as a linear combination of the differentials of \(u\). Conditions for the solvability of (3.19) can be found in e.g. [135]. Note that for constant \(T\), relation (3.19) is equivalent with (3.8) in case \(A = I\). Furthermore, for many nonlinear systems of conservation equations, the eigenvectors \(t_k(\mathbf{u})\) can be scaled such that (3.19) is integrable. Assuming (3.19) has a solution, system (3.18) reduces to

\[
\frac{\partial \tilde{u}}{\partial t} + \Lambda \frac{\partial \tilde{u}}{\partial x} = \tilde{c}.
\]  

(3.20)

Thus, like in the linear case, system (3.18) can be diagonalised. Note however, that system (3.20) is still coupled through the eigenvalue matrix \(\Lambda\), which in general depends on \(u\).

Alternatively, we can premultiply (3.16) by an arbitrary left eigenvector \(t_k\) of \(B\), giving

\[
t_k^T \frac{\partial u}{\partial t} + t_k^T B \frac{\partial u}{\partial x} = t_k^T \left( \frac{\partial u}{\partial t} + \lambda_k \frac{\partial u}{\partial x} \right) = t_k^T c =: \tilde{c}_k.
\]  

(3.21)

This is in fact a linear combination of the equations of (3.16). We would like (3.21) to be equivalent to an ordinary differential equation of the form

\[
t_k^T \frac{du}{ds} = \tilde{c}_k,
\]  

(3.22)

which should hold on some curve \(K := \{(x(s), t(s)) \mid s \in I \subset \mathbb{R}\}\). Since we have

\[
\frac{du}{ds} = \frac{\partial u}{\partial t} \frac{dt}{ds} + \frac{\partial u}{\partial x} \frac{dx}{ds},
\]  

(3.23)
3. FIRST ORDER SYSTEMS

we thus find by comparing equations (3.21) and (3.22) and using relation (3.19)

\[
\frac{dr}{ds} = 1, \quad \frac{dx}{ds} = \lambda_k, \quad \frac{d\tilde{u}_k}{dr} = \tilde{c}_k.
\]  

(3.24)

The curve \( \mathcal{K} \) is apparently the characteristic \( \mathcal{C}_k \) corresponding to the \( k \)th eigenvalue \( \lambda_k \). If system (3.16) is hyperbolic, there exist \( m \) such characteristics. Equation (3.22) is said to be in normal form or characteristic form. The variables \( \tilde{u}_k \) are also called Riemann variables. If \( c = 0 \), the variables \( \tilde{u}_k \) are constant along the corresponding characteristic \( \mathcal{C}_k \) and are therefore often referred to as Riemann invariants.

In many practical applications, the coefficient matrix \( B \) and the right hand side vector \( c \) only depend on \( u \). In the sequel, we will restrict ourselves to this case. Then, system (3.16) can be rewritten as

\[
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = c(u),
\]  

(3.25)

where the flux vector \( f(u) \) is related to \( B(u) \) through

\[
B(u) = \left( \frac{\partial f_i(u)}{\partial u_j} \right).
\]  

(3.26)

So \( B(u) \) is the Jacobi matrix of \( f(u) \). The formulation in (3.25) is again called the conservation form. According to definition 12.21, the system (3.25) is hyperbolic if the Jacobi matrix \( B(u) \) is diagonalisable through its left eigenvectors. Alternatively, \( B(u) \) can be brought onto diagonal form by its right eigenvectors. Indeed, introducing the matrix

\[
S = (s_1, s_2, \ldots, s_m) := T^{-1},
\]  

(3.27)

we readily see from (3.17) that

\[
BS = SA,
\]  

(3.28)

i.e. the \( k \)th column \( s_k \) of \( S \) is the right eigenvector of \( B(u) \) corresponding with the eigenvalue \( \lambda_k \). Note that the right eigenvectors are linearly independent since \( S \) is nonsingular.

Changing to the characteristic variables \( \tilde{u} \), which are now defined by \( d\tilde{u} = S^{-1} du \), we obtain in a similar way as before, the decoupled system (3.20).

Example 12.22 Referring to Chapter 6, we note that the Euler equations for isentropic gas flow can be written in the standard form (3.25), with \( u, f(u) \) and \( c(u) \) given by

\[
u = \left( \begin{array}{c}
\rho \\
\rho u
\end{array} \right), \quad f(u) = \left( \begin{array}{c}
\rho u \\
\rho u^2 + p(\rho)
\end{array} \right), \quad c(u) = 0,
\]  

and where \( \rho, u \) and \( p \) are the density, velocity and pressure, respectively, of the gas flow. For isentropic flow, the pressure is given by the relation

\[
p(\rho) = p_0\rho^\gamma,
\]  

(*)

with \( \gamma = C_P/C_V \) the specific heat ratio and where \( p_0 \) is a reference pressure. The Jacobi matrix is given by

\[
B(u) = \left( \begin{array}{cc}
0 & 1 \\
-u^2 + p'(\rho) & 2u
\end{array} \right).
\]  

and its eigenvalues $\lambda_i(u)$ and (left) eigenvectors $t_i^L(u)$ are given by

$$
\lambda_1(u) = u - c, \quad \lambda_2(u) = u + c, \quad c := \sqrt{\gamma p/\rho}.
$$

$$
t_i^L(u) = (-u - c, 1)^T, \quad t_i^R(u) = (-u + c, 1).
$$

The variable $c$ is the speed of sound. Note that the eigenvectors are determined up to a multiplicative constant. Clearly, the isentropic Euler equations are hyperbolic. To decouple these equations, we introduce the characteristic variables $\tilde{u}$ through the relation (3.19). This way we obtain the system

$$
ad\tilde{u}_1 = -(u + c)d\rho + d(\rho u) = \rho du - cd\rho,
$$

$$
ad\tilde{u}_2 = -(u + c)d\rho + d(\rho u) = \rho du + cd\rho.
$$

Unfortunately, these equations are not integrable. However, scaling the eigenvectors by a factor $1/\rho$, we find the relations

$$
d\tilde{u}_1 = du - \frac{c}{\rho}d\rho, \quad d\tilde{u}_2 = du + \frac{c}{\rho}d\rho.
$$

Using relation ($*$) these equations are easy to integrate, and we find

$$
\tilde{u}_1 = u - \frac{2c}{\gamma - 1}, \quad \tilde{u}_2 = u + \frac{2c}{\gamma - 1}.
$$

Finally, we obtain the decoupled system

$$
\frac{\partial}{\partial t}\left(u - \frac{2c}{\gamma - 1}\right) + (u - c)\frac{\partial}{\partial x}\left(u - \frac{2c}{\gamma - 1}\right) = 0,
$$

$$
\frac{\partial}{\partial t}\left(u + \frac{2c}{\gamma - 1}\right) + (u + c)\frac{\partial}{\partial x}\left(u + \frac{2c}{\gamma - 1}\right) = 0,
$$

implying that the Riemann variables $u - \frac{2c}{\gamma - 1}$ and $u + \frac{2c}{\gamma - 1}$ are constant along the $C_1$ and $C_2$-characteristics, respectively.

### 3.3 Method of characteristics

For systems of two equations, the normal form (3.22) together with the equations for the characteristics (3.24), lend themselves to a (theoretically) simple solution method. We shall work this out below for the general linear case (3.1), which can be written as

$$
\begin{align*}
\alpha_{11} \frac{\partial u_1}{\partial t} + \alpha_{12} \frac{\partial u_2}{\partial t} + b_{11} \frac{\partial u_1}{\partial x} + b_{12} \frac{\partial u_2}{\partial x} &= c_1, \\
\alpha_{21} \frac{\partial u_1}{\partial t} + \alpha_{22} \frac{\partial u_2}{\partial t} + b_{21} \frac{\partial u_1}{\partial x} + b_{22} \frac{\partial u_2}{\partial x} &= c_2.
\end{align*}
$$

(3.29)

Now let $v^T := (v_1, v_2)$ be a left eigenvector corresponding to the eigenvalue $\lambda_\ast$ of (3.2), i.e.

$$
v^T B = \lambda_\ast v^T A.
$$

(3.30)

If we write

$$
v^T A =: (\tilde{v}_1, \tilde{v}_2), \quad v^T c =: \tilde{c}_1,
$$

(3.31)
3. FIRST ORDER SYSTEMS

![Figure 12.9. Illustration of the method of Massau.](image)

and let $\xi$ be the independent variable along the characteristic defined by

$$
\frac{dr}{d\xi} = 1, \quad \frac{dx}{d\xi} = \lambda_\pi,
$$

(3.32)

then we have

$$
\tilde{v}_1 \frac{\partial u_1}{\partial \xi} + \tilde{v}_2 \frac{\partial u_2}{\partial \xi} = \tilde{c}_1.
$$

(3.33)

There also exists a second left eigenvector $w$ corresponding to the other eigenvalue, $\lambda_w$ say. Define

$$
w^T A =: (\tilde{w}_1, \tilde{w}_2), \quad w^T c = \tilde{c}_2,
$$

(3.34)

and finally, let $\eta$ be the independent variable along the corresponding characteristic, with

$$
\frac{dr}{d\eta} = 1, \quad \frac{dx}{d\eta} = \lambda_w.
$$

(3.35)

Then

$$
\tilde{w}_1 \frac{\partial u_1}{\partial \eta} + \tilde{w}_2 \frac{\partial u_2}{\partial \eta} = \tilde{c}_2.
$$

(3.36)

The equations (3.32), (3.33), (3.35) and (3.36) form a complete system that determines both the characteristics and the solutions along them. It lends itself to a numerical method in an obvious way. If we use e.g. forward differences this leads to the method of Massau. In figure 12.9 we have sketched the idea. We denote by $t_P$ the value of the variable $t$ at the point $P$, etc.

First we discretise (3.32) and (3.35), relating it to a step size $\Delta \xi$ and $\Delta \eta$, respectively, which can be chosen to be constant during the process, i.e.

$$
t_Q - t_P \doteq \Delta \xi, \quad x_Q - x_P \doteq \lambda_\pi \Delta \xi,
$$

(3.37a)

$$
t_Q - t_R \doteq \Delta \eta, \quad x_Q - x_R \doteq \lambda_w \Delta \eta.
$$

(3.37b)
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

This then is used in the discretised form of (3.33) and (3.36) to give
\[ \tilde{v}_{1,p}(u_{1,Q} - u_{1,p}) + \tilde{v}_{2,p}(u_{2,Q} - u_{2,p}) = \tilde{c}_1(t_Q - t_p), \tag{3.38a} \]
\[ \tilde{w}_{1,R}(u_{1,Q} - u_{1,R}) + \tilde{w}_{2,R}(u_{2,Q} - u_{2,R}) = \tilde{c}_2(t_Q - t_R), \tag{3.38b} \]
respectively. Suppose we are dealing with a Cauchy problem, so the data are given at \( t = 0 \). Then typically (3.37) determines the point \( Q \), while from (3.38) we can find \( u_{1,Q} \) and \( u_{2,Q} \). From the previous approximation method it immediately follows that we are facing complications if \( m > 2 \).

4 Weak formulation of first order systems

In this section we consider the nonlinear system (3.25). An example of such a system are the shallow water equations, which we will discuss in detail in the next section. In many applications nonlinear systems have discontinuous solutions and that is why we need to take recourse to weak solutions. In Section 4.1 we generalise the concept of weak solutions for hyperbolic systems and in Section 4.2 we investigate (weak) solutions of the Riemann problem.

4.1 Weak solutions

In this section we give the mathematical definition of a weak solution, which is a straightforward generalization of definition 12.8. The discussion is rather concise, since it is very similar to the scalar case.

Let \( \phi \in D^m \) be the space of test functions, with \( D \) defined in (2.6). If we take the inner product of (3.25) with \( \phi \), integrate over \( \mathbb{R} \times [0, \infty) \) and subsequently apply Green’s theorem, we get the relation
\[
\begin{align*}
\int_0^\infty \int_{-\infty}^\infty & (\mathbf{u} \cdot \frac{\partial \phi}{\partial t} + f(\mathbf{u}) \cdot \frac{\partial \phi}{\partial x}) \, dx \, dt = -\int_{-\infty}^\infty \mathbf{u}(x, 0) \cdot \phi(x, 0) \, dx - \int_0^\infty \int_{-\infty}^\infty \mathbf{c} \cdot \phi \, dx \, dt. \tag{4.1}
\end{align*}
\]

In the derivation of (4.1) we have used that \( \phi(x, t) \) vanishes for \( |x| + t \to \infty \). We then have the following definition.

**Definition 12.23.** A function \( \mathbf{u}(x, t) \) is called a weak solution of system (3.25) if relation (4.1) holds for all test functions \( \phi \in (C_0^1(\mathbb{R} \times [0, \infty)))^m \).

A weak solution that often occurs is a piecewise smooth solution, where the smooth parts are connected by discontinuities. These discontinuities cannot be of arbitrary size, as is apparent from the following theorem.

**Theorem 12.24.** Let \( \mathbf{u} \) be a piecewise smooth solution of (3.25), that has a discontinuity across a curve \( E : x = x(t) \). Then \( \mathbf{u} \) satisfies the condition
\[
[f(\mathbf{u})]^+_x = s[\mathbf{u}]^+_x, \tag{4.2}
\]
4. WEAK FORMULATION OF FIRST ORDER SYSTEMS

with $[v]^+_t := v(x(t)-, t) - v(x(t)+, t)$ the jump of $v$ across $\mathcal{E}$ and $s$ the speed of $\mathcal{E}$.

**Proof.** The curve $\mathcal{E}$ separates a domain $\Omega \supset \text{supp}(\phi)$ in a left part $\Omega_\ell$ and a right part $\Omega_r$. The solution $u$ is smooth in both subdomains. Since (3.25) holds in $\Omega_\ell$, we have

$$\int_{\Omega_\ell} \left( \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} \right) \cdot \phi \, dx \, dt = \int_{\Omega_\ell} c \cdot \phi \, dx \, dt,$$

for every test function $\phi \in D^m$. Using the product rule of differentiation, we can rewrite this relation as follows:

$$\int_{\Omega_\ell} \left( \frac{\partial}{\partial t} (u \cdot \phi) + \frac{\partial}{\partial x} (f(u) \cdot \phi) \right) \, dx \, dt =$$

$$\int_{\Omega_\ell} \left( u \cdot \frac{\partial \phi}{\partial t} + f(u) \cdot \frac{\partial \phi}{\partial x} \right) \, dx \, dt + \int_{\Omega_\ell} c \cdot \phi \, dx \, dt.$$

Next, we apply the two-dimensional Gauss theorem to the integral in the left-hand side, and find

$$\int_{\Omega_\ell} \left( \frac{\partial}{\partial t} (u \cdot \phi) + \frac{\partial}{\partial x} (f(u) \cdot \phi) \right) \, dx \, dt =$$

$$\int_{\partial \Omega_\ell} \varphi \cdot (u \, dx - f(u) \, dt)$$

with $\partial \Omega_\ell$ the boundary of $\Omega_\ell$ and $u_\ell := u(x(t)-, t)$ the limit value of $u$ just left of $\mathcal{E}$; see figure 12.3. In the derivation of this relation we have used that $\varphi(x, t) = 0$ for $(x, t) \in \partial \Omega \cap \{t > 0\}$, with $\partial \Omega$ the boundary of $\Omega$. Combining these relations we obtain

$$- \int_{x_a} u(x, 0) \cdot \varphi(x, 0) \, dx - \int_{\mathcal{E}} \varphi \cdot (u_\ell \, dx - f(u_\ell) \, dt)$$

$$= \int_{\Omega_\ell} \left( u \cdot \frac{\partial \phi}{\partial t} + f(u) \cdot \frac{\partial \phi}{\partial x} \right) \, dx \, dt + \int_{\Omega_\ell} c \cdot \phi \, dx \, dt. \quad (\ast)$$

In the same fashion, we find for the right subdomain $\Omega_r$,

$$- \int_{x_b} u(x, 0) \cdot \varphi(x, 0) \, dx + \int_{\mathcal{E}} \varphi \cdot (u_r \, dx - f(u_r) \, dt)$$

$$= \int_{\Omega_r} \left( u \cdot \frac{\partial \phi}{\partial t} + f(u) \cdot \frac{\partial \phi}{\partial x} \right) \, dx \, dt + \int_{\Omega_r} c \cdot \phi \, dx \, dt, \quad (\ast\ast)$$

with $u_r := u(x(t)+, t)$ the limit value of $u$ just right of $\mathcal{E}$. The integral over $\mathcal{E}$ in $\ast\ast$ is evaluated in the same direction as in $(\ast)$, see figure 12.3, and therefore has a $+$ sign in front of it. Adding the relations $(\ast)$ and $(\ast\ast)$ we find

$$- \int_{x_b} u(x, 0) \cdot \varphi(x, 0) \, dx + \int_{\mathcal{E}} \varphi \cdot ([u]^+_t \, dx - [f(u)]^+_t \, dt)$$

$$= \int_{\Omega} \left( u \cdot \frac{\partial \phi}{\partial t} + f(u) \cdot \frac{\partial \phi}{\partial x} \right) \, dx \, dt + \int_{\Omega} c \cdot \phi \, dx \, dt.$$
Combining this last relation with equation (4.1) and taking into account that \( \phi(x, t) \) vanishes outside \( \Omega \), we have
\[
\int_{\Xi} \phi \cdot ([u]^+ dx - [f(u)]^+) dt = 0.
\]
Since this relation holds for arbitrary test functions \( \phi \in \mathcal{D}^m \), we conclude that
\[
[u]^+ dx - [f(u)]^+ dt = 0,
\]
and since \( s = \frac{dx}{dt} \) it is obvious that (4.2) holds.

Written in full, the jump relation (4.2) reads
\[
f(u_r) - f(u_\ell) = s(u_r - u_\ell),
\]
with \( u_r := u(x(t)+, t), \ u_\ell := u(x(t)-, t) \) and is also called the Rankine-Hugoniot jump condition. Relation (4.3) provides \( m \) equations for the \( 2m + 1 \) variables \( u_\ell, u_r \) and \( s \). This result is needed to construct so-called shock wave solutions; see next section.

4.2 The Riemann problem

In this section we will derive the elementary solutions of the Riemann problem for (3.25). We assume the system to be homogeneous, i.e. \( c = 0 \). The canonical form of the Riemann problem for (3.25) reads
\[
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad x \in \mathbb{R}, t > 0,
\]
\[
u(x, 0) = \begin{cases} 
u_\ell & \text{if } x < 0, \\ \nu_r & \text{if } x > 0. \end{cases}
\]
Like in the scalar case, any solution of (4.4) is a similarity solution of the form \( u(x, t) = \hat{u}(x/t) \).

We first consider the linear case, i.e. \( f(u) = Bu \) with \( B \) a constant matrix. Since the eigenvectors \( s_1, s_2, \ldots, s_m \) of the coefficient matrix \( B \) are linearly independent, we can decompose the initial state vectors \( u_\ell \) and \( u_r \) as follows
\[
u_\ell = \sum_{k=1}^{m} \alpha_k s_k, \quad \nu_r = \sum_{k=1}^{m} \beta_k s_k.
\]
Alternatively, the initial condition can be written as
\[
u(x, 0) = S\hat{u}(x, 0) = \sum_{k=1}^{m} \tilde{u}_k(x, 0)s_k.
\]
Comparing (4.4b) and (4.6) and using that the eigenvectors $s_k$ are linearly independent, we see that

$$\tilde{u}_k(x, 0) = \begin{cases} \alpha_k & \text{if } x < 0, \\ \beta_k & \text{if } x > 0. \end{cases}$$ (4.7)

Since the eigenvalues $\lambda_k$ are constant, the variables $\tilde{u}_k$ can be readily computed from the IVP (3.9b) and (4.7), and we find

$$\tilde{u}_k(x, t) = \tilde{u}_k(x - \lambda_k t, 0) = \begin{cases} \alpha_k & \text{if } x/t < \lambda_k, \\ \beta_k & \text{if } x/t > \lambda_k. \end{cases}$$ (4.8)

Inserting this in the relation $u = S\tilde{u}$, we obtain the following solution

$$u(x, t) = \sum_{k=1}^{m} \tilde{u}_k(x, t)s_k = \sum_{x/t < \lambda_k} \alpha_k s_k + \sum_{x/t > \lambda_k} \beta_k s_k.$$ (4.9)

The solution $u$ is thus piecewise constant, because the initial discontinuity at $x = 0$ propagates along all characteristics. The patches of constant $u$ in the $(x, t)$-plane are separated by the characteristics. As an example, we show in figure 12.10 the solution for a $3 \times 3$ system with $\lambda_1 < 0$ and $\lambda_2, \lambda_3 > 0$.

Next, we consider the quasilinear case. Note, that since $B$ depends on $u$ all eigenvalues and eigenvectors depend on $u$ as well. The general solution of a Riemann problem is hard to obtain. Instead, we will derive specific elementary wave solutions corresponding with an eigenvalue, viz. a simple wave, contact discontinuity and shock. In the next section we will solve the Riemann problem for the shallow water equations in full detail. First, we introduce the following definitions.

![Figure 12.10. Similarity solution of the Riemann problem for a $3 \times 3$ linear system. The triple $(\beta_1, \alpha_2, \alpha_3)$ denotes the solution $u = \beta_1 s_1 + \alpha_2 s_2 + \alpha_3 s_3$, etc..](image-url)

17:22 7 Oct 2003 204 version: 26-09-2003
Definition 12.25. An eigenvector \( s_k \) is called genuinely nonlinear, if
\[
\left( \nabla_u \lambda_k(u), s_k(u) \right) \neq 0
\]
for all \( u \), where \( \nabla_u := (\partial/\partial u_1, \partial/\partial u_2, \ldots, \partial/\partial u_m) \). Likewise, an eigenvector \( s_k \) is called linearly degenerate if
\[
\left( \nabla_u \lambda_k(u), s_k(u) \right) = 0
\]
for all \( u \).

Case 1: simple wave solution
Assume that \( \lambda_k(u_\ell) < \lambda_k(u_r) \) and that \( s_k \) is genuinely nonlinear. In this case we can normalize \( s_k \) such that
\[
\left( \nabla_u \lambda_k(u), s_k(u) \right) = 1, \tag{4.11}
\]
for all \( u \). For an arbitrary left state vector \( u_\ell \), we consider the following initial value problem
\[
\frac{d\hat{u}(\eta)}{d\eta} = s_k(\hat{u}(\eta)), \quad \eta > 0, \tag{4.12a}
\]
\[
\hat{u}(0) = u_\ell, \tag{4.12b}
\]
which defines an integral curve in phase space which is everywhere tangent to \( s_k \). Let \( u_r = \hat{u}(\eta_r) \) for some \( \eta_r > 0 \). Since
\[
\frac{d}{d\eta} \left( \lambda_k(\hat{u}(\eta)) \right) = \left( \nabla_u \lambda_k(\hat{u}(\eta)), s_k(\hat{u}(\eta)) \right) = 1,
\]
we have
\[
\lambda_k(\hat{u}(\eta)) = \eta + \lambda_k(u_\ell), \quad \lambda_k(u_\ell) = \eta + \lambda_k(u_\ell) > \lambda_k(u_\ell).
\]
Next, we will show that the solution of the Riemann problem (4.4) is given by
\[
u(x, t) = \begin{cases} 
    u_\ell & \text{if } x/t < \lambda_k(u_\ell), \\
    \hat{u}(x/t - \lambda_k(u_\ell)) & \text{if } \lambda_k(u_\ell) < x/t < \lambda_k(u_r), \\
    u_r & \text{if } \lambda_k(u_r) < x/t.
\end{cases} \tag{4.13}
\]
This solution is called a \( k \)-simple wave or a \( k \)-rarefaction wave. We restrict ourselves to the nontrivial case is \( \lambda_k(u_\ell) < x/t < \lambda_k(u_r) \). We readily see that
\[
\lambda_k(u(x, t)) = \lambda_k(\hat{u}(x/t - \lambda_k(u_\ell))) = x/t - \lambda_k(u_\ell) + \lambda_k(u_\ell) = x/t.
\]
Therefore, using (4.12a) and the previous equation, we have
\[
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \frac{\partial u}{\partial t} + B(u) \frac{\partial u}{\partial x} \\
= -\frac{x}{t^2} s_k(\hat{u}(\eta)) + \frac{1}{t} B(\hat{u}(\eta)) s_k(\hat{u}(\eta)) \\
= \frac{1}{t} \left( -\frac{x}{t} + \lambda_k(\hat{u}(\eta)) \right) s_k(\hat{u}(\eta)) = 0.
\]
4. WEAK FORMULATION OF FIRST ORDER SYSTEMS

![Figure 12.11. Wave pattern of a k-simple wave.](image)

So, \( u(x, t) \) defined in (4.13) is the solution of the Riemann problem (4.4) indeed. An illustration of this solution is given in figure 12.11.

For the computation of \( k \)-simple waves from the initial value problem (4.12), the so-called \( k \)-Riemann invariants are useful. They are defined as follows.

**Definition 12.26.** A \( k \)-Riemann invariant of (4.4a) is a continuously differentiable function \( w_k : \mathbb{R}^m \rightarrow \mathbb{R} \), such that

\[
\left( \nabla_u w_k(u), s_k(u) \right) = 0,
\]

(4.14)

for all \( u \).

Relation (4.14) is a first order equation which usually can be integrated exactly. Let \( \hat{u}(\eta) \) be a solution of (4.12). Then we have

\[
\frac{d}{d\eta} w_k(\hat{u}(\eta)) = \left( \nabla_u w_k(\hat{u}(\eta)), s_k(\hat{u}(\eta)) \right) = 0,
\]

(4.15)

i.e. \( w_k(\hat{u}(\eta)) \) is constant along the integral curve of (4.12). One can prove that there exist \( m - 1 \) such \( k \)-Riemann invariants \( w_k^{(1)}, w_k^{(2)}, \ldots, w_k^{(m-1)} \) with linearly independent gradients [40]. Then it is clear that the integral curve of (4.12) is part of the curve \( K \) given by

\[
K := \{ u \in \mathbb{R}^m | w_k^{(j)}(u) = w_k^{(j)}(u_\ell), \ j = 1, 2, \ldots, m - 1 \}.
\]

(4.16)

We will use this result in the next section to compute \( k \)-simple waves for the shallow water equations.

**Case 2: contact discontinuity**

Assume that \( s_k \) is linearly degenerate. Let \( \hat{u}(\eta) \) be the solution of (4.12) with \( \hat{u}(\eta_r) = u_\ell \).

Then we readily see that

\[
\frac{d}{d\eta} (\lambda_k(\hat{u}(\eta))) = \left( \nabla_u \lambda_k(\hat{u}(\eta)), s_k(\hat{u}(\eta)) \right) = 0.
\]
Chapter 12. Analysis of Hyperbolic Equations

Figure 12.12. Wave pattern of a contact discontinuity.

\[ \frac{\partial}{\partial t} \left( f(\hat{\mathbf{u}}(\eta)) - \lambda_k(\hat{\mathbf{u}}(\eta)) \hat{\mathbf{u}}(\eta) \right) = B(\hat{\mathbf{u}}(\eta)) \frac{d\hat{\mathbf{u}}}{d\eta} - \lambda_k(\hat{\mathbf{u}}(\eta)) \frac{d\hat{\mathbf{u}}}{d\eta} \]

and consequently, \[ f(\hat{\mathbf{u}}(\eta)) - \lambda_k(\hat{\mathbf{u}}(\eta)) \hat{\mathbf{u}}(\eta) = C \quad (C \in \mathbb{R}^m). \]

Case 3: shock wave

Assume that \( \lambda_k(\mathbf{u}_\ell) > \lambda_k(\mathbf{u}_r) \) and that \( s_k \) is genuinely nonlinear. In this case, the solution of the Riemann problem (4.4) is given by the \( k \)-shock wave

\[ \mathbf{u}(x, t) = \begin{cases} \mathbf{u}_\ell & \text{if } x/t < s, \\ \mathbf{u}_r & \text{if } s < x/t. \end{cases} \quad (4.18) \]

where the shock speed \( s \) has to satisfy the Rankine-Hugoniot jump condition (4.3). An illustration of a \( k \)-shock wave is given in figure 12.13.

Like in the scalar case, we need a simple criterion to determine whether a shock wave is physically correct. This is given by the following definition.

Definition 12.27 (Lax entropy condition). The \( k \)-shock wave (4.18) satisfies the entropy
5. THE SHALLOW WATER EQUATIONS

Figure 12.13. Wave pattern of a k-shock.

condition, if the following inequalities hold

\[ \lambda_{k-1}(u_\ell) < s < \lambda_k(u_\ell), \quad (4.19a) \]
\[ \lambda_k(u_\ell) < s < \lambda_{k+1}(u_\ell), \quad (4.19b) \]

with \( s \) the shock speed given by (4.3).

These inequalities imply that \( m - k + 1 \) characteristics move into the shock from the left and \( k \) from the right. Integrating the ODE system (3.24) along these characteristics, we find \( m - k + 1 \) relations between \( u_\ell \) and the initial condition left of the shock and \( k \) relations between \( u_r \) and the initial condition right of the shock. Together with the Rankine-Hugoniot jump conditions these constitute \( m - k + 1 + k + 1 = 2m + 1 \) equations for the same number of unknowns, viz. \( u_\ell, u_r \) and \( s \).

The general solution of a Riemann problem involves the elementary solutions introduced above, as described in the following theorem; for a proof see e.g. [104].

**Theorem 12.28.** Suppose that the system (4.4a) is hyperbolic and that each eigenvector of the Jacobi matrix of \( f(u) \) is either genuinely nonlinear or linearly degenerate. Then for any \( u_\ell \in \mathbb{R}^m \) there exists neighbourhood \( N \) of \( u_\ell \) such that the Riemann problem (4.4) has a unique solution if \( u_r \in N \). This solution consists of at most \( m + 1 \) constant states separated by shocks, simple waves or contact discontinuities.

5 The shallow water equations

In this section we apply the theory of the previous section to the shallow water equations. The one-dimensional shallow water equations describe flow in a straight canal and read [116]

\[ \frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial x}(\varphi u) = 0, \quad (5.1a) \]
\[ \frac{\partial}{\partial t}(\varphi u) + \frac{\partial}{\partial x}(\varphi u^2 + \frac{1}{2}\varphi^2) = 0, \quad (5.1b) \]
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

where \( x \) is the coordinate along the canal, \( u \) the flow velocity and \( \varphi := gh \) the so-called geopotential, with \( h > 0 \) the depth of the canal and \( g \) the gravitational acceleration. The first equation in (5.1) describes conservation of mass and the second conservation of momentum. An alternative formulation is presented in Chapter 15. These equations can be written in the standard form (3.25) with \( c(u) = 0 \) and \( f(u) \) defined by:

\[
\begin{align*}
\mathbf{u} & := \begin{pmatrix} \varphi \\ \varphi u \end{pmatrix}, \\
f(u) & := \begin{pmatrix} \varphi u \\ \varphi u^2 + \frac{1}{2} \varphi^2 \end{pmatrix}.
\end{align*}
\]

(5.2)

The Jacobi matrix \( B(u) \) of the flux is

\[
B(u) = \begin{pmatrix} 0 & 1 \\ \varphi - u^2 & 2u \end{pmatrix},
\]

(5.3)

and its eigenvalues \( \lambda_k(u) \) and eigenvectors \( s_k(u) \) \((k = 1, 2)\) are given by

\[
\begin{align*}
\lambda_1(u) &= u - c, \\
\lambda_2(u) &= u + c, \\
s_1(u) &= \begin{pmatrix} 1 \\ u - c \end{pmatrix}, \\
s_2(u) &= \begin{pmatrix} 1 \\ u + c \end{pmatrix}.
\end{align*}
\]

(5.4a, 5.4b)

The shallow water equations are thus a \( 2 \times 2 \) hyperbolic system of equations. Moreover, by direct substitution in (4.14), we see that \( w_1(u) = u + 2c \) and \( w_2(u) = u - 2c \) are the 1- and 2-Riemann invariants of the shallow water equations, respectively.

Consider the Riemann problem for these equations, i.e.

\[
\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad x \in \mathbb{R}, \quad t > 0,
\]

(5.5a)

\[
\mathbf{u}(x, 0) = \begin{cases} 
\mathbf{u}_\ell & \text{if } x < 0, \\
\mathbf{u}_r & \text{if } x > 0,
\end{cases}
\]

(5.5b)

with \( \mathbf{u} \) and \( f(u) \) as defined in (5.2). Since the eigenvectors \( s_k(u) \) are genuinely nonlinear, the solution of (5.5) consists of (at most) three constant states, separated by shocks and/or rarefaction waves. The possible wave patterns for the Riemann problem (5.5) are shown in figure 12.14.

Suppose that the constant states \( \mathbf{u}_\ell \) and \( \mathbf{u}_r \) are separated by a shock, referred to as the 1-shock. We like to establish a relation between \( \mathbf{u}_\ell \) and \( \mathbf{u}_r \) using the entropy condition and the Rankine-Hugoniot jump conditions. Let \( s_1 \) denote the speed of the shock. For ease of presentation, we introduce the variables

\[
v_1 := u - s_1, \quad m_1 := \varphi v_1,
\]

(5.6)

i.e. \( v_1 \) is the flow velocity relative to the 1-shock. The entropy condition for the 1-shock reads:

\[
\begin{align*}
\lambda_1(\mathbf{u}_\ell) &< s_1 < \lambda_2(\mathbf{u}_r), \\
\lambda_1(\mathbf{u}_r) &< s_1 < \lambda_2(\mathbf{u}_r),
\end{align*}
\]

(5.7a, 5.7b)
5. THE SHALLOW WATER EQUATIONS

![Diagram of wave patterns](image)

**Figure 12.14.** Possible wave patterns of the Riemann problem for the shallow water equations.

which means that three characteristics go into the shock as shown in figure 12.15. From (5.6) and the entropy condition (5.7) we derive the following inequalities

\[ v_{1,\epsilon} > c_{\epsilon}, \quad |v_{1,*}| < c_{*}. \quad (5.8) \]

The 1-shock also satisfies the jump conditions

\[ s_{1}[\psi]_{-}^{+} = [\psi u]_{-}^{+}, \quad (5.9a) \]
\[ s_{1}[\psi u]_{-}^{+} = [\psi u^2 + \frac{1}{2}\psi^2]_{-}^{+}, \quad (5.9b) \]

where \([a]_{-}^{+} := a_{*} - a_{\epsilon}\) for a generic variable \(a\). Introducing the variables \(v_1\) and \(m_1\) into (5.9), these relations simplify to

\[ [m_1]_{-}^{+} = 0, \quad (5.10a) \]
\[ [m_1 v_1 + \frac{1}{2}\psi^2]_{-}^{+} = 0. \quad (5.10b) \]
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

Combining both jump conditions in (5.10) and using that \( m_1 = \phi \ell v_1, \ell > 0 \), we find for \( m_1 \)
\[
m_1 = \sqrt{\frac{1}{2}(\phi_\ell + \phi_\ast)\phi_\ell \phi_\ast}.
\]
(5.11)

Furthermore, from the second jump condition (5.10b) we obtain the relation
\[
u_* - u_\ell = -\frac{1}{2} \phi_\ast - \phi_\ell m_1.
\]
(5.12)

Let the variable \( z_1 \) be defined by
\[
z_1 := \frac{\phi_\ast}{\phi_\ell}.
\]
(5.13)

From the first jump condition (5.10a) and the inequalities in (5.8), we can conclude that \( z_1 > 1 \), i.e. the geopotential increases when crossing the 1-shock from left to right. Substituting (5.11) into (5.12) we get the relation
\[
u_* - u_\ell \frac{c_\ell}{c_\ast} = \frac{1}{2} \sqrt{2(1 - z_1)} \sqrt{1 + \frac{1}{z_1}},
\]
(5.14)

from which we conclude that the flow velocity decreases when passing the 1-shock from left to right.

Alternatively, let the constant states \( u_\ell \) and \( u_* \) be separated by a rarefaction wave, which we will call the 1-rarefaction wave. In this case we have
\[
\lambda_1(u_*) > \lambda_1(u_\ell),
\]
(5.15)

and the rarefaction wave consists of rays \( x/t = \lambda_1(u) \) and is bounded on both sides by the 1-characteristics as shown in figure 12.15. As shown in Section 4.2, the Riemann invariant \( w_1 \) is constant across the 1-rarefaction wave implying that
\[
u + 2c = u_\ell + 2c_\ell = u_* + 2c_\ast.
\]
(5.16)

Figure 12.15. The 1-wave is either a shock (left) or a rarefaction wave (right).
5. THE SHALLOW WATER EQUATIONS

From (5.16) we easily find
\[
\frac{u_* - u_\ell}{c_\ell} = 2(1 - \sqrt{z_1}),
\]
(5.17)
with \(z_1\) defined in (5.13). From (5.15) and (5.16) we see that \(0 < z_1 < 1\), i.e. the geopotential decreases and the flow velocity increases when passing the 1-rarefaction wave from left to right.

Summarizing, we have the following relation for the 1-wave connecting the states \(u_\ell\) and \(u_*\),
\[
\frac{u_* - u_\ell}{c_\ell} = g(z_1),
\]
(5.18)
with \(z_1\) defined in (5.13) and where the function \(g(z)\) is defined by
\[
g(z) := \begin{cases} 
2(1 - \sqrt{z}) & \text{if } 0 < z \leq 1, \\
\frac{\sqrt{2}(1 - z)}{\sqrt{1 + \frac{1}{z}}} & \text{if } z > 1,
\end{cases}
\]
(5.19)
see figure 12.16. The case \(0 < z_1 \leq 1\) corresponds with a 1-rarefaction wave and for \(z_1 > 1\) we have a 1-shock.

Now we consider the 2-wave connecting the constant states \(u_*\) and \(u_r\). First consider the case of a 2-shock. Let \(s_2\) denote the speed of the shock. Analogous to (5.6) we introduce the variables
\[
v_2 := u - s_2, \quad m_2 := \varphi v_2.
\]
(5.20)
The entropy condition for the 2-shock reads:
\[
\lambda_2(u_\ell) < s_2 < \lambda_2(u_*),
\]
(5.21a)
\[
\lambda_1(u_\ell) < s_2 < \lambda_1(u_*),
\]
(5.21b)

Figure 12.16. The function \(g(z)\).
and also in this case, three characteristics go into the shock as shown in figure 12.17. From (5.20) and the entropy condition (5.21) we find the inequalities

\[ v_{2,r} < -c_r, \quad |v_{2,s}| < c_s. \]  

(5.22)

The jump conditions for the 2-shock are identical to (5.9), with \( s_1 \) replaced by \( s_2 \) and where now \([a]^+_1 := a_r - a_s\) for a generic variable \( a \). In a similar way as for the 1-shock we then obtain the jump conditions

\[ \left[ m_2 \right]^+ = 0, \]  

(5.23a)

\[ \left[ m_2 v_2 + \frac{1}{2} \psi^2 \right]^+ = 0. \]  

(5.23b)

Combining both jump conditions in (5.23) and using that \( m_2 = \psi_r \), we find

\[ m_2 = -\sqrt{\frac{1}{2}(\psi_r + \psi_s)\psi_r \psi_s}. \]  

(5.24)

Also, from the second jump condition (5.23b) we can derive the relation

\[ u_r - u_s = -\frac{1}{2} \frac{\psi_r^2 - \psi_s^2}{m_2}. \]  

(5.25)

Substituting (5.24) in (5.25), we get

\[ \frac{u_s - u_r}{c_r} = \frac{1}{2} \sqrt{2(z_2 - 1)} \sqrt{1 + \frac{1}{z_2}}, \]  

(5.26)

with the variable \( z_2 \) defined by

\[ z_2 := \frac{\psi_s}{\psi_r}. \]  

(5.27)

Figure 12.17. The 2-wave is either a shock (left) or a rarefaction wave (right).

5. THE SHALLOW WATER EQUATIONS

From the first jump condition in (5.23) and the inequalities in (5.22) we see that $z_2 > 1$, and consequently both the geopotential and the flow velocity decrease when passing the 2-shock from left to right.

Alternatively, when the constant states $u_*$ and $u_\ell$ are connected by a 2-rarefaction wave, we have

$$\lambda_2(u_*) > \lambda_2(u_\ell),$$

(5.28)

and rays $x/t = \lambda_2(u)$ are bounded by the 2-characteristics as shown in figure 12.17. In this case, the Riemann invariant $\omega_2$ is constant across the rarefaction wave, leading to

$$u - 2c = u_* - 2c_* = u_\ell - 2c_\ell.$$ (5.29)

Rearranging terms, we obtain

$$\frac{u_* - u_\ell}{c_\ell} = 2(\sqrt{z_2} - 1).$$ (5.30)

We conclude from (5.28) and (5.29) that $0 < z_2 < 1$ and consequently both the geopotential and the flow velocity increase when passing the 2-rarefaction wave from left to right.

Summarizing, for the 2-wave connecting the states $u_*$ and $u_\ell$, the following relation holds

$$\frac{u_* - u_\ell}{c_\ell} = -g(z_2),$$ (5.31)

with the variable $z_2$ and the function $g(z)$ defined in (5.27) and (5.19), respectively. For $0 < z_2 < 1$ the 2-wave is a rarefaction wave and for $z_2 > 1$ it is a shock.

By elimination of the intermediate values $\varphi_*$ and $u_*$ from the definitions (5.13) and (5.27) and the equations (5.18) and (5.31), we obtain the equations

$$z_1 \varphi_\ell = z_2 \varphi_\ell,$$ (5.32a)

$$u_\ell + c_\ell g(z_1) = u_\ell - c_\ell g(z_2).$$ (5.32b)

Introducing the auxiliary variables

$$A := \frac{\varphi_\ell}{\varphi_\ell}, \quad B := \frac{u_* - u_\ell}{c_\ell},$$ (5.33)

and eliminating $z_2 = z_1/A$, we find the following nonlinear equation for $z_1$:

$$G(z_1) := g(z_1) + \sqrt{A} g(z_1/A) - B = 0.$$ (5.34)

We can easily verify that the function $G(z_1)$, defined in (5.34), has the following properties: $G(0) = 2(1 + \sqrt{A}) - B$, $G'(z_1) < 0$ and $G''(z_1) > 0$. These conditions imply that the nonlinear equation (5.34) has a unique solution provided $G(0) > 0$. In terms of the variables $u_\ell$ and $u_\ell$, this latter inequality boils down to

$$u_\ell - u_\ell < 2(c_\ell + c_\ell).$$ (5.35)

The Riemann problem (5.5) has a unique solution if the inequality (5.35) holds.

To summarize the Riemann problem (5.5) can be solved as follows:
1. compute $A$ and $B$ from (5.33),
2. solve equation (5.34) for $z_1$,
3. compute $z_2 = z_1/A$,
4. determine $\varphi_*$ and $u_*$ for the intermediate state from e.g. (5.13) and (5.18).

In the last step, we have computed the intermediate state $u_*$ from the left state $u_\ell$. We could have equally well computed $u_*$ from $u_r$ using (5.27) and (5.31). One should note that step 2. above can conveniently be done numerically. Indeed, since $G(z_1) > 0$ and $G''(z_1) > 0$ on the interval $[0, z_1]$, Newton iteration for the numerical solution of equation (5.34) is bound to converge for an initial guess in $[0, z_1]$, cf. [108]. We still have to determine which wave pattern from figure 12.14 is the solution of the Riemann problem. We have seen that the 1-wave is a shock if $z_1 > 1$. This condition is equivalent to $G(1) > 0$, or stated in terms of the variables $A$ and $B$, $B < \sqrt{Ag(1/A)}$. Otherwise, if $B \geq \sqrt{Ag(1/A)}$, the 1-wave is a rarefaction wave. Likewise, the 2-wave is a shock if $z_2 > 1$, or equivalently, $G(A) > 0$. This latter inequality gives the condition $B < g(A)$. On the other hand, if $B \geq g(A)$, the 2-wave is a rarefaction wave.

Putting everything together, we have the following similarity solution $u(x, t) = \hat{u}(x/t, u_\ell, u_r)$ of the Riemann problem (5.5).

1-shock if $B < \sqrt{Ag(1/A)}$:  
\[
\begin{align*}
   u(x, t) &= \begin{cases} 
   u_\ell & \text{if } x/t < s_1, \\
   u_* & \text{if } x/t > s_1,
   \end{cases} 
   \end{align*}
\]  
(5.36a)

with shock speed $s_1$ given by  
\[
  s_1 = u_\ell - c_\ell \sqrt{\frac{1}{2}(1 + z_1)z_1},
\]  
(5.36b)

1-rarefaction wave if $B \geq \sqrt{Ag(1/A)}$:  
\[
\begin{align*}
   u(x, t) &= u_\ell & \text{if } x/t < u_\ell - c_\ell, \\
   u + 2c &= u_\ell + 2c_\ell, & \text{if } u_\ell - c_\ell < x/t < u_* - c_*, \\
   u - c &= x/t, & \text{if } x/t > u_* - c_*,
   \end{align*}
\]  
(5.37)

2-shock if $B < g(A)$:  
\[
\begin{align*}
   u(x, t) &= \begin{cases} 
   u_* & \text{if } x/t < s_2, \\
   u_r & \text{if } x/t > s_2,
   \end{cases} 
   \end{align*}
\]  
(5.38a)

with shock speed $s_2$ given by  
\[
  s_2 = u_r + c_r \sqrt{\frac{1}{2}(1 + z_2)z_2},
\]  
(5.38b)
2-rarefaction wave if \( B \geq g(A) \):

\[
\begin{align*}
\begin{cases}
   u(x, t) = u_* & \text{if } x/t < u_* + c_* , \\
   u - 2c = u_r - 2c_r & \text{if } u_* + c_* < x/t < u_r + c_r , \\
   u + c = x/t & \text{if } x/t > u_r + c_r .
\end{cases}
\end{align*}
\]  

Expression (5.36b) for the shock speed \( s_1 \) follows readily from (5.6) and (5.11), and likewise, expression (5.38b) for \( s_2 \) can be derived from (5.20) and (5.24).

6 The wave equation

A special kind of hyperbolic equations is given by second order problems, as discussed in Section 2.3. In particular the so-called wave equation

\[
\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}.
\]  

(given here in a one-dimensional medium) occurs in the modelling of many phenomena. We shall first consider solutions of this equation in one space dimension in Section 6.1. In Section 6.2 we discuss solutions in more space dimensions.

6.1 One-dimensional problems

We can easily derive a first order system associated with (6.1). Introducing the variables

\[
p := \frac{\partial u}{\partial t}, \quad q := \frac{\partial u}{\partial x},
\]

we obtain

\[
\frac{\partial}{\partial t} \begin{pmatrix} p \\ q \end{pmatrix} - \begin{pmatrix} 0 & a^2 \\ 1 & 0 \end{pmatrix} \frac{\partial}{\partial x} \begin{pmatrix} p \\ q \end{pmatrix} = 0.
\]  

(6.2)

Hence we immediately see that the system matrix has two eigenvalues, \(-a\) and \(a\), corresponding to the two characteristics, \( C_1 \) and \( C_2 \), given by

\[
\begin{align}
C_1 & := \{(x, t) \mid x + at = \text{constant}\}, \\
C_2 & := \{(x, t) \mid x - at = \text{constant}\}.
\end{align}
\]

(6.3a, 6.3b)

These characteristics imply that we should describe two initial-boundary conditions. In the simple case of a Cauchy problem, i.e. data given on the line \( t = 0 \), we can e.g. prescribe \( u \) and \( \frac{\partial u}{\partial t} \). So let for some given functions \( v(x) \) and \( w(x) \)

\[
\begin{align}
   u(x, 0) &= v(x), \\
   \frac{\partial u}{\partial t}(x, 0) &= w(x).
\end{align}
\]  

(6.4a, 6.4b)
There is a nice way to construct a solution, named after d’Alembert. Since (6.2) has constant coefficients we first note that we can find the normal form, cf. (3.24),

\[ \frac{d\tilde{u}_k}{ds} = 0 \quad \text{on} \quad \mathcal{C}_k, \quad (k = 1, 2). \tag{6.5} \]

Because of linearity we have

\[ u(x, t) = \tilde{u}_1(x + at) + \tilde{u}_2(x - at). \tag{6.6} \]

Note that \( \tilde{u}_1 \) and \( \tilde{u}_2 \) are unique, but for a multiplicative constant. From (6.4a) we derive

\[ \tilde{u}_1(x) + \tilde{u}_2(x) = v(x), \tag{6.7} \]

and from (6.4b)

\[ a \left( \tilde{u}_1'(x) - \tilde{u}_2'(x) \right) = w(x), \]

where the prime \((')\) denotes differentiation w.r.t. \(x\). Integration of the latter equation results in

\[ \tilde{u}_1(x) - \tilde{u}_2(x) = \frac{1}{a} \int_0^x w(\xi) \, d\xi + C, \quad C \in \mathbb{R}. \tag{6.8} \]

From (6.7) and (6.8) we then deduce

\[ u(x, t) = \frac{1}{2} \left( v(x + at) + v(x - at) \right) + \frac{1}{2a} \int_{-at}^{at} w(\xi) \, d\xi. \tag{6.9} \]

Formula (6.9) is the \emph{d’Alembert solution} of (6.1) and (6.4) and holds on the whole real line.

Often wave phenomena are defined on a semi-infinite or finite interval. One may e.g. think of a (simplified) model for a string attached at one end or two ends. A natural way to solve such problems is by employing reflections. Consider first (6.1), subject to the following boundary and initial conditions

\[ u(0, t) = 0, \quad t \geq 0 \tag{6.10a} \]

\[ u(x, 0) = v(x), \quad \frac{\partial u}{\partial t}(x, 0) = w(x), \quad x > 0. \tag{6.10b} \]

We then define a problem on \((-\infty, \infty)\) by continuing \(v\) and \(w\) as odd functions for negative values of the argument, i.e. the points in the graph are reflected with respect to the origin; so we have functions \(\tilde{v}\) and \(\tilde{w}\) with

\[ \tilde{v}(x) = -v(-x), \quad \tilde{w}(x) = -w(-x), \quad x < 0. \tag{6.11} \]

We shall omit the bar below again and consider the condition (6.10b) as defined on \((-\infty, \infty)\) now. The solution is then formally given by (6.9). At \(x = 0\) we thus find then

\[ u(0, t) = \frac{1}{2} \left( v(at) + v(-at) \right) + \frac{1}{2a} \int_{-at}^{at} w(\xi) \, d\xi = 0, \tag{6.12} \]

as is required by (6.10a) indeed. The complete solution of (6.1) and (6.10) is given by

\[ u(x, t) = \begin{cases} \frac{1}{2} (v(x + at) + v(x - at)) + \frac{1}{2a} \int_{x-at}^{x+at} w(\xi) \, d\xi & \text{if } x > at, \\ \frac{1}{2} (v(at + x) - v(at - x)) + \frac{1}{2a} \int_{at-x}^{at+x} w(\xi) \, d\xi & \text{if } x < at. \end{cases} \] (6.13)

Assume \( w(x) = 0 \). Then the initial profile \( v(x) \) is split in two parts, one travelling to the right and one to the left. The left travelling part is reflected and 'inverted' at \( x = 0 \). This part of the solution can be interpreted as the 'inverted' profile originating from \(-x\). For a finite interval the procedure is similar, taking a reflection on the right boundary as well.

### 6.2 Solutions in more dimensions

We now turn to the wave equation in three space variables. So consider

\[ \frac{\partial^2 u}{\partial t^2} = a^2 \nabla^2 u, \] (6.14)

subject to the initial conditions

\[ u(x, 0) = v(x), \quad \frac{\partial u}{\partial t}(x, 0) = w(x). \] (6.15)

One can find a d'Alembert type solution to this problem using so-called averaging. Define for \( \xi \in \mathbb{R}^3 \) the average of \( u \) on a sphere \( S(\xi; r) \) with centre \( \xi \) and radius \( r \), i.e.

\[ \bar{u}(r; \xi) := \frac{1}{4\pi r^2} \int_{S(\xi; r)} u(x, t) \, dS, \quad r \neq 0. \] (6.16)

Introducing the variable \( \tilde{r} := (x - \xi)/r = n \), with \( n \) the outward unit normal on \( S(\xi; r) \), we may as well take the average over the unit sphere, i.e.

\[ \bar{u}(r; \xi) := \frac{1}{4\pi} \int_{S(0; 1)} u(\xi + rn, t) \, d\tilde{S}, \] (6.17)

with \( d\tilde{S} = dS/r^2 \). One easily verifies that \( u \) satisfies

\[ u(\xi, t) = \bar{u}(0, t; \xi). \]

For \( \bar{u} \) we have the following property.

**Property 12.29.** The variable \( r\bar{u} \), with \( \bar{u} \) defined in (6.17), satisfies the one-dimensional equation

\[ \frac{\partial^2}{\partial t^2}(r\bar{u}) = a^2 \frac{\partial^2}{\partial r^2}(r\bar{u}). \] (6.18)
CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

Proof. From (6.17) we find
\[
\frac{\partial^2 \tilde{u}}{\partial t^2} = \frac{1}{4\pi} \int_{S(0,1)} \frac{\partial^2 u}{\partial t^2}(\xi + r \mathbf{n}, t) \, d\tilde{S} = \frac{a^2}{4\pi} \int_{S(0,1)} \nabla^2 u(\xi + r \mathbf{n}, t) \, d\tilde{S}. \quad (*)
\]

On the other hand we find from Gauss’ theorem (cf. J)
\[
r^2 \frac{\partial \bar{u}}{\partial r} = \frac{r^2}{4\pi} \int_{S(0,1)} \nabla u(\xi + r \mathbf{n}, t) \cdot n \, d\tilde{S}
= \frac{1}{4\pi} \int_{S(\xi, r)} \nabla u(x, t) \cdot n \, dS
= \frac{1}{4\pi} \int_{\Omega(\xi, r)} \nabla^2 u(x, t) \, dV,
\]
where \( \Omega(\xi; r) \) is the interior of the sphere \( S(\xi; r) \). The latter integral can be rewritten as
\[
\frac{1}{4\pi} \int_0^r \int_{S(\xi, r')} \nabla^2 u(x, t) \, dS'.
\]
Hence we have
\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \bar{u}}{\partial r} \right) = \frac{1}{4\pi r^2} \int_{S(\xi, r')} \nabla^2 u(x, t) \, dS = \frac{1}{4\pi} \int_{S(0,1)} \nabla^2 u(\xi + r \mathbf{n}, t) \, d\tilde{S} \quad (**)
\]
From (*) and (**) we conclude that
\[
\frac{\partial^2 \tilde{u}}{\partial t^2} = \frac{a^2}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \bar{u}}{\partial r} \right).
\]
By straightforward manipulation this then can be rewritten as (6.18). \( \Box \)

From (6.18) we can obtain an expression for the solution \( r \tilde{u} \) as a function of \( r \) and \( t \), viz.
\[
r \tilde{u}(r, t; \xi) = \tilde{v}_1(r + at) + \tilde{v}_2(r - at), \quad (6.19)
\]
for some \( \tilde{v}_1, \tilde{v}_2 \). By substituting \( r = 0 \) we immediately see that \( \tilde{v}_1(at) = -\tilde{v}_2(-at) \), so that (6.19) turns into
\[
r \tilde{u}(r, t; \xi) = \tilde{v}_1(r + at) - \tilde{v}_1(at - r). \quad (6.20)
\]
We can obtain a simpler relation now for \( u(\xi, t) = \tilde{u}(0, t; \xi) \) by differentiating both sides in (6.20) to \( r \) and setting \( r = 0 \) (the prime (') denotes differentiation to the argument of \( \tilde{v}_1 \)), i.e.
\[
u(\xi, t) = 2\tilde{v}'_1(at). \quad (6.21)
\]
We now determine \( \tilde{v}_1 \). First differentiate (6.20) to \( r \) and set \( t = 0 \)
\[
\frac{\partial}{\partial r} (r \tilde{u})(r, 0; \xi) = \tilde{v}'_1(r) + \tilde{v}'_1(-r). \quad (6.22)
\]
Next differentiate (6.20) to \( t \) and set \( t = 0 \)

\[
\frac{\partial}{\partial t}(r \bar{u})(r, 0; \xi) = a(\bar{v}'_1(r) - \bar{v}'_1(-r)).
\]

(6.23)

In (6.22), (6.23) we can eliminate \( \bar{v}'_1(-r) \) to find an expression for \( 2\bar{v}'_1(r) \), and use this to determine \( u(\xi, t) \). We obtain

\[
2\bar{v}'_1(r) = \frac{\partial}{\partial r}(r \bar{u})(r, 0; \xi) + \frac{r}{a} \frac{\partial \bar{u}}{\partial t}(r, 0; \xi)
\]

\[
= \frac{\partial}{\partial r}\left( \frac{r}{4\pi} \int_{S(0;1)} u(\xi + rn; 0) \, d\tilde{S} \right) + \frac{r}{4\pi a} \frac{\partial}{\partial t} \int_{S(0;1)} u(\xi + rn; 0) \, d\tilde{S}
\]

\[
= \frac{\partial}{\partial r}\left( \frac{r}{4\pi} \int_{S(0;1)} v(\xi + rn) \, d\tilde{S} \right) + \frac{r}{4\pi a} \int_{S(0;1)} w(\xi + rn) \, d\tilde{S}.
\]

(6.24)

The anticipated expression for \( u(\xi, t) \) is now obtained from (6.24) by substituting \( r = at \)

\[
u(\xi, t) = \frac{\partial}{\partial t}\left( \frac{t}{4\pi} \int_{S(0;1)} v(\xi + atn) \, d\tilde{S} \right) + \frac{1}{4\pi a^2 t} \int_{S(\xi, at)} w(\xi) \, dS + \frac{1}{4\pi a^2 t} \int_{S(\xi, at)} w(\xi) \, dS
\]

\[
= \frac{\partial}{\partial t}\left( t \bar{v}(at; \xi) \right) + t \bar{w}(at; \xi),
\]

(6.25)

where \( \bar{v} \) and \( \bar{w} \) are defined analogously to (6.16). The expression (6.25) lends itself for obtaining the solution in \( \mathbb{R}^3 \), by employing the fact that \( v \) and \( w \) are depending on two variables only. We leave this as an exercise.

We can also find a solution of an inhomogeneous problem, employing the Duhamel principle. Consider the problem

\[
\frac{\partial^2 u}{\partial t^2} = a^2 \nabla^2 u + s(x, t),
\]

(6.26)

subject to homogeneous initial conditions (6.15) i.e. \( v(x) = w(x) = 0 \). We recall from Section 4.6 that we can solve this IVP if we can find a solution \( \tilde{u}(x, t; \tau) \) such that

\[
\frac{\partial^2 \tilde{u}}{\partial t^2} = a^2 \nabla^2 u, \quad x \in \mathbb{R}^3, \ t > \tau,
\]

(6.27a)

\[
\tilde{u}(x, \tau; \tau) = 0, \quad x \in \mathbb{R}^3,
\]

(6.27b)

\[
\frac{\partial \tilde{u}}{\partial t}(x, \tau; \tau) = s(x, \tau), \quad x \in \mathbb{R}^3.
\]

(6.27c)

Clearly \( \tilde{u} \) follows from (6.25) and is given by

\[
\tilde{u}(\xi, t; \tau) = \frac{1}{4\pi a^2(t - \tau)} \int_{S(\xi, at - \tau)} s(x, \tau) \, dS.
\]

(6.28)

Hence we find the following representation for the solution of (6.26)

\[
u(\xi, t) = \frac{1}{4\pi a^2} \int_0^t \frac{dr}{t - \tau} \int_{S(\xi, at - \tau)} s(x, \tau) \, dS.
\]

(6.29)
Example 12.30 Consider the one-dimensional case. We can directly apply Duhamel’s principle. We then find the representation

\[
u(\xi, t) = \frac{1}{2a} \int_0^t \int_{\xi-a(t-\tau)}^{\xi+a(t-\tau)} s(x, \tau) \, dx.
\]

\[\square\]

7 Boundary conditions

Proper initial and boundary conditions are crucial for having unique solutions of PDEs. As we have seen, for hyperbolic equations they are propagated along characteristics. In order to find out whether the problem is well-posed it is therefore important to know where the latter emanate. As we shall see, the number of conditions we can impose at a boundary, often referred to as physical boundary conditions, depends on (the sign of) the eigenvalues of the system.

We first discuss the linear case. Consider the following \( m \)-dimensional linear initial boundary value problem for \( u(x, t) \):

\[
\begin{align*}
\frac{\partial u}{\partial t} + B \frac{\partial u}{\partial x} &= 0, & 0 < x < 1, \ t > 0, \ (7.1a) \\
u(x, 0) &= v(x), & 0 < x < 1, \ (7.1b) \\
C_\ell u(0, t) &= g_\ell(t), \quad C_r u(1, t) = g_r(t), \quad t > 0. \quad (7.1c)
\end{align*}
\]

Since the system in (7.1) is hyperbolic, the coefficient matrix \( B \) has \( m \) real eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_m \), of which \( p \) say, are positive and \( m - p \) are negative. Without restriction we may assume that they are ordered as

\[
\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{m-p} < 0 < \lambda_{m-p+1} \leq \cdots \leq \lambda_m. \quad (7.2)
\]

Let further \( C_\ell \) be an \( m_\ell \times m \) matrix and \( g_\ell \) an \( m_\ell \)-vector, which means that we impose \( m_\ell \) boundary conditions at the left boundary \( x = 0 \). Likewise, let \( C_r \) be an \( m_r \times m \) matrix and \( g_r \) an \( m_r \)-vector, implying that we have \( m_r \) boundary conditions at \( x = 1 \). Obviously, \( 0 \leq m_\ell, m_r \leq m \).

To investigate the boundary conditions in (7.1), we need to decouple the system and write it in terms of the characteristic variables \( \tilde{u} \). In Section 3.2, we showed that the characteristic variable \( \tilde{u}_k \), corresponding to the \( k \)th eigenvalue, and its characteristic satisfy the differential equations

\[
\begin{align*}
\frac{d\tilde{u}_k}{dt} &= 0, & dx/dr = \lambda_k, \quad (7.3)
\end{align*}
\]

implying that \( \tilde{u}_k(x, t) = C_k (C_k \in \mathbb{R}) \) along the characteristic \( C_k : x - \lambda_k t = C \) (\( C \in \mathbb{R} \)).

Consider the case \( \lambda_k < 0 \) first. From figure 12.18 it is clear that through every point on the left boundary \( x = 0 \) a characteristic passes emanating from either a point on the initial line \( t = 0 \) or the right boundary \( x = 1 \). This means that we have to prescribe \( \tilde{u}_k \) at \( t = 0 \) and \( x = 1 \); then \( \tilde{u}_k \) is completely determined at the boundary \( x = 0 \), where the characteristics leave the domain. Consequently, we may not even impose any boundary condition for \( \tilde{u}_k \) at
7. BOUNDARY CONDITIONS

Figure 12.18. Characteristics of the hyperbolic system (7.1a) corresponding with \( \lambda_k < 0 \) (left) or \( \lambda_k > 0 \) (right).

\( \lambda_k > 0 \) is similar. Characteristics \( C_k \) enter the domain at \( t = 0 \) and \( x = 0 \), and \( \tilde{u}_k \) has to be specified there. On the other hand, characteristics \( C_k \) leave the domain at \( x = 1 \), so that no boundary conditions for \( \tilde{u}_k \) may be given there. Summarizing, the characteristic variable \( \tilde{u}_k \) must be given at the boundary where characteristics \( C_k \) emanate from.

We introduce the following partitioning of the characteristic variables \( \tilde{u} \)

\[
\tilde{u} = \begin{pmatrix} \tilde{u}^- \\ \tilde{u}^+ \end{pmatrix} \quad \text{with} \quad \tilde{u}^- := \begin{pmatrix} \tilde{u}_1 \\ \vdots \\ \tilde{u}_{m-p} \\ \tilde{u}_{m-p+1} \end{pmatrix}, \quad \tilde{u}^+ := \begin{pmatrix} \tilde{u}_{m-p+1} \\ \vdots \\ \tilde{u}_m \end{pmatrix}, \quad (7.4)
\]

i.e. \( \tilde{u}^- \) and \( \tilde{u}^+ \) contain the characteristic variables corresponding with negative and positive eigenvalues, respectively. Likewise, we split the right eigenvector matrix \( S := T^{-1} \) as follows

\[
S = \begin{pmatrix} S^- & S^+ \end{pmatrix} \quad \text{with} \quad S^- := (s_1, \cdots, s_{m-p}), \quad S^+ := (s_{m-p+1}, \cdots, s_m). \quad (7.5)
\]

The matrices \( S^- \) and \( S^+ \) contain the eigenvectors corresponding to negative and positive eigenvalues, respectively. Now consider the boundary condition at \( x = 0 \) for the variable \( u \). In terms of the characteristic variable \( \tilde{u} \), it can be written as

\[
C_{\ell} \tilde{S}_{\ell}(0, t) = g_{\ell}(t). \quad (7.6)
\]

Substituting the partitionings of (7.4) and (7.5) in (7.6), we find

\[
C_{\ell} S^- \tilde{u}^-(0, t) + C_{\ell} S^+ \tilde{u}^+(0, t) = g_{\ell}(t). \quad (7.7)
\]
From the preceding discussion, we conclude that \( \tilde{u}^-(0, t) \) cannot be prescribed. On the other hand, \( \tilde{u}^+(0, t) \) has to be specified. This means that the \( m_\ell \times p \) matrix \( C_\ell S^+ \) in (7.7) has to be invertible. A minimum requirement is then that \( m_\ell = p \), i.e. the number of boundary conditions at \( x = 0 \) is equal to the number of positive eigenvalues. Analogous to (7.7), we obtain for the boundary condition at \( x = 1 \)

\[
C_\ell S^- \tilde{u}^-(1, t) + C_\ell S^+ \tilde{u}^+(1, t) = g_\ell(t). \quad (7.8)
\]

In this case, \( \tilde{u}^- (1, t) \) has to be given, which means that the \( m_\ell \times (m - p) \) matrix \( C_\ell S^- \) in (7.8) has to be invertible. This in turn implies that we must have \( m_\ell = m - p \), thus the number of boundary conditions at \( x = 1 \) is equal to the number of negative eigenvalues. Concluding, the boundary conditions in (7.1) have to satisfy the following conditions

\[
x = 0 : \quad m_\ell = p, \quad C_\ell S^+ \text{ invertible}, \quad (7.9a)
\]
\[
x = 1 : \quad m_\ell = m - p, \quad C_\ell S^- \text{ invertible}. \quad (7.9b)
\]

**Example 12.31** Consider the linear system from example 12.20 describing tidal waves in a canal. We can easily verify that

\[
S = (s_1, s_2) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S^- = s_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad S^+ = s_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix},
\]

Corresponding with the eigenvalues \( \lambda_1 = -a < 0 \) and \( \lambda_2 = a > 0 \), respectively. We thus have to impose one boundary condition at both boundaries \( x = 0 \) and \( x = 1 \). We can write the boundary condition at \( x = 0 \) in the form

\[
e_\ell^T u(0, t) = g_\ell(t) \quad \text{with} \quad e_\ell^T s_1 \neq 0,
\]

or, equivalently,

\[
e_{\ell, 1} \frac{\partial \eta}{\partial t}(0, t) + e_{\ell, 2} \frac{\partial \eta}{\partial x}(0, t) = g_\ell(t) \quad \text{with} \quad e_{\ell, 1} - e_{\ell, 2} \neq 0.
\]

Likewise, we have at \( x = 1 \) the boundary condition

\[
e_{\ell, 1} \frac{\partial \eta}{\partial t}(1, t) + e_{\ell, 2} \frac{\partial \eta}{\partial x}(1, t) = g_\ell(t) \quad \text{with} \quad e_{\ell, 1} + e_{\ell, 2} \neq 0.
\]

Next, consider the nonlinear initial boundary value problem for \( u(x, t) \):

\[
\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad 0 < x < 1, \quad t > 0, \quad (7.10a)
\]
\[
u(x, 0) = v(x), \quad 0 < x < 1, \quad (7.10b)
\]
\[
C_\ell(u)(0, t) = g_\ell(t), \quad C_\ell(u)(1, t) = g_\ell(t), \quad t > 0. \quad (7.10c)
\]

The nonlinear system (7.10a) is hyperbolic, which means that the Jacobi matrix \( B(u) \) has \( m \) real eigenvalues \( \lambda_\epsilon(u) \). The ordering is the same as in (7.2). Since the eigenvalues depend on the solution \( u \), also the number of positive eigenvalues \( p \) depends on \( u \). In particular, \( p\)
7. BOUNDARY CONDITIONS

can be different on the boundaries \(x = 0\) and \(x = 1\). Therefore, we will use the notation \(p(\xi, t)\) \((\xi = 0, 1)\) to denote the number of positive eigenvalues at \(x = \xi\). In (7.10c), \(\mathcal{C}_\ell(u)\) is a vector function mapping the \(m\)-dimensional space onto the \(m\)-dimensional space and \(g_\ell\) is an \(m\)-vector. We thus impose \(m\) boundary conditions at \(x = 0\). At \(x = 1\) we have \(m\) boundary conditions, where \(\mathcal{C}_\ell(u)\) is a vector function mapping the \(m\)-dimensional space onto the \(m\)-dimensional space and \(g_\ell\) is an \(m\)-vector.

To investigate the boundary conditions for (7.10), we introduce the characteristic variables \(\tilde{u}\) by the Pfaffian differential equation

\[
T \tilde{du} = d\tilde{u}. \tag{7.11}
\]

Assuming that (7.11) is integrable, we can determine \(\tilde{u}\) as a function of \(u\). Formally inverting this function, we may write

\[
u = \mathcal{C}(\tilde{u}). \tag{7.12}
\]

In Section 3.2, we have shown that the characteristic variables \(\tilde{u}_k\) satisfy the differential equations (3.24) and consequently, \(\tilde{u}_k\) is constant along the characteristic \(C_\ell\). Repeating the previous discussion on boundary conditions for \(\tilde{u}_k\), we conclude that \(\tilde{u}_k\) must be specified at the boundary where the corresponding characteristic \(C_\ell\) enters the domain. This means that \(\tilde{u}_k(0, t)\) and \(\tilde{u}_k(1, t)\) must be specified if \(\lambda_k(0, t) > 0\) and \(\lambda_k(1, t) < 0\), respectively. Alternatively, in terms of the partitioning (7.4), this means that \(\tilde{u}^+(0, t)\) and \(\tilde{u}^-(1, t)\) must be given.

Now, consider the boundary conditions for \(u\). Substituting the relation (7.12), the boundary condition at \(x = 0\) can be written as

\[
(C_\ell \circ \mathcal{C})(\tilde{u}^-, \tilde{u}^+)(0, t) - g_\ell(t) = 0, \tag{7.13}
\]

where \(C_\ell \circ \mathcal{C}\) is the composition of the mappings \(C_\ell\) and \(\mathcal{C}\), i.e. \((C_\ell \circ \mathcal{C})(\tilde{u}) := C_\ell(\mathcal{C}(\tilde{u}))\). In (7.13), we explicitly distinguish the variables \(\tilde{u}^-, \tilde{u}^+\). From this relation \(\tilde{u}^+(0, t)\) has to be determined, and according to the implicit function theorem, this can be done if the \(m \times p(0, t)\) matrix \(\frac{\partial}{\partial \tilde{u}_k}(C_\ell \circ \mathcal{C})(0, t)\) is invertible. Therefore, we must at least have that \(m \leq p(0, t)\), i.e. the number of boundary conditions at \(x = 0\) equals the number of positive eigenvalues at \(x = 0\). Analogously to (7.13), the boundary condition at \(x = 1\) can be written as

\[
(C_\ell \circ \mathcal{C})(\tilde{u}^-, \tilde{u}^+)(1, t) - g_\ell(t) = 0, \tag{7.14}
\]

from which \(\tilde{u}^-(1, t)\) has to be determined. This leads to the requirement that the \((m \times p(1, t))\) matrix \(\frac{\partial}{\partial \tilde{u}_k}(C_\ell \circ \mathcal{C})(1, t)\) must be invertible. This in turn implies that at least \(m = m - p(1, t)\), i.e. the number of boundary conditions at \(x = 1\) is equal to the number of negative eigenvalues at \(x = 1\). To conclude, we have the following requirements on the boundary conditions in (7.10)

\[
\begin{align*}
  x = 0: & \quad m_\ell = p, \quad \frac{\partial}{\partial \tilde{u}_k}(C_\ell \circ \mathcal{C}) \text{ invertible}, \tag{7.15a} \\
  x = 1: & \quad m_\ell = n - p, \quad \frac{\partial}{\partial \tilde{u}_k}(C_\ell \circ \mathcal{C}) \text{ invertible}. \tag{7.15b}
\end{align*}
\]

CHAPTER 12. ANALYSIS OF HYPERBOLIC EQUATIONS

Example 12.32 Recall from Section 5, that the eigenvalues \( \lambda_2(u) \) and eigenvectors \( s_2(u) \) of the shallow water equations are given by

\[
\begin{align*}
\lambda_1(u) &= u - c, & \lambda_2(u) &= u + c, & c := \sqrt{\gamma}.
\end{align*}
\]

Next, we have to scale the eigenvectors, such that the Pfaffian differential equation \( d\tilde{u} = T d\tilde{u} \) with \( T = S^{-1} \) is integrable. It turn out that we have to choose

\[
T = \frac{1}{c^2} \begin{pmatrix} -(u + c) & 1 \\ -u + c & 1 \end{pmatrix}.
\]

Inserting this matrix into (7.11) we obtain the relations

\[
d\tilde{u}_1 = \left( du - \frac{d\varphi}{\sqrt{\gamma}} \right), \quad d\tilde{u}_2 = \left( du + \frac{d\varphi}{\sqrt{\gamma}} \right).
\]

These equations can be trivially solved, and we find \( \tilde{u}_1 = u - 2c, \quad \tilde{u}_2 = u + 2c. \)

Subsequently, we can give the explicit inverse

\[
u = C(\tilde{u}) := \frac{1}{\sqrt{\gamma}} (\tilde{u}_2 - \tilde{u}_1) \begin{pmatrix} 1 \\ \frac{1}{\gamma}(\tilde{u}_1 + \tilde{u}_2) \end{pmatrix}, \quad (**)
\]

| \( u \geq c \) | \( |u| < c \) | \( u \leq -c \) |
|----------------|----------------|----------------|
| \( x = 0 \) | 2 | 1, \( \tilde{u}_2 \) | 0 |
| \( x = 1 \) | 0 | 1, \( \tilde{u}_1 \) | 2 |

Table 12.1. Number of boundary conditions for the shallow water equations.

The number of boundary conditions for this problem is summarized in table 12.1. We can distinguish three cases at both boundaries. We only discuss the boundary conditions at \( x = 0 \); the situation at \( x = 1 \) is similar. First, when \( u \geq c \), i.e. \( \lambda_2(u) \geq 0 \), we have supercritical inflow and both characteristic variables \( \tilde{u}_1 \) and \( \tilde{u}_2 \) have to be specified; in other words \( \tilde{u}_2^+ = \tilde{u}_2 \).

From (7.15) we deduce that the matrix \( \frac{\partial}{\partial\nu} (C_\ell \circ C) = \frac{\partial}{\partial\nu} C_{\ell} \frac{\partial}{\partial\nu} C \) has to be invertible. Using the mappings (**) we can see that this is equivalent with the requirement

\[
\det \left( \frac{\partial}{\partial\nu} (C_\ell \circ C) \right) = -\frac{1}{c^3} \det \left( \frac{\partial C_\ell}{\partial u} \right) \neq 0.
\]

Secondly, when \( |u| < c \), or equivalently \( \lambda_1(u) < 0 < \lambda_2(u) \), the inflow or outflow is subcritical. The characteristics \( C_2 \) enter the domain and consequently \( \tilde{u}_2 \) has to be imposed; i.e. \( \tilde{u}_2^+ = \tilde{u}_2 \). From (7.15) and (**) we can easily see that \( c_\ell(\varphi, u) \) should satisfy the condition

\[
\frac{\partial c_\ell}{\partial \varphi} + (u + c) \frac{\partial c_\ell}{\partial (\varphi u)} \neq 0.
\]

Third, when \( u \leq -c \), i.e. \( \lambda_2(u) \leq 0 \), we have supercritical outflow and no boundary conditions are required.  

8 Discussion

- Hyperbolic equations describe propagation phenomena, such as the evolution of surfaces or the propagation of waves. Examples of the latter are water waves or electromagnetic waves.

- Quite often, hyperbolic systems are the vanishing viscosity limit of conservation equations from continuum physics. The most well known example of this are the Euler equations.

- The solution of hyperbolic equations need not be smooth, which gives rise to the notion of weak solution. Weak solutions of the Riemann problem are of particular importance, since these are frequently used in numerical schemes. We will address this topic in the next two chapters.

- In Chapter ?? we investigate the mechanical etching of glass by powder erosion. In powder erosion, abrasive particles hit a glass plate at high speed causing the plate to erode. The displacement of the glass surface is modelled by a nonlinear hyperbolic equations, which we solve analytically using the characteristics.

Exercises

12.1. Determine the solution of the problem

\[ x \frac{\partial u}{\partial t} - t \frac{\partial u}{\partial x} = 0, \quad x > 0, \ t > 0, \]

\[ u(x, 0) = x^2, \quad x > 0. \]

12.2. Consider the equation

\[ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = u^2, \]

subject to the condition

\[ u(x, t) = t \quad \text{for} \quad (x, t) \in \mathcal{J} := \{(x, t) \in \mathbb{R}^2 \mid x + t = 0\}. \]

Determine the solution on the relevant domain.

12.3. Determine the solution of the Cauchy problem

\[ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 1, \quad x \in \mathbb{R}, \ t > 0, \]

\[ u(x, 0) = x, \quad x \in \mathbb{R}. \]

12.4. Find the weak solution of the Burgers' equation if the initial condition is given by

\[ u(0, x) = \begin{cases} 
1 & \text{if} \quad x < -1, \\
0 & \text{if} \quad -1 < x < 1, \\
1 & \text{if} \quad x > 1.
\end{cases} \]
12.5. Verify that the shock wave (2.13) is a weak solution of (2.5).

12.6. Verify that the rarefaction wave (2.14) is a weak solution of (2.5).

12.7. Consider the following system of equations

\[
\frac{\partial u_1}{\partial t} + \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial x} = 0,
\]

\[
\frac{\partial u_2}{\partial t} + \frac{\partial u_1}{\partial x} + 2 \frac{\partial u_2}{\partial x} + \frac{\partial u_3}{\partial x} = 0,
\]

\[
\frac{\partial u_3}{\partial t} - \frac{\partial u_3}{\partial x} + 2 \frac{\partial u_2}{\partial x} = 0.
\]

(a) Show that this system is hyperbolic.

(b) Determine the Riemann invariants.

(c) Discuss possible boundary condition on an interval \((0, L)\).

12.8. A one-dimensional model problem from linear acoustics reads

\[
\frac{\partial p}{\partial t} + K_0 \frac{\partial u}{\partial x} = 0,
\]

\[
\rho_0 \frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} = 0,
\]

where \(p\) and \(u\) are small perturbations to the pressure and velocity of the ambient fluid. \(K_0\) and \(\rho_0\) are constants. Repeat the questions from exercise 7.

12.9. A model equation for one-dimensional electromagnetic waves propagating in the \(x\)-direction reads

\[
\frac{\partial E}{\partial t} + \frac{1}{\epsilon_0 \mu_0} \frac{\partial B}{\partial x} = 0,
\]

\[
\frac{\partial B}{\partial t} + \frac{\partial E}{\partial x} = 0,
\]

where \(E\) and \(B\) are the electric field and magnetic inductance, respectively. Repeat the questions from exercise 7.

12.10. A model problem from gas dynamics is the so-called \(p\)-system [40], given by

\[
\frac{\partial v}{\partial t} - \frac{\partial u}{\partial x} = 0,
\]

\[
\frac{\partial u}{\partial t} + \frac{\partial p(v)}{\partial x} = 0,
\]

where \(v, u\) and \(p(v)\) are the specific volume, velocity and pressure of the gas.

(a) Show that this system is hyperbolic if \(p'(v) < 0\) for all \(v > 0\).

(b) Investigate the Riemann problem for this system.

(c) Discuss possible boundary condition on an interval \((0, L)\).

12.11. Show that the Euler equations for a perfect gas are hyperbolic. Compute the Riemann invariants and give the decoupled system.
12.12. Consider the wave equation with a source term, i.e.

\[
\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + 2 - 6x, \quad x \in (0, 1), \quad t > 0,
\]

subject to the following boundary and initial conditions

\[
u(0, t) = 0, \quad u(1, t) = 1, \quad t > 0,
\]

\[
u(x, 0) = x^3, \quad \frac{\partial u}{\partial t}(x, 0) = 2x, \quad x \in (0, 1).
\]

Determine the solution.

12.13. The expression (6.25) lends itself for obtaining the solution in \( \mathbb{R}^2 \), by employing the fact that \( v \) and \( w \) are depending on \( x \) and \( y \) only.

(a) Show, using spherical coordinates, that the average \( \bar{v}(r, t; \xi) \) with \( \xi = (\xi, \eta, 0) \) reduces to

\[
\bar{v}(r, t; \xi) = \frac{1}{2\pi ar} \int_{C(\xi; at)} \left\{ \frac{v(x)}{\sqrt{a^2 r^2 - \|x - \xi\|^2}} \right\} \, dx \, dy,
\]

with \( C(\xi; at) := \{x \in \mathbb{R}^2 \mid \|x - \xi\| \leq at\} \).

(b) Show then that the solution of the two-dimensional wave equation, subject to the initial conditions (6.15), is given by

\[
u(\xi, t) = \frac{\partial}{\partial t} \left( \frac{1}{2\pi a} \int_{C(\xi; at)} \left\{ \frac{v(x)}{\sqrt{a^2 r^2 - \|x - \xi\|^2}} \right\} \, dx \, dy \right) \\
+ \frac{1}{2\pi a} \int_{C(\xi; at)} \left\{ \frac{w(x)}{\sqrt{a^2 r^2 - \|x - \xi\|^2}} \right\} \, dx \, dy.
\]  
(8.16)

12.14. Consider the damped wave equation

\[
\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}.
\]

(a) Write this equation as a linear first order system.

(b) Show that this system is hyperbolic.

(c) Give the decoupled system.
Chapter 15

Perturbation methods

In this chapter we give a concise introduction to a number of important analytical techniques that use available small parameters in a problem. These methods are called perturbation or asymptotic methods, and follow naturally after the ideas of systematic modelling, set forth in Chapter 7. After an introduction, explaining why small parameters occur frequently in practice, we start in Section 2 with stating some principles of asymptotic approximations and expansions. Then we continue in Section 3 with two methods for solving regular perturbation problems, viz. the method of slow variation and the Lindstedt-Poincaré method. Two of the most important methods for singular perturbation problems are considered in Section 4, viz. the method of matched asymptotic expansions and some versions of the method of multiple scales.

1 Introduction

We have seen in Chapter 7 that a real-world problem can be described by a hierarchy of models, such that a higher level model is more comprehensive and more accurate than one from a lower level. Now suppose that we have a fairly good model, describing the dominating phenomena in good order of magnitude. And suppose that we are interested in improving on this model by adding some previously ignored aspects or effects. In general, this implies a very abrupt change in our model. The equations are more complex and more difficult to solve. As an illustration, consider the simple “model" \( x^2 = a^2 \), and the more complete “model" \( x^2 + \varepsilon x^5 = a^2 \). The first one can be solved easily analytically, the second one with much more effort only numerically. So it seems that the relation between solution and model is not continuous in the problem parameters. Whatever small \( \varepsilon \) we take, from a transparent and exact solution of the simple model at \( \varepsilon = 0 \), we abruptly face a far more complicated solution of a model that is just a little bit better. This is a pity, because certain type of useful information (parametric dependencies, trends) become increasingly more difficult to dig out of the more complicated solution of the complex model. This discontinuity of models in the parameter \( \varepsilon \) may therefore be an argument to retain the simpler model.
1. INTRODUCTION

The (complexity of the) model is, however, only discontinuous if we are merely interested in exact or numerically “exact” solutions (for example for reasons of benchmarking or validation of solution methods). This is not always the case. As far as our modelling objectives are concerned, we have to keep in mind that also the improved model is only a next step in the modelling hierarchy and not exact in any absolute sense. So there is no reason to require the solution to be more exact than the corresponding model, as an exact solution of an approximate model is not better than an approximate solution of an exact model. Moreover, the type of information that analytical solutions may provide (functional relationships, etc.) is sometimes so important that numerical accuracy may be worthwhile to sacrifice.

Let us go back to our “fairly good”, improved model. The effects we added are relatively small. Otherwise, the previous lower level model was not fairly good as we assumed, but just completely wrong. Usually, this smallness is quantified by small dimensionless parameters occurring in the equations and (or) boundary conditions. This is the generic situation. The transition from a lower-level to a higher-level theory is characterized by the appearance of one or more modelling parameters, which are (when made dimensionless) small or large, and yield in the limit a simpler description. Examples are infinitely large or small geometries with circular or spherical symmetry that reduce the number of spatial dimensions, small amplitudes allowing linearization, low velocities and long time scales in flow problems allowing incompressible description, small relative viscosity allowing inviscid models, etc. In fact, in any practical problem it is really the rule rather than the exception that dimensionless numbers are either small or large (cf. [66]).

If we accept approximate solutions, where the approximation is based on the inherently small or large modelling parameters, we do have the possibility to gradually increase the complexity of a model, and study small but significant effects in the most efficient way. The methods utilizing this approach systematically are called “perturbation methods”. The approximation constructed is almost always an asymptotic approximation, i.e. where the error reduces with the small or large parameter.

Usually, a distinction is made between regular and singular perturbations. A (loose definition of a) regular perturbation problem is where the approximate problem is everywhere close to the unperturbed problem. This, however, depends of course on the domain of interest and, as we will see, on the choice of coordinates. If a problem is regular without any need for other than trivial reformulations, the construction of an asymptotic solution is straightforward. In fact, it forms the usual strategy in modelling when terms are linearised or effects are neglected. The more interesting perturbation problems are those where this straightforward approach fails.

We will consider here four methods relevant in the presented modelling problems. The first two are examples of regular perturbation methods, but only after a suitable coordinate transformation. The first is called the method of slow variation, where the typical axial length scale is much greater than the transverse length scale. The second is the Lindstedt-Poincaré method or the method of strained coordinates, for periodic processes. Here, the intrinsic time scale (~ the period of the solution) is unknown and has to be found. The other two methods are of singular perturbation type, because there is no coordinate transformation possible that renders the problem into one of regular type. The third
one is the method of matched asymptotic expansions (MAE). To render the problem into one of regular type, different scalings are necessary in spatially distinct regions (boundary layers). The fourth singular perturbation method considered here is the method of multiple scales and may be considered as a combination of the method of slow variation and the method of strained coordinates, as now several (long, short, shorter) length scales occur in parallel. This cannot be repaired by a single coordinate transformation. Therefore, the problem is temporarily reformulated into a higher dimensional problem by taking the various length scales apart. Then the problem is regular again, and can be solved. A refinement of this method is the WKB method, where the coordinate transformation of the fast variable becomes itself slowly varying.

2 Asymptotic approximations and expansions

Before we can introduce the methods, we have to define our terminology of asymptotic approximations and asymptotic expansions.

2.1 Asymptotic approximations

In order to give a qualitative description of the behaviour of a function $f$ with parameter $\varepsilon$ near a point of interest, say $\varepsilon = 0$ (equivalent to any other value by a simple translation), we have the so-called order symbols $O$, $o$, and $O_\varepsilon$; see Appendix A. Often $\varepsilon = 0$ is the lower limit of a parameter range, and we have the tacit assumption that $\varepsilon \downarrow 0$.

**Definition 15.1.** $\varphi(\varepsilon)$ is an asymptotic approximation to $f(\varepsilon)$ as $\varepsilon \to 0$ if

$$ f(\varepsilon) = \varphi(\varepsilon) + o(\varphi(\varepsilon)) \quad \text{as} \quad \varepsilon \to 0, $$

sometimes more compactly denoted by $f \sim \varphi$.

If $f$ and $\varphi$ depend on $x$, this definition remains valid pointwise, i.e. for $x$ fixed. It is, however, useful to extend the definition to uniformly valid approximations.

**Definition 15.2.** Let $f(x; \varepsilon)$ and $\varphi(x; \varepsilon)$ be continuous functions for $x \in \mathcal{D}$ and $0 < \varepsilon < a$. We call $\varphi(x; \varepsilon)$ a uniform asymptotic approximation to $f(x; \varepsilon)$ for $x \in \mathcal{D}$ as $\varepsilon \to 0$, if for any positive number $\delta$ there is an $\varepsilon_1$ (independent of $x$ and $\varepsilon$) such that

$$ |f(x; \varepsilon) - \varphi(x; \varepsilon)| \leq \delta|\varphi(x; \varepsilon)| \text{ for } x \in \mathcal{D} \text{ and } 0 < \varepsilon < \varepsilon_1. $$

We write: $f(x; \varepsilon) = \varphi(x; \varepsilon) + o(\varphi(x; \varepsilon))$ uniformly in $x \in \mathcal{D}$ as $\varepsilon \to 0$. Note that $\mathcal{D}$ may depend on $\varepsilon$.

**Example 15.3** Let $\mathcal{D} = [0, 1]$ and $0 < \varepsilon < 1$. Then we have $\cos(\varepsilon x) = 1 + o(1)$ as $\varepsilon \to 0$ uniformly in $\mathcal{D}$, since for any given $\delta$ we can choose $\varepsilon_1 = \sqrt{\delta}$, such that $|\cos(\varepsilon x) - 1| \leq \varepsilon^3 x^2 \leq \varepsilon^2 \delta = \delta$. $\square$
2. ASYMPTOTIC APPROXIMATIONS AND EXPANSIONS

Example 15.4 Although \( \cos(x/\varepsilon) = O(1) \) uniformly in \( x \in [0, 1] \) for \( \varepsilon \to 0 \), there is no constant \( K \) such that \( \cos(x/\varepsilon) = K + o(1) \). □

Example 15.5 \( x + \sin(\varepsilon x) + e^{-x/\varepsilon} = x + \varepsilon x + O(\varepsilon^3) \) as \( \varepsilon \to 0 \) for all \( x \neq 0 \), but not uniformly in \( x \in [0, 1] \). More precisely, it is not uniformly in \( x \in [\delta(\varepsilon), 1] \) for any \( \delta = O(\varepsilon) \) and uniformly if \( \varepsilon = o(\delta) \). If \( x = O(\varepsilon) \), the otherwise exponentially small term is not small anymore. This is illustrated by the figure 15.1. The difference between the original function and its non-uniform asymptotic approximation is typically large in a neighbourhood of \( x = 0 \), while the size of this neighbourhood is \( x = O(\varepsilon) \). This neighbourhood is an example of a boundary layer. The occurrence and behaviour of boundary layers will be discussed in more detail in Section 4.1. □

2.2 Asymptotic expansions

Asymptotic approximations are usually structured in the form of a series expansion that helps us to construct an approximation systematically.

Definition 15.6. The sequence \( \{\mu_n(\varepsilon)\}_{n=0}^\infty \) is called an asymptotic sequence, if \( \mu_{n+1}(\varepsilon) = o(\mu_n(\varepsilon)) \), as \( \varepsilon \to 0 \), for each \( n = 0, 1, 2, \ldots \).

Example 15.7 Examples of asymptotic sequences (as \( \varepsilon \to 0 \)) are
\[
\mu_n(\varepsilon) = \varepsilon^n, \quad \mu_n(\varepsilon) = \varepsilon^{2n}, \quad \mu_n(\varepsilon) = \tan(\varepsilon), \quad \mu_n(\varepsilon) = \ln(\varepsilon)^{-n}, \quad \mu_n(\varepsilon) = \varepsilon^p \ln(\varepsilon)^q \quad \text{where} \quad p = 0, 1, 2, \ldots, \quad q = 0, \ldots, p \quad \text{and} \quad n = \frac{1}{2} p(p + 3) - q. \]

Definition 15.8. If \( \{\mu_n(\varepsilon)\}_{n=0}^\infty \) is an asymptotic sequence, then \( f(\varepsilon) \) has an asymptotic expansion of \( N \) terms with respect to this sequence, denoted by
\[
f(\varepsilon) \sim \sum_{n=0}^{N-1} a_n \mu_n(\varepsilon),
\]
where the coefficients $a_n$ are independent of $\varepsilon$, if

$$f(\varepsilon) - \sum_{n=0}^{M} a_n \mu_n(\varepsilon) = o(\mu_M(\varepsilon)) \text{ as } \varepsilon \to 0$$

for each $M = 0, \ldots, N - 1$. $\mu_n(\varepsilon)$ is called a gauge-function. If $\mu_n(\varepsilon) = \varepsilon^n$, we call the expansion an asymptotic power series.

**Definition 15.9.** Two functions $f$ and $g$ are asymptotically equal up to $N$ terms, with respect to the asymptotic sequence $\{\mu_n\}$, if $f - g = o(\mu_N)$ as $\varepsilon \to 0$ for each $M = 0, \ldots, N - 1$. $\mu_n(\varepsilon)$ is called a gauge-function. If $\mu_n(\varepsilon) = \varepsilon^n$, we call the expansion an asymptotic power series.

Asymptotic expansions based on the same gauge functions may be added. They may be multiplied if the products of the gauge functions can be asymptotically ordered.

In contrast to ordinary series expansions, defined for an infinite number of terms, in asymptotic expansions only a finite ($N$) number of terms are considered. For $N \to \infty$ the series may either converge or diverge, but this is irrelevant for the asymptotic behaviour. In addition it may be worthwhile to note that it is not necessary for a convergent asymptotic expansion to converge to the expanded function.

For given $\{\mu_n(\varepsilon)\}_{n=0}^{\infty}$, the coefficients $a_n$ can be determined uniquely by the following recursive procedure (provided $\mu_n$ are nonzero for $\varepsilon$ near 0 and each of the limits below exist)

$$a_0 = \lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{\mu_0(\varepsilon)}, \quad a_1 = \lim_{\varepsilon \to 0} \frac{f(\varepsilon) - a_0 \mu_0(\varepsilon)}{\mu_1(\varepsilon)}, \ldots, \quad a_{N-1} = \lim_{\varepsilon \to 0} \frac{f(\varepsilon) - \sum_{n=0}^{N-2} a_n \mu_n(\varepsilon)}{\mu_{N-1}(\varepsilon)}.$$

**Example 15.10** A function may have different asymptotic expansions.

$$\tan(\varepsilon) = \varepsilon + \frac{1}{3!} \varepsilon^3 + \frac{1}{5!} \varepsilon^5 + O(\varepsilon^7)$$

$$= \sin \varepsilon + \frac{1}{2} \cos \varepsilon + \frac{1}{6} \sin^3 \varepsilon + O((\sin \varepsilon)^3)$$

$$= \varepsilon \cos \varepsilon + \frac{1}{2} (\varepsilon \cos \varepsilon)^3 + \frac{1}{60} (\varepsilon \cos \varepsilon)^5 + O((\varepsilon \cos \varepsilon)^7).$$

**Example 15.11** The asymptotic expansion $\sum_{n=1}^{N} n! \varepsilon^n$ diverges as $N \to \infty$ if $\varepsilon \neq 0$.

**Example 15.12** Different functions may have the same asymptotic expansion.

$$\cos(\varepsilon) = 1 - \frac{1}{2} \varepsilon^2 + \frac{1}{24} \varepsilon^4 + O(\varepsilon^6),$$

$$\cos(\varepsilon) + e^{-1/\varepsilon} = 1 - \frac{1}{2} \varepsilon^2 + \frac{1}{24} \varepsilon^4 + O(\varepsilon^6).$$

Note that both asymptotic expansions, considered as regular power series in $\varepsilon$, converge to $\cos(\varepsilon)$ rather than $\cos(\varepsilon) + e^{-1/\varepsilon}$.
Theorem 15.13. An asymptotic expansion vanishes only if the coefficients vanish, i.e.
\[ \{ a_0 \mu_0(\epsilon) + a_1 \mu_1(\epsilon) + a_2 \mu_2(\epsilon) + \ldots = 0 \ (\epsilon \to 0) \} \Leftrightarrow \{ a_0 = a_1 = a_2 = \ldots = 0 \}. \]

**Proof.** The sequence \( \{ \mu_n \} \) is asymptotically ordered, so both \( \mu_0 a_0 = -\mu_1 a_1 - \ldots = O(\mu_1) \) and \( \mu_1 = o(\mu_0) \). So there is a positive constant \( K \) such that for any positive \( \delta \) there is an \( \epsilon \)-interval where \( |a_0 \mu_0| < \delta K |\mu_0| \), which is only possible if \( a_0 = 0 \). This may now be repeated for \( a_1 \), etc. This proves \( \Rightarrow \). The proof of \( \Leftarrow \) is trivial. \( \square \)

2.3 Perturbation problems

The assumed existence of an asymptotic expansion yields a class of methods to solve otherwise intractable problems depending on a typically small parameter. Such methods are called *perturbation methods*.

If \( a(\epsilon) \) is implicitly given as the solution of an algebraic equation

\[ F(a; \epsilon) = 0 \quad (2.1) \]

and both \( a(\epsilon) \) and \( F(a; \epsilon) \) have an asymptotic series expansion with the same gauge functions, \( a(\epsilon) \) may be determined asymptotically by the following perturbation method. We expand \( a \), substitute this expansion in \( F \), and expand \( F \) to obtain

\[ a(\epsilon) = a_0 \mu_0(\epsilon) + a_1 \mu_1(\epsilon) + \ldots, \quad (2.2a) \]

\[ F(a; \epsilon) = F_0(a_0) \mu_0(\epsilon) + F_1(a_1, a_0) \mu_1(\epsilon) + F_2(a_2, a_1, a_0) \mu_2(\epsilon) + \ldots = 0. \quad (2.2b) \]

From theorem 15.13 it follows that that each term \( F_n \) vanishes, and the sequence of coefficients \( (a_n) \) can be determined by induction:

\[ F_0(a_0) = 0, \quad F_1(a_1, a_0) = 0, \quad F_2(a_2, a_1, a_0) = 0, \quad \text{etc.} \quad (2.3) \]

It should be noted that finding the sequence of gauge functions \( (\mu_n) \) is of particular importance. This is done iteratively. First the order of magnitude of \( a \) should be determined by seeking the asymptotic scaling \( a(\epsilon) = \gamma(\epsilon) A(\epsilon) \) which yields a meaningful \( A = O_1(1) \) in the limit \( \epsilon \to 0 \). This is called a *distinguished limit*, while the reduced equation for \( A(0) \), i.e. \( F_0(A) = 0 \), is called a *significant degeneration* (there may be more than one.) The first gauge function that occurs is now \( \mu_0(\epsilon) = \gamma(\epsilon) \), while \( a_0 = A(0) \). The procedure may be repeated for the new unknown \( a(\epsilon) - \mu_0(\epsilon) a_0 \), and so on. It is not unusual that the rest of the sequence \( (\mu_n) \) can be guessed from the structure of the defining equation \( F = 0 \).

We illustrate this procedure by the following example.

**Example 15.14** Consider the roots for \( \epsilon \to 0 \) of the equation

\[ x^3 - \epsilon x^2 + 2\epsilon^3 x + 2\epsilon^6 = 0. \]
From the structure of the equation it seems reasonable to assume that the solutions $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ have an asymptotic expansion in powers of $\varepsilon$. However, the order of magnitude of the leading order term is not immediately clear.

$$x(\varepsilon) = \varepsilon^n (X_0 + \varepsilon X_1 + \varepsilon^2 X_2 + \mathcal{O}(\varepsilon^3))$$

Therefore, we have to determine $n$ first. This is done by balancing terms. We scale

$$x = \varepsilon^n X(\varepsilon), \quad X = \mathcal{O}(1)$$

and we seek such $n$ that produce a non-trivial limit under the limit $\varepsilon \to 0$. We compare asymptotically the coefficients in the equation that remain after scaling

$$\varepsilon^{3n} X^3 - \varepsilon^{1+2n} X^2 + 2\varepsilon^{3+n} X + 2\varepsilon^6 = 0.$$ 

In order to have a meaningful (or “significant”) degenerate solution $X(0) = \mathcal{O}(1)$ at least two terms of the equation should be asymptotically equivalent, and at the same time of leading order when $\varepsilon \to 0$. So this leaves us with the task to compare the powers $3n$, $1+2n$, $3+n$, $6$ as a function of $n$. Consider the figure 15.2. The solid lines denote the logarithm of the powers of $\varepsilon$, that occur in the coefficients of the equation considered. At the intersections of these lines, denoted by the open and closed circles, we find the candidates of distinguished limits, i.e. the points where at least two coefficients are asymptotically equivalent. Finally, only the closed circles are the distinguished limits, because these are located along the lower envelope (thick solid line) and therefore correspond to leading order terms when $\varepsilon \to 0$. We have now three cases.

$n = 1$.

$$\varepsilon^3 X^3 - \varepsilon^3 X^2 + 2\varepsilon^4 X + 2\varepsilon^6 = 0, \quad \text{or} \quad X^3 - X^2 + 2\varepsilon X + 2\varepsilon^3 = 0.$$
If we assume the expansion \( X = X_0 + \varepsilon X_1 + \ldots \), we finally have
\[
X_0^3 - X_0^2 = 0, \quad 3X_0^2X_1 - 2X_0X_1 + 2X_0 = 0, \quad \text{etc.}
\]
and so \( X_0 = 1 \), and \( X_1 = -2 \), etc. leading to \( x(\varepsilon) = \varepsilon - 2\varepsilon^2 + \ldots \). Note that solution \( X_0 = 0 \) is excluded because that would change the order of the scaling!

\[ n = 2. \]
\[
\varepsilon^6 X_0^3 - \varepsilon^5 X_0^2 + 2\varepsilon^6 X_0 X_1 - 2\varepsilon X_0 X_1 + 2\varepsilon^6 X_0 = 0, \quad \text{or} \quad \varepsilon X_0^3 - X_0^2 + 2X_0 + 2\varepsilon = 0.
\]
If we assume the expansion \( X = X_0 + \varepsilon X_1 + \ldots \), we finally have
\[
-X_0^2 + 2X_0 = 0, \quad \text{etc.}
\]
and so \( X_0 = 2 \), etc. leading to \( x(\varepsilon) = 2\varepsilon^2 + \ldots \)

\[ n = 3. \]
\[
\varepsilon^9 X_0^3 - \varepsilon^7 X_0^2 + 2\varepsilon^6 X_0 X_1 - 2\varepsilon X_0 X_1 + 2\varepsilon^6 X_0 = 0, \quad \text{or} \quad \varepsilon^3 X_0^3 - \varepsilon X_0^2 + 2X_0 + 2\varepsilon = 0.
\]
If we assume the expansion \( X = X_0 + \varepsilon X_1 + \ldots \), we finally have
\[
2X_0 + 2 = 0, \quad \text{etc.}
\]
and so \( X_0 = -1 \), etc. leading to \( x(\varepsilon) = -\varepsilon^3 + \ldots \)

It is not always so easy to guess the general form of the gauge functions. Then all terms have to be estimated iteratively by a similar process of balancing as for the leading order term. See exercise 2.

### 2.4 Asymptotic expansions of Poincaré type

An asymptotic expansion of a function of variable \( x \) and small parameter \( \varepsilon \) of the following form
\[
f(x; \varepsilon) = \sum_{n=0}^{N-1} a_n(x; \varepsilon) \mu_n(\varepsilon) + O(\mu_N), \quad a_n = O(1),
\]
appears to be too general to be of practical use. The restriction that the coefficients \( a_n \) depend on \( x \) only appears to be fruitful. This is called a Poincaré expansion, or more precisely

**Definition 15.15.** If \( \{\mu_n(\varepsilon)\}_{n=0}^{\infty} \) is an asymptotic sequence, and \( f(x; \varepsilon) \) has an asymptotic expansion of \( N \) terms with respect to this sequence, given by
\[
f(x; \varepsilon) \sim \sum_{n=0}^{N-1} a_n(x) \mu_n(\varepsilon),
\]
where the shape functions \( a_n(x) \) are independent of \( \varepsilon \), this expansion is called a Poincaré expansion.

**Definition 15.16.** If a Poincaré expansion is uniform in \( x \) on a given domain \( D \) (Definition 15.2) this expansion is called a regular expansion. Else, the expansion is called a singular expansion.
CHAPTER 15. PERTURBATION METHODS

(Note, that there is no uniformity in the literature on the definition of regular and singular expansions.) Regular expansions may be differentiated to the independent variable x.

Example 15.17 The following asymptotic power series expansion
\[
\cos(x + \epsilon) = \cos(x) - \epsilon \sin(x) - \frac{1}{2} \epsilon^2 \cos(x) + \frac{1}{6} \epsilon^3 \sin(x) + \mathcal{O}(\epsilon^4),
\]
with \( D = \mathbb{R} \) and the gauge functions \( \mu_n(\epsilon) = \epsilon^n, \) is uniform since \( \cos(x) \) and \( \sin(x) \) are bounded for all \( x \in \mathbb{R} \). It follows that it is a regular expansion.

Example 15.18 The following expansion
\[
\cos(x + \epsilon) e^{-x/\epsilon} + \sin(x + \epsilon) = \sin x + \epsilon \cos x + \mathcal{O}(\epsilon^2)
\]
is a uniform, and therefore regular, expansion on any interval \([A, \infty)\), where \( A > 0 \). However, it is a non-uniform, and therefore singular, expansion on \([0, \infty)\). In fact, on any interval \([A\epsilon^\alpha, \infty)\) it is regular if \( \alpha < 1 \), and singular if \( \alpha \geq 1 \).

It is vital for the understanding of the perturbations methods to be considered below, to appreciate the central role of the choice of the independent variable \( x \) in a Poincaré expansion. By suitable linear coordinate transformations of the type \( x = \lambda(\epsilon) + \delta(\epsilon) \xi \) we can change and optimize the domain of uniformity, and filter out asymptotically specific behaviour that belongs to one length scale. This filter property is especially useful when asymptotically analysing models.

Example 15.19

1. \( \sin(x + \epsilon x + \epsilon^2 x) = \sin(x) + \epsilon x \cos(x) + \epsilon^2 (\cos x - \frac{1}{2} x^2 \sin x) + \mathcal{O}(\epsilon^3) \), which is only uniform on an interval \([0, A]\), but if we introduce \( \xi = (1 + \epsilon) x \), we have \( \sin(\xi + \epsilon^2 x) = \sin(\xi) + \mathcal{O}(\epsilon) \) uniform in \( x \in [0, Ae^{-1}] \) for any positive constant \( A \).

2. \( \sin(\epsilon x + \epsilon) = \epsilon x + \epsilon + \mathcal{O}(\epsilon^3) \), which is only uniform on a finite interval and, moreover, does not show any of the inherent periodicity. If we introduce \( X = \epsilon x \), we get the much better \( \sin(\epsilon x + \epsilon) = \sin(\epsilon x) + \mathcal{O}(\epsilon) \) which is even uniform in \( \mathbb{R} \).

3. \( e^{-x/\epsilon} = 0 + \mathcal{O}(\epsilon^n) \) is a singular expansion on \( x > 0 \), but if we introduce \( \xi = x/\epsilon \) it becomes the regular expansion \( e^{-\xi} = \mathcal{O}(1) \) on \( \xi > 0 \).

4. On \( x > 0 \) we have \( \frac{\pi}{2} \arctan(\epsilon) + \sin(x)/(1 + x^2) = 1 + \epsilon x/(1 + x^2) - 2/\pi x + \mathcal{O}(\epsilon^3) \) in \( x, = 1 + \epsilon x/\sin(x)/X^2 - 2/\pi X + \mathcal{O}(\epsilon^3) \) in \( X = \epsilon x \), and \( = \frac{\pi}{2} \arctan(x) + \epsilon x + \mathcal{O}(\epsilon^3) \) in \( \xi = x/\epsilon \).

5. \( e^{-x} \sin(x/\sqrt{\epsilon + 1}) = \sin x + \epsilon x (\frac{\pi}{2} \cos x - \sin x) + \mathcal{O}(\epsilon^2) \), which is only uniform on an interval \([0, A]\). This cannot be improved by a single other choice of independent variable. However, if we introduce two variables, \( x_1 = \epsilon x \) and \( x_2 = (1 + 1/\epsilon) x \), we get the much better \( e^{-x_1} \sin x_2 - \frac{1}{2} \epsilon^2 x_2^2 e^{-x_1} \cos x_2 + \mathcal{O}(\epsilon^4) \).

3 Regular perturbation problems

If a function \( \Phi(x; \epsilon) \) is implicitly given by an equation (usually a differential equation with boundary conditions), say
\[
\mathcal{L}[\Phi](x; \epsilon) = 0 \quad \text{on a domain } D
\]

3. REGULAR PERTURBATION PROBLEMS

and both $\Phi$ and $\mathcal{L}[\Phi](x; \varepsilon)$ have a regular asymptotic expansion on $\mathcal{D}$ with the same gauge functions, (3.1) is called a regular perturbation problem [57]. The shape functions $\Phi_n$ are determined as follows. We expand $\mathcal{L}[\Phi]$

$$
\mathcal{L}[\Phi](x; \varepsilon) = \mu_0(\varepsilon) L_0[\Phi_0](x) + \mu_1(\varepsilon) L_1[\Phi_1, \Phi_0](x) + \mu_2(\varepsilon) L_2[\Phi_2, \Phi_1, \Phi_0](x) + \ldots = 0. \quad (3.2)
$$

According to theorem 15.13 each term vanishes, and the sequence $(\Phi_n)$ can be determined by induction:

$$
L_0[\Phi_0](x) = 0, \quad L_1[\Phi_1, \Phi_0](x) = 0, \quad L_2[\Phi_2, \Phi_1, \Phi_0](x) = 0, \ldots \quad (3.3)
$$

It should be noted that in many interesting cases the problem is only regular after a suitable coordinate transformation. The major task when solving the problem is then to find this scaled or shifted coordinate. Practically important solution methods of this type are the method of slow variation, for geometrically stretched or slowly varying configurations, and the Lindstedt-Poincaré method, for solutions which are periodic in time with an unknown, $\varepsilon$-dependent period.

If (3.1) is not a regular perturbation problem, we call it a singular perturbation problem. Practically important solution methods for singular perturbation problems are the method of matched asymptotic expansions, where regular expansions exist locally but not in the whole region considered, and the method of multiple scales, where 2 or more distinct long and short length scales occur intertwined.

3.1 Method of slow variation

Suppose we have a function $\varphi(x; \varepsilon)$ of spatial coordinates $x$ and a small parameter $\varepsilon$, such that the typical variation in one direction, say $x$, is of the order of length scale $\varepsilon^{-1}$. We can express this behaviour most conveniently by writing $\varphi(x, y, z; \varepsilon) = \Phi(\varepsilon x, y, z; \varepsilon)$. Now if we were to expand $\Phi$ for small $\varepsilon$, we might, for example, get something like

$$
\Phi(\varepsilon x, y, z; \varepsilon) = \Phi(0, y, z; 0) + \varepsilon(\varphi_1(0, y, z; 0) + \varphi_2(0, y, z; 0)) + \ldots,
$$

which is only uniform in $x$ on an interval $[0, L]$ if $L = O(1)$, and the inherent slow variation on the longer scale of $x = O(\varepsilon^{-1})$ would be masked. It is clearly much better to introduce the scaled variable $X = \varepsilon x$, and a (assumed) regular expansion of $\Phi(X, y, z; \varepsilon)$

$$
\Phi(X, y, z; \varepsilon) = \mu_0(\varepsilon) \varphi_0(X, y, z) + \ldots \quad (3.4)
$$

now retains the slow variation in $X$ in the shape functions of the expansion.

This situation frequently happens when the geometry involved is slender [120]. The theory of one dimensional gas dynamics, lubrication flow, or sound propagation in horns (Webster’s equation) are important examples, although they are usually derived not systematically, without explicit reference to the slender geometry. We will illustrate the method both for heat flow in a varying bar, quasi 1-D gas flow and the shallow water problem.
Example 15.20 (Heat flow in a bar.) Consider the stationary problem of the temperature distribution $T$ in a long heat-conducting bar with outward surface normal $n$ and slowly varying cross section $A$. The bar is kept at a temperature difference such that a given heat flux is maintained, but is otherwise isolated. As there is no leakage of heat, the flux is constant. With spatial coordinates made dimensionless on a typical bar cross section, we have the following equations and boundary conditions

$$\nabla^2 T = 0, \quad \nabla T \cdot n = 0, \quad \int_A \frac{\partial T}{\partial x} dS = Q.$$

After integrating $\nabla^2 T$ over a slice $x_1 \leq x \leq x_2$, and applying Gauss' theorem, we find that the axial flux $Q$ is indeed independent of $x$. The typical length scale of diameter variation is assumed to be much larger than a diameter. We introduce the ratio between a typical diameter and this length scale as the small parameter $\varepsilon$, and write for the bar surface $S(X, r, \theta)$:

$$S(X, r, \theta) := r - R(X, \theta) = 0, \quad X = \varepsilon x,$$

where $(x, r, \theta)$ form a cylindrical coordinate system (see figure 15.3). By writing $R$ as a continuous function of slow variable $X$, rather than $x$, we have made our formal assumption of slow variation explicit in a convenient and simple way, since $R_x = \varepsilon R_X = O(\varepsilon)$.

The crucial step will now be the assumption that the temperature is only affected by the geometric variation induced by $R$. Any initial or entrance effects are ignored or have disappeared. As a result the temperature field $T$ is a function of $X$, rather than $x$, and its axial gradient scales on $\varepsilon$, as $T_x = O(\varepsilon)$.

Introduce the gradient $\nabla S$ and the transverse gradient $\nabla_\perp S$

$$\nabla S = -\varepsilon R_x e_x + e_r - r^{-1} R_\theta e_\theta, \quad \nabla_\perp S := S_r e_r + r^{-1} S_\theta e_\theta = e_r - r^{-1} R_\theta e_\theta.$$

At the bar surface $S = 0$ the gradient $\nabla S$ is a vector normal to the surface, while the transverse gradient $\nabla_\perp S$, directed in the plane of a cross section $X = \text{const.}$, is normal to the circumference $S(X = c, r, \theta) = 0$. Inside the bar we have the rescaled heat equation

$$\varepsilon^2 T_{XX} + \nabla_\perp^2 T = 0. \quad (\ast)$$

At the wall the boundary condition of vanishing heat flux is

$$\nabla T \cdot \nabla S = \varepsilon^2 T_X S_X + \nabla_\perp T \cdot \nabla_\perp S = 0 \quad \text{at} \quad S = 0. \quad (\dagger)$$

Figure 15.3. Slowly varying bar
The flux condition, for lucidity rewritten with $Q = \varepsilon q$, is given by
\[ \iint_{A} \frac{\partial T}{\partial X} \, d\sigma = q. \]
This problem is too difficult in general, so we try to utilize the small parameter $\varepsilon$ in a systematic manner. Since the perturbation terms are $O(\varepsilon^2)$, we assume the asymptotic expansion
\[ T(X, r, \theta; \varepsilon) = T_0(X, r, \theta) + \varepsilon^2 T_1(X, r, \theta) + O(\varepsilon^3). \]

After substitution in equation (\ref{eq:flux}) and boundary condition (\ref{eq:bc}), further expansion in powers of $\varepsilon^2$ and equating like powers of $\varepsilon$, we obtain to leading order a Laplace equation in $(r, \theta)$
\[ \nabla \cdot T_0 = 0 \quad \text{with} \quad \nabla \cdot \nabla S = 0 \quad \text{at} \quad S = 0. \]

An obvious solution is $T_0(X, r, \theta) \equiv 0$. Since solutions of Laplace’s equation with vanishing normal derivatives at the boundary are unique up to a constant (here: a function of $X$), we have
\[ T_0 = T_0(X). \]

We could substitute this directly in the flux condition, to find that $AT_0 = q$, where $A(X)$ is the area of cross section $A(X)$. For the present exposition, however, it is of interest to show that this result also emerges from the equations as follows. To obtain an equation for $T_0$ in $X$ we continue with the $O(\varepsilon^0)$-equation and corresponding boundary condition
\[ \nabla^2 T_0 + T_{0,xx} = 0, \quad \nabla \cdot T_0 \cdot \nabla S = -T_0 S_X. \quad (\dagger) \]

The boundary condition can be rewritten as
\[ \nabla \cdot T_0 \cdot \mathbf{n}_\perp = \frac{T_{0,x} R_X}{|\nabla S|} = \frac{T_{0,x} RR_X}{\sqrt{R^2 + R_0^2}} \]

where $\mathbf{n}_\perp = \nabla S/|\nabla S|$ is the transverse unit normal vector. By integrating equation (\ref{eq:laplace}) over a cross section $A$ of area $A(X)$, using Gauss' theorem, and noting that $A = \int_0^{2\pi} \frac{1}{2} R^2 \, d\theta$, and that a circumferential line element is given by $d\ell = (R^2 + R_0^2)^{1/2} \, d\theta$, we obtain
\[ \iint_{A} \left( \nabla^2 T_0 + T_{0,xx} \right) \, dS = \int_{A} \nabla \cdot T_0 \cdot \mathbf{n}_\perp \, d\ell + AT_{0,xx} \\
= T_{0,x} \int_0^{2\pi} RR_X \, d\theta + AT_{0,xx} = A_x T_{0,x} + AT_{0,xx} = \frac{d}{dX} \left( A \frac{d}{dX} T_0 \right) = 0. \]

The finally obtained equation can be solved easily. Note that we recovered the conservation law of heat flux $AT_{0,xx} = q$. Finally we have
\[ T_0(X) = \int_X^Z \frac{q}{A(z)} \, dz + T_{\text{ref}}. \]

It should be noted that we did not include in our analysis any boundary conditions at the ends of the bar. It is true that the present method fails here. The found solution is uniformly valid on $\mathbb{R}$ (since $R(X)$ is assumed continuous and independent of $\varepsilon$), but only as long as we stay away from any end. Near the ends the boundary conditions induce transverse gradients of $O(1)$ which makes the prevailing length scale again $x$, rather than $X$. This region is asymptotically of boundary layer type, and should be treated differently (see below). \qed
CHAPTER 15. PERTURBATION METHODS

Example 15.21 (Quasi 1-D gas dynamics). Consider a slowly varying duct with irrotational inviscid isentropic flow, described (in dimensionless form) by the velocity potential $\varphi$ and density $\rho$ satisfying the equation for mass conservation and the compressible form of Bernoulli's equation (see Chapter 7), i.e.

$$\nabla \cdot (\rho \nabla \varphi) = 0, \quad \frac{1}{2} |\nabla \varphi|^2 + \frac{\rho^{\gamma-1}}{\gamma - 1} = E.$$

The parameter $\gamma$ is a gas constant (1.4 for air) and $E$ is a constant of the problem. Using the same notation as in the previous example, the duct wall is given by $S(x, z, \theta) = 0$, while at the impermeable wall $\nabla \varphi \cdot \nabla S = 0$. The mass flux, the same at any cross section $A$, is given by

$$\int_A \rho \varphi \, dS = F.$$

Introduce the slow variable $X = \varepsilon x$, and assume $\varphi$ and $\rho$ to depend essentially on $X$, rather than $x$. The dimensionless axial flow velocity $\varphi_0$, the density $\rho$, the cross sectional area $A$, the flux $F$ and the thermodynamical constant $E$ are $\mathcal{O}(1)$. So we have to rescale $\varphi$ and write

$$\varphi(x, y, z; \varepsilon) = \varepsilon^{-1} \Phi(X, y, z; \varepsilon).$$

The equations for $\Phi$ and $\rho$ become

$$\varepsilon^2 \left( \frac{1}{\varepsilon^2} \rho \frac{\partial^2 \Phi}{\partial x^2} \right) + \nabla \cdot (\rho \nabla \Phi) = 0, \quad \frac{1}{2} \Phi^2 + \frac{1}{2} \varepsilon^{-2} |\nabla_{\varepsilon} \Phi|^2 + \frac{\rho^{\gamma-1}}{\gamma - 1} = E,$$

with boundary condition

$$\nabla \Phi \cdot \nabla S = \varepsilon^2 \Phi X S_X + \nabla_{\varepsilon} \Phi \cdot \nabla_{\varepsilon} S = 0 \quad \text{at} \quad S = 0.$$

We assume the expansions

$$\Phi(X, y, z; \varepsilon) = \Phi_0(X, y, z) + \mathcal{O}(\varepsilon^2), \quad \rho(X, y, z; \varepsilon) = \rho_0(X, y, z) + \mathcal{O}(\varepsilon^2).$$

From Bernoulli's equation it follows that $|\nabla_{\varepsilon} \Phi_0|^2 = 0$, so $\Phi_0 = \Phi_0(X)$, and hence $\rho_0 = \rho_0(X)$. From the mass flux equation we get $\Phi_0 X = F / \rho_0 A$, while the algebraic equation

$$\frac{F^2}{2 \rho_0^2 A^2} + \frac{\rho_0^{\gamma-1}}{\gamma - 1} = E$$

finally determines $\rho_0$ (in general to be solved numerically).

Example 15.22 (Shallow water equations.) The irrotational motion under gravity of a horizontal layer of inviscid, incompressible water is described by the equation for mass conservation and the Bernoulli equation (see Eq. (7.4.11))

$$\nabla^2 \varphi = 0, \quad \frac{\partial \varphi}{\partial t} + \frac{1}{2} |\nabla \varphi|^2 + \frac{p}{\rho_0} + gz = C(t),$$

where $\varphi$ is the velocity potential with velocity $v = \nabla \varphi$, $\rho_0$ the density, $p$ the pressure, $g$ the gravitational acceleration, and $C$ an unimportant function of time. The boundary conditions are provided by the impermeability of the bottom at $z = 0$, the assumption that the free surface $z = h(x, y, t)$ consists of streamlines (particles remain there), and the fact that the pressure is uniformly constant along the free surface (the big difference between the density of water
3. REGULAR PERTURBATION PROBLEMS

and air makes the water insensitive to any air motion. As any constant in pressure may be absorbed by \( C \), we may assume that the surface pressure is zero. These conditions result into

\[
\begin{align*}
  z = 0 : & \quad \frac{\partial \varphi}{\partial z} = 0, \\
  z = h : & \quad \frac{\partial \varphi}{\partial z} = \frac{\partial h}{\partial t} + \frac{\partial \varphi}{\partial x} \frac{\partial h}{\partial x} + \frac{\partial \varphi}{\partial y} \frac{\partial h}{\partial y}, \quad \text{and} \quad p = 0.
\end{align*}
\]

We assume that the typical horizontal velocities \( U \) are so large and frequencies \( f \) are so low, that the corresponding typical length scale \( L = U/f \) is large compared to the typical water depth \( D \) (for example tidal motion). To quantify this slenderness we introduce the small parameter \( \varepsilon = D/L \). We are interested in the situation where pressure is both coupled to the inertia of the flow and to the effects of gravity. This corresponds to the assumption that \( p \) scales on \( \rho_0 U^2 \) and the inverse Froude number \( \gamma = gD/U^2 = O(1) \). If we scale and make dimensionless as follows

\[
\begin{align*}
  x = LX, \quad y = LY, \quad z = DZ, \quad t = LU^{-1} \tau, \quad \varphi = UL \psi, \quad p = \rho_0 U^2 P, \quad h = DH,
\end{align*}
\]

and introduce the operator \( \nabla_\perp := e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y} \), which is \( \nabla \) restricted to \( X \) and \( Y \), we obtain

\[
\frac{\partial^2 \psi}{\partial Z^2} + \varepsilon^2 \nabla_\perp^2 \psi = 0, \quad \frac{\partial \psi}{\partial \tau} + \frac{\varepsilon}{2} |\nabla_\perp \psi|^2 + \frac{1}{2} \varepsilon^{-2} \left( \frac{\partial \psi}{\partial Z} \right)^2 + P + \gamma Z = \tilde{C}(t),
\]

with boundary conditions

\[
\begin{align*}
  Z = 0 : & \quad \frac{\partial \psi}{\partial Z} = 0, \\
  Z = H : & \quad \frac{\partial \psi}{\partial Z} = \varepsilon^2 \left( \frac{\partial H}{\partial \tau} + \nabla_\perp \psi \cdot \nabla_\perp H \right), \quad \text{and} \quad P = 0.
\end{align*}
\]

Assuming no interfering \( O(\varepsilon) \)-effects (e.g. from initial or boundary conditions), we expand in powers of \( \varepsilon^2 \), the only small parameter that occurs,

\[
\psi = \Psi_0 + \varepsilon^2 \Psi_1 + O(\varepsilon^4), \quad H = H_0 + \varepsilon^2 H_1 + O(\varepsilon^4), \quad P = P_0 + \varepsilon^2 P_1 + O(\varepsilon^4),
\]

to obtain to leading order \( \Psi_{0,zz} = 0 \), which integrates to \( \Psi_{0,z} = B_0(X, Y, \tau) = 0 \) because of the boundary conditions. So we finally have

\[
\Psi_0 = A_0(X, Y, \tau).
\]

To first order we have \( \Psi_{1,zz} = -\nabla_\perp^2 \Psi_0 \), which integrates to

\[
\frac{\partial}{\partial Z} \Psi_1 = -Z \nabla_\perp^2 \Psi_0
\]

when we take into account the boundary condition at \( Z = 0 \). Next we expand and substitute these results into Bernoulli’s equation (note that \( \varepsilon^{-2} |\psi_x|^2 = O(\varepsilon^2) \)), and get

\[
\frac{\partial}{\partial \tau} \Psi_0 + \frac{1}{2} |\nabla_\perp \Psi_0|^2 + P_0 + \gamma Z = \tilde{C}(t).
\]

By the pressure boundary condition at \( Z = H_0 \) this yields

\[
\frac{\partial}{\partial \tau} \Psi_0 + \frac{1}{2} |\nabla_\perp \Psi_0|^2 + \gamma H_0 = \tilde{C}(t).
\]

(\ast)
Finally, the streamline condition produces
\[ \frac{\partial}{\partial Z} \Psi_1 = -H_0 \nabla_\perp^2 \Psi_0 = \frac{\partial}{\partial \tau} H_0 + \nabla_\perp \Psi_0 \cdot \nabla_\perp H_0, \]
or
\[ \frac{\partial}{\partial \tau} H_0 + \nabla_\perp \cdot (H_0 \nabla_\perp \Psi_0) = 0 \quad (\dagger) \]
Equations (\( \ast \)) with (\( \dagger \)) are known as a form of the shallow water equations (see also Chapter 12). They are not generally solvable, and their behaviour requires extensive analysis.

A family of simple wave solutions (i.e. along a characteristic) may be found as follows (cf. [127]). We look for a plane wave in, say, \( \xi = X \cos \theta + Y \sin \theta \)-direction, so we have after rewriting the equations in terms of \( \xi \), velocity \( V_0 = \Psi_{0,\xi} \) and height \( H_0 \)
\[ \frac{\partial}{\partial \tau} H_0 + \frac{\partial}{\partial \xi} (V_0 H_0) = 0, \quad \frac{\partial}{\partial \tau} V_0 + \frac{\partial}{\partial \xi} \left( \frac{1}{2} V_0^2 + \gamma H_0 \right) = 0. \]
Evidently, both equations may be written in the same characteristic form if
\[ \frac{d}{dH_0} (V_0 H_0) = \frac{d}{dV_0} \left( \frac{1}{2} V_0^2 + \gamma H_0 \right). \]
This is satisfied if \( (\frac{d}{dH_0} V_0)^2 H_0 = \gamma \), or
\[ V_0 = \pm 2\sqrt{\gamma H_0} + c_0, \]
where \( c_0 \) is a constant. We can now assume that \( H_0 = H_0(\eta) \) and the corresponding \( V_0 = V_0(\eta) \), where \( \eta = \eta(x, t) \) satisfies
\[ \frac{\partial}{\partial \tau} \eta + (V_0 \pm \sqrt{\gamma H_0}) \frac{\partial}{\partial \xi} \eta = 0. \]
Using the result of Exercise (1.5), this has solutions implicitly described by
\[ \eta = F(\xi - f(\eta) \tau), \quad \text{where} \quad f(\eta) = V_0(\eta) \pm \sqrt{\gamma H_0(\eta)} \]
and \( F \) is any suitable differentiable function. \( \square \)

3.2 Lindstedt-Poincaré method

When we have a function \( f \), depending on a small parameter \( \varepsilon \), and periodic in \( t \) with fundamental frequency \( \omega(\varepsilon) \), we can write \( f \) as a Fourier series
\[ f(t; \varepsilon) = \sum_{n=-\infty}^{\infty} A_n(\varepsilon) e^{i\omega(\varepsilon)t} \quad (3.5) \]
If amplitudes and frequency have an asymptotic expansion for small \( \varepsilon \), say
\[ A_n(\varepsilon) = A_{n,0} + \varepsilon A_{n,1} + \ldots, \quad \omega(\varepsilon) = \omega_0 + \varepsilon \omega_1 + \ldots, \quad (3.6) \]
we have a natural asymptotic series expansion for \( f \) of the form
\[ f(t; \varepsilon) = \sum_{n=-\infty}^{\infty} A_{n,0} e^{i\omega_{0n}t} + \varepsilon \sum_{n=-\infty}^{\infty} (A_{n,1} + i\omega_1 A_{n,0}) e^{i\omega_{0n}t} + \ldots \quad (3.7) \]
3. REGULAR PERTURBATION PROBLEMS

This expansion, however, is only uniform in \( t \) on an interval \([0, T]\), where \( T = o(\varepsilon^{-1}) \). On a larger interval, for example \([0, \varepsilon^{-1}]\), the asymptotic hierarchy in the expansion becomes invalid, because \( \varepsilon t = \mathcal{O}(1) \). This happens because of the occurrence of algebraically growing oscillatory terms, called “secular terms”. Secular = occurring once in a century, and saeculum = generation, referring to their astronomical origin.

**Definition 15.23.** The terms proportional to \( t^m \sin(n \omega_0 t) \), \( t^m \cos(n \omega_0 t) \) are called “secular terms”.

It is therefore far better to apply first a coordinate transformation \( \tau = \omega(\varepsilon) t \), introduce \( F(\tau; \varepsilon) = f(t; \varepsilon) \), and expand \( F \), rather than \( f \), asymptotically. We get the uniformly valid approximation

\[
F(\tau; \varepsilon) = \sum_{n=-\infty}^{\infty} A_n(\varepsilon) e^{i n \tau} = \sum_{n=-\infty}^{\infty} A_{n,0} e^{i n \tau} + \varepsilon \sum_{n=-\infty}^{\infty} A_{n,1} e^{i n \tau} + \ldots \quad (3.8)
\]

The method is called the Lindstedt-Poincaré method or the method of strained coordinates. In practical situations, the frequency \( \omega \) is of course unknown, and has to be found. Therefore, when constructing the solution we have to allow for an unknown coordinate transformation. In order to construct the unknown \( \omega(\varepsilon) \) we expand this, for example like

\[
\tau = (\omega_0 + \varepsilon \omega_1 + \varepsilon^2 \omega_2 + \ldots) t \quad (3.9)
\]

but this depends of course on the problem. Note that the purpose of the scaling is to render the asymptotic expansion of \( F \) regular, so it is no restriction to assume \( \omega_0 = 1 \). The other coefficients are determined from the additional condition that the asymptotic hierarchy should be respected as long as possible. In other words, secular terms should not occur. We will illustrate this with the following example.

**Example 15.24** Consider the motion of the pendulum, described by the initial value problem (see Example 7.2)

\[
\ddot{\theta} + K^2 \sin(\theta) = 0, \quad \text{with} \quad \theta(0) = \varepsilon, \ \theta'(0) = 0,
\]

where \( 0 < \varepsilon \ll 1 \). After the transformation \( \tau = \omega t \) and noting that \( \theta = \mathcal{O}(\varepsilon) \), we have

\[
\omega^2 \theta'' + K^2 (\theta - \frac{1}{6} \theta^3 + \ldots) = 0.
\]

We expand

\[
\omega = 1 + \varepsilon^2 \omega_1 + \ldots, \quad \theta = \varepsilon \theta_0 + \varepsilon^3 \theta_1 + \ldots,
\]

and find, after substitution, the equations for the first two orders

\[
\theta_0'' + K^2 \theta_0 = 0, \quad \theta_0(0) = 1, \ \theta_0'(0) = 0, \quad \theta_1'' + K^2 \theta_1 = -2\omega_1 \theta_0'' + \frac{1}{6} K \theta_0^3, \quad \theta_1(0) = 0, \ \theta_1'(0) = 0.
\]

The solution \( \theta_0 \) is obviously given by

\[
\theta_0 = \cos(K \tau),
\]
leading to (see Appendix B) the following equation for $\theta_1$

$$\theta''_1 + K^2 \theta_1 = 2K(\omega_1 K + \frac{1}{16}) \cos(K\tau) + \frac{1}{24} K \cos(3K\tau).$$

At this point it is essential to observe that the right-hand-side consists of two forcing terms: one with frequency $3K$ and one with $K$, the resonance frequency of the left-hand-side. This resonance would lead to secular terms, as the solutions will behave like $\tau \sin(K\tau)$ and $\tau \cos(K\tau)$. Therefore, in order to suppress the occurrence of secular terms, the amplitude of the resonant forcing term should vanish, which yields the next order terms $\omega_1$ and $\theta_1$.

$$\omega_1 = -\frac{1}{16K} \quad \text{leading to} \quad \theta_1 = \frac{1}{192K} (\cos(K\tau) - \cos(3K\tau)). \quad \Box$$

4 Singular perturbation problems

If the solution of the problem considered does not allow a regular expansion, the problem is singular and the solution has no uniform Poincaré expansion in the same variable. We will consider two classes of problems. In the first one the singular behaviour is of boundary layer type and the solution can be built up from locally regular expansions. The solution method is called “method of matched asymptotic expansions”. In the other one more time or length scales occur together and a solution is constructed by considering these length scales as if they were independent. The solution method is called “method of multiple scales”.

4.1 Matched asymptotic expansions

Very often it happens that a simplifying limit applied to a more comprehensive model gives a correct approximation for the main part of the domain, but not everywhere: the limit is non-uniform. This non-uniformity may be in space, in time, or in any other variable. For the moment we think of non-uniformity in space, let’s say a small region near $x = 0$. If this region of non-uniformity is crucial for the problem, for example because it contains a boundary condition, or a source, the primary reduced problem (which does not include the region of non-uniformity) is not sufficient. This, however, does not mean that no use can be made of the inherent small parameter. The local nature of the non-uniformity itself gives often the possibility of another reduction. In such a case we call this a couple of limiting forms, “inner and outer problems”, and are evidence of the fact that we have apparently physically two connected but different problems as far as the dominating mechanism is concerned. Depending on the problem, we now have two simpler problems, serving as boundary conditions to each other via continuity or matching conditions.

Non-uniform asymptotic approximations

Suppose that a given sufficiently smooth function $\Phi(x; \varepsilon)$, with $0 \leq x \leq 1$, $0 < \varepsilon \leq \varepsilon_1$, does not have a uniform limit $\varepsilon \to 0$, $x \to 0$. Typically, such a function will depend, apart from $x$, on combinations like $x/\delta(\varepsilon)$, where $\delta = o(1)$. 
Assume that this function does not have a regular asymptotic expansion on the whole interval \([0, 1]\) but only on partial intervals \(x \in [\eta(\epsilon), 1]\), where \(\eta = o(1)\) and \(\delta = o(\eta)\). We call this expansion the outer expansion, principally valid in the “\(x = \mathcal{O}(1)\)”-outer region, but extendible to \([\mathcal{O}(\eta), 1]\).

\[
\Phi(x; \epsilon) = \sum_{k=0}^{n} \mu_k(\epsilon) \phi_k(x) + o(\mu) \quad \epsilon \to 0, \ x = \mathcal{O}(1). \tag{4.1}
\]

Consider now the transformation to the stretched coordinate

\[
\xi = \frac{x}{\delta(\epsilon)}. \tag{4.2}
\]

Assume that the function \(\Phi = \Psi(\xi; \epsilon)\) has a non-trivial regular asymptotic expansion on partial intervals \(\xi \in [0, \zeta(\epsilon)/\delta(\epsilon)]\), where \(\eta(\epsilon) < \zeta(\epsilon)\). We call this expansion the inner expansion, principally valid in the “\(\xi = \mathcal{O}(1)\)”-inner region, but extendible to \([0, \mathcal{O}(\zeta)]\).

\[
\Psi(\xi; \epsilon) = \sum_{k=0}^{m} \lambda_k(\epsilon) \psi_k(\xi) + o(\lambda) \quad \epsilon \to 0, \ \xi = \mathcal{O}(1). \tag{4.3}
\]

**Example 15.25**

\[
\Phi(x; \epsilon) = \arctan(x) + \sin(x + \epsilon) = \frac{\pi}{2} + \sin x + \epsilon \cos x - \frac{\pi}{2} + \mathcal{O}(\epsilon^{3/2}) \quad \text{on} \quad \frac{1}{2} \epsilon^{1/2} \leq x \leq 1
\]

\[
\Psi(\xi; \epsilon) = \arctan(\xi) + \sin(\epsilon \xi + \epsilon) = \arctan(\xi) + \epsilon(\xi + 1) + \mathcal{O}(\epsilon^{3/2}) \quad \text{on} \quad 0 \leq \xi \leq 2\epsilon^{-1/2}
\]

The adjective “non-trivial” is essential: the expansion must be significant, i.e. different from the outer-expansion in \(\phi_n\) rewritten in the inner variable \(\xi\). This determines the choice of the inner variable \(\xi = x/\delta(\epsilon)\). The scaling \(\delta(\epsilon)\) is the asymptotically largest gauge function with this property. We call the expansion for \(\Psi\) the inner expansion or boundary layer expansion, the region \(\xi = \mathcal{O}(1)\) or \(x = \mathcal{O}(\delta)\) being the boundary layer with thickness \(\delta\), and \(\xi\) the boundary layer variable. Boundary layers may be nested, and may occur at internal points of the domain of \(\Phi\). Then they are called internal layers. The assumption \(\eta < \zeta\), i.e. that inner and outer expansion may be extended to regions that overlap, is called the overlap hypothesis.

Suppose, \(\Phi(x; \epsilon)\) has an outer-expansion

\[
\Phi(x; \epsilon) = \sum_{k=0}^{n} \mu_k(\epsilon) \phi_k(x) + o(\mu)
\]

and a boundary layer \(x = \mathcal{O}(\delta)\) with inner-expansion

\[
\Psi(\xi; \epsilon) = \sum_{k=0}^{m} \lambda_k(\epsilon) \psi_k(\xi) + o(\lambda).
\]

Suppose that both expansions are complementary, i.e. there is no other boundary layer in between \(x = \mathcal{O}(1)\) and \(x = \mathcal{O}(\delta)\), then the overlap-hypothesis says that both expansions
CHAPTER 15. PERTURBATION METHODS

represent the same function in an intermediate region of overlap. This overlap region may be described by a stretched variable \( x = \eta(\varepsilon)\sigma \), asymptotically in between \( O(1) \) and \( O(\delta) \), so: \( \delta \ll \eta \ll 1 \). In the overlap region both expansions match, which means that asymptotically both expansions are equivalent and reduce to the same expressions. A widely used and relatively simple procedure is Van Dyke’s matchings rule [119, 31]: the outer-expansion, rewritten in the inner-variable, has a regular series expansion, which is equal to the regular asymptotic expansion of the inner-expansion, rewritten in the outer-variable. Suppose that

\[
\sum_{k=0}^{n} \mu_k(\varepsilon) \phi_k(\delta \xi) = \sum_{k=0}^{m} \lambda_k(\varepsilon) \eta_k(\xi) + o(\lambda_m),
\]

(4.6a)

\[
\sum_{k=0}^{n} \lambda_k(\varepsilon) \psi_k(x/\delta) = \sum_{k=0}^{n} \mu_k(\varepsilon) \theta_k(x) + o(\mu_n),
\]

(4.6b)

then the expansion of \( \eta_k \) back to \( x \)

\[
\sum_{k=0}^{n} \lambda_k(\varepsilon) \eta_k(x/\delta) = \sum_{k=0}^{n} \mu_k(\varepsilon) \xi_k(x) + o(\mu_n),
\]

(4.7)

is such that \( \xi_k = \theta_k \) for \( k = 0, \ldots, n \).

The idea of matching is very important because it allows one to move smoothly from one regime into the other. The method of constructing local, but matching, expansions is therefore called “Matched Asymptotic Expansions” (MAE) [67].

Constructing asymptotic solutions

The most important application of this concept of inner- and outer-expansions is that approximate solutions of certain differential equations can be constructed for which the limit under a small parameter is apparently non-uniform. The main lines of argument for constructing a MAE solution to a differential equation satisfying some boundary conditions are as follows. Suppose \( \Phi \) is given by the equation

\[
D[\Phi', \Phi](x; \varepsilon) = 0 \quad + \text{boundary conditions},
\]

(4.8)

where \( \Phi' = \frac{d}{d\varepsilon} \Phi \). Then we try to construct an outer solution by looking for “non-trivial degenerations” of \( D \) under \( \varepsilon \to 0 \), that is, find \( \mu_0(\varepsilon) \) and \( v_0(\varepsilon) \) such that

\[
\lim_{\varepsilon \to 0} v_0^{-1}(\varepsilon) D[\mu_0, \mu_0\phi_0](x; \varepsilon) = D_0[\phi_0, \phi_0](x) = 0
\]

(4.9)

has a non-trivial solution \( \phi_0 \). A series \( \psi = \mu_0\psi_0 + \mu_1\psi_1 + \cdots \) is constructed by repeating the process for \( D - v_0 D_0 \), etc.

Suppose, the approximation is non-uniform. For example, not all boundary conditions can be satisfied. Then we start looking for an inner-expansion if we have reasons to believe that the non-uniformity is of boundary-layer type. Presence, location and size of the boundary layer(s) are now found by the correspondence principle, that is the (heuristic)
idea that if $\Phi$ behaves somehow differently in the boundary layer, the defining equation must also be essentially different. Therefore, we search for significant degenerations or distinguished limits of $\mathcal{D}$. These are degenerations of $\mathcal{D}$ under $\varepsilon \to 0$, with scaled $x$ and $\Phi$, that contain the most information, and without being contained in other, richer, degenerations.

**Example 15.26** Under the limit $\varepsilon \to 0$, the equation $\varepsilon y' + y = \sin x$, $y(0) = 1$ reduces to $y = \sin x$ with $y(0) \neq 1$. After the scaling $x = \varepsilon \xi$, the equation reduces to the essentially different $y + y' = 0$. □

The next step is then to select from these distinguished limits the one(s) allowing a solution that matches with the outer solution and satisfies any applicable boundary conditions. Symbolically:

$$\text{find } x_0, \delta(\varepsilon), \lambda(\varepsilon), \kappa(\varepsilon)$$

$$\text{with } x = x_0 + \delta \xi, \quad \Phi(x; \varepsilon) = \lambda(\varepsilon) \Psi(\xi; \varepsilon),$$

$$\text{such that } B_0(\psi', \psi_0, \xi) = \lim_{\varepsilon \to 0} \kappa^{-1} \mathcal{D}[\delta^{-1} \lambda \Psi', \lambda \Psi](x_0 + \delta \xi; \varepsilon) \text{ has the "richest" structure,}$$

$$\text{while } B_0(\psi', \psi_0, \xi) = 0 \text{ has a solution satisfying boundary and matching conditions.}$$

Again, an asymptotic expansion may be constructed inductively, by repeating the argument. It is of practical importance to note that the order estimate $\lambda$ of $\Phi$ in the boundary layer is often determined a posteriori by boundary or matching conditions. We can illustrate some of the main arguments by considering

$$\mathcal{D}[\psi, \psi](x; \varepsilon) = \varepsilon \frac{d^2 \psi}{dx^2} + \frac{d \psi}{dx} - 2x = 0, \quad \psi(0) = \psi(1) = 2. \quad (4.10)$$

The leading order outer-equation is evidently (with $\mu_0 = \nu_0 = 1$)

$$\mathcal{D}_0 = \frac{d \psi_0}{dx} - 2x = 0, \quad (4.11)$$

with solution

$$\psi_0 = x^2 + A. \quad (4.12)$$

The integration constant $A$ can be determined by the boundary condition $\psi_0(0) = 2$ at $x = 0$ or $\psi_0(1) = 2$ at $x = 1$, but not both, so we expect a boundary layer at either end. By trial and error we find that no solution can be constructed if we assume a boundary layer at $x = 1$, so, inferring a boundary layer at $x = 0$, we have to use the boundary condition at $x = 1$ and find

$$\psi_0 = x^2 + 1. \quad (4.13)$$

The structure of the equation suggests a correction of $\mathcal{O}(\varepsilon)$, so we try the expansion

$$\psi = \psi_0 + \varepsilon \psi_1 + \varepsilon^2 \psi_2 + \cdots. \quad (4.14)$$
CHAPTER 15. PERTURBATION METHODS

For \( \varphi_1 \) this results into the equation

\[
\frac{d\varphi_1}{dx} + \frac{d^2\varphi_0}{dx^2} = 0,
\]

with \( \varphi_1(1) = 0 \) (the \( \mathcal{O}(\varepsilon) \)-term of the boundary condition), which has the solution

\[
\varphi_1 = 2 - 2x.
\]

Higher orders are straightforward:

\[
\frac{d\varphi_n}{dx} = 0, \quad \text{with} \quad \varphi_n(1) = 0,
\]

leading to solutions \( \varphi_n \equiv 0 \). We find for the outer expansion

\[
\varphi = x^2 + 1 + 2\varepsilon(1 - x) + \mathcal{O}(\varepsilon^N).
\]

We continue with the inner expansion, and find with \( x_0 = 0 \), \( \varphi = \lambda \psi, x = \delta \xi \)

\[
\frac{\varepsilon \lambda}{\delta^2} \frac{d^2\psi}{d\xi^2} + \frac{\lambda}{\delta} \frac{d\psi}{d\xi} - 2\delta \xi = 0.
\]

Both from the matching (\( \varphi_{\text{outer}} \to 1 \) for \( x \downarrow 0 \)) and from the boundary condition (\( \varphi(0) = 2 \)) we have to conclude that \( \varphi_{\text{inner}} = \mathcal{O}(1) \) and so \( \lambda = 1 \). Furthermore, the boundary layer has only a reason for existence if it comprises new effects, not described by the outer solution. From the correspondence principle we expect that new effects are only included if \( (d^2\psi/d\xi^2) \) is included. So \( \varepsilon \delta^{-2} \) must be at least as large as \( \delta^{-1} \), the largest of \( \delta^{-1} \) and \( \delta \). From the principle that we look for the equation with the richest structure, it must be exactly as large, implying a boundary layer thickness \( \delta = \varepsilon \). Thus we have \( \kappa = \varepsilon^{-1} \), and the inner equation

\[
\frac{d^2\psi}{d\xi^2} + \frac{d\psi}{d\xi} - 2\varepsilon^2 \xi = 0.
\]

From this equation it would seem that we have a series expansion without the \( \mathcal{O}(\varepsilon) \)-term, since the equation for this order would be the same as for the leading order. However, from matching with the outer solution:

\[
\varphi_{\text{outer}} \to 1 + 2\varepsilon + \varepsilon^2 (\xi^2 - 2\xi) + \cdots \quad (x = \varepsilon \xi, \xi = \mathcal{O}(1)),
\]

we see that an additional \( \mathcal{O}(\varepsilon) \)-term is to be included. So we substitute the series expansion:

\[
\varphi = \psi_0 + \varepsilon \psi_1 + \varepsilon^2 \psi_2 + \cdots.
\]

It is a simple matter to find

\[
\frac{d^2\psi_0}{d\xi^2} + \frac{d\psi_0}{d\xi} = 0, \quad \psi_0(0) = 2 \quad \Rightarrow \quad \psi_0 = 2 + A_0(e^{-\xi} - 1),
\]

\[
\frac{d^2\psi_1}{d\xi^2} + \frac{d\psi_1}{d\xi} = 0, \quad \psi_1(0) = 0 \quad \Rightarrow \quad \psi_1 = A_1(e^{-\xi} - 1),
\]

\[
\frac{d^2\psi_2}{d\xi^2} + \frac{d\psi_2}{d\xi} = 2\xi, \quad \psi_2(0) = 0 \quad \Rightarrow \quad \psi_2 = \xi^2 - 2\xi + A_2(e^{-\xi} - 1),
\]
where the constants $A_0, A_1, A_2, \cdots$ are to be determined from the matching condition that inner and outer solution should be asymptotically equivalent in the region of overlap. We can follow the method of intermediate variables and rewrite outer expansion (4.18) and inner expansion (4.22) in an intermediate variable $\xi = \eta(e)\sigma$ where $e \ll \eta \ll 1$ and re-expand as follows.

$$
\begin{align*}
\xi^2 + 1 + 2e(1 - \xi) + O(e^3) & \simeq 1 + 2\xi + \xi^2 - 2\xi^2 + O(e^3) & (4.24a) \\
2 + A_0(e^{-\xi} - 1) + eA_1(e^{-\xi} - 1) + e^2(\xi^2 - 2\xi + A_2(e^{-\xi} - 1)) + O(e^3) & \simeq 2 - A_0 - \xi A_1 + \xi^2\sigma^2 - 2\xi\eta\sigma - \xi^2 A_2 + O(\epsilon^3) & (4.24b)
\end{align*}
$$

Alternatively, we can follow Van Dyke’s matching rule, and rewrite outer expansion (4.18) in inner variable $\xi$, inner expansion (4.22) in outer variable $x$, re-expand and rewrite the result in $x$. This results into

$$
\begin{align*}
\xi^2 + 1 + 2e(1 - \xi) + O(e^3) & \simeq 1 + 2\xi + x^2 - 2\xi x + O(e^3) & (4.24c) \\
2 + A_0(e^{-\xi} - 1) + eA_1(e^{-\xi} - 1) + e^2(\xi^2 - 2\xi + A_2(e^{-\xi} - 1)) + O(e^3) & \simeq 2 - A_0 - \xi A_1 + x^2 - 2\xi x - \xi^2 A_2 + O(\epsilon^3) & (4.24d)
\end{align*}
$$

In either case, the resulting reduced expressions, (4.24a) and (4.24b), respectively (4.24c) and (4.24d), must be functionally equivalent. A full matching is thus obtained if we choose $A_0 = 1, A_1 = -2, A_2 = 0$.

**Example 15.27 Prandtl’s boundary layer analysis.** The start of modern boundary layer theory is Prandtl’s analysis of uniform incompressible low-viscous flow along a flat plate. Consider the stationary 2D version of equations (7.4.3), with $\epsilon = Re^{-1}$ small,

$$
\begin{align*}
u_x + v_y &= 0, \quad uu_x + uu_y = -p_x + \epsilon(u_{xx} + u_{yy}), \quad uu_x + vv_y = -p_y + \epsilon(v_{xx} + v_{yy}),
\end{align*}
$$

subject to boundary conditions $u = v = 0$ at $y = 0$, $0 < x < 1$, and outer solution for $y = O(1)$ to leading order given by $(u, v, p) = (1, 0, 0)$. When we scale $x = X, y = \epsilon^m Y$, $u = U, v = \epsilon^n V$, and $p = \epsilon^k P$, we find

$$
\begin{align*}
U_X + \epsilon^m \cdot \epsilon^{-n} V_Y &= 0, \quad UUU_X + \epsilon^{m-n} V U_Y = -e^k P_X + \epsilon U_{XX} + \epsilon^{1-2n} U_{YY}, \\
\epsilon^n U V_X + \epsilon^{2m-n} V V_Y &= -e^{k-n} P_Y + \epsilon^{1+m} V_{XX} + \epsilon^{1+m-2n} V_{YY}.
\end{align*}
$$

This yields the distinguished limits $m = n = \frac{1}{2}$ and $k = 1$, with the significant degeneration

$$
U_X + V_Y = 0, \quad UUU_X + V U_Y = U_{YY},
$$

known as Prandtl’s boundary layer equations. The same equations, but with other boundary conditions, are valid in Goldstein’s viscous wake $x > 1, y = O(\epsilon^{1/2})$. The trailing edge region around $x = 1, y = 0$, is far more complicated. Here the boundary layer structure consists of three layers $y = O(\epsilon^{5/8}), O(\epsilon^{4/8}), O(\epsilon^{3/8})$ within $x - 1 = O(\epsilon^{3/8})$. This is known as Stewarton’s triple deck.
Logarithmic switchback

It is not always evident from just the structure of the equation what the necessary expansion will look like. Sometimes it is well concealed and we are only made aware of an invalid initial choice by a matching failure. In fact, it is also the matching process itself that reveals us the required sequence of scaling functions. An example of such a back reaction is known as logarithmic switch back.

Consider the following problem for $y = y(x; \varepsilon)$ on the unit interval.

$$
\varepsilon y'' + x (y' - y) = 0, \quad 0 < x < 1, \quad y(0; \varepsilon) = 0, \quad y(1; \varepsilon) = e. \tag{4.25}
$$

The outer solution appears to have the expansion

$$
y(x; \varepsilon) = y_0(x) + \varepsilon y_1(x) + O(\varepsilon^2). \tag{4.26}
$$

By trial and error, the boundary layer appears to be located near $x = 0$, so the governing equations and boundary conditions are then

$$
y_0' - y_0 = 0, \quad y_0(1) = e, \tag{4.27a}
y_n' - y_n = -x^{-1} y_{n-1}'' , \quad y_n(1) = 0, \tag{4.27b}
$$

with general solution

$$
y_n(x) = A_n e^x + \int_x^1 z^{-1} e^{x-z} y_{n-1}'' (z) \, dz, \tag{4.28a}
$$

such that

$$
y_0(x) = e^x, \tag{4.28b}
y_1(x) = -e^x \ln(x), \tag{4.28c}
$$

eq. The boundary layer thickness is found from the assumed scaling $x = \varepsilon^m t$ and noting that $y = O(1)$ because of the matching with the outer solution. This leads to the significant degeneration of $m = \frac{1}{2}$, or $x = \varepsilon^{1/2} t$. The boundary layer equation for $y(x; \varepsilon) = Y(t; \varepsilon)$ is thus

$$
Y'' + t Y' - \varepsilon^{1/2} t Y = 0, \quad Y(0; \varepsilon) = 0. \tag{4.29}
$$

The obvious choice of expansion of $Y$ in powers of $\varepsilon^{1/2}$ is not correct, as the found solution does not match with the outer solution. Therefore, we consider the outer solution in more detail for small $x$. When $x = \varepsilon^{1/2} t$, we have for the outer solution

$$
y(\varepsilon^{1/2} t; \varepsilon) = 1 + \varepsilon^{1/2} t - \frac{1}{2} \varepsilon \ln(\varepsilon) + \varepsilon (\frac{1}{4} t^2 - \ln t)
- \frac{1}{2} \varepsilon^{1/2} \ln(\varepsilon) t + \varepsilon^{1/2} (\frac{1}{6} t^3 - t \ln t) + O(\varepsilon^2). \tag{4.30}
$$

So we apparently need at least

$$
Y(t; \varepsilon) = Y_0(t) + \varepsilon^{1/2} Y_1(t) + \varepsilon \ln(\varepsilon) Y_2(t) + \varepsilon Y_3(t) + o(\varepsilon), \tag{4.31}
$$
4. SINGULAR PERTURBATION PROBLEMS

with equations and boundary conditions
\[ Y_0'' + tY_0' = 0, \quad Y_0(0) = 0, \]  
\[ Y_1'' + tY_1' = tY_0, \quad Y_1(0) = 0, \]  
\[ Y_2'' + tY_2' = 0, \quad Y_2(0) = 0, \]  
\[ Y_3'' + tY_3' = tY_1, \quad Y_3(0) = 0. \]  

Hence, the solution is given by
\[ Y_0(t) = A_0 \text{ erf} \left( \frac{t}{\sqrt{2}} \right), \]  
\[ Y_1(t) = A_1 \text{ erf} \left( \frac{t}{\sqrt{2}} \right) + Y_0'(0) - Y_0(t) + \int_0^t Y_0(z) \, \text{d}z \]  
\[ = (A_1 + A_0 t) \text{ erf} \left( \frac{t}{\sqrt{2}} \right) + 2 \left( \frac{2}{\sqrt{\pi}} \right) A_0 (e^{-\frac{t^2}{2}} - 1), \]  
\[ Y_2(t) = A_2 \text{ erf} \left( \frac{t}{\sqrt{2}} \right). \]

Unfortunately, \( Y_3 \) cannot be expressed in closed form. However, for demonstration it is sufficient to derive the behaviour of \( Y_3 \) for large \( t \). As \( \text{erf}(z) \to 1 \) exponentially fast for \( z \to \infty \), we obtain
\[ tY_1(t) = A_0 t^2 + (A_1 - 2 \left( \frac{2}{\sqrt{\pi}} \right) A_0) t + \text{exponentially small terms.} \]  
After some simple manipulations we obtain
\[ Y_3(t) = \frac{1}{2} A_0 t^2 + (A_1 - 2 \left( \frac{2}{\sqrt{\pi}} \right) A_0) t - A_0 \ln(t) + \ldots. \]

For matching of the inner solution, we compare with expression (4.30), and have
\[ A_0 + \varepsilon \frac{1}{2} (A_1 + A_0 t - 2 \left( \frac{2}{\sqrt{\pi}} \right) A_0) + \varepsilon \ln(\varepsilon) A_2 + \varepsilon \left( \frac{1}{2} A_0 t^2 + (A_1 - 2 \left( \frac{2}{\sqrt{\pi}} \right) A_0) t - A_0 \ln(t) \right) \]  
\[ = 1 + \varepsilon \frac{1}{2} t - \frac{1}{4} \varepsilon \ln(\varepsilon) + \varepsilon \left( \frac{1}{2} t^2 - \ln(t) \right) \]  
and indeed, a full matching is achieved with
\[ A_0 = 1, \quad A_1 = 2 \left( \frac{2}{\sqrt{\pi}} \right), \quad A_2 = -\frac{1}{2}. \]

The rôle of matching

It is important to note that a matching is possible at all! Only a part of the terms can be matched by selection of the undetermined constants. Other terms are already equal, without free constants, and there is no way to repair a possibly incomplete matching here. This is an important consistency check on the found solution, at least as long as no real proof is available. If no matching appears to be possible, almost certainly one of the assumptions made with the construction of the solution has to be reconsidered. Particularly notorious are logarithmic singularities of the outer solution, as we saw above. See for other examples [71].

Summarizing, matching of inner- and outer expansion plays an important rôle in the following ways:
i) it provides information about the sequence of order (gauge) functions \( \{ \mu_k \} \) and \( \{ \lambda_k \} \) of the expansions;

ii) it allows us to determine unknown constants of integration;

iii) it provides a check on the consistency of the solution, giving us confidence in the correctness.

4.2 Multiple scales

Suppose a function \( \varphi(x; \varepsilon) \) depends on more than one length scale acting together, for example \( x, \varepsilon x, \) and \( \varepsilon^2 x \). Then the function does not have a regular expansion on the full domain of interest, \( x \leq \mathcal{O}(\varepsilon^{-2}) \) say. It is not possible to bring these different length scales together by a simple coordinate transformation, like in the method of slow variation or the Lindsted-Poincaré method, or to split up our domain in subdomains like in the method of matched asymptotic expansions. Therefore we have to find another way to construct asymptotic expansions, valid in the full domain of interest. The approach that is followed in the *method of multiple scales* is at first sight rather radical: the various length scales are temporarily considered as independent variables: \( x_1 = x, x_2 = \varepsilon x, x_3 = \varepsilon^2 x, \) and the original function \( \varphi \) is identified with a more general function \( \psi(x_1, x_2, x_3; \varepsilon) \) depending on a higher dimensional independent variable.

**Example 15.28**

\[
\varphi(x; \varepsilon) = A(\varepsilon) e^{-\varepsilon x} \cos(x + \theta(\varepsilon)) \quad \text{becomes} \quad \psi(x_1, x_2; \varepsilon) = A(\varepsilon) e^{-x_2} \cos(x_1 + \theta(\varepsilon)).
\]

Since this identification is not unique, we may add constraints such that this auxiliary function \( \psi \) does have a Poincaré expansion on the full domain of interest. After having constructed this expansion, it may be associated to the original function along the line \( x_1 = x, x_2 = \varepsilon x, x_3 = \varepsilon^2 x \).

The technique, utilizing this difference between small scale and large scale behaviour is the method of multiple scales. As with most approximation methods, this method has grown out of practice, and works well for certain types of problems. Typically, the multiple scale method is applicable to problems with on the one hand a certain global quantity (energy, power), which is conserved or almost conserved, controlling the amplitude, and on the other hand two rapidly interacting quantities (kinetic and potential energy), controlling the phase. Usually, this describes slowly varying waves, affected by small effects during a long time. Intuitively, it is clear that over a short distance (a few wave lengths) the wave only sees a constant conditions and will propagate approximately as in the constant case, but over larger distances it will somehow have to change its shape in accordance with its new environment.

We will illustrate the method by considering a damped harmonic oscillator

\[
\frac{d^2y}{dr^2} + 2\varepsilon \frac{dy}{dr} + y = 0 \quad (r \geq 0), \quad y(0) = 0, \quad \frac{dy(0)}{dr} = 1
\]
with $0 < \varepsilon \ll 1$. The exact solution is readily found to be

$$y(t) = e^{-\varepsilon t} \sin(\sqrt{1 - \varepsilon^2} t).$$  

(4.37)

A naive approximation of this $y(t)$, for small $\varepsilon$ and fixed $t$, would give

$$y(t) = \sin t - \varepsilon t \sin t + O(\varepsilon^2),$$  

(4.38)

which appears to be useful for $t = O(1)$ only. For large $t$ the approximation becomes incorrect:

1) if $t \geq O(\varepsilon^{-1})$ the second term is of equal importance, or larger, as the first term and nothing is left over of the slow exponential decay;

2) if $t \geq O(\varepsilon^{-2})$ the phase has an error of $O(1)$, or larger, giving an approximation of which even the sign may be in error.

We would obtain a far better approximation if we adopted two different time variables, viz. $T = \varepsilon t$ and $\tau = \sqrt{1 - \varepsilon^2} t$, and changed to $y(t; \varepsilon) = Y(\tau, T; \varepsilon)$ where

$$Y(\tau, T; \varepsilon) = e^{-T} \frac{\sin(\tau)}{\sqrt{1 - \varepsilon^2}}.$$  

It is easily verified that a Taylor series of $Y$ in $\varepsilon$ yields a regular expansion for all $t$.

If we construct a straightforward approximate solution directly from equation (4.36), we would get the same approximation as in (4.38), which is too limited for most applications. However, knowing the character of the error, we may try to avoid them and look for the auxiliary function $Y$, instead of $y$. As we, in general, do not know the occurring time scales, their determination becomes part of the problem.

Suppose we can expand

$$y(t; \varepsilon) = y_0(t) + \varepsilon y_1(t) + \varepsilon^2 y_2(t) + \cdots.$$  

(4.39)

Substituting in (4.36) and collecting equal powers of $\varepsilon$ gives

$$O(\varepsilon^0) : \frac{d^2 y_0}{dt^2} + y_0 = 0 \quad \text{with} \quad y_0(0) = 0, \quad \frac{dy_0}{dt}(0) = 1,$$

$$O(\varepsilon^1) : \frac{d^2 y_1}{dt^2} + y_1 = -2 \frac{dy_0}{dt} \quad \text{with} \quad y_1(0) = 0, \quad \frac{dy_1}{dt}(0) = 0.$$  

We then find

$$y_0(t) = \sin t, \quad y_1(t) = -t \sin t, \quad \text{etc.}$$

which reproduces indeed expansion (4.38). The straightforward, Poincaré type, expansion (4.39) breaks down for large $t$, when $\varepsilon t \geq O(1)$. It is important to note that this caused by the fact that any $y_n$ is excited in its eigenfrequency (by the “source”-terms $-2dy_{n-1}/dt$), resulting in resonance. We recognise the generated algebraically growing terms of the type $t^n \sin t$ and $t^n \cos t$, called secular terms (definition 15.23). Apart from being of limited validity, the expansion reveals nothing of the real structure of the solution, and we change
our strategy to looking for an auxiliary function dependent on different time scales. We start with the hypothesis that, next to a fast time scale $t$, we have the slow time scale

$$T := \varepsilon t.$$  \hfill (4.40)

Then we identify the solution $y$ with a suitably chosen other function $Y$ that depends on both variables $t$ and $T$

$$Y(t, T; \varepsilon) := y(t; \varepsilon).$$  \hfill (4.41)

There exist infinitely many functions $Y(t, T; \varepsilon)$ that are equal to $y(t, \varepsilon)$ along the line $T = \varepsilon t$ in $(t, T)$-space. So we have now some freedom to prescribe additional conditions. With the unwelcome appearance of secular terms in mind it is natural to think of conditions, to be chosen such that no secular terms occur when we construct an approximation.

Since the time derivatives of $y$ turn into partial derivatives of $Y$, i.e.

$$\frac{dy}{dt} = \frac{\partial Y}{\partial t} + \varepsilon \frac{\partial Y}{\partial T},$$  \hfill (4.42)

equation (4.36) becomes for $Y$

$$\frac{\partial^2 Y}{\partial t^2} + Y + 2\varepsilon \left( \frac{\partial Y}{\partial t} + \frac{\partial^2 Y}{\partial t \partial T} \right) + \varepsilon^2 \left( \frac{\partial^2 Y}{\partial T^2} + 2 \frac{\partial Y}{\partial T} \right) = 0.$$  \hfill (4.43)

Assume the expansion

$$Y(t, T; \varepsilon) = Y_0(t, T) + \varepsilon Y_1(t, T) + \varepsilon^2 Y_2(t, T) + \cdots$$  \hfill (4.44)

and substitute this into (4.43) to obtain to leading orders

$$\frac{\partial^2 Y_0}{\partial t^2} + Y_0 = 0,$$

$$\frac{\partial^2 Y_1}{\partial t^2} + Y_1 = -2 \frac{\partial Y_0}{\partial t} - 2 \frac{\partial^2 Y_0}{\partial t \partial T},$$

with initial conditions

$$Y_0(0, 0) = 0, \quad \frac{\partial}{\partial t} Y_0(0, 0) = 1,$$

$$Y_1(0, 0) = 0, \quad \frac{\partial}{\partial t} Y_1(0, 0) = -\frac{\partial}{\partial T} Y_0(0, 0).$$

The solution for $Y_0$ is easily found to be

$$Y_0(t, T) = A_0(T) \sin(t + \theta_0(T)) \quad \text{with} \quad A_0(0) = 1, \theta_0(0) = 0.$$  \hfill (4.45)

This gives a right-hand side for the $Y_1$-equation of

$$-2 \left( A_0 + \frac{\partial A_0}{\partial T} \right) \cos(t + \theta_0) + 2 A_0 \frac{\partial \theta_0}{\partial T} \sin(t + \theta_0).$$

No secular terms occur (no resonance between $Y_1$ and $Y_0$) if these terms vanish:

$$A_0 + \frac{\partial A_0}{\partial T} = 0 \quad \text{yielding} \quad A_0 = e^{-T}, \quad \frac{\partial \theta_0}{\partial T} = 0 \quad \text{yielding} \quad \theta_0 = 0. \quad (4.46)$$

Together we have indeed constructed an approximation of (4.37), valid for $t \leq O(\epsilon^{-1})$.

$$y(t; \epsilon) = e^{-\epsilon t} \sin t + O(\epsilon). \quad (4.47)$$

Note (this is typical of this approach), that we determined $Y_0$ only on the level of $Y_1$, but without having to solve $Y_1$ itself.

The present approach is by and large the multiple scale technique in its simplest form. Variations on this theme are sometimes necessary. For example, we have not completely got rid of secular terms. On a longer time scale ($t = O(\epsilon^{-2})$) we have again resonance in $Y_2$ because of the “source” $e^{-T} \sin t$, yielding terms $O(\epsilon^2 t)$. We see that a second time scale $T_2 = \epsilon^2 t$ is necessary. From the exact solution we may infer that these longer time scales are not really independent and it may be worthwhile to try a fast time of strained coordinates type: $\tau = \omega(\epsilon)t = (1 + \epsilon^2 \omega_1 + \epsilon^4 \omega_4 + \ldots)t$. In the present example we would recover $\omega(\epsilon) = \sqrt{1 - \epsilon^2}$.

The method fails when the slow variation is due to external effects, like a slowly varying problem parameter, as is demonstrated by the next example.

**Example 15.29** Consider the problem

$$\ddot{x} + \kappa(\epsilon t)^2 x = 0, \quad x(0; \epsilon) = 1, \quad \dot{x}(0; \epsilon) = 0,$$

where $\kappa = O(1)$. It seems plausible to assume 2 time scales: a fast one $O(\kappa^{-1}) = O(1)$ and a slow one $O(\epsilon^{-1})$. So we introduce next to $t$ the slow scale $T = \epsilon t$, and rewrite $x(t; \epsilon) = X(t, T; \epsilon)$. We expand $X = X_0 + \epsilon X_1 + \ldots$, and obtain $X_0 = A_0(T) \cos(\kappa T - \theta_0(T))$. Suppressing secular terms in the equation for $X_1$ requires $A_0 = \kappa \tau - \theta'_0 = 0$, which is impossible.

Here, the fast time scale is slowly varying itself and the fast variable is to be strained locally by a suitable strain function, as follows

$$\tau = \int^t \omega(t'; \epsilon) \, dt' = \frac{1}{\epsilon} \int^T \omega(z; \epsilon) \, dz, \quad \text{where} \quad T = \epsilon t, \quad (4.48)$$

while for $x(t; \epsilon) = X(t, T; \epsilon)$ we have

$$\dot{x} = \omega X_T + \epsilon X_T \quad \text{and} \quad \ddot{x} = \omega^2 X_{TT} + \epsilon \omega X_{TT} + 2\epsilon \omega X_T + 2\epsilon^2 X_{TT} \quad (4.49)$$

**Example 15.30** Reconsider the problem 15.29. After expanding $X = X_0 + \epsilon X_1 + \ldots$ and $\omega = \omega_0 + \epsilon \omega_1 + \ldots$ we obtain

$$\omega^2_0 X_{0TT} + \kappa^2 X_0 = 0$$

and

$$\omega^2_0 X_{1TT} + \kappa^2 X_1 = -2\omega_0 \omega_1 X_{0TT} - \omega_0 \omega_T X_{0T} - 2\omega_0 X_{0T} \quad (*)$$

CHAPTER 15. PERTURBATION METHODS

The leading order solution is \( X_0 = A_0(T) \cos(\lambda(T) \tau - \theta_0(T)) \), where \( \lambda = \kappa/\omega_0 \). The right-hand side of (\( \ast \)) is then

\[
2 \omega_0 A_0 \lambda (\omega_1 \lambda + \lambda \tau - \theta_0) \cos(\lambda \tau - \theta_0) + (A_0 \lambda)^{-1} (\omega_0 A_0^4 \lambda) \sin(\lambda \tau - \theta_0).
\]

Suppression of secular terms requires \( \lambda \tau = 0 \). Without loss of generality we can take \( \lambda = 1 \), or \( \omega_0 = \kappa \). Then we need \( \omega_1 = \theta_0 \), which just yields that \( \lambda \tau - \theta_0 = \tau - \theta_0 = \varepsilon^{-1} \int^T \omega(z) \, dz - \int^T \omega_1(z) \, dz = \varepsilon^{-1} \int^T \omega_1(z) \, dz + O(\varepsilon) \). In other words, we may just as well take \( \omega_1 = 0 \) and \( \theta_0 = \) a constant. Finally we have \( \omega_0 A_0^4 \lambda^2 = \kappa A_0^2 = \) a constant. \( \Box \)

For linear wave-type problems we may anticipate the structure of the solution and assume the so-called WKB hypothesis (after Wentzel, Kramers and Brillouin)

\[
y(t; \varepsilon) = A(T; \varepsilon) e^{i \int_0^T \omega(\tau; \varepsilon) \, d\tau}.
\] (4.50)

The method is again illustrated by the example of the damped oscillator (4.36). After substitution and suppressing the exponential factor, we get

\[
(1 - \omega^2) A + i \varepsilon \left( \frac{\partial A}{\partial T} + \frac{\partial \omega}{\partial T} A + 2 \omega A \right) + \varepsilon^2 \left( \frac{\partial^2 A}{\partial T^2} + 2 \frac{\partial A}{\partial T} \right) = 0.
\]

Note that the secular terms are not explicitly suppressed now. The necessary additional condition here is that the solution of the present type \textit{exists} and that each higher order correction is no more secular than its predecessor. The solution is expanded as

\[
A(T; \varepsilon) = A_0(T) + \varepsilon A_1(T) + \varepsilon^2 A_2(T) + \cdots
\]

(4.51)

\[
\omega(T; \varepsilon) = \omega_0(T) + \varepsilon^2 \omega_2(T) + \cdots.
\]

Note that \( \omega_1 \) may be set to zero since the factor \( \exp(i \int_0^T \omega_1(\tau) \, d\tau) \) may be incorporated in \( A \). Substitute and collect equal powers of \( \varepsilon \)

\[
O(\varepsilon^0) : \quad (1 - \omega^2_0) A_0 = 0 \quad \rightarrow \quad \omega_0 = 1,
\]

\[
O(\varepsilon^1) : \quad \frac{\partial A_0}{\partial T} + A_0 = 0 \quad \rightarrow \quad A_0 = e^{-T},
\]

\[
O(\varepsilon^2) : \quad 2i \left( \frac{\partial A_1}{\partial T} + A_1 \right) = (1 + 2 \omega_2) e^{-T} \quad \rightarrow \quad \omega_2 = -\frac{1}{2}, \quad A_1 = 0.
\]

The solution that emerges is indeed consistent with the exact solution.

\textbf{Example 15.31 (The air-damped resonator).} In dimensionless form this is given by

\[
\frac{d^2 y}{d z^2} + \varepsilon \frac{dy}{dr} \left| \frac{dy}{dr} \right| + y = 0, \quad \text{with} \quad y(0) = 1, \quad \frac{dy(0)}{dr} = 0. \quad (\ast)
\]

By rewriting the equation into the form

\[
\frac{d}{dr} \left[ \frac{1}{2} (y')^2 + \frac{1}{4} y^2 \right] = -\varepsilon (y')^2 |y'|
\]

and assuming that $y$ and $y' = \mathcal{O}(1)$, it may be inferred that the damping acts on a time scale of $\mathcal{O}(\varepsilon^{-1})$. So we conjecture the presence of the slow time variable $T = \varepsilon t$ and introduce a new dependent variable $Y$ that depends on both $t$ and $T$. We have

$$T = \varepsilon t,$$

and obtain for equation (2)

$$\frac{\partial^2 Y}{\partial t^2} + Y + \varepsilon \left(2 \frac{\partial^2 Y}{\partial t \partial T} + \frac{\partial Y}{\partial T} \frac{\partial Y}{\partial t} \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$Y(0, 0; \varepsilon) = 1, \quad \left(\frac{\partial}{\partial T} + \varepsilon \frac{\partial}{\partial T}\right)Y(0, 0; \varepsilon) = 0.$$

The error of $\mathcal{O}(\varepsilon^2)$ results from the approximation $\frac{\partial}{\partial t} + \varepsilon \frac{\partial}{\partial T} = \frac{\partial}{\partial t} + \mathcal{O}(\varepsilon)$, and is of course only valid outside a small neighbourhood of the points where $\frac{\partial}{\partial t} Y = 0$. We expand

$$Y(t, T; \varepsilon) = Y_0(t, T) + \varepsilon Y_1(t, T) + \mathcal{O}(\varepsilon^2),$$

to find for the leading order

$$\frac{\partial^2 Y_0}{\partial t^2} + Y_0 = 0, \quad \text{with} \quad Y_0(0, 0) = 1, \quad \frac{\partial}{\partial t} Y_0(0, 0) = 0.$$

The solution is given by

$$Y_0 = A_0(T) \cos(t - \Theta_0(T)), \quad \text{where} \quad A_0(0) = 1, \quad \Theta_0(0) = 0.$$

For the first order we have the equation

$$\frac{\partial^2 Y_1}{\partial t^2} + Y_1 = -2 \frac{\partial^2 Y_0}{\partial t \partial T} - \frac{\partial Y_0}{\partial T} \frac{\partial Y_0}{\partial t}$$

$$= 2 \frac{dA_0}{dT} \sin(t - \Theta_0) - A_0 \frac{d\Theta_0}{dT} \cos(t - \Theta_0) + A_0^2 \sin(t - \Theta_0) \sin(t - \Theta_0),$$

with corresponding initial conditions. The secular terms are suppressed if the first harmonics of the right-hand side cancel. For this we use the Fourier series expansion

$$\sin(t) | \sin(t)| = \frac{8}{\pi} \sum_{n=0}^{\infty} \frac{\sin(2n + 1)t}{(2n + 1)(2n + 1)(2n + 3)},$$

We obtain the equations

$$2 \frac{dA_0}{dT} + \frac{8}{3\pi} A_0^2 = 0 \quad \text{and} \quad \frac{d\Theta_0}{dT} = 0,$$

with solution $\Theta_0(T) = 0$ and $A_0(T) = \frac{1}{1 + \frac{4}{3\pi} T}$.

Altogether we have the approximate solution

$$y(t; \varepsilon) = \frac{\cos(t)}{1 + \frac{4}{3\pi} \varepsilon t} + \mathcal{O}(\varepsilon).$$

This approximation appears to be remarkably accurate. See figure 15.4 where plots, made for a parameter value of $\varepsilon = 0.1$, of the approximate and a numerically “exact” solution are hardly distinguishable. An maximum difference is found of 0.03. \qed
CHAPTER 15. PERTURBATION METHODS

Figure 15.4. Plots of the approximate and a numerically “exact” solution \( y(t; \varepsilon) \) of the air-damped resonator problem for \( \varepsilon = 0.1 \).

Example 15.32 (Sound propagation in a slowly varying duct). Consider a hard-walled circular cylindrical duct with a slowly varying diameter described in polar coordinates \((x, r, \theta)\) as \( r = R(\varepsilon x) \) with \( \varepsilon \) a dimensionless small parameter. In this duct we have an acoustic medium with constant mean pressure and sound speed \( c \). Sound waves of circular frequency \( \omega \) are most conveniently written in complex form, so the physical pressure perturbation is given by \( \text{Re}(p(x) e^{i\omega t}) \), where \( p \) is the complex pressure amplitude described by Helmholtz’s equation

\[
\nabla^2 p + \kappa^2 p = 0. \quad (\ast)
\]

\( \kappa = \omega/c \) is the free-field wave number, and the boundary condition of a vanishing normal gradient at the wall yields

\[
\frac{\partial p}{\partial r} - \varepsilon R'(\varepsilon x) \frac{\partial p}{\partial x} = 0 \quad \text{at } r = R(\varepsilon x).
\]

For constant \( R \) and constant \( \kappa \) the general solution can be built up from a sum of right- and left-running modes of the following type

\[
p = \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{\mu=1}^{\infty} J_m(\alpha_{m\mu} r) \left( A_{m\mu} e^{-i\kappa_{m\mu} x} + B_{m\mu} e^{i\kappa_{m\mu} x} \right),
\]

\( \alpha_{m\mu} = j'_{m\mu}/R, \quad \kappa_{m\mu} = \kappa^2 - \alpha_{m\mu}^2, \quad \text{Re}(\kappa_{m\mu}) \geq 0, \quad \text{Im}(\kappa_{m\mu}) \leq 0. \)

\( J_m \) denotes the \( m \)-th order Bessel function [6] and \( J'_m(j'_{m\mu}) = 0 \). For the present problem we consider only a single mode and we assume, following the previous section, that the solution for the straight duct is locally close to the one for the varying duct. We introduce the slow variable \( X = \varepsilon x \) so that \( r = r(X) \), and we seek a solution of slowly varying modal type:

\[
p = A(X, r; \varepsilon) e^{-im\theta} e^{-i\gamma X} e^{-i\int_{X}^{x} \gamma(\xi; \varepsilon) d\xi}
\]

Since

\[
\frac{\partial^2 p}{\partial x^2} = \left( -\gamma^2 A - 2i\gamma \frac{\partial A}{\partial x} - i\gamma A' + \varepsilon^2 \frac{\partial^2 A}{\partial x^2} \right) \exp(\cdots)
\]

we have for (\ast)

\[
\left[ -\gamma^2 A - 2i\gamma \frac{\partial A}{\partial x} - i\gamma A' + \varepsilon^2 \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial r^2} + \frac{1}{r} \frac{\partial A}{\partial r} - \frac{m^2}{r^2} A + \kappa^2 A \right] \exp(\cdots) = 0.
\]

4. SINGULAR PERTURBATION PROBLEMS

After suppressing the exponential factor, this is up to order $\mathcal{O}(\varepsilon)$

$$\mathcal{L}[A] = \frac{i\varepsilon}{A} \frac{\partial}{\partial X} (\gamma A^2), \quad \frac{\partial A}{\partial r} + i\varepsilon R\gamma A = 0 \quad \text{at} \quad r = R(X). \tag{\dagger}$$

Here we introduced for short the Bessel-type operator

$$\mathcal{L}[A] := \frac{\partial^2 A}{\partial r^2} + \frac{1}{r} \frac{\partial A}{\partial r} + \left(\kappa^2 - \gamma^2 - \frac{m^2}{r^2}\right)A$$

and rewrote the right-hand side in a form that will turn out to be convenient later. Expand

$$A(X, r; \varepsilon) = A_0(X, r) + \varepsilon A_1(X, r) + \mathcal{O}(\varepsilon^2), \quad \gamma(X; \varepsilon) = \gamma_0(X) + \mathcal{O}(\varepsilon^2),$$

substitute in (\dagger), and collect like powers of $\varepsilon$, to obtain

$$\mathcal{O}(1) : \mathcal{L}[A_0] = 0, \quad \frac{\partial A_0}{\partial r} = 0 \quad \text{at} \quad r = R(X) \tag{\ddagger}$$

$$\mathcal{O}(\varepsilon) : \mathcal{L}[A_1] = \frac{i}{A_0} \frac{\partial}{\partial X} \gamma_0 A_0^2, \quad \frac{\partial A_1}{\partial r} = -i R\gamma_0 A_0 \quad \text{at} \quad r = R(X). \tag{\ddagger\ddagger}$$

Since the variable $X$ plays no other rôle in (\ddagger) than that of a parameter, we have for $A_0$ the “almost-mode”

$$A_0(X, r) = P_0(X) J_\nu(\alpha(X)r), \quad \alpha(X) = j_{m\mu}/R(X),$$

$$\gamma_0^2(X) = \kappa^2 - \alpha^2(X), \quad \text{Re}(\gamma_0) \geq 0, \quad \text{Im}(\gamma_0) \leq 0.$$ 

The amplitude $P_0$ is still undetermined and follows from a solvability condition for $A_1$. As before, amplitude $P_0$ is determined at the level of $A_1$, without $A_1$ necessarily being known. We multiply left- and right-hand side of (\ddagger\ddagger) with $rA_0$ and integrate to $r$ from 0 to $R(X)$. For the left-hand side we utilize the self-adjointness of $\mathcal{L}$.

$$\int_0^R rA_0 \mathcal{L}[A_1] \, dr = \int_0^R rA_0 \mathcal{L}[A_1] - r A_1 \mathcal{L}[A_0] \, dr$$

$$= \left[ rA_0 \frac{\partial A_1}{\partial r} - r A_1 \frac{\partial A_0}{\partial r} \right]_0^R = -i\gamma_0 RR' A_0^2.$$

For the right-hand side we apply Leibniz’s rule, i.e.

$$\int_0^R \frac{\partial}{\partial X} (i\gamma_0 A_0^2) r \, dr = \frac{d}{dX} \int_0^R i\gamma_0 A_0^2 r \, dr - i\gamma_0 RR' A_0^2.$$ 

Hence

$$\int_0^R r\gamma_0 A_0^2 \, dr = \left[ \frac{1}{2} \gamma_0 P_0^2 (r^2 - \frac{m^2}{\alpha^2}) J_\nu(\alpha r)^2 \right]_0^R$$

$$= \frac{1}{2} \gamma_0 P_0^2 R^2 \left( 1 - \frac{m^2}{j_{m\mu}^2} \right) J_\nu(j_{m\mu})^2 = C \quad \text{(where $C$ denotes any constant)}.$$ 

$$P_0(X) = \frac{C}{R(X) \sqrt{\gamma_0(X)}}. \quad \Box$$

In more dimensions, the assumed form of (4.50), where an integral occurs in the argument of the exponential, is not practical. In this case it is more convenient to write

$$\varphi(x, t; \varepsilon) := A(X, T; \varepsilon) \exp^{i \Omega(X, T; \varepsilon)},$$

(4.52)

while for clarity of notation we may introduce the slowly varying frequency and wave vector

$$\omega := \frac{\partial \Omega}{\partial T}, \quad \kappa := -\nabla \Omega,$$

where $\nabla := \frac{\partial}{\partial x} e_x + \frac{\partial}{\partial y} e_y + \frac{\partial}{\partial z} e_z$. Consider the following example of a one-dimensional wave equation with slowly varying coefficients.

$$\frac{\partial}{\partial t} \left( m(x, t) \frac{\partial}{\partial t} \varphi \right) = \frac{\partial}{\partial x} \left( C(x, t) \frac{\partial}{\partial x} \varphi \right) + B(x, t) \varphi,$$

(4.53)

where $X = \varepsilon x$ and $T = \varepsilon t$ are slow variables. We assume the solution $\varphi$ to take the form given by (4.52). This yields the equation

$$-\omega^2 mA + \frac{i \varepsilon}{A} \frac{\partial}{\partial T} (\omega m A^2) = -\kappa^2 CA - \frac{i \varepsilon}{A} \frac{\partial}{\partial X} (\kappa C A^2) + BA + \mathcal{O}(\varepsilon^2).$$

(4.54)

As before, we expand

$$A = A_0 + \varepsilon A_1 + \mathcal{O}(\varepsilon^2), \quad \Omega = \Omega_0 + \mathcal{O}(\varepsilon^2), \quad \omega = \omega_0 + \mathcal{O}(\varepsilon^2), \quad \kappa = \kappa_0 + \mathcal{O}(\varepsilon^2).$$

After substitution and collecting equal powers of $\varepsilon$, we get to leading order the slowly varying dispersion relation for $\omega_0$ and $\kappa_0$, or eikonal-type equation for $\Omega_0$

$$\omega_0^2 m = \kappa_0^2 C - B.$$

(4.55)

The next order yields a conservation-type equation for $A_0$

$$\frac{\partial}{\partial T} \left( \omega_0 m A_0^2 \right) + \frac{\partial}{\partial X} \left( \kappa_0 C A_0^2 \right) = 0.$$

(4.56)

(It should be noted that this result reflects the underlying physics, and therefore depends on the original equation. In general the resulting equation is not of conserved type.) The pair $\omega_0 m A_0^2$ and $\kappa_0 C A_0^2$ are called adiabatic invariants, because they correspond to the density and the flux of a quantity that is conserved, on the level of approximation, on the slow time and length scales. This is seen as follows. When we integrate (4.56) between the moving boundaries $X = X_1(T)$ and $X = X_2(T)$, we obtain

$$\int_{X_1}^{X_2} \frac{\partial}{\partial T} \left( \omega_0 m A_0^2 \right) + \frac{\partial}{\partial X} \left( \kappa_0 C A_0^2 \right) \, dX = \int_{X_1}^{X_2} \frac{d}{dT} \omega_0 m A_0^2 \, dX$$

$$- V_2 [\omega_0 m A_0^2]_{X_2} + V_1 [\omega_0 m A_0^2]_{X_1} + [\kappa_0 C A_0^2]_{X_2} - [\kappa_0 C A_0^2]_{X_1} = 0$$

where $V_1 = \frac{d}{dT} X_1$ and $V_2 = \frac{d}{dT} X_2$. This reduces to

$$\frac{d}{dT} \int_{X_1}^{X_2} \omega_0 m A_0^2 \, dX = 0.$$
if the velocity of either end point is equal to
\[ V = \frac{\kappa_0 C}{\omega_0 m} = \frac{\omega_0}{\kappa_0} \]  
(4.57)

In other words, \( \omega_0 m A_0^2 \) is conserved and propagates with group velocity \( V \) (see (3.4.8)) of waves that satisfy the slowly varying dispersion relation (4.55).

**Example 15.33 (Ray acoustics in a temperature gradient).** A very important application of waves in a slowly varying medium is the theory of rays. When a sound wave propagates in free space through a medium that varies on a much larger scale than the typical wave length (typically: temperature gradients, or wind with shear), the same ideas of multiple scales may be applied. In contrast to the duct, where the wave is confined by the duct walls, the waves may now freely refract and follow curved paths. These paths are called rays.

Consider an infinite 3D medium with varying temperature (typical length scale \( L \)) but otherwise with a constant mean pressure, so that we have for the acoustic pressure perturbations the equation (7.4.7), i.e.
\[ \frac{\partial^2}{\partial t^2} p = \nabla \cdot (c(X)^2 \nabla p), \]  
(\*)

where \( X = \varepsilon x \) for small \( \varepsilon \) and sound speed \( c \) is slowly varying when compared to the acoustic length scale. This means that \( \omega_0 \), the frequency considered, is high enough for the corresponding wave length \( \lambda \sim 2\pi c/\omega \) to be small compared to \( L \), so \( \varepsilon \sim \lambda/L \). Assuming the field to be locally plane we try an approximate solution
\[ p(x, t) = A e^{i\kappa \cdot x - \tau t/\varepsilon} \]  
(\dagger)

having the form of a plane wave but with slowly varying amplitude \( A = A(X; \varepsilon) \) and phase \( \varepsilon^{-1}(\omega_0 T - \tau(X; \varepsilon)) \), where \( T = \varepsilon t \). The surfaces \( \tau - \omega_0 T = \text{constant} \) describe the propagating wave front, with normal vectors \( \nabla \tau \). Define the operator \( \nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) \) so that \( \nabla = \varepsilon \nabla \). Define the local wave vector \( \nabla \tau = \kappa \). Since
\[ \nabla p = \left( -i \kappa A + \varepsilon \nabla A \right) e^{i\kappa \cdot x - \tau t/\varepsilon}, \]
\[ \nabla^2 p = \left( -|\kappa|^2 A - 2i \kappa \cdot \nabla A - i \varepsilon (\nabla \cdot \kappa) A + \varepsilon^2 \nabla^2 A \right) e^{i\kappa \cdot x - \tau t/\varepsilon}, \]

we obtain by substituting (\dagger) in (\*) the equation
\[ (\omega_0^2 - c^2|\kappa|^2)A = \frac{i\varepsilon}{A} \nabla \cdot (c^2 |\kappa| A) + O(\varepsilon^2). \]  
(\ddag)

Expand
\[ A = A_0 + \varepsilon A_1 + O(\varepsilon^2), \quad \tau = \tau_0 + O(\varepsilon^2), \quad \kappa = \kappa_0 + O(\varepsilon^2), \]
and collect like powers in (\ddag). We find to leading orders for \( \kappa_0 \) and \( A_0 \):
\[ c^2|\kappa_0|^2 = \omega_0^2, \]  
(2)
\[ \nabla \cdot (c^2 \kappa_0 A_0) = 0. \]  
(\ddag)

Written in \( \tau_0 \), equation (2) is called the *eikonal equation*, which determines the wave fronts and the ray paths. Equation (\ddag) is called the *transport equation* and describes the conservation of wave action, which is here equivalent to conservation of energy. It relates the amplitude variation to diverging or converging rays (see the problem considered in section ?? of chapter ??). The eikonal equation is a nonlinear first order partial differential equation, of hyperbolic type, which can always be reduced to a system of ordinary differential equations along characteristics (see Chapter 2).
CHAPTER 15. PERTURBATION METHODS

5 Discussion

- Inherent in any modelling is the hierarchy in importance of the various effects that constitute the model. Therefore, certain effects in any modelling will be small. Sometimes small but not small enough to be ignored, and sometimes small but in a non-uniform way such that they are important locally.

- For an efficient solution, and to obtain qualitative insight, it makes sense to utilize this "smallness". Methods that systematically exploit such intrinsic smallness are called "perturbation methods".

- The basis of a systematic approach is formed by the concept of asymptotic approximations and expansions of Poincaré type. If the approximation is uniform on the domain of interest it is called regular and singular otherwise. The rôle of the chosen independent variable in this respect cannot be overemphasized.

- Perturbation methods have a long history. Before the time of the numerical methods and the computer, perturbation methods were the only way to increase the applicability of available exact solutions to difficult, otherwise intractable problems. Nowadays, perturbation methods have their use as a natural step in the process of systematic modelling, the insight it provides in the nature of singularities occurring in the problem and typical parameter dependencies, and sometimes the speed of practical calculations.

Exercises

15.1. Derive asymptotic solutions (for $\varepsilon \to 0$) of the equation

$$\varepsilon x^3 - x + 2 = 0.$$ 

15.2. Derive step by step, by iteratively scaling $x(\varepsilon) = \mu_0(\varepsilon)x_0 + \mu_1(\varepsilon)x_1 + \ldots$ and balancing, that a third order asymptotic solution (for $\varepsilon \to 0$) of the equation

$$\log(\varepsilon x) + x = a,$$

is given by

$$x(\varepsilon) = \log \varepsilon^{-1} - \log(\log \varepsilon^{-1}) + a + o(1).$$

Find a more efficient expansion based on an alternative asymptotic sequence of gauge functions by combining $e^{-\alpha\varepsilon}$.

15.3. Derive the so-called Webster’s equation for sound of long wave length propagating in slowly varying horns, by the method of slender approximation [71]. The reduced wave equation for pressure perturbations $p$ and wavenumber $k$ is given by

$$\nabla^2 p + k^2 p = 0,$$
Exercises

within a duct given in cylindrical polar coordinates by

\[ S = r - R(X, \theta) = 0, \quad X = \varepsilon x, \quad \varepsilon \text{ is small.} \]

The wave number is \( \mathcal{O}(\varepsilon) \), so we scale \( k = \varepsilon k \). The duct wall is hard, so we have the boundary condition

\[ \nabla p \cdot \nabla S = 0 \text{ at } S = 0. \]

15.4. Consider the incompressible Navier-Stokes equations to describe lubrication flow in a two-dimensional narrow and slowly varying channel, with prescribed volume flux. (In actual practice this flux is created by a pressure difference.)

(a) Make dimensionless on the channel height and volume flux, and scale the pressure gradient such that viscous forces are balanced by the pressure gradient, so the Reynolds number \( \text{Re} \leq \mathcal{O}(1) \). Verify that we obtain in dimensionless form

\[ \text{Re}(v \cdot \nabla v) + \nabla p = \nabla^2 v, \quad \nabla \cdot v = 0 \]

for the velocity \( v = (u, v)^T \) and pressure \( p \) in the channel given by \(-\infty < x < \infty \) and \( g(\varepsilon x) \leq y \leq h(\varepsilon x) \) where \( \varepsilon \) is a small parameter. (End conditions in \( x \) are not important.) Boundary conditions are: no slip at the walls, i.e. \( u = v = 0 \) at \( y = g(\varepsilon x) \) and \( y = h(\varepsilon x) \), and a flux

\[ \int_{g(\varepsilon x)}^{h(\varepsilon x)} u(x, y) \, dy = 1. \]

(b) We rewrite \( X = \varepsilon x \) and assume that the field varies slowly in \( X \) (any end-effects are local and irrelevant for the \( x \)'s considered). Rescale \( u, v, p \). The order of magnitude of the pressure can be found from the observation that a pressure gradient is necessary to have a flow. The crosswise velocity \( v \) will be much smaller than the axial velocity \( u \).

(c) Assume for rescaled \( u, v, p \) an obvious asymptotic expansion in \( \varepsilon \), and solve up to leading order.

15.5. Reconsider equation (†) of Example 7.4 to describe a stationary suspended flexible bar of length \( L \).

(a) First we consider a cable with clamped ends at equal height. This is described by adding boundary conditions \( x = y = \phi = 0 \) at \( s = 0 \) and \( x = D, y = \phi = 0 \) at \( s = L \), where \( 1 - D/L \) is positive and not small. Note that for given \( D \), the necessary horizontal force \( H \) is unknown and to be determined. Make the problem dimensionless by scaling lengths on \( L \) and forces on \( QL \). Introduce the parameters \( \varepsilon = \sqrt{(EI/QL^3)} \), and \( h = h(\varepsilon) = H_0/QL \) and \( v = v(\varepsilon) = V_0/QL \). If \( \varepsilon \) is small the equation describes a suspended cable. Find the solution asymptotically to leading order for small \( \varepsilon \).

(b) Do the same for a cable with hinged ends, i.e. with \( \phi' (0) = \phi(L)' = 0 \).

(c) The same differential equation represents a model for laying submarine gas and oil pipelines from a laybarge. The pipe is freely suspended over an unknown length \( L \), with prescribed curvature \( R \) at the lift-off point at height \( y = W \) and a prescribed horizontal tension \( H \) in order to avoid buckling of the pipe. Both the
angle $\phi$ and the curvature $\phi'$ vanish at the touch-down point $s = 0$. We have thus $\phi(0) = \phi'(0) = 0$, $\phi'(L) = -R^{-1}$, $y(0) = 0$, $y(L) = W$, while $L$ is unknown. Make dimensionless and solve the resulting problem asymptotically for small $\varepsilon$.

15.6. Determine the asymptotic approximation of solution $y(x; \varepsilon)$ (1st or 1st and 2nd order terms for positive small parameter $\varepsilon \to 0$) of the following singularly perturbed problems. Let $\alpha$ and $\beta$ be non-zero constants, independent of $\varepsilon$. Provide arguments for the determined boundary layer thickness and location, and show how free constants are determined by the matching procedure.

(a) $\varepsilon y'' - y' = 2x$, $y(0; \varepsilon) = \alpha$, $y(1; \varepsilon) = \beta$.
(b) $\varepsilon y' + y^2 = \cos x$, $y(0; \varepsilon) = 0$, $0 \leq x \leq 1$.
(c) $\varepsilon y'' + (2x + 1)y' + y^2 = 0$, $y(0; \varepsilon) = \alpha$, $y(1; \varepsilon) = \beta$.
(d) $\varepsilon y'' + x(y' - y) = 0$, $y(0; \varepsilon) = 0$, $y(1; \varepsilon) = e$.

15.7. When we stir a cup of tea, the surface of the fluid deforms until equilibrium is attained between gravity, centrifugal force and surface tension. This last force is only important near the wall. Consider for this problem the following model.

A cylinder (radius $a$, axis vertically) with fluid (density $\rho$, surface tension $\sigma$) rotates around its axis $e_z$ (angular velocity $\Omega_1$) in a gravity field $-g e_z$. By the gravity and the centrifugal force the surface deforms to something that looks like a paraboloid. Within a small neighbourhood of the cylinder wall the contact angle $\alpha$ is felt by means of the surface tension (see Eq. (8.14)). Because of symmetry we can describe the surface by a radial tangent angle $\psi$ with the horizon, parametrized by arc length $s$, such that $s = 0$ corresponds with the axis, and $s = L$ with the wall of the cylinder. $L$ is unknown. Select the origin on the axis at the surface, such that the vertical and radial coordinate are given by $Z(s) = \int_0^s \sin \psi(s') \, ds'$ and $R(s) = \int_0^s \cos \psi(s') \, ds'$.

The balance between hydrostatic pressure and surface tension yields the equation

$$p_0 - \rho g Z + \frac{1}{2} \rho \Omega^2 R^2 = -\sigma \left( \frac{d\psi}{ds} + \frac{\sin \psi}{R} \right)$$

with unknown $p_0$. Boundary conditions are $\psi(0) = 0$, $\psi(L) = \alpha$, and $R(L) = a$. 

Exercises

(a) Scale lengths on a: \( s = at, \ R = ar, \ Z = az, \ L = a\lambda \), and introduce \( \beta = p_0/\rho ga \), and the dimensionless parameters \( \varepsilon^2 = \sigma/\rho ga^2 \), and \( \mu = \Omega^2 a/g \). Can you identify common names of these dimensionless numbers.

(b) Solve the resulting problem asymptotically for \( \varepsilon \to 0 \), while \( \mu = O(1) \).

15.8. Consider the \textit{van der Pol} equation for variable \( y = y(t; \varepsilon) \) in \( t \) and small parameter \( \varepsilon \)

\[ y'' + y - \varepsilon (1 - y^2) y' = 0. \]

Construct by using the Lindstedt-Poincaré method a first order approximation of a periodic solution.

15.9. Following Example 15.32, derive a multiple scales solution of sound waves in a slowly varying duct while also the sound speed is a slowly varying function of \( x \). The pertaining equation is therefore Eq. (\ast) of Example 15.33.

15.10. (a) Rewrite the eikonal equation (\sharp) of Example 15.33 in characteristic form by using Theorem (12.6).

(b) Prove that in a medium with a linearly varying sound speed the path of rays are circles.

15.11. Analyse the error \( u = y - \tilde{y} \) of Eq. (3.4.13), described by Eq. (3.4.14), but now including the effects of the slowly varying coefficients \( a, b, \) and \( c \). Formulate your result in the form adiabatic invariants.

15.12. Derive an approximate solution for large \( c \) of the Fisher travelling wave problem (10.6.70)

\[ U'' + c^2 U' + c^2 U(1 - U) = 0, \]

(a) on \( (-\infty, \infty) \) with \( U(-\infty) = 1, \ U(\infty) = 0 \). It is no restriction to assume that \( U(0) = \frac{1}{2} \).

(b) on \([0, \infty)\) and \( U(0) = 0 \), while the previous solution is the outer solution.
Appendix

Useful definitions and properties

A  Asymptotic order symbols

A function $f(\varepsilon)$ may be expressed asymptotically for small $\varepsilon$ by another (usually simpler) function $\varphi(\varepsilon)$ as follows.

Definition A.1.

1. $f(\varepsilon) = O(\varphi(\varepsilon))$ as $\varepsilon \to 0$ if there are positive constants $K$ and $\varepsilon_1$ (both independent of $\varepsilon$) such that

   $$|f(\varepsilon)| \leq K|\varphi(\varepsilon)|$$

   for $0 < \varepsilon < \varepsilon_1$.

   We say: “$f$ is big-O of $\varphi$ as $\varepsilon$ tends to zero.”.

2. $f(\varepsilon) = o(\varphi(\varepsilon))$ as $\varepsilon \to 0$ if for every $\delta > 0$, there is an $\varepsilon_1$ (independent of $\varepsilon$) such that

   $$|f(\varepsilon)| \leq \delta|\varphi(\varepsilon)|$$

   for $0 < \varepsilon < \varepsilon_1$.

   We say: “$f$ is little-o of $\varphi$ as $\varepsilon$ tends to zero.”.

3. $f(\varepsilon) = O_s(\varphi(\varepsilon))$ as $\varepsilon \to 0$ if $f(\varepsilon) \neq o(\varphi(\varepsilon))$. Note that quite often $O$ is used where $O_s$ is actually meant. Further, there is no uniform terminology. We could say: “$f$ is big-O sharp of $\varphi$ as $\varepsilon$ tends to zero”.

Theorem A.2.

1. If $f(\varepsilon) = o(\varphi(\varepsilon))$ as $\varepsilon \to 0$, then also $f(\varepsilon) = O(\varphi(\varepsilon))$.

2. If the limit $\lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{\varphi(\varepsilon)}$ exists as a finite number, then $f(\varepsilon) = O(\varphi(\varepsilon))$ as $\varepsilon \to 0$. 

3. If the limit \( \lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{\varphi(\varepsilon)} \) exists as a finite number \( \neq 0 \), then \( f(\varepsilon) = \mathcal{O}_s(\varphi(\varepsilon)) \) as \( \varepsilon \to 0 \).

4. If \( \lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{\varphi(\varepsilon)} = 0 \), then \( f(\varepsilon) = o(\varphi(\varepsilon)) \) as \( \varepsilon \to 0 \).

**Proof.** Trivial. \( \square \)

**Example A.3**

\[
\begin{align*}
\varepsilon \sin(\varepsilon) &= \mathcal{O}_s(\varepsilon^2), \quad \varepsilon \to 0, \\
\varepsilon \cos(\varepsilon) &= \mathcal{O}(1), \quad \varepsilon \to 0, \\
\varepsilon^n &= o(1), \quad \varepsilon \to 0, \text{ for any positive } n, \\
e^{-1/\varepsilon} &= o(\varepsilon^n), \quad \varepsilon \to 0, \text{ for any positive } n.
\end{align*}
\]

From this last example, \( e^{-1/\varepsilon} \) is called a transcendentally (TST) or exponentially small term (EST) and can be ignored asymptotically against any power of \( \varepsilon \). \( \square \)

### B Trigonometric relations

The real or imaginary parts of the binomial series \((e^{ix} \pm e^{-ix})^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k e^{i(n-2k)x}\)

easily yield trigonometric relations, useful for recognising resonance terms:

\[
\begin{align*}
\sin^2x &= \frac{1}{2}(1 - \cos 2x), & \cos^2x \sin^2x &= \frac{1}{4}(1 - \cos 4x), \\
\sin x \cos x &= \frac{1}{2} \sin 2x, & \sin x \cos^2x &= \frac{1}{4}(2 \sin 2x + \sin 4x). \\
\cos^2x &= \frac{1}{2}(1 + \cos 2x), & \cos x \cos^2x &= \frac{1}{4}(3 + 4 \cos 2x + \cos 4x), \\
\sin^3x &= \frac{1}{4}(3 \sin x - \sin 3x), & \sin^5x &= \frac{1}{16}(10 \sin x - 5 \sin 3x + \sin 5x), \\
\sin^2x \cos x &= \frac{1}{4}(\cos x - \cos 3x), & \sin^4x \cos x &= \frac{1}{16}(2 \cos x - 3 \cos 3x + \cos 5x), \\
\sin x \cos^2x &= \frac{1}{4}(\sin x + \sin 3x), & \sin^3x \cos^2x &= \frac{1}{16}(2 \sin x + \sin 3x - \sin 5x), \\
\cos^3x &= \frac{1}{4}(3 \cos x + \cos 3x), & \sin^3x \cos^3x &= \frac{1}{16}(2 \cos x - \cos 3x - \cos 5x), \\
\sin^4x &= \frac{1}{8}(3 - 4 \cos 2x + \cos 4x), & \sin x \cos^4x &= \frac{1}{16}(2 \sin x + 3 \sin 3x + \sin 5x), \\
\sin^3x \cos x &= \frac{1}{8}(2 \sin 2x - \sin 4x), & \cos^5x &= \frac{1}{16}(10 \cos x + 5 \cos 3x + \cos 5x).
\end{align*}
\]
C Convergence of series

The series
\[ S(x) := \sum_{n=0}^{\infty} c_n f_n(x), \quad x \in \Omega, \quad (C.1) \]
is said to converge pointwise in \( x \in \Omega \) if we can find for any given \( \varepsilon > 0 \) a sufficiently large number \( N \in \mathbb{N} \), such that the remaining part of the series is smaller than \( \varepsilon \), i.e.
\[ \| S(x) - \sum_{n=0}^{N} c_n f_n(x) \| < \varepsilon. \]

In general, \( N \) depends on \( \varepsilon \) and \( x \). If \( N \) may be chosen independent of any \( x \in \Omega \), the series is said to converge uniformly in \( \Omega \). A sufficient condition for uniform convergence is Weierstrass’s M-test. If
\[ \| c_n f_n(x) \| \leq M_n \quad \text{for all } x \in \Omega \]
and
\[ \sum_{n=0}^{\infty} M_n \]
converges, then (C.1) is uniformly convergent.

The concept of uniform convergence is important because of the following properties. If (C.1) is uniformly convergent and
(i) all \( f_n(x) \) are continuous, then \( S(x) \) is continuous and
\[ \lim_{x \to a} S(x) = \sum_{n=0}^{\infty} c_n \lim_{x \to a} f_n(x). \]
(ii) all \( f_n(x) \) are continuous, then \( S(x) \) is integrable and
\[ \int S(x) \, dx = \sum_{n=0}^{\infty} c_n \int f_n(x) \, dx. \]
(iii) all \( f_n(x) \) are differentiable, then \( S(x) \) is differentiable and
\[ \frac{d}{dx} S(x) = \sum_{n=0}^{\infty} c_n \frac{d}{dx} f_n(x). \]
D Multistep formulae

If the grid points $\xi_i$ are equispaced, a simple relation exists between interpolating polynomials and backward difference operators. They in turn give explicit expressions for the coefficients of multistep methods.

The backward difference operators $\nabla^k$ for $k = 0, 1, \ldots$ are defined for a set $f_1, \ldots, f_m$ by

$$\nabla^0 f_j = f_j \quad \nabla^l f_j = \nabla^{l-1} f_j - \nabla^{l-1} f_{j-1}, \quad l = 1, 2, \ldots.$$  \hfill (D.1)

So we have e.g.

$$\nabla^1 f_j = f_j - f_{j-1} \quad \nabla^2 f_j = (f_j - f_{j-1}) - (f_{j-1} - f_{j-2}) = f_j - 2f_{j-1} + f_{j-2}.$$  \hfill (D.2)

From this we find

$$f_{m-1} = \sum_{j=0}^{m-1} (-1)^j \binom{l}{j} \nabla^j f_m.$$  \hfill (D.3)

The backward difference notation is the same as the nabla operator. However, no confusion should occur as we only use backward difference operators in this appendix D.

Let $\xi_1, \ldots, \xi_m$ be equispaced grid points with grid size $h$ and let $f_j$ be a function value at the point $\xi_j$. Then we can rewrite the interpolation polynomial in terms of the $\nabla^j f_m$

$$p(\xi) = \sum_{j=0}^{m-1} (-1)^j \binom{l}{j} \nabla^j f_m,$$  \hfill (D.3)

where $s := \frac{\xi - \xi_m}{h}$.

Now consider the ODE

$$\frac{dx}{dt} = f(x, t).$$

We like to find an approximation of the solution $x(t_j)$ on a set of equispaced grid points $t_0, t_1, \ldots, t_i, t_{i+1}, \ldots$ At the point $t_{i+1}$ we can easily find an approximation of $\frac{dx}{dt}(t_{i+1})$ at from (D.3) by differentiation. As a result we find, identifying the points $\xi_j$ with $t_{i-m+1+j}$ and denoting the numerical approximation of $x(t_{i+1})$ by $x_{i+1}$

$$\sum_{j=0}^{m-1} \gamma_j \nabla^j x_{i+1} = hf(x_{i+1}, t_{i+1}).$$

The coefficients in (D.4a) have a nice form. One can check that

$$\gamma_0 = 0, \quad \gamma_j = \frac{1}{j}, \quad j \geq 0.$$  \hfill (D.4b)
APPENDIX . USEFUL DEFINITIONS AND PROPERTIES

The formulae found in (D.4) are the Backward Difference Formulae (BDF). In standard form (cf. (? ? ? ?)) they are given by

\[ \sum_{j=0}^{k} a_j x_{i-j+1} \doteq hf(t_{i+1},x_{i+1}), \quad (D.5) \]

For \( k = 1, 2, 3, 4 \) the coefficients \( a_j \) are given in table .1. Note that for \( k = 1 \) we have Euler backward. Also for other multistep methods the formulation in terms of backward differences enables one to find the coefficients in a simple way

\[ x(t_{i+1}) - x(t_i) = \int_{t_i}^{t_{i+1}} f(x(\tau), \tau) \, d\tau. \quad (D.6) \]

If we apply (D.3) on the interval \( (t_{i-k+1}, t_i) \), i.e. approximate \( f(x(t), t) \) by such a polynomial \( p \) of degree \( k - 1 \) there, we obtain

\[ x(t_{i+1}) - x(t_i) \doteq \sum_{j=0}^{k-1} \hat{\gamma}_j \nabla^j f_i, \quad (D.7a) \]

where

\[ \hat{\gamma}_j := (-1)^j \int_{0}^{1} \tau^j \left( \begin{array}{c} \tau \\ j \end{array} \right) d\tau. \quad (D.7b) \]

This is the \( k \)-step Adams-Bashforth formula. The \( \hat{\gamma}_j \) are simply calculated by recursion,

\[ \hat{\gamma}_0 = 1 \]

\[ \hat{\gamma}_1 = 1 - \frac{1}{2} \hat{\gamma}_0, \quad \hat{\gamma}_j = 1 - \frac{1}{j+1} \hat{\gamma}_0 - \frac{1}{j} \hat{\gamma}_1 - \cdots - \frac{1}{2} \hat{\gamma}_{j-1}, \quad j \geq 1. \quad (D.7c) \]

In standard form we would have

\[ x(t_{i+1}) - x(t_i) \doteq h \sum_{j=1}^{k} b_j f_{t_{i-j+1}}. \quad (D.8) \]

For \( k = 1, 2, 3, 4 \) the coefficients are given in table .2. If we apply (D.3) on the interval

\[ 17:22 \quad 7 \text{ Oct } 2003 \quad 271 \quad \text{version: 26-09-2003} \]
\[ (t_{i-k+1}, t_{i+1}), \text{i.e. approximate } \rho \text{ by a polynomial of degree } k \text{ there, we obtain} \]

\[ x(t_{i+1}) - x(t_i) = \sum_{j=0}^{k} \bar{\gamma}_j \nabla^j f_i, \quad (D.9a) \]

where

\[ \bar{\gamma}_j = (-1)^j \int_{-1}^{0} \left( -\tau \right)^j \frac{d^j}{d\tau^j} \quad (D.9b) \]

This is the \( k \)-step Adams-Moulton formula.

From this we find

\[ \bar{\gamma}_0 = 1 \]

\[ \bar{\gamma}_j = -\frac{1}{j+1} \bar{\gamma}_0 - \frac{1}{j} \bar{\gamma}_1 - \cdots - \frac{1}{2} \bar{\gamma}_{j-1}, \quad j \geq 1. \quad (D.9c) \]

The coefficients \( \hat{\gamma}_j \) and the \( \bar{\gamma}_j \) are independent of \( k \). They are related by

\[ \sum_{j=0}^{k} \bar{\gamma}_j = \hat{\gamma}_k. \quad (D.10) \]

We can write the Adams-Moulton formulae in the standard formulation as

\[ x^h_{i+1} - x^h_i = h \sum_{j=0}^{k} \beta_j f_{i-j+1} \quad (D.11) \]

For \( k = 0, 1, 2, 3 \) the coefficients \( b_j \) are given in table .3. Note that for \( k = 0 \) we obtain Euler backward and for \( k = 1 \) we find the trapezoidal formula (both one-step!).

---

### Table .2.

<table>
<thead>
<tr>
<th>( k )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( b_3 )</th>
<th>( b_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( \frac{3}{2} )</td>
<td>-( \frac{1}{2} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( \frac{23}{12} )</td>
<td>-( \frac{16}{12} )</td>
<td>( \frac{5}{12} )</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( \frac{25}{24} )</td>
<td>-( \frac{19}{24} )</td>
<td>( \frac{37}{24} )</td>
<td>-( \frac{9}{24} )</td>
</tr>
</tbody>
</table>
### E Solution of recursions

Consider the recursion

\[ x_{i+1} = A_i x_i + b_i, \quad (E.1) \]

where \( \{A_i\} \) is a set of square matrices (in particular it may be scalars) and \( \{b_i\} \) a set of vectors. Then the solution of (E.1) is given by

\[ x_i = \left( \prod_{j=0}^{i-1} A_j \right) x_0 + \sum_{m=0}^{i-1} \left( \prod_{j=m+1}^{i-1} A_j \right) b_j. \quad (E.2) \]

In (E.2) one has to interpret \( \prod_{i-1}^{m+1} A_j \) as \( A_i - 1 A_{i-2} \ldots A_{m+1} \) and as \( I \) if \( i - 1 < m + 1 \).

Now consider the second order scalar recursion

\[ x_{i+1} = a_i x_i + b_i x_{i-1} + c_i. \quad (E.3) \]

If \( c_i \equiv 0 \) and \( a_i \) en \( b_i \) are constant, then the solution is given by

\[ x_i = \alpha t_{i}^{\alpha} + \beta t_{i}^{\beta}, \quad \alpha, \beta \in \mathbb{R}, \quad (E.4) \]

if the characteristic equation

\[ \lambda^2 - a\lambda - b = 0 \quad (N.B. \forall a_i = a, \forall b_i = b), \quad (E.5) \]

has two different roots, \( \lambda_1, \lambda_2 \). If it has a double root, the solution can be written as

\[ x_i = (\alpha + \beta i)\lambda_{i}^{\lambda_{1}}. \quad (E.6) \]

The constants in (E.4) and (E.6) are to be found from initial or boundary values.

In general the homogeneous part of (E.3) has two independent basis solutions \( \{f_i\} \) en \( \{g_i\} \).

In terms of these we can give a formal solution of (E.3). We find

\[ x_i = \sum_{j=1}^{i} \left[ \frac{f_j-1 f_j - g_j-1 f_j}{f_j-1 g_j - g_j-1 f_j} c_j \right] + \alpha f_i + \beta g_i, \quad (E.7) \]

where \( \alpha \) en \( \beta \) are to be determined from initial or boundary conditions. In particular for constant coefficients and two different roots \( \lambda_1, \lambda_2 \) we obtain

\[ x_i = \sum_{j=1}^{i} c_j \frac{\lambda_{i-j} - \lambda_{i-j}}{\lambda_2 - \lambda_1} \alpha \lambda_j + \beta \lambda_j. \quad (E.8) \]
F EIGENVALUES AND EIGENVECTORS OF A TRIDIAGONAL MATRIX

Consider the matrix \( A \in \mathbb{R}^{n \times n} \)

\[
A := \begin{pmatrix}
  b & c & \emptyset \\
  a & b & c & \emptyset \\
  \quad & \ddots & \ddots & \ddots \\
  \emptyset & \quad & a & b
\end{pmatrix}.
\] (F.1)

Let \( \mu \) be an eigenvalue of \( A \) and \( x = (x_1, \ldots, x_N)^T \) a corresponding eigenvector. Then we find from \( Ax = \mu x \)

\[
(b - \mu)x_1 + cx_2 = 0 \\
ax_j + (b - \mu)x_j + cx_{j+1} = 0, \quad j = 2, \ldots, N - 1 \\
ax_{N-1} + (b - \mu)x_N = 0.
\] (F.2a, b, c)

Let us take \( x_0 = x_{N+1} = 0 \). Then we find from (F.2) that \( \{x_j\}_{j=0}^{N+1} \) is a solution of the three-term recurrence equation

\[
ax_{j-1} + (b - \mu)x_j + cx_{j+1} = 0, \quad j = 1, \ldots, N,
\] (F.3)

with boundary values

\[
x_0 = x_{N+1} = 0
\] (F.4)

The recursion (F.3) has the characteristic polynomial

\[
\lambda^2 + (b - \mu)\lambda + c = 0,
\] (F.5)

having roots \( \lambda_1 \) and \( \lambda_2 \) say, so that a solution is formally given by

\[
x_j = \alpha \lambda_1^j + \beta \lambda_2^j,
\] (F.6)

where \( \alpha \) and \( \beta \) can be found from the boundary values (F.4). This gives \( \beta = -\alpha \) and eventually the relation

\[
\left( \frac{\lambda_1}{\lambda_2} \right)^{N+1} = 1,
\] (F.7)

i.e. \( \frac{\lambda_1}{\lambda_2} \) is an \((N + 1)\)st power unit root. Hence

\[
\frac{\lambda_1}{\lambda_2} = e^{\frac{2\pi i l}{N+1}}, \quad l = 1, \ldots, N.
\] (F.8)

Since the product of the roots equals \( c \), we obtain

\[
\lambda_1 = \left( \frac{c}{a} \right)^{1/2} e^{\frac{\pi i l}{N+1}}, \quad l = 1, \ldots, N
\] (F.9a)

\[
\lambda_2 = \left( \frac{c}{a} \right)^{1/2} e^{-\frac{\pi i l}{N+1}}.
\] (F.9b)

17:22 7 Oct 2003 274 version: 26-09-2003
We now use the fact that \( \lambda_1 + \lambda_2 = -(b - \mu)/a \), so that (F.9) gives for the eigenvalues \( \mu_l \)
\[
\mu_l = b + \sqrt{ac} \left( e^{\frac{\mu_l}{N + 1}} + e^{-\frac{\mu_l}{N + 1}} \right) = b + 2\sqrt{ac} \cos \frac{l\pi}{N + 1}, \quad l = 1, \ldots, N. \quad (\text{F.10})
\]
For the \( j \)th component \( x_j^l \) of the corresponding eigenvector, \( \mathbf{x}_l \), we apparently have
\[
x_j^l = \alpha (\lambda_1^l - \lambda_2^l) = \alpha \left( \frac{c}{a} \right)^{i/2} \left( e^{\frac{\mu_l}{N + 1}} + e^{-\frac{\mu_l}{N + 1}} \right) = 2i\alpha \left( \frac{c}{a} \right)^{i/2} \sin \frac{\pi jl}{N + 1}. \quad (\text{F.11})
\]
Note that \( \alpha \) can be chosen arbitrarily, so in particular, if \( \text{sign}(a) = \text{sign}(b) \), we may assume \( x_j \) to be real for all \( l \) and \( j \). If, moreover, \( a = c \) we can choose \( \alpha \) such that \( \sum x_j^2 = 1 \), resulting in an orthogonal set of eigenvectors, with
\[
x_j^l = \frac{2}{\sqrt{N}} \sin \frac{\pi jl}{N + 1}. \quad (\text{F.12})
\]

**Remark 1.** We still have to check whether the representation (F.6) is correct. Indeed, if \( \lambda_1 = \lambda_2 \) (so we have a double root) the foregoing is not correct. This situation cannot occur, however. Suppose we would have \( \lambda_1 = \lambda_2 \), then we would have in stead of (F.6),
\[
x_j = (\alpha + \beta j) t_j^l. \quad (\text{F.13})
\]
From (F.4) we then immediately see that \( \alpha = \beta = 0 \), which is not interesting, of course.

**Remark 2.** We can can find out more generally whether a tridiagonal matrix has (geometrically) multiple eigenvalues and real eigenvectors (so being diagonisable) or not. Indeed consider in stead of (F.1) the matrix
\[
\mathbf{A} = \begin{pmatrix}
b_1 & c_1 & 0 & \cdots & 0 \\
a_2 & b_2 & c_2 & & 0 \\
& \ddots & \ddots & \ddots & \ddots \\
0 & & c_{N-1} & \cdots & 0 \\
0 & & 0 & & a_N & b_N
\end{pmatrix}. \quad (\text{F.14})
\]
Let \( \mathbf{D} \) be a diagonal matrix \( \mathbf{D} = \text{diag}(d_1, \ldots, d_N) \), with \( d_1, \ldots, d_N \) still to be chosen. Define
\[
\tilde{\mathbf{A}} = \mathbf{D} \mathbf{A} \mathbf{D}^{-1}. \quad (\text{F.15})
\]
We now require \( \mathbf{D} \) to be such that \( \tilde{\mathbf{A}} \) is a symmetric matrix. This is so if
\[
\left( \frac{d_{j+1}}{d_j} \right)^2 = \frac{c_j}{a_{j+1}}, \quad 1 \leq j \leq N - 1. \quad (\text{F.16})
\]
Hence if e.g. \( \text{sign}(c_j) = \text{sign}(a_{j+1}) \), such a matrix \( \mathbf{D} \) certainly exists. Apparently \( \tilde{\mathbf{A}} \) has simple eigenvalues. Due to the fact that \( \mathbf{A} \) and \( \tilde{\mathbf{A}} \) are similar, this property carries over to \( \mathbf{A} \).
Let $V$ be a linear vector space.

**Definition G.1.** A norm on $V$, denoted by $\| \cdot \|$, satisfies the following four conditions

(i) $\| x \| \in \mathbb{R}$ and $\| x \| \geq 0$.

(ii) $\| x \| = 0$ if and only if $x = 0$.

(iii) $\| \gamma x \| = |\gamma| \| x \|$, for any $\gamma \in \mathbb{R}$.

(iv) $\| x + y \| \leq \| x \| + \| y \|$, for any $x, y \in V$ (triangular inequality).

If $V = \mathbb{R}^n$ and $x \in \mathbb{R}^N$ denotes a vector with coordinates $x_1, \ldots, x_N$, say, then we often use so called Hölder norms

$$\| x \|_1 = \sum_{j=1}^N |x_j| \quad \text{(G.1a)}$$

$$\| x \|_2 = \left( \sum_{j=1}^N |x_j|^2 \right)^{1/2} \quad \text{(G.1b)}$$

$$\| x \|_{\infty} = \max_{1 \leq j \leq N} |x_j| \quad \text{(G.1c)}$$

The norms $\| \cdot \|_\alpha$ and $\| \cdot \|_\beta$, for some $\alpha$ and $\beta$ are called equivalent if there exist $c_1, c_2 \in \mathbb{R}_+$ such that $\forall x, c_1 \| x \|_\alpha \leq \| x \|_\beta \leq c_2 \| x \|_\alpha$. For $N < \infty$ all norms are equivalent. In particular we have

$$\| x \|_2 \leq \| x \|_1 \leq \sqrt{N} \| x \|_2 \quad \text{(G.2)}$$

The bounds above are attainable. Hence, there is no longer an equivalence if $N \to \infty$.

A consequence of equivalence is that a series, which converges in one norm, also converges in an equivalent norm. If $V = \mathbb{R}^n$ this implies that convergence considerations are norm independent. See, e.g. [1].

A useful property is the inequality of Cauchy-Schwartz: For all $x, y \in \mathbb{R}^N$ we have

$$| \langle x, y \rangle | \leq \| x \|_2 \| y \|_2 \quad \text{(G.3)}$$

Next matrix norms are considered. Let $V$ be the linear space consisting of matrices. A norm on $V$ satisfies the conditions similar to those in Definition G.1 (i), ..., (iv). A vector norm induces a so called associated matrix norm in a natural way as follows:

$$\| A \| := \max_{x \neq 0} \frac{\| Ax \|}{\| x \|} \quad \text{(G.4)}$$

As one can easily verify we have

$$\| A \| = \max_{x = 1} \| Ax \|.$$
On top of (i), . . . , (iv) such an associated norm apparently also has a multiplicativity property

\[(v) \|AB\| \leq \|A\|\|B\|\]

The most often used associated matrix norms are

\[
\|A\|_1 = \max_{j} \sum_{i=1}^{n} |a_{ij}|, \quad \tag{G.5a}
\]

\[
\|A\|_\infty = \max_{i} \sum_{j=1}^{n} |a_{ij}|, \quad \tag{G.5b}
\]

\[
\|A\|_2 = \left( \rho(A^T A) \right)^{\frac{1}{2}}, \quad \tag{G.5c}
\]

where \( \rho(B) \) is the absolutely largest eigenvalue of \( B \), see Definition H.3.

The 2-norm is especially interesting because it is orthogonally invariant, i.e. if \( Q_1 \) en \( Q_2 \) are orthogonal matrices, the the following holds

\[
\|Q_1 A Q_2\|_2 = \|A\|_2.
\]

Also for matrices one can show equivalence of norms, at least for finite dimension. In particular we have

\[
\frac{1}{\sqrt{N}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{N}\|A\|_1 \leq \sqrt{N}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{N}\|A\|_\infty. \quad \tag{G.6}
\]

If \( V \) is the linear space of scalar functions \( x(t) \), defined on an interval \([\alpha, \beta]\) say, we can introduce analogues of the Hölder norm for the continuous case:

\[
\|x\|_p := \left( \int_{\alpha}^{\beta} |x(t)|^p \, dt \right)^{\frac{1}{p}}. \quad \tag{G.7}
\]

Clearly, if \( V \) is a space of vector functions one has to replace the modulus by a suitable vector norm in (G.7). For \( p = \infty \) we have

\[
\|x\|_\infty = \sup_{t \in [\alpha, \beta]} |x(t)|. \quad \tag{G.8}
\]

Each norm on a linear space \( V \) generates a metric \( d \) on \( V \) by \( d(x, y) := \|x - y\|, \ x, y \in V \). This metric has the property of translation invariance: \( d(x + z, y + z) = d(x, y) \).

H Similarity

Let \( A \) be a matrix. If

\[
A x = \lambda x, \quad \tag{H.1}
\]
for some vector \( \mathbf{x} \) and scalar \( \lambda \), then \( \lambda \) is called an eigenvalue and \( \mathbf{x} \) an eigenvector (belonging to \( \lambda \)) of \( \mathbf{A} \).

**Property H.1.**

(i) An eigenvalue \( \lambda \) is a zero of the characteristic polynomial \( \det(\mathbf{A} - \lambda \mathbf{I}) \).

(ii) The product of the eigenvalues of \( \mathbf{A} \) is equal to \( \det(\mathbf{A}) \).

(iii) The sum of the eigenvalues of \( \mathbf{A} \) is equal to \( \sum_{j=1}^{n} a_{jj} \), the so-called trace of \( \mathbf{A} \).

For every matrix \( \mathbf{A} \) a nonsingular matrix \( \mathbf{T} \) exists such that

\[
\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{J},
\]

where \( \mathbf{J} \) is a bidagonal matrix, consisting of blocks containing the same eigenvalues and of which the dimensions correspond with the (algebraic) multiplicity of those eigenvalues.

\[
\mathbf{J} := \begin{pmatrix}
\lambda_1 & 1 & & \\
& \ddots & \ddots & \\
& & \lambda_j & \\
& & & \ddots & \ddots & 1 \\
& & & & & \ddots & \ddots & \ddots & \ddots \\
& & & & \ddots & & & & \\
& & & & & & & & & \\
\end{pmatrix}
\]

The matrices \( \mathbf{T} \) in (H.2) are often chosen such that the “dots” in (H.3) are all equal to 1. From a numerical point of view this may not be so meaningful (as it may cause \( \mathbf{T} \) to be very skew). The form (H.2) is called the Jordan normal form. The geometric multiplicity of an eigenvalue is the dimension of the space of independent eigenvectors. If, in particular, the algebraic and geometric multiplicity are the same for all eigenvalues \( \mathbf{J} \) is a diagonal matrix and each column of \( \mathbf{T} \) then is an eigenvector.

If \( \mathbf{A} \) is symmetric, i.e. \( \mathbf{A} = \mathbf{A}^T \), or skewsymmetric, i.e. \( \mathbf{A} = -\mathbf{A}^T \), the transformation matrix \( \mathbf{T} \) is orthogonal.

The transformation (H.2) is a special instance of a so-called matrix similarity transformation. If \( \mathbf{S} \) is nonsingular, the matrix \( \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{B} \) is called being similar to \( \mathbf{A} \). Of course, it corresponds with viewing a mapping on a different basis.

**Property H.2.** Similar matrices have the same eigenvalues.

**Definition H.3.** The absolute value of the absolutely largest eigenvalue of a matrix is called the spectral radius of \( \mathbf{A} \) and is denoted as \( \rho(\mathbf{A}) \).

We find

Property H.4. If $A$ is symmetric then $\|A\|_2 = \rho(A)$. If $A$ is not symmetric then $\|A\|_2 = (\rho(A^T A))^{\frac{1}{2}}$.

Estimates of eigenvalues and consequences

The spectral radius of a matrix is a useful means for deciding a number of properties. In particular it is sometimes crucial to know whether this radius is at least smaller than 1. This is why we have the following theorems

Theorem I.1. $\rho(A) < 1$ if and only if $\lim_{i \to \infty} A^i = 0$

Proof.

If : $(\rho(A))^i = \rho(A^i) \leq \|A^i\|^2$.

Only if : Consider the Jordan normal form. Then $A^i = TJ^T$. By studying a single Jordan block and splitting into diagonal and codiagonal the proof is simple to complete.

Theorem I.2 (Neumann series). $\lim_{n \to \infty} \sum_{i=0}^{n} A^i$ exists if and only if $\rho(A) < 1$. The following holds: $\sum_{i=0}^{\infty} A^i = (I - A)^{-1}$.

Proof. If $\{\sum_{i=0}^{n} A^i\}$ converges then $A^i \to 0$, which implies that $\rho(A) < 1$.

If, on the other hand, $\rho(A) < 1$ then $\det(I - A) \neq 0$, so $(I - A)^{-1}$ exists.

For each $n$ we have $(I + A + \cdots + A^n)(I - A) = I - A^{n+1}$.

Corollary I.3. The matrix $I - A$ is nonsingular if for any norm $\|A\| < 1$.

Definition I.4. A matrix is called stable if $\rho(A) \leq 1$ and all eigenvalues with modulus 1 are simple.

Property I.5. If $A$ is stable then there exists constant $\kappa$, say, such that $\forall x \in \mathbb{R}^n \|A^i x\| \leq \kappa \|x\|$ (this $\kappa$ is depending on $N$ and the norm chosen).

Finally we give an important theorem to estimate the eigenvalues.

Theorem I.6. (Gershgorin) An eigenvalue of $A$ lies in at least one of the closed discs with centre point $a_{ii}$ and radius $\sum_{j \neq i} |a_{ij}|$. (N.B. the discs are in $\mathbb{C}$).
Proof. Let \( \lambda \) be an eigenvalue and \( x = (x_1, \ldots, x_N)^T \) a corresponding eigenvector. Then

\[
Ax = \lambda x \Rightarrow (\lambda - a_{rr})x_r = \sum_{j \neq r} a_{rj}x_j.
\]

Hence

\[
|\lambda - a_{rr}|x_r \leq \sum_{j \neq r} |a_{rj}| |x_j| \quad \text{for all } r.
\]

If \( m \) is such that \( |x_m| = \max_r |x_r| \) (N.B. \( x_m \neq 0 \)), then

\[
|\lambda - a_{mm}| \leq \sum_{j \neq m} |a_{mj}| |x_j| \leq \sum_{j \neq m} |a_{mj}|.
\]

Note that if the matrix is symmetric the discs are in fact segments of the real axis.

---

### J. Theorems from Vector Calculus

Let \( a, b \) and \( c \) be 3D vectors, let \( v \) and \( w \) be well-behaved vector functions from \( \mathbb{R}^3 \to \mathbb{R}^3 \), and let \( \phi \) and \( \psi \) be a well-behaved scalar functions \( \mathbb{R}^3 \to \mathbb{R} \). Let \( \Omega \) be a three-dimensional volume with volume element \( dV \), and \( \partial \Omega \) a closed two-dimensional surface bounding \( \Omega \) with area element \( dS \) and associated unit outward vector \( n \). Let \( S \) denote an open surface with the oriented contour \( C \), with line element \( d\ell \), bounding it. The normal \( n \) to \( S \) is defined according to the right-hand-screw rule applied to \( C \). Then we have the following vector and integral relations.

\[
a \cdot (b \times c) = b \cdot (c \times a) = c \cdot (a \times b) \quad (J.1)
\]

\[
a \times (b \times c) = b(a \cdot c) - c(a \cdot b) \quad (J.2)
\]

\[
(a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c) \quad (J.3)
\]

\[
\nabla (v \cdot w) = v \cdot \nabla w + w \cdot \nabla v + v \times (\nabla \times w) + w \times (\nabla \times v) \quad (J.4)
\]

\[
v \cdot \nabla v = \frac{1}{2} \nabla (v^2) + (\nabla \times v) \times v \quad (J.5)
\]

\[
a \cdot \nabla (v \cdot w) = v \cdot (a \cdot \nabla w) + w \cdot (a \cdot \nabla v) \quad (J.6)
\]

\[
\nabla \cdot (v \times w) = w \cdot (\nabla \times v) - v \cdot (\nabla \times w) \quad (J.7)
\]

\[
\nabla \times (v \times w) = v(\nabla \cdot w) - w(\nabla \cdot v) - v \cdot \nabla w + w \cdot \nabla v \quad (J.8)
\]

\[
\nabla \cdot (\nabla \times v) = 0 \quad (J.9)
\]

\[
\nabla \times (\nabla \phi) = 0 \quad (J.10)
\]

\[
\nabla \times (\nabla \cdot v) = \nabla (\nabla \cdot v) - \nabla^2 v \quad (J.11)
\]

Gauss’ or divergence theorem:

\[
\int_{\Omega} \nabla \cdot v \, dV = \oint_{\partial \Omega} v \cdot n \, dS \quad (J.12)
\]

\[
\int_{\Omega} \nabla \phi \, dV = \oint_{\partial \Omega} \phi \, n \, dS \quad (J.13)
\]

**APPENDIX . USEFUL DEFINITIONS AND PROPERTIES**

\[ \int_{\Omega} \nabla \times \mathbf{v} \, dV = \oint_{\partial \Omega} \mathbf{n} \times \mathbf{v} \, dS \] (J.14)

Green's first identity: \[ \int_{\Omega} \left( \phi \nabla^2 \psi + \nabla \phi \cdot \nabla \psi \right) \, dV = \oint_{\partial \Omega} \phi \nabla \psi \cdot \mathbf{n} \, dS \] (J.15)

Green's second identity: \[ \int_{\Omega} \left( \phi \nabla^2 \psi - \psi \nabla^2 \phi \right) \, dV = \oint_{\partial \Omega} \left( \phi \nabla \psi - \psi \nabla \phi \right) \cdot \mathbf{n} \, dS \] (J.16)

Stokes’s theorem: \[ \int_{S} (\nabla \times \mathbf{v}) \cdot \mathbf{n} \, dS = \oint_{C} \mathbf{v} \cdot d\ell \] (J.17)

\[ \int_{S} \mathbf{n} \times \nabla \phi \, dS = \oint_{C} \phi \, d\ell \] (J.18)

Let \( q(x, t) \) be a quantity per unit volume of a fluid. Consider a material volume \( \Omega(t) \) moving with the flow. Then we have the

**Transport theorem:**

\[ \frac{d}{dt} \int_{\Omega(t)} q(x, t) \, dV = \int_{\Omega(t)} \left( \frac{\partial}{\partial t} q(x, t) + \nabla \cdot (q \mathbf{v})(x, t) \right) \, dV. \] (J.19)

**K Properties of cartesian, cylindrical and spherical co-ordinate systems**

**Cartesian.** Let \( \mathbf{e}_x, \mathbf{e}_y, \) and \( \mathbf{e}_z \) be the orthogonal unit vectors associated with the cartesian \( x, y \) and \( z \) co-ordinates, and \( \mathbf{E} = \mathbf{e}_x E_x + \mathbf{e}_y E_y + \mathbf{e}_z E_z \) and \( \phi \) are smooth functions of \((x, y, z)\). Then

\[ (d\mathbf{e})^2 = (dx)^2 + (dy)^2 + (dz)^2, \] (K.1a)

\[ \nabla \phi = \mathbf{e}_x \frac{\partial \phi}{\partial x} + \mathbf{e}_y \frac{\partial \phi}{\partial y} + \mathbf{e}_z \frac{\partial \phi}{\partial z}, \] (K.1b)

\[ \nabla^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \] (K.1c)

\[ \nabla \cdot \mathbf{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}, \] (K.1d)

\[ \nabla \times \mathbf{E} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_x & E_y & E_z \end{vmatrix} \] (K.1e)

**Cylindrical.** Let \( \mathbf{e}_r, \mathbf{e}_\phi, \) and \( \mathbf{e}_z \) be the orthogonal unit vectors associated with the cylindrical \( r, \phi \) and \( z \) co-ordinates, and \( \mathbf{E} = \mathbf{e}_r E_r + \mathbf{e}_\phi E_\phi + \mathbf{e}_z E_z \) and \( \phi \) are smooth functions of \((r, \phi, z)\). Then

\[ x = r \cos \phi, \quad y = r \sin \phi, \quad z = z, \] (K.2a)
\[(d\ell)^2 = (dr)^2 + (r \, d\phi)^2 + (dz)^2. \quad \text{(K.2b)}\]
\[\nabla \varphi = e_r \frac{\partial \varphi}{\partial r} + e_\theta \frac{1}{r} \frac{\partial \varphi}{\partial \theta} + e_\phi \frac{1}{r \sin \theta} \frac{\partial \varphi}{\partial \phi}. \quad \text{(K.2c)}\]
\[\nabla^2 \varphi = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \varphi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \varphi}{\partial \phi^2}. \quad \text{(K.2d)}\]
\[\nabla \cdot \mathbf{E} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 E_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta E_\theta) + \frac{1}{r \sin \theta} \frac{\partial E_\phi}{\partial \phi}. \quad \text{(K.3a)}\]
\[\nabla \times \mathbf{E} = \frac{1}{r^2 \sin \theta} \begin{vmatrix} e_r & e_\theta r & e_\phi r \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ E_r & E_\theta r & E_\phi r \end{vmatrix}. \quad \text{(K.3b)}\]

### Spherical

Let \(e_r, e_\theta, \text{ and } e_\phi\) be the orthogonal unit vectors associated with the spherical \(r, \theta\) and \(\phi\) co-ordinates, and \(E = e_r E_r + e_\theta E_\theta + e_\phi E_\phi\) and \(\varphi\) are smooth functions of \((r, \theta, \phi)\). Then
\[
x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta, \quad \text{(K.3a)}
\]
\[
(d\ell)^2 = (dr)^2 + (r \, d\phi)^2 + (r \sin \theta \, d\phi)^2, \quad \text{(K.3b)}
\]
\[
\nabla \varphi = e_r \frac{\partial \varphi}{\partial r} + e_\theta \frac{1}{r} \frac{\partial \varphi}{\partial \theta} + e_\phi \frac{1}{r \sin \theta} \frac{\partial \varphi}{\partial \phi}, \quad \text{(K.3c)}
\]
\[
\nabla^2 \varphi = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \varphi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \varphi}{\partial \phi^2}. \quad \text{(K.3d)}
\]
\[
\nabla \cdot \mathbf{E} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 E_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta E_\theta) + \frac{1}{r \sin \theta} \frac{\partial E_\phi}{\partial \phi}. \quad \text{(K.3e)}
\]
\[
\nabla \times \mathbf{E} = \frac{1}{r^2 \sin \theta} \begin{vmatrix} e_r & e_\theta r & e_\phi r \sin \theta \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ E_r & E_\theta r & E_\phi r \sin \theta \end{vmatrix}. \quad \text{(K.3f)}
\]

## L Tensors

A three-dimensional vector space over the field of real numbers, equipped with an inner, dot, or scalar product \(a \cdot b\) and an outer, cross, or vector product \(a \times b\) is called a Euclidean vector space.

A tensor – strictly speaking: of order 2 – is a linear transformation of a Euclidean vector space into itself. The identity tensor is denoted by \(\mathcal{I}\).

If the tensor \(\mathbf{A}\) is written as a \(3 \times 3\) matrix \((a_{ij})\) on the standard basis \([e_1, e_2, e_3]\) of \(\mathbb{R}^3\), we have the contraction or trace of \(\mathbf{A}\), given by \(\text{tr}(\mathbf{A}) = a_{11} + a_{22} + a_{33}\), and the determinant of \(\mathbf{A}\), given by \(\det(\mathbf{A}) = \mathbf{A}e_1 \cdot (\mathbf{A}e_2 \times \mathbf{A}e_3)\). Both are invariants of \(\mathbf{A}\) under rotation of axes.
and therefore may be expressed in the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of matrix $(a_{ij})$. Altogether, $\mathbf{A}$ has three invariants.

A (second order) tensor $\mathbf{A}$ is completely determined by its invariants

$$\text{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \lambda_3,$$

$$\frac{1}{2}[\text{tr}(\mathbf{A})^2 - \text{tr}(\mathbf{A}^2)] = \lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1,$$

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \lambda_3.$$  

(L.1a)  
(L.1b)  
(L.1c)

To each $\mathbf{A}$ there is a traceless tensor (the deviator or deviatoric part of $\mathbf{A}$)

$$\mathbf{A}' := \mathbf{A} - \frac{1}{3} \text{tr}(\mathbf{A}) \mathbf{I},$$  

(L.2)

with the same invariants except the first one, which is zero.

The inner product of two tensors $\mathbf{A}$ and $\mathbf{B}$ produces a tensor $\mathbf{A} \cdot \mathbf{B}$, whose components are given by

$$(\mathbf{A} \cdot \mathbf{B})_{ij} := \sum_{k=1}^{3} A_{ik} B_{kj}. $$

(L.3)

The double inner product of two tensors $\mathbf{A}$ and $\mathbf{B}$ produces a scalar $\mathbf{A} : \mathbf{B}$, which can be evaluated as the sum of the 9 products of the tensor components

$$\mathbf{A} : \mathbf{B} = \sum_{i=1}^{3} \sum_{j=1}^{3} A_{ij} B_{ij}. $$

(L.4)

The dyadic product of two vectors $\mathbf{a}$ and $\mathbf{b}$ produces a tensor $\mathbf{ab}^T$, given by

$$(\mathbf{ab}^T)_{ij} = a_i b_j, \quad \text{or} \quad \mathbf{ab}^T = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix}.$$  

(L.5)
M. Dimensionless numbers

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Expression</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archimedes</td>
<td>Ar</td>
<td>$g \Delta \rho L^3 / \rho v^2$</td>
<td>particles, drops or bubbles</td>
</tr>
<tr>
<td>Arrhenius</td>
<td>Arr</td>
<td>$E / RT$</td>
<td>chemical reactions</td>
</tr>
<tr>
<td>Biot</td>
<td>Bi</td>
<td>$hL / \kappa$</td>
<td>heat transfer</td>
</tr>
<tr>
<td>Biot</td>
<td>Bi</td>
<td>$h_p L / D$</td>
<td>mass transfer</td>
</tr>
<tr>
<td>Bodenstein</td>
<td>Bo</td>
<td>$VL / D_{ax}$</td>
<td>mass transfer with axial dispersion</td>
</tr>
<tr>
<td>Bond</td>
<td>Bo</td>
<td>$\rho g L^2 / \sigma$</td>
<td>gravity against surface tension</td>
</tr>
<tr>
<td>Dean</td>
<td>De</td>
<td>$(VL / \nu)(L / 2r)^{1/2}$</td>
<td>flow in curved channels</td>
</tr>
<tr>
<td>Eckert</td>
<td>Ec</td>
<td>$V^2 / C_P \Delta T$</td>
<td>kinetic energy against enthalpy difference</td>
</tr>
<tr>
<td>Euler</td>
<td>Eu</td>
<td>$\Delta p / \rho V^2$</td>
<td>pressure resistance</td>
</tr>
<tr>
<td>Fourier</td>
<td>Fo</td>
<td>$\alpha t / L^2$</td>
<td>heat conduction</td>
</tr>
<tr>
<td>Fourier</td>
<td>Fo</td>
<td>$Dt / L^2$</td>
<td>diffusion</td>
</tr>
<tr>
<td>Froude</td>
<td>Fr</td>
<td>$V / \sqrt{gL}$</td>
<td>gravity waves</td>
</tr>
<tr>
<td>Grashof</td>
<td>Gr</td>
<td>$\beta \Delta T g L^3 \nu^3$</td>
<td>natural convection</td>
</tr>
<tr>
<td>Helmholtz</td>
<td>He</td>
<td>$fL / c$</td>
<td>acoustic wave number</td>
</tr>
<tr>
<td>Kapitza</td>
<td>Ka</td>
<td>$g \mu^4 / \rho \sigma^3$</td>
<td>film flow</td>
</tr>
<tr>
<td>Knudsen</td>
<td>Kn</td>
<td>$\lambda / L$</td>
<td>low density flow</td>
</tr>
<tr>
<td>Lewis</td>
<td>Le</td>
<td>$\alpha / D$</td>
<td>combined heat and mass transfer</td>
</tr>
<tr>
<td>Mach</td>
<td>M</td>
<td>$V / c$</td>
<td>compressible flow</td>
</tr>
<tr>
<td>Nusselt</td>
<td>Nu</td>
<td>$hL / \kappa$</td>
<td>convective heat transfer</td>
</tr>
<tr>
<td>Péclet</td>
<td>Pe</td>
<td>$VL / \alpha$</td>
<td>forced convection heat transfer</td>
</tr>
<tr>
<td>Péclet</td>
<td>Pe</td>
<td>$VL / D$</td>
<td>forced convection mass transfer</td>
</tr>
<tr>
<td>Prandtl</td>
<td>Pr</td>
<td>$\nu / \alpha = C_P \mu / \kappa$</td>
<td>convective heat transfer</td>
</tr>
<tr>
<td>Rayleigh</td>
<td>Ra</td>
<td>$\beta \Delta T g L^3 \alpha v$</td>
<td>natural convection heat transfer</td>
</tr>
<tr>
<td>Reynolds</td>
<td>Re</td>
<td>$\rho VL / \mu$</td>
<td>fluid flow</td>
</tr>
<tr>
<td>Schmidt</td>
<td>Sc</td>
<td>$\nu / D$</td>
<td>convective mass transfer</td>
</tr>
<tr>
<td>Sherwood</td>
<td>Sh</td>
<td>$h_p L / D$</td>
<td>convective mass transfer</td>
</tr>
<tr>
<td>Stanton</td>
<td>St</td>
<td>$h / \rho C_P V$</td>
<td>forced convection heat transfer</td>
</tr>
<tr>
<td>Stanton</td>
<td>St</td>
<td>$h_p / V$</td>
<td>forced convection mass transfer</td>
</tr>
<tr>
<td>Stokes</td>
<td>S</td>
<td>$v / f L^2$</td>
<td>viscous damping in unsteady flow</td>
</tr>
<tr>
<td>Strouhal</td>
<td>Sr</td>
<td>$fL / V$</td>
<td>hydrodynamic wave number</td>
</tr>
<tr>
<td>Weber</td>
<td>We</td>
<td>$\rho V^2 L / \sigma$</td>
<td>film flow, bubble formation, droplet breakup</td>
</tr>
</tbody>
</table>
### Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_p$</td>
<td>specific heat</td>
<td>J/kg K</td>
</tr>
<tr>
<td>$D$</td>
<td>diffusion coefficient</td>
<td>m²/s</td>
</tr>
<tr>
<td>$D_{ax}$</td>
<td>axial dispersion coefficient</td>
<td>m²/s</td>
</tr>
<tr>
<td>$E$</td>
<td>activation energy</td>
<td>J/mol</td>
</tr>
<tr>
<td>$f$</td>
<td>frequency</td>
<td>1/s</td>
</tr>
<tr>
<td>$g$</td>
<td>gravitational acceleration</td>
<td>m/s²</td>
</tr>
<tr>
<td>$h$</td>
<td>heat transfer coefficient</td>
<td>W/m² K</td>
</tr>
<tr>
<td>$h_D$</td>
<td>mass transfer coefficient</td>
<td>m/s</td>
</tr>
<tr>
<td>$L$</td>
<td>length</td>
<td>m</td>
</tr>
<tr>
<td>$p$, $\Delta p$</td>
<td>pressure</td>
<td>Pa</td>
</tr>
<tr>
<td>$R$</td>
<td>universal gas constant</td>
<td>J/mol K</td>
</tr>
<tr>
<td>$r$</td>
<td>radius of curvature</td>
<td>m</td>
</tr>
<tr>
<td>$T$, $\Delta T$</td>
<td>temperature</td>
<td>K</td>
</tr>
<tr>
<td>$t$</td>
<td>time</td>
<td>s</td>
</tr>
<tr>
<td>$V$</td>
<td>velocity</td>
<td>m/s</td>
</tr>
<tr>
<td>$c$</td>
<td>sound speed</td>
<td>m/s</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>thermal conductivity</td>
<td>W/m K</td>
</tr>
<tr>
<td>$\alpha = \kappa / \rho C_p$</td>
<td>thermal diffusivity</td>
<td>m²/s</td>
</tr>
<tr>
<td>$\beta$</td>
<td>coef. of thermal expansion</td>
<td>K⁻¹</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>molecular mean free path</td>
<td>m</td>
</tr>
<tr>
<td>$\mu$</td>
<td>dynamic viscosity</td>
<td>Pa s</td>
</tr>
<tr>
<td>$\nu = \mu / \rho$</td>
<td>kinematic viscosity</td>
<td>m²/s</td>
</tr>
<tr>
<td>$\rho$, $\Delta \rho$</td>
<td>density</td>
<td>kg/m³</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>surface tension</td>
<td>N/m</td>
</tr>
</tbody>
</table>
Bibliography

[1] Groetsch DIT MOET NOG
[2] H. Kreiss, ???iets over schattingen bij laxstabiliteit (H. Kreiss)
[3] van Rooij, Analyse voor Beginners DIT MOET NOG


Bibliography


Bibliography

Bibliography

Index

k-Riemann invariants, 206
absence of free magnetic poles, 89
adiabatic invariants, 261
adjoint operator, 63
Ampère-Maxwell’s law, 88
associated matrix norm, 276
averaging, 218
backward difference, 270
Backward Difference Formulae (BDF), 271
barotropic flow, 118
base characteristic, 18
block-wave function, 41
boundary condition, 122
boundary layer, 246
Buckley-Leverett equation, 192
Cauchy problem, 201
Cauchy’s equation of motion, 81
causal, 45
causality condition, 45
characteristic, 18
characteristic polynomial, 278
characteristic equations, 18
characteristic form, 198
characteristic variable, 21
characteristic variables, 197
characteristics, 176
complementary error function, 110, 152
conservation equation, 179
conservation equations, 75
conservation form, 179, 198
conservation of angular momentum, 80
conservation of electric charge, 89
conservation of energy, 81
conservation of linear momentum, 79
conservative, 50
constitutive equations, 83
constitutive relations, 75
contact angle, 86
contact discontinuity, 207
continuum physics, 75
convergence: pointwise, 269
convergence: uniform, 269
convolution theorem, 43
correspondence principle, 247
Coulomb’s law, 88
d’Alembert solution, 217
decibel, 112
delta function, 58
dimensionally homogeneous, 103
dimensionless parameter, 103
Dirichlet boundary condition, 122
dispersion, 50
displacement vector, 76
distinguished limit, 234
distinguished limits, 248
distribution, 58
regular, 58
divergence theorem, 280
Duhamel integral, 68, 69
Eckert number, 114
eigenvalue, 278
eigenvector, 278
elasticity, 86
electromagnetic continuity conditions, 93
energy, 253
energy velocity, 50
enthalpy, 84, 164
entropy, 84
entropy condition, 192
Index

entropy solution, 192
equivalent norms, 276
error function, 152
Euclidean vector space, 282
Euler number, 114
expansion shock, 190
Faraday-Henry’s law, 89
Fick’s law, 84
Fisher’s equation, 168
flux function, 178
flux vector, 198
Fourier
   mode, 38
cosine series, 41
discrete transform, 47
series, 38, 40
sine series, 41
transform, 43
Fourier number, 114
Fourier’s law, 84
Fredholm’s alternative, 66
frequency, 38
Gauss’ theorem, 280
generalised eigenvalue problem, 194
genuinely nonlinear, 205
geometric multiplicity, 278
Green’s first identity, 280
Green’s function, 137
Green’s functions, 66
group velocity, 50
harmonic function, 122
heat capacity, 85
heat equation, 28, 149
Helmholtz equation, 121
Helmholtz number, 114
Hooke’s law, 86
hyperbolic system, 195, 196
inequality of Cauchy-Schwartz, 276
inherent scaling, 103
initial curve, 18
inner expansion, 246
integral surface, 18
internal layer, 246
Jordan normal form, 278
Joule heating, 90
k-rarefaction wave, 205
k-shock wave, 207
k-simple wave, 205
Kelvin’s Theorem, 92
Kelvin-Voigt model, 87
Laplace equation, 28, 121
latent heat, 85
left eigenvectors, 194
Lindstedt-Poincaré method, 243
linearly degenerate, 205
logarithmic switchback, 251
Lorentz force, 91
lubrication flow, 264
Mach number, 114
matched asymptotic expansions, 245, 247
matching, 247, 252
material surface, 77
material volume, 77
mathematizing, 96
matrix
   similarity transformation, 278
   stable, 279
   trace, 278
matrix pencil, 194
Maxwell model, 88
Maxwell’s equations, 89
method of Massau, 200
method of strained coordinates, 244, 256
model, 97
   asymptotic, 97
   building block, 98
   canonical, 101
   characteristic, 101
   constructing, 98
   lumped parameter, 98
   quintessential, 101
   reducing, 97
   systematic, 97
multiple scales, 253, 256

Index

Navier-Stokes equations, 85
Neumann boundary condition, 122
normal form, 198
objective, 83
observer transformation, 83
Ohm’s law, 90
orthogonally invariant, 277
outer expansion, 246
overlap hypothesis, 246
Péclet number, 114
Parseval’s identity, 39
particular, 66
particular solution, 136
perfect gas, 85
permafrost, 120
perturbation methods, 234
Pfaffian differential equation, 197
phase velocity, 49
physical model, 83
piecewise continuous, 39
piecewise smooth, 39
Poincaré expansion, 236
Poisson equation, 121
Poisson’s Formula, 44
Poisson’s ratio, 87
Poynting’s theorem, 90
Prandtl number, 114
Prandtl’s boundary layer equations, 250
pressure
    mechanical, 86
    thermodynamic, 86
quasilinear, 21
quasilinear equation, 178
radiative, 160
Rankine-Hugoniot jump condition, 186, 203
ray, 262
reality condition, 45
reflections, 217
regular matrix pencil, 194
resonance, 254, 256
Reynolds number, 114
Riemann invariants, 198
Riemann problem, 187
Riemann variables, 198
Robin boundary condition, 122
saw-tooth function, 40
secular terms, 244, 245, 254, 255, 257, 258
self-adjoint, 127
semilinear, 177
shallow water equations, 208, 243
shock speed, 188
shock wave, 188
signal velocity, 50
significant degeneration, 234
significant degenerations, 248
similarity solutions, 30
singular matrix pencil, 194
skewsymmetric, 278
slowly varying fast time scale, 256
solitary wave, 171
soliton, 171
specific energy, 82
specific heat, 85
specific internal energy, 82
specific kinetic energy, 82
specific-heat ratio, 86
spectral radius, 278
stationary elliptic equation, 165
Stefan constant, 162
Stokes flow, 116
Stokes’s theorem, 281
Stokeslets, 142
stretched coordinate, 246
Strouhal number, 114
subcritical flow, 225
supercritical flow, 225
superposition, 66, 136
support, 57
surface tension, 86
symmetric, 278
telegraph equation, 9
tensor, 282
    contraction, 282
deformation velocity, 77

Index

determinant, 282
deviator, 283
double inner product, 283
dyadic product, 283
inner product, 283
invariants, 283
linear deformation, 76
linear strain, 76
rate of deformation, 77
trace, 282
transport equation, 179
transport theorem, 77, 281
travelling wave, 29
travelling-wave solution, 169
triangular inequality, 276
vanishing viscosity solution, 182
viscoelasticity, 87
viscosity
  bulk, 86
  dynamic, 86
  expansion, 86
  kinematic, 86
  second, 86
viscous stress tensor, 84
wave equation, 28, 216
wave number, 38
wave speed, 29, 49
waverarefaction wave, 189
WKB hypothesis, 257
Young’s modulus, 87