UNITEXT 154

Alberto Valli

A Compact Course on Linear PDEs

Second Edition

Springer

UNITEXT

La Matematica per il 3+2

Volume 154

Editor-in-Chief

Alfio Quarteroni, Politecnico di Milano, Milan, Italy
École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Series Editors

Luigi Ambrosio, Scuola Normale Superiore, Pisa, Italy

Paolo Biscari, Politecnico di Milano, Milan, Italy

Ciro Ciliberto, Università di Roma "Tor Vergata", Rome, Italy

Camillo De Lellis, Institute for Advanced Study, Princeton, NJ, USA

Victor Panaretos, Institute of Mathematics,École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Lorenzo Rosasco, DIBRIS, Università degli Studi di Genova, Genova, Italy
Center for Brains Mind and Machines, Massachusetts Institute of Technology, Cambridge, Massachusetts, US
Istituto Italiano di Tecnologia, Genova, Italy

The **UNITEXT - La Matematica per il 3+2** series is designed for undergraduate and graduate academic courses, and also includes books addressed to PhD students in mathematics, presented at a sufficiently general and advanced level so that the student or scholar interested in a more specific theme would get the necessary background to explore it.

Originally released in Italian, the series now publishes textbooks in English addressed to students in mathematics worldwide.

Some of the most successful books in the series have evolved through several editions, adapting to the evolution of teaching curricula.

Submissions must include at least 3 sample chapters, a table of contents, and a preface outlining the aims and scope of the book, how the book fits in with the current literature, and which courses the book is suitable for.

For any further information, please contact the Editor at Springer: francesca.bonadei@springer.com

THE SERIES IS INDEXED IN SCOPUS

***

UNITEXT is glad to announce a new series of free webinars and interviews handled by the Board members, who rotate in order to interview top experts in their field.

Access this link to subscribe to the events:

https://cassyni.com/events/TPQ2UgkCbJvvz5QbkcWXo3

Alberto Valli

A Compact Course on Linear PDEs

Second Edition

Springer

Alberto Valli
Department of Mathematics
University of Trento
Trento, Italy

ISSN 2038-5714 ISSN 2532-3318 (electronic)
UNITEXT
ISSN 2038-5722 ISSN 2038-5757 (electronic)
La Matematica per il 3+2
ISBN 978-3-031-35975-0 ISBN 978-3-031-35976-7 (eBook)
https://doi.org/10.1007/978-3-031-35976-7

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2020, 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: Henrique Oliveira: "Radice comune". Copyright Arte Sella 2019. Photo by Giacomo Bianchi.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

To Jarno and Beatrice,
the future

Preface

It don't mean a thing
(if it ain't got that swing)[1]
Edward "Duke" Ellington

This book stems from a 45-hour course that I delivered for the Master degree at the Department of Mathematics of the University of Trento.

Partial differential equations (PDEs) are an extremely wide topic, and it is not possible to include them into a single course, no matter how many lessons are assigned to it. Thus, the first question I had to face was the viewpoint I wanted to adopt and choice of the arguments.

I decided to focus on linear equations. It is well known to everyone that the mathematical description of natural phenomena is mainly based on nonlinear models; however, in many cases, a reasonable approximation is obtained by a linear formulation, and, moreover, the knowledge of linear problems is the first step for dealing with more complex nonlinear cases.

The second choice I made is to limit the presentation to the so-called weak formulation of partial differential equations. This means that our point of view is the following: solving a linear partial differential equation is interpreted as the solution of a problem associated to a linear operator acting between *suitable* infinite dimensional vector spaces.

The path for arriving at this abstract formulation needs some tools that were not available in the classical theory. In a nutshell, the four main missing ingredients are the following:

- Weak derivatives
- Weak solutions
- Sobolev spaces
- A bit of functional analysis

[1] If you are curious, take a look on the web: you can find nice videos on YouTube with this title.

The first results in this direction date back to the 1930s, with the pioneering works of Jean Leray,[2] Sergei L. Sobolev,[3] and others. In the same period, the study of infinite dimensional vector spaces and of functional analysis attracted the attention of many researchers: let us only mention the milestone book by Stefan Banach.[4]

Still speaking about concepts not present in the classical theory, I decided not introducing the distributions and the distributional derivatives, as they are not essential for the presentation. In fact, as it is well known, the distributional derivative of a function essentially coincides with its weak derivative, and dealing with spaces of functions permits to avoid further generalizations.

The determination of the weak formulation is essentially performed by transforming the original problem into a set of infinitely many integral equations, one for each "test" function belonging to a suitable vector space. In several points of the book, I have tried to motivate the various steps of this approach starting from the analysis of finite dimensional linear systems, then enlightening analogies and differences when passing to the infinite dimensional case. In particular, the third chapter is devoted to results of functional analysis that show some typical differences between a finite dimensional and an infinite dimensional vector space. Another section of that chapter has the aim to clarify that suitable spaces for the new approach are those endowed with a scalar product, more precisely those for which the orthogonal projection on a closed subspace is well-defined: in other words, this means Hilbert spaces.

This recurring comparison between algebraic linear systems and weak formulations of linear PDEs has the aim of making clear that for the latter, subject functional analysis plays the role of linear algebra, namely, it is a basic tool for its study; however, as it has been observed by Lawrence C. Evans, this does not mean that it is a good idea to transform the whole topic into a too abstract branch of functional analysis itself.

When I started to teach the course, I suggested a couple of books to the students: those by Evans [8] and Salsa [24]. For this reason, I cannot hide that the structure of these books has influenced what I presented then to my students and what is included now in this book. However, I hope that the reader can find here at least a different flavor (together with some new topics).

The book is organized as follows. Chapter 1 is a very brief introduction to the subject, in which some definitions are given and a list of examples are presented.

In the second chapter, many important items already appear: second order elliptic equations and related boundary value problems, weak solutions, and finally also the Lax–Milgram theorem. However, the functional analysis framework is not made

[2] Leray [18]. It seems that Leray has been the first one to speak about weak solutions ("solutions turbulentes") and weak derivatives ("quasi-dérivées").

[3] Sobolev [25]. The functional framework where we describe and analyze the problems is given by Sobolev spaces, a name on which there is agreement since the middle of last century.

[4] Banach [2].

clear, and for that the reader is referred to following results included in Chaps. 3 and 4.

Chapter 3 is devoted to analogies and differences between finite dimensional and infinite dimensional vector spaces, and to the motivation that makes useful the introduction of Hilbert spaces.

In Chap. 4, some core topics are introduced and analyzed: weak derivatives and Sobolev spaces.

The fifth chapter is a central part of the book: a systematic presentation of weak formulations of elliptic boundary value problems is there included. Moreover, the properties of the bilinear forms which describe the problems are presented in full detail. A section is also devoted to the boundary value problems associated to the biharmonic equation.

Chapter 6 is devoted to several technical results that have been used in the previous chapters: approximation in Sobolev spaces, Poincaré and trace inequalities, Rellich compactness theorem, and du Bois-Reymond lemma.

In Chap. 7, a rich variety of additional results is presented: Fredholm alternative, spectral theory for elliptic operators, maximum principle, regularity results and Sobolev embedding theorems, and finally Galerkin numerical approximation.

The eighth chapter deals with constrained minimization and Lagrange multipliers in the infinite dimensional case. A general theory for saddle point problems is presented, and two specific examples are described: second order elliptic equations rewritten as a first order system of two equations, and the Stokes problem. The Galerkin approximation of saddle point problems is also described and analyzed.

Chapter 9 is focused on parabolic problems, starting from the abstract evolution theory in Hilbert spaces and then arriving to its application to specific problems, among them non-stationary linear Navier–Stokes equations. The proof of maximum principle is also included.

A similar presentation is given in Chap. 10 for hyperbolic problems, including Maxwell equations, ending with the proof of the property of finite propagation speed.

The book finishes with some appendices, devoted to technical results: a detailed construction of a partition of unity; the precise definition of the regularity of the boundary of a domain; integration by parts formulas; the Reynolds transport theorem; the Gronwall lemma; a general well-posedness theorem for weak problems.

Each chapter of the book is complemented by some exercises: they have different difficulty, and in some case could be more properly intended as an additional in-depth analysis. For the ease of the reader, I decided to present the complete solution of all of them.

At the end, a few words about the sentence by "Duke" Ellington that I chose as an incipit: a book is not a course, even though the title seems to suggest it. Thus, for the delight of the students, a colleague who will decide to follow this presentation should find the way to add some swing to these barren pages: I tried my best, but it is never enough.

This book would not have been written without my former Master students Federico Bertacco and Laura Galvagni, who a day (but after the exam!) entered my office with the Latex file of my unrefined handwritten notes. This has been the irresistible push for rearranging everything into a better structured textbook. I am also grateful to Gabriele Dalla Torre, who suggested the best way for drawing the figures, to my old friend Paolo Acquistapace, who helped me in clarifying the proof of uniqueness in Theorem 10.1, to Nicolò Drago, who furnished the proof of Exercise 6.11, and to Arte Sella and Giacomo Bianchi for having permitted the reproduction of the photo on the cover.

Finally, I want to thank the editors Luigi Ambrosio, Paolo Biscari, Ciro Ciliberto, Camillo De Lellis, Victor Panaretos, and Lorenzo Rosasco and the editor-in-chief Alfio Quarteroni for having accepted to publish this book in the Springer series UNITEXT: La Matematica per il 3+2. Special thanks to Francesca Bonadei from Springer, who encouraged me to undertake this project and with great experience and enthusiasm has followed me along its realization.

In the second edition of this book some sections and exercises have been added. In particular, Sect. 2.2.2 (on a general strategy for solving linear problems in an infinite dimensional vector space), Sect. 5.6 (on the biharmonic equation), Sect. 9.2.2 (on non-stationary linear Navier–Stokes equations), and Sect. 10.1.2 (on Maxwell equations) are new, as well as Exercises 1.2, 4.1, 6.11, 7.18, 7.20, 7.21, 7.22, 8.4, 9.3, 9.5, and 10.1; moreover, the solutions of Exercises 6.4, 7.8, and 9.1 have been largely rewritten. Theorem 7.16, Corollary A.1, and Remark 2.6 are also new, while the proofs of Theorems 6.7, 7.7, 7.10, 7.12, 9.4, and 10.1, as well as the solutions of Exercises 7.2 and 7.3, have been modified for better understanding. Finally, many misprints have been corrected and many sentences have been rephrased.

Some final considerations: I recently read the nice book by Vladimir Maz'ya and Tatyana Shaposhnikova on the life and scientific activity of Jacques Hadamard,[5] and I learnt that the last work of Hadamard is a book on partial differential equations, completed at the age of 97 years (!) and published in Beijing in 1964,[6] just after his death. This book is a treasure trove of very interesting and elegantly presented results on partial differential equations and shows the deep knowledge that one of the most relevant mathematicians of the last century had of the subject. However, as Maz'ya and Shaposhnikova also remarked, it looks dramatically old, even for that time: weak derivatives, Sobolev spaces, and functional analysis are completely missing and the variational approach, the one systematically adopted here, is not even mentioned. Without any doubt, for a young mathematician of that time, the theory of partial differential equations was already a different thing.

Without comparing me to Hadamard, which would be clearly a nonsense, I vaguely feel that today we are in a similar situation. In this book, I put the knowledge of the subject that I achieved and metabolized in 50 years: but it is possible that all this material becomes rapidly old, like the classical theory presented by Hadamard in the 1960s.

---

[5] Maz'ya and Shaposhnikova [20].

[6] Hadamard [13].

Indeed, it is a rather general opinion that the main importance of partial differential equations lies in the fact that they are, since Galileo and Newton, the way in which we have modeled natural phenomena; and being able to solve partial differential equations, both at the theoretical and the numerical level, gives the possibility of finding answers on the behavior of those phenomena. In other words, it opens the road to predict the future.

What we see now arriving is a different paradigm: for getting answers on natural phenomena modeling seems no longer strictly necessary, as artificial neural networks and machine learning methods could furnish an efficient alternative. In front of us, we see a turning point: huge amount of data versus equations. The question is: will these approaches live together, or in about a dozen of years (this period of time could be enough, as changes are running faster than 60 years ago. . .) the new strategy will cancel the old one? “Ai posteri l’ardua sentenza”.[7]

Povo, Italy Alberto Valli
May 2023

[7] A. Manzoni, Il cinque maggio.

Contents

Chapter 1
Introduction

Very often the description that we give of natural phenomena is based on physical laws that express the conservation of some quantity (mass, momentum, energy, ...). In addition, some experimental relations are also taken into account (how the pressure is related to the density, how the heat flux is related to the variation of temperature, ...).

Conservation and variation are thus basic ingredients: in mathematical words, the latter one means *derivatives*. More precisely, very often the description we want to devise involves many variables: therefore we have to play with *partial derivatives* and with *equations* involving unknown quantities and their partial derivatives.

Definition 1.1 A partial differential equation (PDE) is an equation involving an unknown function $u = u(x)$ of two or more variables $x = (x_1, \ldots, x_n)$, $n \geq 2$, and certain of its partial derivatives. An expression of the form

$$F(x_1, \ldots, x_n, u, \mathcal{D}u, \mathcal{D}^2 u, \ldots, \mathcal{D}^k u) = 0$$

is called a kth order PDE, where $k \geq 1$ is an integer and we have denoted by $\mathcal{D}^k u$ a generic partial derivative of order k.

Equivalently, keeping on the left all the terms involving the unknown u and putting on the right all the other terms, we can write a PDE in the form

$$L(x, u) = f \,,$$

where L is called *partial differential operator* and f turns out to be a given datum.

Definition 1.2 A PDE is said to be non-linear if it is not linear.

The reason of this apparently meaningless definition is that we want to enlighten the fact that the crucial point is to understand the definition of what is a linear PDE.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_1

Definition 1.3 A PDE in the form $L(x, u) = f$ is said to be linear if the operator L is linear, i.e., $L(x, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 L(x, w_1) + \alpha_2 L(x, w_2)$ for all $\alpha_1, \alpha_2 \in \mathbb{R}$ and all functions w_1, w_2.

This definition is a little bit inaccurate, as the operator L has not a meaning for *all* functions w: it is necessary that the derivatives appearing in L do exist for these functions.

Definition 1.4 Let the operator L be linear; then the linear equation $Lu = f \neq 0$ is said to be non-homogeneous, while the linear equation $Lu = 0$ it is said to be homogeneous.

We use the notation $\mathcal{D}_i u$ for indicating the partial derivative $\frac{\partial u}{\partial x_i}$. Other equivalent notations are u_{x_i}, $\mathcal{D}_{x_i} u$, $\partial_{x_i} u$.

Remark 1.1 The general form of a linear operator of first order ($k = 1$) is:

$$L(x, w) = \sum_{i=1}^{n} \hat{b}_i(x)\mathcal{D}_i w + a_0(x)w .$$

The general form of a linear operator of second order ($k = 2$) is:

$$L(x, w) = \sum_{i,j=1}^{n} \hat{a}_{ij}(x)\mathcal{D}_i\mathcal{D}_j w + \sum_{i=1}^{n} \hat{b}_i(x)\mathcal{D}_i w + a_0(x)w .$$

We will see in the sequel that very often a second order linear operator will be written in the variational form

$$L(x, w) = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}(x)\mathcal{D}_j w) + \sum_{i=1}^{n} b_i(x)\mathcal{D}_i w + a_0(x)w .$$

Clearly, for smooth coefficients a_{ij} it is easy to return to the previous form.

1.1 Examples of Linear Equations

Transport equation: $u_t + b \cdot \nabla u = f$, where $\nabla = (\mathcal{D}_1, \dots, \mathcal{D}_n)$.
Laplace equation/Poisson equation: $-\Delta u = 0 \,/\, -\Delta u = f$, where $\Delta = \sum_{i=1}^{n} \mathcal{D}_i^2$ is the Laplace operator. A solution u of the Laplace equation is called harmonic function.
Helmholtz equation: $-\Delta u - \omega^2 u = 0$, with $\omega \neq 0$.
Biharmonic equation: $\Delta^2 u = 0$, where $\Delta^2 = \Delta\Delta$.

Heat equation: $u_t - k\Delta u = f$, with $k > 0$ (thermal conductivity). A solution u has an infinite speed of propagation.
Schrödinger equation: $-i\hbar u_t - \frac{\hbar}{2m}\Delta u + Vu = 0$, with $\hbar > 0$ (reduced Planck constant), $m > 0$ (mass).
Wave equation: $u_{tt} - c^2\Delta u = f$, with $c > 0$ (speed of propagation). A solution u has the finite speed of propagation c.
Damped wave equation: $u_{tt} - c^2\Delta u + \sigma u_t = f$, with $c > 0$, $\sigma > 0$.
Klein–Gordon equation: $u_{tt} - c^2\Delta u + \frac{m^2c^4}{\hbar^2}u = 0$, with $c > 0$, $\hbar > 0$, $m > 0$.
Telegraph equation: $u_{tt} - \tau^2 u_{xx} + d_1 u_t + d_2 u = 0$, with $\tau > 0$, $d_1 > 0$, $d_2 > 0$ (the three constants being related to resistance, inductance, capacitance, conductance).
Plate equation: $u_{tt} + \Delta^2 u = f$.

1.2 Examples of Non-linear Equations

Burgers equation: $u_t + uu_x = \varepsilon u_{xx}$ (viscous: $\varepsilon > 0$; inviscid: $\varepsilon = 0$).
Korteweg–de Vries equation: $u_t + cuu_x + u_{xxx} = 0$, with $c \neq 0$.
Cahn–Hilliard equation: $u_t + \nu\Delta^2 u - \Delta(\beta u^3 - \alpha u) = 0$, with $\nu > 0$, $\alpha > 0$, $\beta > 0$.
Minimal surface equation: $\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0$, where $\operatorname{div} w = \nabla \cdot w = \sum_{i=1}^n \mathcal{D}_i w_i$.
Monge–Ampere equation: $\det(\mathcal{H}u) = f(x, u, \nabla u)$, where $\mathcal{H}$ is the Hessian matrix of second order derivatives.

1.3 Examples of Systems

Elasticity system: $-\mu\Delta u - \nu\nabla \operatorname{div} u = f$, where $\mu > 0$, $\nu > 0$ (Lamé coefficients).
Incompressible Navier–Stokes/Euler system:

$$\begin{cases} \partial_t u + (u \cdot \nabla)u - \nu\Delta u + \nabla p = f \\ \operatorname{div} u = 0 \qquad \text{(incompressibility condition)}\,, \end{cases}$$

where $(u \cdot \nabla)u$ is the vector with components $[(u \cdot \nabla)u]_i = \sum_{j=1}^n u_j \mathcal{D}_j u_i$, and $\nu > 0$ (viscosity per unit density) for Navier–Stokes, $\nu = 0$ for Euler.

Compressible Navier–Stokes/Euler system (barotropic case):

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho u) = 0 \\ \rho\big(\partial_t u + (u \cdot \nabla)u\big) - \mu \Delta u - (\zeta + \frac{n-2}{n}\mu)\nabla \operatorname{div} u + \nabla P = \rho f \\ P = p_*(\rho) \qquad \text{(barotropic condition)}\,, \end{cases}$$

where $\mu > 0$ (kinematic viscosity) and $\zeta > 0$ (bulk viscosity) for Navier–Stokes, $\mu = 0$ and $\zeta = 0$ for Euler.

Maxwell system:

$$\begin{cases} \partial_t B + \operatorname{curl} E = 0 \quad , & \operatorname{div} B = 0 \\ \partial_t D - \operatorname{curl} H = -J_e \quad , & \operatorname{div} D = \rho \\ B = \mu H \\ D = \epsilon E\,, \end{cases}$$

where $\mu > 0$ (magnetic permeability), $\epsilon > 0$ (electric permittivity),

$$\operatorname{curl} E = \nabla \times E = \det \begin{bmatrix} i & j & k \\ \mathcal{D}_1 & \mathcal{D}_2 & \mathcal{D}_3 \\ E_1 & E_2 & E_3 \end{bmatrix} .$$

Eddy current system:

$$\begin{cases} \partial_t B + \operatorname{curl} E = 0 \quad , & \operatorname{div} B = 0 \\ \operatorname{curl} H = J \quad , & \chi_I \operatorname{div} D = \rho \\ B = \mu H \\ D = \epsilon E \\ J = \chi_C \sigma E + J_e\,, \end{cases}$$

where $\sigma > 0$ (electric conductivity), χ_I and χ_C are the characteristic functions of Q_I and Q_C, respectively, and Q_I and Q_C are two subsets which furnish a splitting of the whole domain. This is an approximation of Maxwell system for slow varying electromagnetic fields.

1.4 Exercises

Exercise 1.1 Write the Poisson equation $-\Delta u = f$ as a first order system in terms of u and $q = -\nabla u$.

Solution Since $\Delta u = \mathrm{div}\nabla u$, we have

$$\begin{cases} q + \nabla u = 0 \\ \mathrm{div}\, q = f\,. \end{cases}$$

The problem in this form will be analyzed in Chap. 8 (see Example 8.3).

Exercise 1.2

(i) Write the wave equation $u_{tt} - c^2 \Delta u = f$ as a first order system in terms of $w = u_t$ and $q = c\nabla u$.
(ii) Note that the first order system obtained in (i) can be written as a *symmetric* system of the form $U_t + \sum_{j=1}^n A_i \mathcal{D}_i U = F$, with $A_i = A_i^T$ and $U = (w, q_1, \ldots, q_n)$.

Solution

(i) Setting $w = u_t$ and $q = c\nabla u$ we have $w_t = u_{tt} = c^2 \mathrm{div}\nabla u + f = c\,\mathrm{div}\, q + f$. Moreover, $q_t = c\nabla u_t = c\nabla w$. Thus we have obtained

$$\begin{cases} w_t - c\,\mathrm{div}\, q = f \\ q_t - c\nabla w = 0\,. \end{cases}$$

Since $-\mathrm{div}$ is the (formal) adjoint of ∇, the above system has the anti-symmetric form

$$U_t + c \begin{pmatrix} 0 & \nabla^T \\ -\nabla & 0 \end{pmatrix} U = F\,,$$

with $U = (w, q_1, \ldots, q_n)$, $F = (f, 0, \ldots, 0)$. This can be a little bit surprising when looking at the second part of the exercise.

(ii) Expanding the expressions just derived, it is straightforward to check that the first order system obtained in (i) can be written as $U_t + \sum_{j=1}^n A_i \mathcal{D}_i U = F$ with U and F as above and A_i, $i = 1, \ldots, n$, the $(n+1) \times (n+1)$ symmetric matrices given by

$$A_1 = \begin{pmatrix} 0 & -c & 0 & \ldots & 0 \\ -c & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}, \quad \ldots\,, \quad A_n = \begin{pmatrix} 0 & 0 & 0 & \ldots & -c \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -c & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Exercise 1.3

(i) Determine the second order system that is obtained for the electric field E by applying the backward Euler scheme to the Maxwell system (assume that μ and ϵ are constants).
(ii) Determine the second order system that is obtained for the magnetic field H by applying the backward Euler scheme to the Maxwell system (assume that μ and ϵ are constants).
(iii) Note that the two systems have the same structure $\operatorname{curl}\operatorname{curl} + \alpha I$, with $\alpha > 0$.

Solution

(i) Approximating the time derivatives by the difference quotients

$$\partial_t B \approx \frac{B - B^{old}}{\tau} \quad , \quad \partial_t D \approx \frac{D - D^{old}}{\tau}$$

where $\tau > 0$ is the time step, and remembering that $B = \mu H$ and $D = \epsilon E$ we find

$$\begin{cases} \mu H + \tau \operatorname{curl} E = B^{old} \\ \epsilon E - \tau \operatorname{curl} H = -\tau J + D^{old} . \end{cases} \tag{1.1}$$

Applying the curl operator to the first equation and using the second equation for expressing $\operatorname{curl} H$ we easily find

$$\operatorname{curl}\operatorname{curl} E + \frac{\mu\epsilon}{\tau^2} E = \frac{1}{\tau} \operatorname{curl} B^{old} + \frac{\mu}{\tau^2} D^{old} - \frac{\mu}{\tau} J .$$

(ii) Applying the curl operator to the second equation in (1.1) and using the first equation in (1.1) for expressing $\operatorname{curl} E$ we have

$$\operatorname{curl}\operatorname{curl} H + \frac{\mu\epsilon}{\tau^2} H = -\frac{1}{\tau} \operatorname{curl} D^{old} + \frac{\epsilon}{\tau^2} B^{old} + \operatorname{curl} J .$$

(iii) Evident from (i) and (ii).

Exercise 1.4 Let u be a smooth solution in $\mathbb{R}^3$ of the equation $u - \nabla \operatorname{div} u = f$.

(i) Show that $\operatorname{div} u$ is a solution in $\mathbb{R}^3$ of the equation $p - \Delta p = \operatorname{div} f$.
(ii) If $\operatorname{curl} f = 0$ in $\mathbb{R}^3$, show that $u = \nabla\psi$ for a suitable function ψ.
(iii) If $\operatorname{curl} f = 0$ in $\mathbb{R}^3$, $\operatorname{div} f = 0$ in $\mathbb{R}^3$ and the derivatives of u decay fast enough at infinity, say, $|\operatorname{div} u| + |\nabla \operatorname{div} u| \le C_* |x|^{-\alpha}$ for $\alpha > \frac{3}{2}$ and $|x| \ge q_*$ large enough, then $u = \nabla\psi$ for a suitable harmonic function ψ.

Solution

(i) Taking into account that $\mathrm{div}\nabla\mathrm{div} = \Delta\mathrm{div}$, the result follows at once by applying the div operator to the equation.

(ii) Taking into account that $\mathrm{curl}\,\nabla = 0$, applying the curl operator to the equation we find $\mathrm{curl}u = \mathrm{curl}\,f = 0$. Since $\mathbb{R}^3$ is a simply-connected domain, we deduce that there exists a function ψ such that $u = \nabla\psi$ in $\mathbb{R}^3$.

(iii) Multiply the equation $\mathrm{div}u - \Delta\mathrm{div}u = \mathrm{div}\,f = 0$ by $\mathrm{div}u$ and integrate over the ball $B_s = \{x \in \mathbb{R}^3 \mid |x| < s\}$, $s > q_*$. It holds

$$\begin{aligned} 0 &= \int_{B_s} [(\mathrm{div}u)^2 - (\Delta\mathrm{div}u)\mathrm{div}u]dx \\ &= \int_{B_s} [(\mathrm{div}u)^2 + \nabla\mathrm{div}u \cdot \nabla\mathrm{div}u]dx - \int_{\partial B_s} \nabla\mathrm{div}u \cdot n\, \mathrm{div}u\, dS_x\,, \end{aligned} \tag{1.2}$$

where we have used the integration by parts formula (C.5). The boundary integral can be estimated as follows

$$\Big|\int_{\partial B_s} \nabla\mathrm{div}u \cdot n\, \mathrm{div}u\, dS_x\Big| \le C_* s^{-2\alpha} 4\pi s^2\,,$$

and moreover

$$\begin{aligned} \textstyle\int_{B_s} (\mathrm{div}u)^2 dx &= \underbrace{\int_{B_{q*}} (\mathrm{div}u)^2 dx}_{=C_0} + \int_{B_s\setminus B_{q*}} (\mathrm{div}u)^2 dx \\ &\le C_0 + C_* \int_{B_s\setminus B_{q*}} |x|^{-2\alpha} dx = C_0 + 4\pi C_* \int_{q*}^{s} r^2 r^{-2\alpha} dr \le Q_0\,, \end{aligned}$$

where Q_0 is independent of $s > q_*$, as $\alpha > \frac{3}{2}$. Similarly,

$$\int_{B_s} |\nabla\mathrm{div}u|^2 dx \le Q_1\,.$$

Passing to the limit as $s \to +\infty$ in (1.2) we find

$$\int_{\mathbb{R}^3} [(\mathrm{div}u)^2 + |\nabla\mathrm{div}u|^2] dx = 0\,,$$

therefore $\mathrm{div}u = 0$ in $\mathbb{R}^3$. Since from (ii) we already know that $u = \nabla\psi$, it follows that $\mathrm{div}\nabla\psi = \Delta\psi = 0$ in $\mathbb{R}^3$.

Exercise 1.5 Let u be a smooth solution in $\mathbb{R}^3$ of the equation $u + \mathrm{curl}\,\mathrm{curl}u = f$.

(i) Show that $\mathrm{curl}u$ is a solution in $\mathbb{R}^3$ of the equation $q + \mathrm{curl}\,\mathrm{curl}q = \mathrm{curl}\,f$.

(ii) If $\mathrm{div}\,f = 0$ in $\mathbb{R}^3$, show that $u = \mathrm{curl}\Psi$ for a suitable function Ψ.

(iii) If $\text{curl} f = 0$ in $\mathbb{R}^3$, $\text{div} f = 0$ in $\mathbb{R}^3$ and the derivatives of u decay fast enough at infinity, say, $|\text{curl} u| + |\text{curl}\,\text{curl} u| \le C_* |x|^{-\alpha}$ for $\alpha > \frac{3}{2}$ and $|x| \ge q_*$ large enough, then $u = \text{curl}\Psi$ for a suitable function Ψ that satisfies $\text{curl}\,\text{curl}\Psi = 0$.

Solution

(i) This a sort of "curl" version of the previous exercise. The first result follows at once by applying the curl operator to the equation.

(ii) Taking into account that div curl = 0, applying the div operator to the equation we find $\text{div} u = \text{div} f = 0$. It is well-known that this condition $\mathbb{R}^3$ is equivalent to the fact that there exists a function Ψ such that $u = \text{curl}\Psi$ in $\mathbb{R}^3$.

Note that, if we know that the vector potential Ψ decays sufficiently fast at infinity, we can apply the classical Helmholtz decomposition and write $\Psi = \nabla\phi + \text{curl} Q$. Thus $\Psi_\star = \Psi - \nabla\phi$ satisfies $\text{curl}\Psi_\star = u$ and $\text{div}\Psi_\star = 0$: in other words, we have found a divergence free vector potential $\Psi_\star$.

(iii) Take the scalar product of the equation $\text{curl} u + \text{curl}\,\text{curl}\,\text{curl} u = \text{curl} f = 0$ by $\text{curl} u$ and integrate over the ball $B_s = \{x \in \mathbb{R}^3 \mid |x| < s\}$, $s > q_*$. It holds

$$\begin{aligned} 0 &= \int_{B_s} [|\text{curl} u|^2 + \text{curl}\,\text{curl}\,\text{curl} u \cdot \text{curl} u] dx \\ &= \int_{B_s} [|\text{curl} u|^2 + \text{curl}\,\text{curl} u \cdot \text{curl}\,\text{curl} u] dx \\ &\quad - \int_{\partial B_s} n \times \text{curl}\,\text{curl} u \cdot \text{curl} u \, dS_x \,, \end{aligned} \tag{1.3}$$

where we have used the integration by parts formula (C.8). The boundary integral can be estimated as follows

$$\left| \int_{\partial B_s} n \times \text{curl}\,\text{curl} u \cdot \text{curl} u \, dS_x \right| \le C_* |s|^{-2\alpha} 4\pi s^2 \,,$$

and moreover

$$\begin{aligned} \textstyle\int_{B_s} |\text{curl} u|^2 dx &= \underbrace{\int_{B_{q*}} |\text{curl} u|^2 dx}_{= C_0} + \int_{B_s \setminus B_{q*}} |\text{curl} u|^2 dx \\ &\le C_0 + C_* \int_{B_s \setminus B_{q*}} |x|^{-2\alpha} dx = C_0 + 4\pi C_* \int_{q*}^{s} r^2 r^{-2\alpha} dr \le Q_0 \,, \end{aligned}$$

where Q_0 is independent of $s > q_*$, as $\alpha > \frac{3}{2}$. Similarly,

$$\int_{B_s} |\text{curl}\,\text{curl} u|^2 dx \le Q_1 \,.$$

Passing to the limit as $s \to +\infty$ in (1.3) we find

$$\int_{\mathbb{R}^3} (|\mathrm{curl}\, u|^2 + |\mathrm{curl}\,\mathrm{curl}\, u|^2)\, dx = 0\,,$$

therefore $\mathrm{curl}\, u = 0$ in $\mathbb{R}^3$. Since from (ii) we already know that $u = \mathrm{curl}\,\Psi$, it follows that $\mathrm{curl}\,\mathrm{curl}\,\Psi = 0$ in $\mathbb{R}^3$.

As in case (ii), if we know that the vector potential Ψ decays sufficiently fast at infinity, we can modify it and find a vector potential $\Psi_\star$ such that $\mathrm{curl}\,\Psi_\star = u$ and $\mathrm{div}\,\Psi_\star = 0$. Thus

$$0 = \mathrm{curl}\,\mathrm{curl}\,\Psi_\star = -\Delta\Psi_\star + \nabla \mathrm{div}\,\Psi_\star = -\Delta\Psi_\star\,;$$

in other words, all the components of $\Psi_\star$ are harmonic functions.

Chapter 2
Second Order Linear Elliptic Equations

This chapter is concerned with a general presentation of second order linear elliptic equations and of some of the most popular boundary value problems associated to them (Dirichlet, Neumann, mixed, Robin).

Before introducing the concept of weak solution and of weak formulation we briefly describe the general ideas behind two quite classical methods for finding the solution of partial differential equations: the Fourier series expansion in terms of an orthonormal basis given by the eigenvectors of the operator, and the representation of the solution by integral formulas, using the fundamental solution of the operator as integral kernel.

The approach leading to the weak formulation is then described without giving all the technical details, but only trying to specify which steps are needed for obtaining the desired result. Though the complete functional framework is not yet clarified, nonetheless we end the chapter with the proof of the fundamental existence and uniqueness result: the Lax–Milgram theorem.

2.1 Elliptic Equations

In this chapter we will study the boundary value problem

$$\begin{cases} Lu = f & \text{in } D \\ \text{BC} & \text{on } \partial D \, , \end{cases} \tag{2.1}$$

where D is an open, connected and bounded subset of $\mathbb{R}^n$, $u : \overline{D} \mapsto \mathbb{R}$ is the unknown, and BC stands for "boundary condition". Here $f : D \mapsto \mathbb{R}$ is given and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_2

L denotes a second order partial differential operator having the form

$$Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w) + \sum_{i=1}^{n} b_i \mathcal{D}_i w + a_0 w \,. \tag{2.2}$$

The second order term $-\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w)$ is called the *principal part* of L. The reason of the (mysterious) minus sign will be clear in the sequel (see Remark 2.3).

Remark 2.1 In physical models, u in general represents the density of some quantity, for instance a chemical concentration. In the operator L, the principal part represents the diffusion of u within D. The first order term represents advection (transport) of u within D. The term of order zero describes the local reactions that occur in D.

We will focus on four different types of boundary condition (here below n is the unit outward normal vector on ∂D):

Dirichlet BC : $u = 0$ on ∂D [homogeneous case].

Neumann BC : $\sum_{i,j=1}^{n} n_i a_{ij}\mathcal{D}_j u = g$ on ∂D.

Mixed BC : $u = 0$ on Γ_D and $\sum_{i,j=1}^{n} n_i a_{ij}\mathcal{D}_j u = g$ on Γ_N, where $\partial D = \overline{\Gamma_D} \cup \overline{\Gamma_N}$, $\Gamma_D \cap \Gamma_N = \emptyset$ [homogeneous case on Γ_D].

Robin BC : $\sum_{i,j=1}^{n} n_i a_{ij}\mathcal{D}_j u + \kappa u = g$ on ∂D, where $\kappa \geq 0$ almost everywhere (a.e. henceforth) on ∂D and $\int_{\partial D} \kappa dS_x \neq 0$.

Remark 2.2 In the case of a non-homogeneous Dirichlet boundary condition

$$u = u_\sharp \quad \text{on } \partial D$$

(and, similarly, of the non-homogeneous mixed boundary condition $u = u_\sharp$ on Γ_D) we proceed as follows:

1. find $\widehat{u} : \overline{D} \mapsto \mathbb{R}$ such that $\widehat{u}_{|\partial D} = u_\sharp$;
2. setting $\omega = u - \widehat{u}$, we see that $\omega_{|\partial D} = 0$ and $L\omega = Lu - L\widehat{u} = f - L\widehat{u}$. Then the second step is: find ω, a solution of the homogeneous Dirichlet boundary value problem $L\omega = f - L\widehat{u}$, $\omega_{|\partial D} = 0$;
3. finally define $u = \omega + \widehat{u}$.

For arriving at the definition of elliptic equation we need now to give a deeper look at the matrix $\{a_{ij}(x)\}_{i,j=1}^{n}$ of the coefficients of the principal part of L.

Definition 2.1 A (real) matrix A is said to be positive definite if $Av \cdot v > 0$ for every $v \in \mathbb{R}^n$, $v \neq 0$.

Exercise 2.1 A matrix A is positive definite if and only if it exists $\alpha > 0$ such that $Av \cdot v \geq \alpha |v|^2$ for every $v \in \mathbb{R}^n$.

Exercise 2.2 Consider a positive definite matrix A (thus satisfying $Av \cdot v \geq \alpha |v|^2$ for every $v \in \mathbb{R}^n$, for a suitable $\alpha > 0$). Then the real part of an eigenvalue of A is greater than or equal to α; in particular, a positive definite matrix is non-singular.

Exercise 2.3

(i) A matrix A is positive definite if and only if $\frac{A+A^T}{2}$ is positive definite.
(ii) A matrix A is positive definite if and only if all the eigenvalues λ_i of $\frac{A+A^T}{2}$ are strictly positive.

Definition 2.2 The partial differential operator L is said to be (uniformly) elliptic in D if the matrix $\{a_{ij}(x)\}_{i,j=1}^n$ is (uniformly) positive definite, i.e., if there exists a constant $\alpha_0 > 0$ such that

$$\sum_{i,j=1}^n a_{ij}(x)\eta_j\eta_i \geq \alpha_0|\eta|^2$$

for almost all $x \in D$, for every $\eta \in \mathbb{R}^n$.

Exercise 2.4

(i) Show that the operator

$$Lw = -\mathcal{D}_1\big((1 + x_1x_2)\mathcal{D}_1 w\big) - \mathcal{D}_1(x_1\mathcal{D}_2 w) - \mathcal{D}_2(x_2\mathcal{D}_1 w) - \mathcal{D}_2\mathcal{D}_2 w \, ,$$

is uniformly elliptic in $D = \{x \in \mathbb{R}^2 \,|\, 0 < x_1 < 1/2, 0 < x_2 < 1\}$.
(ii) Show that the operator $Lw = -\sum_{i,j=1}^3 D_i(a_{ij}D_j w)$, with

$$\{a_{ij}\} = \begin{pmatrix} 1 & -x_3 & x_2 \\ x_3 & 1+x_1^2 & x_1 \\ -x_2 & x_2 & 1+x_3^2 \end{pmatrix}$$

is uniformly elliptic in $D = \{x \in \mathbb{R}^3 \,|\, |x| < 1\}$.

2.2 Weak Solutions

Before speaking about a different idea of what is the solution of a partial differential equation, let us spend a few words about a couple of "classical" approaches concerning this question (say, in use throughout nineteenth century and after).

2.2.1 Two Classical Approaches

A first approach is based on series expansion. Suppose we want to solve the problem

$$\begin{cases} -\Delta u = f & \text{in } D \\ u_{|\partial D} = 0 & \text{on } \partial D\,, \end{cases} \tag{2.3}$$

and we have a countable basis $\{\omega_k\}_{k=1}^{\infty}$, with $\omega_k : \overline{D} \mapsto \mathbb{R}$ and $\omega_{k|\partial D} = 0$. We can expand u and f as $u = \sum_{k=1}^{\infty} u_k \omega_k$ and $f = \sum_{k=1}^{\infty} f_k \omega_k$, with $u_k, f_k \in \mathbb{R}$, and impose equation (2.3)$_1$ (we are not making precise here in which sense these series are convergent…). This formally gives

$$\sum_{k=1}^{\infty} f_k \omega_k = f = -\Delta u = \sum_{k=1}^{\infty} u_k(-\Delta\omega_k)\,. \tag{2.4}$$

Expanding also $-\Delta\omega_k$ (and admitting that this is possible…) we find

$$-\Delta\omega_k = \sum_{j=1}^{\infty} q_j^k \omega_j\,,$$

and inserting this result in (2.4) we obtain

$$\sum_{j=1}^{\infty} f_j \omega_j = \sum_{k=1}^{\infty} u_k \Big(\sum_{j=1}^{\infty} q_j^k \omega_j\Big) = \sum_{j=1}^{\infty} \Big(\sum_{k=1}^{\infty} q_j^k u_k\Big)\omega_j\,. \tag{2.5}$$

Thus we have to solve the infinite dimensional linear system

$$\sum_{k=1}^{\infty} q_j^k u_k = f_j \quad , \quad j = 1, 2, \ldots \tag{2.6}$$

This simplifies a lot if ω_k are eigenvectors of the $-\Delta$ operator: $-\Delta\omega_k = \lambda_k\omega_k$, with $\lambda_k \in \mathbb{R}$ the associated eigenvalues. In this case the coefficients q_j^k have to satisfy

$$\lambda_k\omega_k = \sum_{j=1}^{\infty} q_j^k \omega_j\,,$$

hence we infer

$$q_j^k = \lambda_k \delta_{kj} \quad , \quad k, j = 1, 2, \ldots$$

where δ_{kj} is the Kronecker symbol, defined by $\delta_{kj} = 0$ if $k \neq j$, $\delta_{kj} = 1$ if $k = j$. Then (2.6) can be easily solved by setting

$$u_j = \frac{f_j}{\lambda_j} \ , \quad j = 1, 2, \dots$$

provided that $\lambda_j \neq 0$. In particular, if the eigenvectors ω_j are an orthonormal basis with respect to some scalar product $(\cdot, \cdot)$, one has $f_j = (f, \omega_j)$, the classical Fourier coefficients.

We have thus solved the problem via *Fourier series expansion*. This procedure requires that we are able to find an orthonormal basis given by eigenvectors of the operator which satisfy the boundary condition. Clearly, one has to check that the formal procedure we have described can be rigorously justified: the series expansions hold, the series can be differentiated term by term, the eigenvalues λ_j are different from 0. Some answers concerning these points can be found in Sect. 7.2.

The following exercise furnishes an example of orthonormal system of eigenvectors in $L^2(D)$ (the proof that it is a orthonormal basis, namely, that any function $f \in L^2(D)$ can be expressed by a convergent Fourier series requires some additional work: for this, see Theorem 7.7):

Exercise 2.5 Consider $D = (0, a) \times (0, b)$. Determine the eigenvalues and the eigenvectors associated to the operator $-\Delta$ with homogeneous Dirichlet boundary condition, and verify that, after a suitable normalization, the eigenvectors are an orthonormal system in $L^2(D)$. [Hint: use the method of separation of variables.]

Still referring to problem (2.3), a second approach we want to describe is the following: suppose we know a function $K(x, \xi) : \overline{D} \times \overline{D} \mapsto \mathbb{R}$ satisfying, for $x \in D$,

$$\int_D (-\Delta_x K)(x, \xi) f(\xi) d\xi = f(x) \,. \tag{2.7}$$

Before proceeding, let us see in which way such a function K could be determined. Fix $x \in D$ and for $m \geq 1$ set

$$\rho_m(\xi; x) = \frac{1}{\text{meas}(B(x, \frac{1}{m}))} \chi_{B(x, \frac{1}{m})}(\xi) \,,$$

where $B(x, \frac{1}{m}) = \{\xi \in \mathbb{R}^n \mid |x - \xi| < \frac{1}{m}\}$ and $\chi_{B(x, \frac{1}{m})}$ is the characteristic function of $B(x, \frac{1}{m})$. It is readily verified that $\int_{B(x,t)} \rho_m(\xi; x) d\xi = 1$ for each $t > 0$ and $m > 1/t$. Moreover, it is well-known that, if f is continuous at x, then

$$\lim_{m \to \infty} \int_D \rho_m(\xi; x) f(\xi) d\xi = f(x) \,.$$

Thus one could try to find a function $K(x,\xi)$ such that $-(\Delta_x K)(x,\xi) = -(\Delta_\xi K)(\xi,x)$ and

$$-(\Delta_\xi K)(\xi,x) = \lim_{m\to\infty} \rho_m(\xi;x)\,.$$

Clearly, the weak point here is that $\lim_{m\to\infty} \rho_m(\xi;x) = 0$, in the pointwise sense for all $\xi \neq x$, and moreover in the limit the condition saying that the average on $B(x,t)$ is equal to 1 is lost. A surrogate of this choice can be to look for $K(x,\xi)$ such that $-(\Delta_x K)(x,\xi) = -(\Delta_\xi K)(\xi,x) = 0$ for $\xi \neq x$ and satisfying

$$-\int_{\partial B(x,t)} (\nabla_\xi K)(\xi,x)\cdot n(\xi) dS_\xi = 1\,,$$

where n is the unit outward normal vector on $\partial B(x,t)$. The reason of this condition is that by the divergence theorem (see Theorem C.3) we have $-\int_{B(x,t)} \Delta g(\xi)d\xi = -\int_{\partial B(x,t)} \nabla g(\xi)\cdot n(\xi)dS_\xi$ for a smooth function g.

This procedure is indeed feasible (in Exercise 2.6 we give an example of the construction of a function with these two properties: which however is just the starting point for saying that (2.7) is satisfied in some suitable sense).

Exercise 2.6

(i) Find a function $K_0 = K_0(\xi)$ defined in $\mathbb{R}^2 \setminus \{0\}$ and such that

$$-\Delta K_0 = 0 \text{ in } \mathbb{R}^2 \setminus \{0\} \quad \text{and} \quad -\int_{\partial B(0,t)} \nabla K_0 \cdot n dS_\xi = 1$$

for any $t > 0$. [Hint: look for a radial function $K_0 = K_0(|\xi|)$.]

(ii) Verify that a function $K(x,\xi)$ satisfying $-(\Delta_x K)(x,\xi) = -(\Delta_\xi K)(\xi,x) = 0$ for $\xi \neq x$ and $-\int_{\partial B(x,t)} (\nabla_\xi K)(\xi,x)\cdot n(\xi)dS_\xi = 1$ for each $t > 0$ is given by $K(x,\xi) = K_0(|x-\xi|)$.

Let us go back to (2.7). Being $K(x,\xi)$ available, we set

$$u(x) = \int_D K(x,\xi) f(\xi) d\xi \tag{2.8}$$

and proceeding formally from (2.7) we have

$$-\Delta u(x) = \int_D (-\Delta_x K)(x,\xi) f(\xi)d\xi = f(x)\,.$$

What is missing is the fact that u satisfies the boundary condition. This difficulty can be overcome if we know a function $G(x,\xi) : \overline{D}\times\overline{D} \mapsto \mathbb{R}$ satisfying (2.7) and also $G(x, \xi)_{|x \in \partial D} = 0$ for each $\xi \in \overline{D}$. Then setting

$$u(x) = \int_D G(x, \xi) f(\xi) d\xi$$

furnishes a solution of (2.3).

The possibility of finding the function K introduced above depends on the properties of the operator $-\Delta$, while the possibility of finding G also depends on the properties of the domain D. Therefore, it could be useful to devise a procedure only based on the knowledge of K. Given a function $v : \overline{D} \mapsto \mathbb{R}$, by integration by parts (see Theorem C.2) we obtain

$$\begin{aligned} &\int_D (-\Delta_\xi v)(\xi) K(\xi, x) d\xi - \int_D v(\xi)(-\Delta_\xi K)(\xi, x) d\xi \\ &\quad = \int_{\partial D} \Big(- \nabla_\xi v(\xi) \cdot n(\xi) K(\xi, x) + v(\xi) \nabla_\xi K(\xi, x) \cdot n(\xi) \Big) dS_\xi \, . \end{aligned} \tag{2.9}$$

If $K(x, \xi) = K(\xi, x)$, so that $(\Delta_\xi K)(\xi, x) = (\Delta_x K)(x, \xi)$, and we select $v = u$, where u satisfies $-\Delta u = f$ in D, from (2.9) and (2.7) we find for $x \in D$

$$\begin{aligned} &\int_D f(\xi) K(\xi, x) d\xi - u(x) \\ &\quad = \int_{\partial D} \Big(- \nabla_\xi u(\xi) \cdot n(\xi) K(\xi, x) + u(\xi) \nabla_\xi K(\xi, x) \cdot n(\xi) \Big) dS_\xi \, . \end{aligned} \tag{2.10}$$

This is a representation formula for $u(x)$, $x \in D$, in terms of K, f and the values of $\nabla u \cdot n$ and u on the boundary ∂D. If we are considering the Dirichlet or the Neumann boundary value problems, on the boundary ∂D we know only one of the two functions $\nabla u \cdot n$ and u: thus we cannot conclude our argument. But if a similar formula can be obtained for $x \in \partial D$ (to be more precise, what it is known to hold is the same formula with the only modification given by the replacement at the left hand side of $u(x)$ with $p(x)u(x)$, for a suitable function p), and we assume that u is a solution of the Dirichlet boundary value problem with boundary datum $u_\sharp$, then we finally obtain

$$\begin{aligned} \int_{\partial D} \nabla_\xi u(\xi) \cdot n(\xi) K(\xi, x) dS_\xi &= - \int_D f(\xi) K(\xi, x) d\xi + p(x) u_\sharp(x) \\ &\quad + \int_{\partial D} u_\sharp(\xi) \nabla_\xi K(\xi, x) \cdot n(\xi) dS_\xi \quad , \quad x \in \partial D \, . \end{aligned} \tag{2.11}$$

This is a boundary integral equation for the boundary unknown $\nabla u \cdot n$. If we are able to solve it, we can put the obtained value of $\nabla u \cdot n$ in (2.10) and we have found a representation formula for the solution $u(x)$, $x \in D$. Note that a similar dual result is obtained if we assume that u satisfies the Neumann boundary condition: in that case the unknown function of the boundary integral equation is $u_{|\partial D}$, while $(\nabla u \cdot n)_{|\partial D}$ becomes a known datum.

With this procedure we have thus transformed the original boundary value problem into a boundary integral equation. Also in this case we need to show that this formal process gives indeed the solution we are looking for. This means that we have to show that all the integrals appearing in (2.10) and (2.11) have a meaning, that the function given by (2.10) is differentiable as many times as we need and satisfies the equation, and that as $x \to \hat{x} \in \partial D$ the given boundary condition is achieved at $\hat{x}$.

The theory related to this method is called *potential theory*: indeed, the function $x \to K(x, \xi)$, up to a normalization, is the potential of the electric field generated by a point charge placed at ξ. The function $K(x, \xi)$ satisfying (2.7) is called the *fundamental solution* of the partial differential operator (in our presentation, of the operator $-\Delta$). A classical (and a little bit old fashioned) reference on this topic is the textbook by Kellogg [14] (originally printed in 1929, and several times reprinted); for a more recent one see McLean [21].

2.2.2 An Infinite Dimensional Linear System?

When it is looked from far enough, a linear partial differential equation is essentially an infinite dimensional linear system:

- the solution we look for is a function, thus an object depending on infinitely many independent information (say, its values in all the points of the domain where it is defined, or, in more specific cases, the coefficients of a series expansion which represents it);
- the relations between these unknowns are expressed by a linear operator.

Therefore it could be reasonable trying to extend to the infinite dimensional case the theory of existence and uniqueness that is known for a linear system of m equations with m unknowns. This problem can always be associated to a square $m \times m$-matrix, say Q, and takes the form

$$Qq = p \,, \tag{2.12}$$

with $q, p \in \mathbb{R}^m$.

From linear algebra we know various necessary and sufficient conditions that imply existence and uniqueness of a solution q for any given p. By far the most famous is assuming that Q is non-singular, namely, that $\det Q \neq 0$. Unfortunately, when Q has infinitely many rows and columns, it does not seem so easy to translate this condition in something of simple use.

From a more abstract perspective a simple answer is that the problem is well-posed if and only if the map $r \mapsto Qr$, $r \in \mathbb{R}^m$, is one-to-one and onto. In this respect a simplification indeed occurs: in fact the so-called "rank–nullity" theorem states that

$$\dim N(Q) + \dim R(Q) = m \,, \tag{2.13}$$

where $N(Q) = \{v \in \mathbb{R}^m \mid Qv = 0\}$ and $R(Q) = \{Qv \in \mathbb{R}^m \mid v \in \mathbb{R}^m\}$ are the kernel and the range of Q, respectively. Therefore it follows that $N(Q) = \{0\}$ (namely, the map $r \mapsto Qr$ is one-to-one) implies $R(Q) = \mathbb{R}^m$ (namely, the map $r \mapsto Qr$ is onto) and vice versa: in other words, from uniqueness one obtains existence and vice versa. However, also in this case a direct extension to the infinite dimensional case of the "rank–nullity" theorem does not seem immediate, as an equation like (2.13) loses its meaning when $m = +\infty$.

Instead, another interesting and well-known result seems to be much more promising: the characterization of the range of Q given by $R(Q) = N(Q^T)^\perp$, where $N(Q^T)^\perp$ denotes the subspace orthogonal to $N(Q^T)$ (see Exercise 7.2). Here we are not playing with infinite quantities, but with simple space relations. In particular, existence and uniqueness follows from the two conditions $N(Q) = \{0\}$ and $R(Q)^\perp = N(Q^T) = \{0\}$.

For obtaining that the kernels of Q and Q^T are trivial it is sufficient that Q is positive definite, as this implies that Q^T is positive definite, too. Clearly, we already know that a positive definite matrix is non-singular (see Exercise 2.2); but here we are interested in conditions that can have a simple extension to the infinite dimensional case, continuing to be sufficient for existence and uniqueness also in that case. We will see in Sect. 2.3 that, with a slight modification, the condition Q positive definite will be the right one.

Let us conclude this section with two additional remarks about the strategy described above. The first one is that in an infinite dimensional vector space the range of a linear and bounded operator is not always closed (see Sect. 3.1, item 5), and that the correct relation between the range of Q and the kernel of Q^T is $\overline{R(Q)} = N(Q^T)^\perp$ (see Exercise 7.3). Therefore it will be necessary to find conditions assuring that the range of Q is closed. What is nice here is that a positiveness condition is also sufficient for this result.

The second remark is that a linear differential operator typically does not act from a vector space V into itself, but from V into its dual space (i.e., the space of linear and bounded functionals from V to $\mathbb{R}$), that will be denoted by V'. Clearly, by using the Riesz representation theorem 3.1 one could go back to an operator from V to V: this will be done, for instance, in the proof of the Lax–Milgram theorem 2.1. However, if this step is not performed, we have at hand an operator Q from V to V', and therefore the relations between the range and the kernel must be reconsidered, as $R(Q) \subset V'$ while $N(Q^T) \subset V$; the orthogonal subspace $N(Q^T)^\perp$ will be replaced by the polar set $N(Q^T)_\sharp$ (see Theorem 8.4).

In conclusion, we have seen that the following conditions seem to be sufficient for the existence and uniqueness of the solution of an infinite dimensional linear problem $Qq = p$:

- $N(Q) = \{0\}$
- $R(Q)$ is closed
- $N(Q^T) = \{0\}$.

We are saying "seem" instead of "are" as we still have to clarify which properties the infinite dimensional vector space V has to satisfy in order that for a linear and bounded operator $Q : V \mapsto V'$ it holds $\overline{R(Q)} = N(Q^T)_\sharp$.

2.2.3 *The Weak Approach*

After the two examples in Sect. 2.2.1 and the general presentation in Sect. 2.2.2, the aim now is to completely describe a different point of view, based on the definition of what is called a *weak* solution u of (2.1).

Let us start again from the finite dimensional linear problem. System (2.12) is equivalent to

$$(Qq, r) = (p, r) \qquad \forall\, r \in \mathbb{R}^m\,, \tag{2.14}$$

where we have denoted by $(\cdot,\cdot)$ a scalar product in $\mathbb{R}^m$. In fact, from (2.14) we have $(Qq - p, r) = 0$ for each $r \in \mathbb{R}^m$, and taking $r = Qq - p$ the result follows. We can also remark that the same holds true if (2.14) is valid for all r in a set $\mathcal{V}$ that is dense in $\mathbb{R}^m$: it is enough to recall the continuity of the scalar product due to Cauchy–Schwarz inequality.

Noting that the new form (2.14) of problem (2.12) has at the left hand side a bilinear form and at the right hand side a linear functional, one is led to analyze the problems that can be written in this form: find the solution $q \in \mathbb{R}^m$ of

$$b(q, r) = F(r) \qquad \forall\, r \in \mathbb{R}^m\,, \tag{2.15}$$

where $b(\cdot,\cdot)$ is a bilinear form on $\mathbb{R}^m \times \mathbb{R}^m$ and $F(\cdot)$ is a linear functional on $\mathbb{R}^m$.

It is straightforward to check that this can be easily rewritten in the matrix form $Qq = p$, by setting $Q_{ij} = b(\omega_j, \omega_i)$ and $p_i = F(\omega_i)$, where ω_i are basis vectors of $\mathbb{R}^m$, $i = 1, \ldots, m$. Then we could go back to the analysis of a linear system associated to a matrix that has been constructed in terms of the bilinear form $b(\cdot,\cdot)$ and a basis of $\mathbb{R}^m$. However this is not so enlightening, and it is better to introduce a more abstract approach, which avoids the use of a basis and which will be easily extended to the infinite dimensional case. Applying to (2.15) the *finite dimensional* Riesz representation theorem we know that, for each fixed $w \in \mathbb{R}^m$, we can represent the linear functional $r \mapsto b(w, r)$ by means of the scalar product of a unique element $\omega_w \in \mathbb{R}^m$ and r, namely, $b(w, r) = (\omega_w, r)$ for each $r \in \mathbb{R}^m$. The same happens for $r \mapsto F(r)$, say, $F(r) = (g_F, r)$ for each $r \in \mathbb{R}^m$. The map $w \mapsto \omega_w$ is clearly linear, thus ω_w can be represented as Mw for a suitable $m \times m$ matrix M. Then solving (2.15) is equivalent to finding the solution $q \in \mathbb{R}^m$ of the linear system $Mq = g_F$. In particular, well-posedness of (2.15) is satisfied if and only if the map $r \mapsto Mr$ is one-to-one and onto from $\mathbb{R}^m$ to $\mathbb{R}^m$: this will be the strategy employed in the proof of Lax–Milgram theorem 2.1.

Having clarified this correspondence between the matrix formulation (2.12) and formulation (2.15), let us come back to our elliptic boundary value problem. We assume in the following that

$$a_{ij},\ b_i,\ a_0 \in L^\infty(D) \quad (i, j = 1, \ldots, n) \tag{2.16}$$

and

$$f \in L^2(D)\,, \tag{2.17}$$

and, for the sake of definiteness, in the rest of this section we will consider the Dirichlet boundary value problem.

When solving (2.1), we are looking for an element in an infinite dimensional vector space (loosely speaking, functions are elements of a vector space, as we can add them and we can multiply them by a real number; moreover, for identifying each one of them we need infinitely many information, namely, its value in all the points of the domain D: thus they live in a infinite dimensional vector space). If we can play with a scalar product, we could repeat what has been done here above for a finite dimensional linear system.

We know that in an infinite dimensional vector space we can have infinitely many scalar products, and they are not equivalent to each other. Thus we must choose the scalar product to be employed for mimicking the finite dimensional case, and the natural choice is the simplest scalar product we use when dealing with functions: the $L^2(D)$-scalar product, i.e.,

$$(w, v)_{L^2(D)} = \int_D wv dx\,. \tag{2.18}$$

Let us start now from (2.1). We know that the space of smooth functions with compact support $C_0^\infty(D)$ is dense in $L^2(D)$, thus it could play the role of the dense subspace $\mathcal{V}$. With this in mind, for each function v (we will call it a test function) Eq. (2.1) could be rewritten as

$$(Lu, v)_{L^2(D} = (f, v)_{L^2(D} \quad \forall\, v \in C_0^\infty(D)$$

(we are admitting, for the moment, that $u \in C^2(\overline{D})$ and the coefficients $a_{ij} \in C^1(\overline{D})$, so that all the three terms defining Lu belong to $L^2(D)$). This reads

$$\int_D -\sum_{i,j=1}^n \mathcal{D}_i(a_{ij}\mathcal{D}_j u)v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx = \int_D f v dx \, .$$

The term associated to the principal part can be balanced in a better way. In fact, integrating it by parts and remembering that $v_{|\partial D} = 0$, we obtain

$$\begin{aligned}\int_D -\sum_{i,j=1}^n \mathcal{D}_i(a_{ij}\mathcal{D}_j u)v dx &= \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx \\ &\quad \underbrace{- \int_{\partial D} \sum_{i,j=1}^n n_i a_{ij} \mathcal{D}_j u v_{|\partial D} dS_x}_{=0}\end{aligned}$$

and so

$$\begin{aligned}\int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx \\ = \int_D f v dx \quad \forall\, v \in C_0^\infty(D) \, .\end{aligned}$$

Definition 2.3 The bilinear form $B_L(\cdot\ ,\ \cdot)$ associated with the elliptic operator L introduced in (2.2) is defined by

$$B_L(w, v) = \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx \, . \tag{2.19}$$

Remark 2.3 Having chosen the minus sign in (2.2) has as a consequence that in the definition of the bilinear form (2.19) we have the plus sign!

We indicate by $F_D(\cdot)$ the linear functional associated to the right hand side f, namely, we set

$$F_D(v) = \int_D f v dx \, . \tag{2.20}$$

With this notation, problem (2.1) has been rephrased as follows: find u (in which space?) such that

$$B_L(u, v) = F_D(v) \quad \forall\, v \in C_0^\infty(D) \, . \tag{2.21}$$

Remark 2.4 Let us note, from the very beginning, that the weak problem is the *right* problem to face and we can focus on it without being afraid of considering something that is not meaningful. In fact, suppose we have a classical solution u to problem (2.1). We have just seen that u is also a solution to problem (2.21). If we know that for problem (2.21) a uniqueness result holds, then solving (2.21) furnishes the solution to (2.1). Furthermore, if the classical problem (2.1) has not a solution (for instance, the right hand side f has a jump discontinuity, so that a twice differentiable solution u cannot exist), it is still possible that the solution to (2.21) does exist (for example, the definition of the right hand side just needs $f \in L^2(D)$), and that it has a correct physical meaning. In this respect, remember that physical models are based on conservation principles, where the balance between integral quantities is required, and the process leading to pointwise partial differential equations is a limit process as volumes shrink at a point.

As we have remarked, the missing point in (2.21) is that we have to devise a suitable infinite dimensional vector space V where looking for u (and possibly also selecting the test functions v). The analogy with the finite dimensional matrix problem suggests that V should enjoy the following properties:

1. V is a subspace of $L^2(D)$ and is endowed with a *scalar product* (possibly, stronger than the $L^2(D)$-scalar product);
2. the bilinear form $B_L(\cdot, \cdot)$ and the linear functional $F_D(\cdot)$ are defined and *bounded* in $V \times V$ and V, respectively;
3. the (infinite dimensional) Riesz representation theorem holds in V. This essentially says that V must be a *Hilbert space*: namely, any Cauchy sequence in V is convergent to an element of V. (See Sect. 3.2 for the proof of Riesz theorem and also for some other interesting remarks.)
4. $C_0^\infty(D)$ is a subspace of V. (We will see that relaxing the assumption that $C_0^\infty(D)$ is a subspace of V is possible, but one must be careful: see Sect. 5.5 and the second part of Sect. 5.6.)
5. $C_0^\infty(D)$ is *dense* in V with respect to the convergence in V.

Let us note that in a finite dimensional vector space a linear functional is always bounded, while this is not true in the infinite dimensional case (see Sect. 3.1). Therefore in property 2 we have explicitly assumed boundedness. As shown in Exercise 2.7 this means that $B_L(\cdot, \cdot)$ and $F_D(\cdot)$ are continuous, thus by a density argument we see that (2.21) is satisfied also for all $v \in V$.

Note also that in property 5 the assumption that $C_0^\infty(D)$ is dense in V is related to the fact that we are considering the homogeneous Dirichlet boundary value problem; we will see that for the other boundary value problems this assumption could refer to other subspaces of $C^\infty(\overline{D})$.

Another remark about the space $C_0^\infty(D)$ is in order: here we have not yet shown which is its essential role (we only underlined the fact that it is suitable to assume that it is a subspace of V, as we have built our procedure by using test functions in $C_0^\infty(D)$). We will see in Chap. 4 that its use will permit us to introduce new relevant concepts and new Banach spaces, giving a solid ground to our analysis.

The following exercise clarifies the relation between boundedness and continuity for a linear functional.

Exercise 2.7 Let V be a Hilbert space (indeed, a normed space would be enough), and $F : V \mapsto \mathbb{R}$ a linear functional. Then F is bounded if and only if it is continuous.

An inspection of the terms in $B_L(u, v)$ shows that the principal part of it is defined if $\nabla u, \nabla v$ belong to $(L^2(D))^n$ (and the assumption $a_{ij} \in L^\infty(D)$ is sufficient); for the lower order terms we must add the assumption $u, v \in L^2(D)$. Thus we could choose $V = \{v \in C^1(\overline{D}) \mid v_{|\partial D} = 0\}$, but the choice of the scalar product $(w, v)_{L^2(D)}$ would not be enough (there is not a control of the integrals where first order derivatives appear). Therefore, we could endow V with the scalar product

$$(w, v)_1 = \int_D (wv + \nabla w \cdot \nabla v) dx \, . \tag{2.22}$$

However, it is easy to check that with these choices of V and $(\cdot, \cdot)_1$ property 3 here above is not satisfied. In fact, let us consider this exercise:

Exercise 2.8

(i) Consider $D = (-1, 1)$ and for $x \in D$ define $f(x) = 1 - |x|$, $g(x) = -\text{sign}(x)$. Show that there exists a sequence $v_k \in V = \{v \in C^1(\overline{D}) \mid v_{|\partial D} = 0\}$ such that $v_k \to f$ in $L^2(D)$ and $v_k' \to g$ in $L^2(D)$.

(ii) Show that V is not a Hilbert space with respect to the scalar product $(\cdot, \cdot)_1$ defined in (2.22).

Thus a new problem is enlightened: on one side, the scalar product $(\cdot, \cdot)_1$, that seems to be quite reasonable, requires that the gradient is defined (and square-summable); on the other side, the sequence v_n constructed in Exercise 2.8, part (i), is a Cauchy sequence with respect to the scalar product $(\cdot, \cdot)_1$, and, if we could obtain that $f'(x) = g(x)$ (which is definitely not true in the standard sense, but also does not seem to be completely meaningless), then we would have that v_n converges to f with respect to the scalar product $(\cdot, \cdot)_1$.

Summing up, here there is something to do: we need derivatives (and that they belong to $L^2(D)$), and we also need that a "corner" function admits a derivative (belonging to L^2). Therefore a natural question arises: is it the time to introduce a different definition of derivative?

We will see: for the moment, assume that we will be able to overcome these difficulties, and let us analyze how to solve a general problem of the form:

$$\text{find } u \in V : B(u, v) = F(v) \quad \forall\, v \in V \, , \tag{2.23}$$

where V is a Hilbert space, endowed with the scalar product $(\cdot, \cdot)_V$ and the norm $\| \cdot \|_V$, and the bilinear form $B(\cdot, \cdot)$ and the linear functional $F(\cdot)$ are defined and bounded in $V \times V$ and V, respectively.

A particular interesting and at the same time simple situation arises when $B(\cdot,\cdot)$ satisfies $|B(w,v)| \leq \gamma\|w\|_V\|v\|_V$ for each $w, v \in V$ (boundedness), $B(v,v) \geq \alpha\|v\|_V^2$ for each $v \in V$ (coerciveness) and is symmetric, i.e., $B(w,v) = B(v,w)$ for each $w, v \in V$ (for the bilinear form $B_L(\cdot,\cdot)$ introduced in (2.19) this means that the coefficients of the operator L satisfy $a_{ij} = a_{ji}$ and $b_i = 0$ for each $i, j = 1, \ldots, n$). In this case $B(\cdot,\cdot)$ is a scalar product in V, and the induced norm is equivalent to the original one: in fact from boundedness and coerciveness we have

$$\alpha\|v\|_V^2 \leq B(v,v)) \leq \gamma\|v\|_V^2 .$$

Thus solving problem (2.23) is a direct consequence of the (infinite dimensional) Riesz representation theorem (see Theorem 3.1).

Let us note, however, that in the finite dimensional case the linear system $Qq = p$ has a unique solution if and only if $\det Q \neq 0$. Hence, as shown in Exercise 2.2, a sufficient condition to have a unique solution is that Q is positive definite, i.e.,

$$(Qr, r) \geq \alpha|r|^2 \quad \forall\, r \in \mathbb{R}^m$$

for some $\alpha > 0$. Therefore symmetry does not seem to be essential: we could hope that the well-posedness of (2.23) is true even if $B(\cdot,\cdot)$ is not symmetric, but still bounded and such that $B(v,v) \geq \alpha\|v\|_V^2$ for each $v \in V$.

The answer is in the quite important result presented in next section.

2.3 Lax–Milgram Theorem

In this section we assume V is a (real) Hilbert space, with norm $\|\cdot\|_V$ and scalar product $(\cdot,\cdot)_V$ (note however that the result below, with easy modification, is also true for a complex Hilbert space).

Theorem 2.1 (Lax–Milgram Theorem) *Let $B : V \times V \mapsto \mathbb{R}$ and $F : V \mapsto \mathbb{R}$ be a bilinear form and a linear functional, respectively. Assume that $B(\cdot,\cdot)$ is bounded and coercive in $V \times V$, i.e., there exist constants $\gamma > 0$, $\alpha > 0$ such that*

$$|B(w,v)| \leq \gamma\|w\|_V\|v\|_V \quad \forall\, w, v \in V \tag{2.24}$$

and

$$B(v,v) \geq \alpha\|v\|_V^2 \quad \forall\, v \in V\,, \tag{2.25}$$

and that $F : V \mapsto \mathbb{R}$ is bounded in V, i.e., there exists a constant $M > 0$ such that

$$|F(v)| \leq M\|v\|_V \quad \forall\, v \in V\,. \tag{2.26}$$

Then there exists a unique element $u \in V$ *such that*

$$B(u, v) = F(v) \quad \forall\, v \in V .$$

Moreover the stability estimate $\|u\|_V \leq \frac{M}{\alpha}$ *holds true.*

Proof The proof presented here is based on the Riesz representation theorem 3.1 and the projection theorem (see Yosida [28, Theorem 1, p. 82]), two well-known results of functional analysis. Another proof, using the less known closed range theorem 8.4, can be found in Exercise 8.4.

The proof is divided into 6 steps. The first three have the aim to rewrite the problem as an equation in the Hilbert space V for a suitable linear and bounded operator.

1. For each fixed element $w \in V$, the mapping $v \mapsto B(w, v)$ is a bounded linear functional on V; hence the Riesz representation theorem 3.1 asserts the existence of a unique element $\omega_w \in V$ satisfying

$$B(w, v) = (\omega_w, v)_V \quad \forall\, v \in V .$$

Let us write $Aw = \omega_w$, so that for $w, v \in V$ it holds

$$B(w, v) = (Aw, v)_V .$$

2. Similarly, once more from the Riesz representation theorem 3.1 we observe that we can write

$$F(v) = (g_F, v)_V \quad \forall\, v \in V$$

for a unique element $g_F \in V$. Then problem (2.23) reduces to finding a unique $u \in V$ satisfying $Au = g_F$, namely, to show that $A : V \mapsto V$ is one-to-one and onto.
3. We first claim A is a bounded linear operator. Indeed if $\lambda_1, \lambda_2 \in \mathbb{R}$ and $w_1, w_2 \in V$, for each $v \in V$ we see that

$$\begin{aligned}
(A(\lambda_1 w_1 + \lambda_2 w_2), v)_V &= B(\lambda_1 w_1 + \lambda_2 w_2, v) \\
&= \lambda_1 B(w_1, v) + \lambda_2 B(w_2, v) \\
&= \lambda_1 (Aw_1, v)_V + \lambda_2 (Aw_2, v)_V \\
&= (\lambda_1 Aw_1 + \lambda_2 Aw_2, v)_V .
\end{aligned}$$

This equality is true for each $v \in V$, thus we have proved that A is linear. Furthermore

$$\|Av\|_V^2 = (Av, Av)_V = B(v, Av) \le \gamma \|v\|_V \|Av\|_V .$$

Consequently $\|Av\|_V \le \gamma \|v\|_V$ for all $v \in V$ and A is bounded.

4. Next we assert

$$\begin{cases} A \text{ is one-to-one} \\ \text{and} \\ R(A), \text{ the range of } A, \text{ is closed in } V. \end{cases} \tag{2.27}$$

To prove this, let us compute

$$\alpha \|v\|_V^2 \le B(v, v) = (Av, v)_V \le \|Av\|_V \|v\|_V .$$

Hence $\alpha \|v\|_V \le \|Av\|_V$. This inequality easily implies that A is one-to-one. Moreover, take a sequence $Av_n \in R(A)$ such that $Av_n \to \omega_0 \in V$. Since Av_n is convergent, it is a Cauchy sequence; using the linearity of A and the last inequality we also have $\|v_n - v_m\|_V \le \alpha \|Av_n - Av_m\|_V$, thus v_n is a Cauchy sequence, too. Being V a Hilbert space we have that $v_n \to w_0 \in V$, and since A is bounded it follows $Av_n \to Aw_0$. The uniqueness of the limit yields $\omega_0 = Aw_0$, thus $R(A)$ is a closed subspace.

5. We prove now that

$$R(A) = V . \tag{2.28}$$

By the projection theorem (see Yosida [28, Theorem 1, p. 82]), it is enough to prove that $R(A)^\perp = \{0\}$. Let us take $w \in R(A)^\perp$; then

$$\alpha \|w\|_V^2 \le B(w, w) = (Aw, w)_V = 0 ,$$

hence $w = 0$. In conclusion, A is onto.

6. Finally we have that

$$\alpha \|u\|_V^2 \le B(u, u) = F(u) \le M \|u\|_V ,$$

thus $\|u\|_V \le \frac{M}{\alpha}$.

□

Remark 2.5 As already said, the dual space of V (i.e., the space of linear and bounded functionals from V to $\mathbb{R}$) will be denoted by V'. Following this notation, in Lax–Milgram theorem we have assumed $F \in V'$.

Remark 2.6 It is well known that if V is a Hilbert space then V' is a Hilbert space, too. Its scalar product is given by $(\Psi, \Phi)_{V'} = (\omega_\Psi, \omega_\Phi)_V$, where by the Riesz representation theorem 3.1 it holds $\Psi(v) = (\omega_\Psi, v)_V$ and $\Phi(v) = (\omega_\Phi, v)_V$ for each $v \in V$ (see, e.g., Yosida [28, Corollary 1, p. 91]).

Remark 2.7 Necessary and sufficient conditions for a general existence and uniqueness result are presented in Theorem F.1.

Remark 2.8 For the sake of simplicity, in the sequel we will often say that a bilinear form $B(\cdot, \cdot) : V \times V \mapsto \mathbb{R}$ is bounded or coercive in V, instead of in $V \times V$.

2.4 Exercises

Exercise 2.1 A matrix A is positive definite if and only if it exists $\alpha > 0$ such that $Av \cdot v \geq \alpha|v|^2$ for every $v \in \mathbb{R}^n$.

Solution

($\Leftarrow$) Trivial.

($\Rightarrow$) The map $v \mapsto Av \cdot v$ is positive for all $v \neq 0$ and it is continuous. On the subset $|v| = 1$, which is bounded and closed, it has a minimum $\alpha > 0$ and a minimum point v^* such that $Av^* \cdot v^* = \alpha$. Now take $v \neq 0$ and let $v^\sharp = \frac{v}{|v|}$, $|v^\sharp| = 1$. Therefore we have that $Av^\sharp \cdot v^\sharp \geq \alpha > 0$, that is

$$\alpha \leq A\frac{v}{|v|} \cdot \frac{v}{|v|} = \frac{1}{|v|^2}Av \cdot v \implies Av \cdot v \geq \alpha|v|^2 \,.$$

Exercise 2.2 Consider a positive definite matrix A (thus satisfying $Av \cdot v \geq \alpha|v|^2$ for every $v \in \mathbb{R}^n$, for a suitable $\alpha > 0$). Then the real part of an eigenvalue of A is greater than or equal to α; in particular, a positive definite matrix is non-singular.

Solution Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, with (unit) eigenvector $\omega = v+iw \in \mathbb{C}^n$, $v, w \in \mathbb{R}^n$. We have

$$\lambda = \lambda|\omega|^2_{\mathbb{C}^n} = (\lambda\omega, \omega)_{\mathbb{C}^n} = (A\omega, \omega)_{\mathbb{C}^n} = Av \cdot v - iAv \cdot w + iAw \cdot v + Aw \cdot w \,,$$

thus

$$\mathrm{Re}\,\lambda = Av \cdot v + Aw \cdot w \geq \alpha(|v|^2 + |w|^2) = \alpha \,.$$

As a consequence, all the eigenvalues of A are different from 0 and $\det A \neq 0$, thus A is non-singular.

Exercise 2.3

(i) A matrix A is positive definite if and only if $\frac{A+A^T}{2}$ is positive definite.
(ii) A matrix A is positive definite if and only if all the eigenvalues λ_i of $\frac{A+A^T}{2}$ are strictly positive.

Solution

(i) We have

$$\frac{A+A^T}{2}v \cdot v = \frac{1}{2}(Av \cdot v + A^T v \cdot v) = Av \cdot v \, ,$$

thus (i) is proved.
(ii) It is enough to note that $\frac{A+A^T}{2}$ is a symmetric matrix, thus being positive definite is equivalent to say that its minimum eigenvalue is strictly positive.

Exercise 2.4

(i) Show that the operator

$$Lw = -\mathcal{D}_1((1+x_1x_2)\mathcal{D}_1 w) - \mathcal{D}_1(x_1\mathcal{D}_2 w) - \mathcal{D}_2(x_2\mathcal{D}_1 w) - \mathcal{D}_2\mathcal{D}_2 w \, ,$$

is uniformly elliptic in $D = \{x \in \mathbb{R}^2 \,|\, 0 < x_1 < 1/2, 0 < x_2 < 1\}$.
(ii) Show that the operator $Lw = -\sum_{i,j=1}^{3} D_i(a_{ij}D_j w)$, with

$$\{a_{ij}\} = \begin{pmatrix} 1 & -x_3 & x_2 \\ x_3 & 1+x_1^2 & x_1 \\ -x_2 & x_2 & 1+x_3^2 \end{pmatrix}$$

is uniformly elliptic in $D = \{x \in \mathbb{R}^3 \,|\, |x| < 1\}$.

Solution Using Exercise 2.3, it is enough to show that the minimum eigenvalue $\lambda_1(x)$ of the matrix $\{\frac{1}{2}(a_{ij} + a_{ji})\}$ satisfies $\inf_D \lambda_1(x) > 0$.

(i) Writing $A = \{a_{ij}\}$ we have

$$A = \begin{pmatrix} 1+x_1x_2 & x_1 \\ x_2 & 1 \end{pmatrix}$$

and

$$\frac{1}{2}(A + A^T) = \begin{pmatrix} 1+x_1x_2 & \frac{1}{2}(x_1+x_2) \\ \frac{1}{2}(x_1+x_2) & 1 \end{pmatrix} .$$

A simple calculation shows that

$$\lambda_1(x) = \frac{1}{2}\Big(2 + x_1 x_2 - \sqrt{x_1^2 x_2^2 + (x_1 + x_2)^2}\Big)\,.$$

Since for $a \geq 0$, $b \geq 0$ we have $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, it follows that $\lambda_1(x) \geq \frac{1}{2}\big(2 - x_1 - x_2\big) \geq \frac{1}{4}$ for $x \in \overline{D}$.

(ii) We have

$$\frac{1}{2}(A + A^T) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 + x_1^2 & \frac{1}{2}(x_1 + x_2) \\ 0 & \frac{1}{2}(x_1 + x_2) & 1 + x_3^2 \end{pmatrix}\,.$$

Clearly one of the eigenvalues is equal to 1, while the minimum of the other two is given by

$$\lambda_1(x) = \frac{1}{2}\Big(2 + x_1^2 + x_3^2 - \sqrt{(x_1^2 - x_3^2)^2 + (x_1 + x_2)^2}\Big)\,.$$

Using again the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, we find

$$\lambda_1(x) \geq \frac{1}{2}\Big(2 + x_1^2 + x_3^2 - |x_1^2 - x_3^2| - |x_1 + x_2|\Big) \geq \frac{1}{2}\Big(2 - |x_1 + x_2|\Big) \geq 1 - \frac{\sqrt{2}}{2}$$

for $x \in \overline{D}$.

Exercise 2.5 Consider $D = (0, a) \times (0, b)$. Determine the eigenvalues and the eigenvectors associated to the operator $-\Delta$ with homogeneous Dirichlet boundary condition, and verify that, after a suitable normalization, the eigenvectors are an orthonormal system in $L^2(D)$. [Hint: use the method of separation of variables.]

Solution We must find functions $\omega = \omega(x, y)$ and numbers λ such that $-\Delta\omega = \lambda\omega$ in $(0, a) \times (0, b)$ and $\omega_{|\partial D} = 0$. Using the technique of separation of variables we look for $\omega(x, y) = p(x)q(y)$, with $p(0) = p(a) = 0$ and $q(0) = q(b) = 0$. Imposing the equation we find

$$-\Delta\omega = -p''q - pq'' = \lambda pq = \lambda\omega \quad \text{in } (0, a) \times (0, b)\,,$$

and dividing by pq (this is justified for $p \neq 0$ and $q \neq 0$, but let us go on...) we obtain

$$-\frac{p''}{p} - \frac{q''}{q} = \lambda\,.$$

Since $\frac{p''}{p}$ is a function of the variable x only and $\frac{q''}{q}$ is a function of the variable y only, this equation can be satisfied if and only if $\frac{p''}{p}$ and $\frac{q''}{q}$ are both equal to a constant.

Let us write $\frac{p''}{p} = -\mu$ (thus $\frac{q''}{q} = \mu - \lambda$). The ordinary differential equation $p'' + \mu p = 0$ has a general solution given by $p(x) = c_1 \exp(\sqrt{-\mu}x) + c_2 \exp(-\sqrt{-\mu}x)$ for $\mu < 0$, by $p(x) = c_1 + c_2 x$ for $\mu = 0$ and by $p(x) = c_1 \sin(\sqrt{\mu}x) + c_2 \cos(\sqrt{\mu}x)$ for $\mu > 0$. In the first two cases imposing the boundary conditions $p(0) = p(a) = 0$ readily yields $c_1 = c_2 = 0$, thus p is vanishing and it is not an eigenvector; in the third case from $p(0) = 0$ it follows $c_2 = 0$, thus we have to impose $p(a) = c_1 \sin(\sqrt{\mu}a) = 0$ without setting $c_1 = 0$. The condition to be satisfied is therefore

$$\sin(\sqrt{\mu}a) = 0 \Longrightarrow \sqrt{\mu}a = m\pi \quad \text{for } m \geq 1 .$$

We have thus found the sequence $\mu_m = \frac{m^2\pi^2}{a^2}$, $m \geq 1$, and the corresponding functions $p_m(x) = \sin(\frac{m\pi}{a}x)$. Setting $\nu = \lambda - \mu$, a similar computation for the other factor q yields $\nu_l = \frac{l^2\pi^2}{b^2}$ and $q_l(y) = \sin(\frac{l\pi}{b}y)$, for $l \geq 1$.

We have thus determined

$$\lambda_{ml} = \frac{m^2\pi^2}{a^2} + \frac{l^2\pi^2}{b^2} \ , \ \widehat{\omega}_{ml}(x,y) = \sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{l\pi}{b}y\right) \ , \ m \geq 1 \, , \, l \geq 1 \, .$$

From $\int_0^a \sin(\frac{m\pi}{a}x)\sin(\frac{m'\pi}{a}x)dx = 0$ for $m \neq m'$ and $\int_0^a \sin^2(\frac{m\pi}{a}x)dx = \frac{a}{2}$ it is readily seen that $\omega_{ml} = \frac{2}{\sqrt{ab}}\widehat{\omega}_{ml}$ is an orthonormal system in $L^2((0,a)\times(0,b))$.

Exercise 2.6

(i) Find a function $K_0 = K_0(\xi)$ defined in $\mathbb{R}^2 \setminus \{0\}$ and such that

$$-\Delta K_0 = 0 \text{ in } \mathbb{R}^2 \setminus \{0\} \ \text{ and } \ -\int_{\partial B(0,t)} \nabla K_0 \cdot n dS_\xi = 1$$

for any $t > 0$. [Hint: look for a radial function $K_0 = K_0(|\xi|)$.]

(ii) Verify that a function $K(x,\xi)$ satisfying $-(\Delta_x K)(x,\xi) = -(\Delta_\xi K)(\xi,x) = 0$ for $\xi \neq x$ and $-\int_{\partial B(x,t)} (\nabla_\xi K)(\xi,x)\cdot n(\xi) dS_\xi = 1$ for each $t > 0$ is given by $K(x,\xi) = K_0(|x-\xi|)$.

Solution

(i) Let us write $|\xi| = r$ and look for $K_0(r)$. The Laplace operator in polar coordinates is given by

$$\Delta = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2$$

(see Exercise 7.14). Therefore we have to solve, for $r > 0$,

$$0 = K_0''(r) + \frac{1}{r} K_0'(r) = \frac{1}{r}(r K_0'(r))' ,$$

thus we have $r K_0'(r) = c_0$, a constant. Consequently we find $K_0(r) = c_0 \log r + c_1$, and, for simplicity, we can choose $c_1 = 0$. Then let us compute ∇K_0. We obtain

$$\mathcal{D}_i K_0(r) = c_0 \mathcal{D}_i \log r = c_0 \frac{1}{r} \mathcal{D}_i r = c_0 \frac{1}{r} \frac{\xi_i}{r} .$$

On the other hand, on $\partial B(0,t)$ we have $n_i = \frac{\xi_i}{t}$. Thus on $\partial B(0,t)$ we obtain $\nabla K_0 \cdot n = c_0 \frac{1}{t} \frac{\xi}{t} \cdot \frac{\xi}{t} = c_0 \frac{1}{t^3} |\xi|^2 = c_0 \frac{1}{t}$. Let us integrate this function on $\partial B(0,t)$:

$$\int_{\partial B(0,t)} \nabla K_0 \cdot n dS_\xi = c_0 \frac{1}{t} \text{meas}(\partial B(0,t)) = c_0 \frac{1}{t} 2\pi t = 2\pi c_0 .$$

In conclusion we have found $c_0 = -\frac{1}{2\pi}$ and $K_0(\xi) = -\frac{1}{2\pi} \log |\xi|$.

(ii) The result is straightforward as $K(x,\xi)$ given by $K_0(|x - \xi|)$ is symmetric with respect to x and ξ, and then radial with center at x.

Exercise 2.7 Let V be a Hilbert space (indeed, a normed space would be enough) and $F : V \mapsto \mathbb{R}$ a linear operator. Then F is bounded if and only if it is continuous.

Solution If F is bounded, namely, $|F(v)| \le \gamma \|v\|_V$ for a suitable $\gamma > 0$, from linearity we readily obtain that $F(v_k) \to F(v_0)$ if $v_k \to v_0$ in V.
Conversely, assume that F is continuous. Since F is linear we have $F(0) = 0$; then there exists $\delta > 0$ such that $|F(v)| \le 1$ for $\|v\|_V \le \delta$. Take now $v \in V$, $v \ne 0$. Define $w = \delta \frac{v}{\|v\|_V}$, so that $\|w\|_V = \delta$. We have $|F(w)| \le 1$, hence $|F(v)| \le \frac{1}{\delta} \|v\|_V$.

Exercise 2.8

(i) Consider $D = (-1, 1)$ and for $x \in D$ define $f(x) = 1 - |x|$, $g(x) = -\text{sign}(x)$. Show that there exists a sequence $v_k \in V = \{v \in C^1(\overline{D}) \,|\, v_{|\partial D} = 0\}$ such that $v_k \to f$ in $L^2(D)$ and $v_k' \to g$ in $L^2(D)$.

(ii) Show that V is not a Hilbert space with respect to the scalar product $(v, w)_1$ defined in (2.22).

Solution

(i) Take v_k defined as follows:

$$v_k(x) = \begin{cases} 1 - |x| & \text{for } -1 < x < -\frac{1}{k} \\ 1 - \frac{1}{2k} - \frac{k}{2} x^2 & \text{for } -\frac{1}{k} \le x \le \frac{1}{k} \\ 1 - |x| & \text{for } \frac{1}{k} < x < 1 . \end{cases}$$

It is easily seen that $v_k \in V$ and that

$$v_k'(x) = \begin{cases} 1 & \text{for } -1 < x < -\frac{1}{k} \\ -kx & \text{for } -\frac{1}{k} \leq x \leq \frac{1}{k} \\ -1 & \text{for } \frac{1}{k} < x < 1\,. \end{cases}$$

Then

$$\int_{-1}^{1} (v_k'(x) + \operatorname{sign}(x))^2 dx = \int_{-\frac{1}{k}}^{0} (-kx - 1)^2 dx + \int_{0}^{\frac{1}{k}} (-kx + 1)^2 dx$$
$$= \frac{(kx+1)^3}{3k}\Big|_{-\frac{1}{k}}^{0} + \frac{(kx-1)^3}{3k}\Big|_{0}^{\frac{1}{k}} = \frac{2}{3k}\,.$$

On the other hand

$$\int_{-1}^{1} (v_k(x) - 1 + |x|)^2 dx = \int_{-\frac{1}{k}}^{0} (1 - \frac{1}{2k} - \frac{k}{2}x^2 - 1 + x)^2 dx$$
$$+ \int_{0}^{\frac{1}{k}} (1 - \frac{1}{2k} - \frac{k}{2}x^2 - 1 - x)^2 dx$$
$$= 2\int_{0}^{\frac{1}{k}} (\frac{1}{2k} + \frac{k}{2}x^2 + x)^2 dx \leq 2\,\frac{1}{k}\,\frac{4}{k^2} = \frac{8}{k^3}\,.$$

(ii) Part (i) says that v_k and v_k' are convergent sequences, therefore Cauchy sequences in $L^2(D)$. Thus v_k is a Cauchy sequence with respect to norm induced by the scalar product $(\cdot,\cdot)_1$. Assume, by contradiction, that v_k converges with respect to this norm to a function $v_0 \in V$. Since the scalar product $(\cdot,\cdot)_1$ is stronger than the scalar product $(\cdot,\cdot)_{L^2(D)}$, one also has that v_k converges to v_0 in $L^2(D)$, therefore $v_0 = f$. Since $f \notin V$, a contradiction is produced.

Chapter 3
A Bit of Functional Analysis

For the ease of the reader, in this chapter we present some results of functional analysis: in particular, we show how a finite dimensional normed vector space and a infinite dimensional normed vector space enjoy different properties, and which are some basic points that make a Hilbert space different from a pre-Hilbertian space.

3.1 Why Is Life in an Infinite Dimensional Normed Vector Space V Harder than in a Finite Dimensional One?

1. The boundedness (continuity) of a linear functional must be explicitly required. In fact:
   If $\dim V < +\infty$ a linear functional is bounded.
   If $\dim V = +\infty$ this is not true anymore.

   Example 3.1 Let's take the space of trigonometric polynomials

$$V = \Big\{ v : [0, 2\pi] \mapsto \mathbb{R} \mid \exists\, N \geq 0,\ \exists\, \{a_k, b_k\}_{k=0}^{N} \text{ such that } v = \sum_{k=0}^{N} \big(a_k \cos(kx) + b_k \sin(kx)\big) \Big\},$$

   endowed with the scalar product $(v, w)_V = \int_0^{2\pi} vw\,dx$. Set $Lv = v'$ and take $v_m = \sin(mx)$, $m \geq 1$, then

$$\int_0^{2\pi} v_m^2 dx = \int_0^{2\pi} (\sin(mx))^2 dx = \pi$$
$$\int_0^{2\pi} (Lv_m)^2 dx = \int_0^{2\pi} (m\cos(mx))^2 dx = m^2\pi$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_3

and

$$\frac{\|Lv_m\|_V}{\|v_m\|_V} = \frac{m\sqrt{\pi}}{\sqrt{\pi}} = m \to \infty \,.$$

Hence the functional L is linear but not bounded.

2. The precompactness of a bounded set must be explicitly proved. In fact:
   If $\dim V < +\infty$ from a bounded sequence you can extract a convergent subsequence (Bolzano–Weierstrass Theorem).
   If $\dim V = +\infty$ this is not true anymore.

   Example 3.2 Let's take w_m an orthonormal system in $L^2(0, 2\pi) = V$ (for instance $w_m(x) = \frac{1}{\sqrt{\pi}} \sin(mx)$). Then

$$\|w_m\|_V = 1$$

   and, for $k \neq m$,

$$\begin{aligned} \|w_m - w_k\|_V^2 &= (w_m - w_k, w_m - w_k)_V \\ &= \|w_m\|_V^2 + \|w_k\|_V^2 - \underbrace{2(w_m, w_k)_V}_{=0} = 2 \,. \end{aligned}$$

   Thus any subsequence extracted by w_m is not convergent, as it is not a Cauchy sequence.

3. The convergence of Cauchy sequences must be explicitly proved. In fact:
   If $\dim V < +\infty$ any Cauchy sequence in V is convergent to an element in V. [Indeed a Cauchy sequence is bounded (see Exercise 3.1 (i)) and from point 2. you can extract a convergent subsequence; if a Cauchy sequence has a convergent subsequence then the whole sequence is convergent (see Exercise 3.1 (ii)).]
   If $\dim V = +\infty$ this is not true anymore.

   Example 3.3 Let us take $V = C^0([-1, 1])$ endowed with the scalar product $(v, w)_V = \int_{-1}^{1} vw dx$ and consider

$$v_m(x) = \begin{cases} 0 & x \in [-1, 0] \\ mx & x \in (0, 1/m) \\ 1 & x \in [1/m, 1] \end{cases} \tag{3.1}$$

   (see Fig. 3.1).

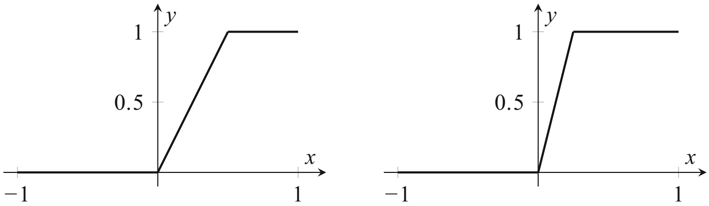


Fig. 3.1 The graph of the function v_m in (3.1) for $m = 2$ (left) and $m = 4$ (right)

Then setting

$$v(x) = \begin{cases} 0 & x \in [-1, 0] \\ 1 & x \in (0, 1] , \end{cases}$$

we have that

$$\int_{-1}^{1} |v_m - v|^2 dx = \int_{0}^{1/m} (1 - mx)^2 dx \leq \frac{1}{m} \to 0 .$$

Therefore v_m is a Cauchy sequence in V, but it is not convergent to an element in V, as $v \notin C^0([-1, 1])$.

4. The closure of a vector subspace must be explicitly proved. In fact;
If $\dim V < +\infty$ a subspace is always closed.
If $\dim V = +\infty$ this is not true anymore.

Example 3.4 Let us take $V = L^2(-1, 1)$ with $(v, w)_V = \int_{-1}^{1} vw dx$. As a subspace of V take $W = C^0([-1, 1])$ and choose v_m as in the previous example. Then $v_m \in W$, $v_m \to v$ in V but $v \notin W$.

5. The closure of the range of a linear and bounded operator $A : V \mapsto W$, V and W Hilbert spaces, must be explicitly proved. In fact;
If $\dim W < +\infty$ the range of A, being a subspace, is always closed.
If $\dim W = +\infty$ this is not true anymore.

Example 3.5 Take $V = W = L^2(-1, 1)$ and $A : v \mapsto Av$ where $(Av)(x) = \int_{-1}^{x} v(t) dt$. Clearly A is a linear operator, $Av \in V$ and finally A is a bounded operator, as by Cauchy–Schwarz inequality

$$\int_{-1}^{1} (Av)^2(x) dx = \int_{-1}^{1} \left(\int_{-1}^{x} v(t) dt \right)^2 dx \leq \int_{-1}^{1} (x + 1) \left(\int_{-1}^{x} v(t)^2 dt \right) dx$$
$$\leq \left(\int_{-1}^{1} v(t)^2 dt \right) \frac{(x+1)^2}{2} \Big|_{-1}^{1} = 2 \int_{-1}^{1} v(t)^2 dt .$$

Indeed, we have a further regularity result, as any $Av \in R(A)$ is uniformly continuous in $[-1, 1]$. In fact, for $x_1, x_2 \in [-1, 1]$, $x_1 < x_2$ it holds

$$|(Av)(x_2) - (Av)(x_1)| = \left| \int_{x_1}^{x_2} v(t)dt \right| \underbrace{\leq}_{\text{Cauchy–Schwarz}} \sqrt{x_2 - x_1} \left(\int_{-1}^{1} v(t)^2 dt \right)^{1/2} .$$

Choose now $\omega_m \in V$ as follows:

$$\omega_m(x) = \begin{cases} 0 & \text{for } -1 \leq x \leq 0 \\ m & \text{for } 0 < x < 1/m \\ 0 & \text{for } 1/m \leq x \leq 1 \,. \end{cases}$$

As a consequence we have that $A\omega_m$ is given by

$$(A\omega_m)(x) = \begin{cases} 0 & \text{for } -1 \leq x \leq 0 \\ mx & \text{for } 0 < x < 1/m \\ 1 & \text{for } 1/m \leq x \leq 1 \,, \end{cases}$$

thus $A\omega_m$ are equal to the functions v_m in Example 3.3, (3.1). There we have seen that $A\omega_m = v_m$ converges to

$$v(x) = \begin{cases} 0 & x \in [-1, 0] \\ 1 & x \in (0, 1] \,. \end{cases}$$

Since v is discontinuous, it follows that $v \notin R(A)$ and therefore the range of A is not closed.

3.2 Why Is Life in a Hilbert Space Better than in a Pre-Hilbertian Space?

Definition 3.1 A pre-Hilbertian space is a space endowed with a scalar product.

It is clearly difficult to express which is the main basic difference between a pre-Hilbertian space and a Hilbert space. A possible answer, the one on which we first focus here, is that in a Hilbert space we have the Riesz representation theorem, whereas in a pre-Hilbertian space that is not true. We will see later that we can make more precise this assertion.

Theorem 3.1 (Riesz Representation) *Let V be a Hilbert space, and let $\mathcal{F} : V \mapsto \mathbb{R}$ be a linear and bounded functional. Then there exists a unique $\omega \in V$ such that $\mathcal{F}(v) = (\omega, v)_V$ for each $v \in V$.*

Let us give a proof of Riesz theorem. If an element $\omega \in V$ satisfies $\mathcal{F}(v) = (\omega, v)_V$ for all $v \in V$, then $\omega \in N^\perp = \{w \in V \,|\, (w, v)_V = 0 \;\forall\, v \in N\}$, where $N = \{v \in V \,|\, \mathcal{F}(v) = 0\}$. If $\mathcal{F}(v) = 0$ for all $v \in V$, take $\omega = 0$. Otherwise take $\hat{\omega} \neq 0$, $\hat{\omega} \in N^\perp$, and look for ω in the form $\omega = \alpha\hat{\omega}$ for a suitable $\alpha \in \mathbb{R}$. Imposing that the representation formula is true for $v = \hat{\omega}$, namely, that we have $\mathcal{F}(\hat{\omega}) = (\alpha\hat{\omega}, \hat{\omega})_V$, it follows

$$\alpha = \frac{\mathcal{F}(\hat{\omega})}{\|\hat{\omega}\|_V^2}\,.$$

We claim that

$$\omega = \frac{\mathcal{F}(\hat{\omega})}{\|\hat{\omega}\|_V^2}\hat{\omega}\,.$$

We have to prove that such ω satisfies $\mathcal{F}(v) = (\omega, v)_V$ for each $v \in V$. It holds

$$\begin{aligned}\mathcal{F}(v) \stackrel{?}{=} \frac{\mathcal{F}(\hat{\omega})}{(\hat{\omega}, \hat{\omega})_V}(\hat{\omega}, v)_V \iff & \mathcal{F}(v)(\hat{\omega}, \hat{\omega})_V - \mathcal{F}(\hat{\omega})(\hat{\omega}, v)_V \stackrel{?}{=} 0 \\ \iff & (\mathcal{F}(v)\hat{\omega} - \mathcal{F}(\hat{\omega})v, \hat{\omega})_V \stackrel{?}{=} 0\,,\end{aligned}$$

thus it is sufficient to prove that $(\mathcal{F}(v)\hat{\omega} - \mathcal{F}(\hat{\omega})v) \in N$. Indeed by linearity we have

$$\begin{aligned}\mathcal{F}(\hat{\omega}\mathcal{F}(v) - v\mathcal{F}(\hat{\omega})) &= \mathcal{F}(\hat{\omega}\mathcal{F}(v)) - \mathcal{F}(v\mathcal{F}(\hat{\omega})) \\ &= \mathcal{F}(\hat{\omega})\mathcal{F}(v) - \mathcal{F}(v)\mathcal{F}(\hat{\omega}) = 0\,.\end{aligned}$$

We have thus completed the proof of the Riesz representation theorem. But where did we use the assumption that V is a Hilbert space and not simply a pre-Hilbertian space? At a first look it is not so evident...

The point is that we have assumed that there exists $\hat{\omega} \neq 0$, $\hat{\omega} \in N^\perp$. But we only know that there exists $\omega^* \neq 0$ such that $\mathcal{F}(\omega^*) \neq 0$, namely, $\omega^* \neq 0$, $\omega^* \notin N$. In a pre-Hilbertian space this does not mean that we can find $\hat{\omega} \neq 0$, $\hat{\omega} \in N^\perp$. It is possible that $N^\perp = \{0\}$ even if $N \neq V$! On the contrary this is not possible for a Hilbert space, as we have the *projection theorem* (see Yosida [28, Theorem 1, p. 82]) and therefore if $N \neq V$ we know that $N^\perp$ is not trivial, because we can split $V = N \oplus N^\perp$, writing $\omega^* \neq 0$ as

$$\omega^* = \underbrace{P_N\omega^*}_{\in N} + \underbrace{P_{N^\perp}\omega^*}_{\in N^\perp}$$

with $P_{N^\perp}\omega^* \neq 0$ if $\omega^* \notin N$.

Example 3.6 Let us give an example of $N \neq V$, $N^\perp = \{0\}$ for a pre-Hilbertian space V. Take $V = C_0^\infty(D)$ with D an open, connected, bounded set, and endow V with the scalar product $(v, w)_V = \int_D vw dx$. Consider $\mathcal{F}(v) = \int_D v dx$ and note that $\mathcal{F}$ is linear and continuous, as by the Cauchy–Schwarz inequality

$$|\mathcal{F}(v)| = \left| \int_D v dx \right| \leq \int_D |v| dx \leq (\text{meas}(D))^{1/2} \left(\int_D v^2 dx \right)^{1/2} \quad \forall\, v \in V\,.$$

It is also clear that

$$N = \left\{ v \in C_0^\infty(D) \,\Big|\, \int_D v dx = 0 \right\} \tag{3.2}$$

is a subspace with $N \neq V$, as there are $C_0^\infty(D)$ functions that are positive and not identically 0, thus satisfying $\int_D v dx > 0$. It is also easy to show that N is a *closed* subspace, namely, if a sequence $v_m \in N$ converges to $v^\star \in V$ with respect to the norm associated to $(\cdot, \cdot)_V$, then $\int_D v^\star dx = 0$, thus $v^\star \in N$. If $\omega \in N^\perp$ (orthogonality in V, thus $\omega \in C_0^\infty(D) \ldots$), for each $v \in N$ it follows

$$0 = \int_D \omega v dx = \int_D (\omega - \omega_D) v dx + \omega_D \underbrace{\int_D v dx}_{=0} = \int_D (\omega - \omega_D) v dx\,,$$

where

$$\omega_D = \frac{1}{\text{meas}(D)} \int_D \omega dx\,.$$

If we prove that N is dense in

$$L^2_*(D) = \left\{ v \in L^2(D) \,\Big|\, \int_D v dx = 0 \right\}$$

(see below, Exercise 3.2), then by a density argument we can also write

$$0 = \int_D (\omega - \omega_D) v dx \quad \forall\, v \in L^2_*(D)\,.$$

Taking $v = \omega - \omega_D$, which satisfies $v \in C^\infty(\overline{D})$ with $\int_D v dx = 0$, therefore belongs to $L^2_*(D)$, it follows that

$$\int_D (\omega - \omega_D)^2 dx = 0 \implies \omega - \omega_D = 0 \text{ in } D\,.$$

As a consequence ω is constant in D, and from $\omega \in C_0^\infty(D)$ it follows $\omega = 0$.

Example 3.7 In particular, we can also see that in $V = C_0^\infty(D)$, endowed with the scalar product $(v, w)_V = \int_D vw dx$, the Riesz theorem is false. If we had $\omega \in V$ such that

$$\mathcal{F}(v) = \int_D v dx = (\omega, v)_V \quad \forall\, v \in V \,,$$

then we would have

$$(\omega, v)_V = 0 \quad \forall\, v \in N \,,$$

hence $\omega \in N^\perp$. From what we have seen above we would obtain $\omega = 0$, and this is a contradiction as there exists $v \in V$ with $\mathcal{F}(v) = \int_D v dx \neq 0$.

As a final comment, let us come back to the main basic difference between a pre-Hilbertian space and a Hilbert space. We can conclude that, in our context, it is the fact that for a Hilbert space the projection theorem holds, and, as a consequence, the Riesz theorem is valid.

3.3 Exercises

Exercise 3.1 Let V be a normed vector space.

(i) A Cauchy sequence $v_k \in V$ is bounded.
(ii) A Cauchy sequence $v_k \in V$ with a convergent subsequence is convergent.

Solution

(i) Fix $\epsilon_0 > 0$ and consider $N_* \in \mathbb{N}$ such that $\|v_k - v_s\|_V \leq \epsilon_0$ for $k, s \geq N_*$. Then for $k \geq N_*$ it holds

$$\|v_k\|_V \leq \|v_k - v_{N_*}\|_V + \|v_{N_*}\|_V \leq \epsilon_0 + \|v_{N_*}\|_V \,,$$

thus v_k is bounded as there are only a finite number of terms v_k for $k < N_*$.

(ii) Let v_{k_s} be a subsequence convergent to $v_* \in V$. Fix $\epsilon > 0$: we know that there exists $N_\epsilon \in \mathbb{N}$ such that

$$\|v_{k_m} - v_*\|_V \leq \epsilon \quad , \quad \|v_s - v_r\|_V \leq \epsilon$$

for $m \geq N_\epsilon$ and $s, r \geq N_\epsilon$. Since the sequence of integers k_m is strictly increasing (definition of a subsequence...) it holds $k_m \geq m$; thus taking $m \geq N_\epsilon$ it follows

$$\|v_m - v_*\|_V \leq \|v_m - v_{k_m}\|_V + \|v_{k_m} - v_*\|_V \leq 2\epsilon \,.$$

Exercise 3.2 N, defined as in (3.2), is dense in $L^2_*(D)$.

Solution Take $v \in L^2_*(D)$. Since $C_0^\infty(D)$ is dense in $L^2(D)$, we have $\varphi_m \in C_0^\infty(D)$ with $\|\varphi_m - v\|_V \to 0$ as $m \to \infty$. Take $\psi \in C_0^\infty(D)$ with $\int_D \psi dx \neq 0$ and define

$$\hat{\psi} = \frac{\psi}{\int_D \psi dx} ;$$

thus $\hat{\psi} \in C_0^\infty(D)$ and $\int_D \hat{\psi} dx = 1$. Define $I_m = \int_D \varphi_m dx$ and take

$$\tilde{\varphi}_m = \varphi_m - I_m \hat{\psi} .$$

We have $\tilde{\varphi}_m \in C_0^\infty(D)$ and

$$\int_D \tilde{\varphi}_m dx = \int_D \varphi_m dx - I_m \int_D \hat{\psi} dx = 0 ,$$

thus $\tilde{\varphi}_m \in N$. Moreover

$$\|\tilde{\varphi}_m - v\|_V = \|\varphi_m - I_m \hat{\psi} - v\|_V \leq \|\varphi_m - v\|_V + |I_m| \, \|\hat{\psi}\|_V .$$

Since

$$|I_m| = \left| \int_D \varphi_m dx \right| = \left| \int_D (\varphi_m - v) dx \right| \leq (\text{meas}(D))^{1/2} \|\varphi_m - v\|_V ,$$

the result follows.

Chapter 4
Weak Derivatives and Sobolev Spaces

The functional spaces defined in terms of classical derivatives are unfortunately not a suitable setting for a PDEs theory based on weak formulations, as we are not usually able to prove that weak solutions actually belong to such spaces. Therefore other kind of spaces are needed: we must weaken the requirement of smoothness for the functions belonging to them. On the other hand, the bilinear form determined in (2.19) contains derivatives. Summing up, we need to speak about derivatives, but this is not possible in the classical sense: we have to introduce a new concept.

The aim of the next section is to extend the meaning of partial derivative. On the basis of this new idea, in Sect. 4.2 we define the functional spaces that will be used for the variational formulation of the boundary value problems we are interested in.

4.1 Weak Derivatives

Let us start with some preliminaries.

Remark 4.1 (Motivation for Definition of Weak Derivatives) Assume we are given a function $u \in C^1(D)$. Then if $\varphi \in C_0^\infty(D)$ (we will call a function φ belonging to $C_0^\infty(D)$ a *test function*), we see from the integration by parts formula (see Theorem C.2) that

$$\int_D u\mathcal{D}_i\varphi dx = -\int_D \mathcal{D}_i u\varphi dx \quad \forall\, i = 1, \ldots, n\,. \tag{4.1}$$

There are no boundary terms, since φ has a compact support in D and thus vanishes near ∂D. More generally, if k is a positive integer, $u \in C^k(D)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_4

is a multi-index of order $|\alpha| = \alpha_1 + \cdots + \alpha_n = k$, then

$$\int_D u \mathcal{D}^\alpha \varphi dx = (-1)^{|\alpha|} \int_D \mathcal{D}^\alpha u \varphi dx \tag{4.2}$$

This equality holds since

$$\mathcal{D}^\alpha \varphi = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \varphi$$

and we can apply (4.1) $|\alpha|$ times.

We next examine if (4.2) can be generalized to functions u that are not k times continuously differentiable. The left hand side of (4.2) makes sense if u is only locally summable: the problem is rather that if u is not C^k, then the expression $\mathcal{D}^\alpha u$ on the right hand side of (4.2) has no obvious meaning. We overcome this difficulty by asking that there exists a locally summable function ω_α for which formula (4.2) is valid, with ω_α replacing $\mathcal{D}^\alpha u$. (We remember that a function v is locally summable, written $v \in L^1_{\text{loc}}(D)$, if for every measurable subset E that is bounded and satisfies $\overline{E} \subset D$, written $E \subset\subset D$, we have that $v \in L^1(E)$.)

Definition 4.1 Let $D \subset \mathbb{R}^n$ be an open set. Suppose $u, \omega_\alpha \in L^1_{\text{loc}}(D)$, and α is a multi-index. We say that ω_α is the αth-weak partial derivative of u, written

$$\mathcal{D}^\alpha u = \omega_\alpha ,$$

if

$$\int_D u \mathcal{D}^\alpha \varphi dx = (-1)^{|\alpha|} \int_D \omega_\alpha \varphi dx \tag{4.3}$$

for all test functions $\varphi \in C_0^\infty(D)$.

Remark 4.2 Note that, for the sake of simplicity, we are using the same notation $\mathcal{D}^\alpha u$ for weak derivatives and for classical derivatives. However, we believe that in the sequel it will be easy to understand from the context which type of derivative we refer at.

Remark 4.3 Let us not that in the classical sense differentiation is a local concept: we define the derivative of a function u at a point $x_0 \in D \subset \mathbb{R}$, and we say that u is differentiable in D if its derivative exists at each point $x \in D$. Here the concept of weak derivative is global: the weak derivative is a function defined in D.

Proposition 4.1 (Uniqueness of Weak Derivatives) *A weak αth-partial derivatives of u, if it exists, is uniquely defined up to a set of measure zero.*

Proof Assume that $\omega_\alpha, \tilde{\omega}_\alpha \in L^1_{loc}(D)$ satisfy

$$\int_D u\mathcal{D}^\alpha \varphi dx = (-1)^{|\alpha|} \int_D \omega_\alpha \varphi dx = (-1)^{|\alpha|} \int_D \tilde{\omega}_\alpha \varphi dx$$

for all $\varphi \in C_0^\infty(D)$. Then

$$\int_D (\omega_\alpha - \tilde{\omega}_\alpha)\varphi dx = 0$$

for all $\varphi \in C_0^\infty(D)$; whence, since $\omega_\alpha - \tilde{\omega}_\alpha \in L^1_{loc}(D)$, we have that $\omega_\alpha - \tilde{\omega}_\alpha = 0$ almost everywhere by du Bois-Reymond lemma (see Lemma 6.1). □

Remark 4.4 Note that if a function u is continuously differentiable in D, then its classical derivative $\mathcal{D}_i u$ coincides with its weak derivative, as it is a function which belongs to $L^1_{loc}(D)$ and satisfies (4.3). Hence the concept of weak derivative is a generalization of the concept of classical derivative.

However, take into account that there are differentiable functions (but not *continuously* differentiable) for which the classical derivatives are not the weak derivatives, as they do not belong to $L^1_{loc}(D)$ (see Exercise 4.1).

Exercise 4.1 Find a function $u : (-1, 1) \mapsto \mathbb{R}$ which is differentiable and whose classical derivative u' does not belong to $L^1_{loc}(-1, 1)$ (therefore u' is not the weak derivative of u).

Proposition 4.2 *The map $u \mapsto \omega_\alpha$, where ω_α is the αth-weak partial derivatives of u, is linear.*

Proof Straightforward from the definition. □

Exercise 4.2 Set $X_\alpha = \{v \in L^2(D) \,|\, \mathcal{D}^\alpha v \in L^2(D)\}$, where α is a multi-index. The operator $\mathcal{D}^\alpha : u \mapsto \mathcal{D}^\alpha u$ defined in X_α is a closed operator from $L^2(D)$ to $L^2(D)$, namely, if for $u_m \in X_\alpha$ one has $u_m \to u$ in $L^2(D)$ and $\mathcal{D}^\alpha u_m \to w_\alpha$ in $L^2(D)$ then it follows $w_\alpha = \mathcal{D}^\alpha u$.

Example 4.1 Let $n = 1$, $D = (0, 2)$, and

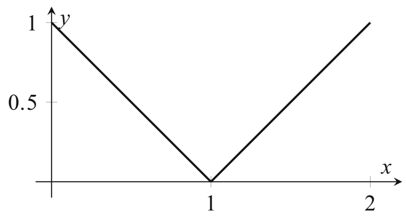
$$u(x) = \begin{cases} 1 - x & \text{if } 0 < x \le 1 \\ x - 1 & \text{if } 1 < x < 2 \end{cases} \tag{4.4}$$

(see Fig. 4.1).

Define

$$\omega(x) = \begin{cases} -1 & \text{if } 0 < x \le 1 \\ 1 & \text{if } 1 < x < 2\,. \end{cases} \tag{4.5}$$

Fig. 4.1 The graph of the function u in (4.4)



Let us show that $u' = \omega$ in the weak sense. To see this, we must prove that

$$\int_0^2 u\varphi' dx = -\int_0^2 \omega\varphi dx$$

for each $\varphi \in C_0^\infty(D)$. We easily compute, integrating by parts in (0, 1) and in (1, 2),

$$\begin{aligned}\int_0^2 u\varphi' dx &= \int_0^1 (1-x)\varphi' dx + \int_1^2 (x-1)\varphi' dx \\ &= \int_0^1 \varphi dx - \underbrace{\varphi(0)}_{=0} - \int_1^2 \varphi dx + \underbrace{\varphi(2)}_{=0} \\ &= -\int_0^2 \omega\varphi dx\,,\end{aligned}$$

as required.

Example 4.2 Let $n = 1$, $D = (0, 2)$, and

$$u(x) = \begin{cases} 1 & \text{if } 0 < x \le 1 \\ 2 & \text{if } 1 < x < 2 \end{cases} \tag{4.6}$$

(see Fig. 4.2). We claim that u' does not exist in the weak sense. To check this, we must show that it is not possible to find any function $\omega \in L^1_{\mathrm{loc}}(D)$ satisfying

$$\int_0^2 u\varphi' dx = -\int_0^2 \omega\varphi dx \tag{4.7}$$

for all $\varphi \in C_0^\infty(D)$. Suppose, by contradiction, that (4.7) is valid for some $\omega \in L^1_{\mathrm{loc}}(D)$ and all $\varphi \in C_0^\infty(D)$. Then, taking into account that $\varphi(0) = \varphi(2) = 0$,

$$\begin{aligned}-\int_0^2 \omega\varphi dx = \int_0^2 u\varphi' dx &= \int_0^1 \varphi' dx + 2\int_1^2 \varphi' dx \\ &= \varphi(1) - 2\varphi(1) = -\varphi(1)\,.\end{aligned} \tag{4.8}$$

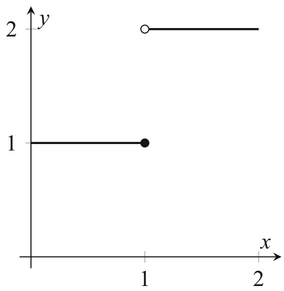


Fig. 4.2 The graph of the function u in (4.6)

Choose in $C_0^\infty(D)$ a sequence $\{\varphi_m\}_{m=1}^\infty$ satisfying

$$0 \le \varphi_m \le 1,! \quad \varphi_m(1) = 1, \quad \varphi_m(x) \to 0 \text{ for all } x \neq 1, \quad \text{supp}\varphi_m \subset K \subset\subset (0, 2)\,.$$

Replacing φ by φ_m in (4.8) and sending $m \to \infty$, we discover, by the Lebesgue dominated convergence theorem,

$$1 = \lim_{m\to\infty} \varphi_m(1) = \lim_{m\to\infty} \int_0^2 \omega\varphi_m dx = 0\,,$$

a contradiction. Note that we can apply the Lebesgue dominated convergence theorem, as $\int_0^2 \omega\varphi_m dx = \int_K \omega\varphi_m dx$ and $|\omega\varphi_m| \le |\omega|$, with $\omega \in L^1(K)$.

Remark 4.5 The computations in Example 4.2 in particular show that the functional $\varphi \mapsto \varphi(1)$, $\varphi \in C_0^\infty(0, 2)$, cannot be represented by $\int_0^2 \omega\varphi dx$ for a function $\omega \in L^1_{\text{loc}}(0, 2)$. In other words, the Dirac δ "function" is not a function.

An example of sequence $\varphi_m \in C_0^\infty(0, 2)$ with the required properties is given by

$$\varphi_m(x) = \begin{cases} e^{1-\frac{1}{1-4m^2|x-1|^2}} & \text{if } |x-1| < \frac{1}{2m} \\ 0 & \text{if } |x-1| \ge \frac{1}{2m} \end{cases} \tag{4.9}$$

(see Fig. 4.3).

Exercise 4.3 Let φ_m as in (4.9) and set $\psi_m(x) = I_m^{-1}\varphi_m(x)$, $x \in (0, 2)$, where $I_m = \int_0^2 \varphi_m dx$. Show that $\int_0^2 \psi_m\varphi dx \to \varphi(1)$ for each $\varphi \in C_0^\infty(0, 2)$. Repeat the proof for each $\varphi \in C^0(0, 2)$.

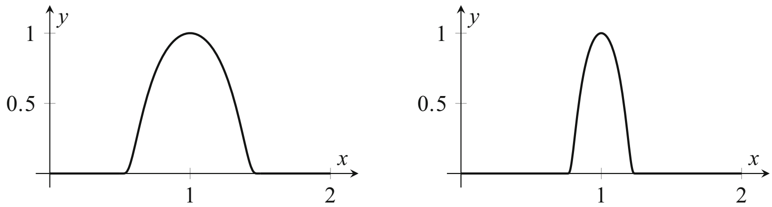


Fig. 4.3 The graph of the function φ_m in (4.9) for $m = 1$ (left) and $m = 2$ (right)

4.2 Sobolev Spaces

In this section we finally introduce the infinite dimensional vector spaces that furnish the "right" framework for the weak formulation of partial differential equations. In some particular case, these spaces had been considered since the beginning of the last century, but their systematic definition and use dates back to the thirties, especially in the papers by Sergei L. Sobolev[1 2].

Take $1 \le p \le +\infty$ and let k be a non-negative integer. Now we define certain functional spaces, whose elements have weak derivatives of some order lying in L^p.

Definition 4.2 Let $D \subset \mathbb{R}^n$ be an open set. The Sobolev space

$$W^{k,p}(D)$$

consists of all locally summable function $u : D \mapsto \mathbb{R}$ such that for each multi-index α with $|\alpha| \le k$ the derivative $\mathcal{D}^\alpha u$ exists in the weak sense and belongs to $L^p(D)$.

Remark 4.6

(i) If $p = 2$, we usually write

$$W^{k,2}(D) = H^k(D)\,.$$

In particular, $W^{0,2}(D) = H^0(D) = L^2(D)$.

(ii) From the definition it is clear that we identify functions in $W^{k,p}(D)$ if they agree almost everywhere.

[1] S.L. Sobolev [26].

[2] S.L. Sobolev [27].

Definition 4.3 If $v \in W^{k,p}(D)$, with $1 \leq p < +\infty$ we define its norm to be

$$\|v\|_{W^{k,p}(D)} := \Big(\sum_{|a|\leq k} \int_D |\mathcal{D}^\alpha v|^p dx\Big)^{1/p} = \Big(\sum_{|a|\leq k} \|\mathcal{D}^\alpha v\|^p_{L^p(D)}\Big)^{1/p}.$$

If $p = +\infty$ the norm is defined as

$$\|v\|_{W^{k,\infty}(D)} := \max_{|\alpha|\leq k} \|\mathcal{D}^\alpha v\|_{L^\infty(D)}.$$

For $1 \leq p \leq +\infty$ we may also use the equivalent norm defined as

$$\sum_{|\alpha|\leq k} \|\mathcal{D}^\alpha v\|_{L^p(D)}.$$

Definition 4.4 We denote by

$$W_0^{k,p}(D)$$

the closure of $C_0^\infty(D)$ in $W^{k,p}(D)$.

Thus $v \in W_0^{k,p}(D)$ if and only if there exist functions $v_m \in C_0^\infty(D)$ such that $v_m \to v$ in $W^{k,p}(D)$. We will se later (see Remark 6.5) that we can interpret $W_0^{k,p}(D)$ as the space of those functions $v \in W^{k,p}(D)$ such that

$$\text{“}\mathcal{D}^\alpha v = 0 \text{ on } \partial D\text{”} \quad \text{for all } |\alpha| \leq k-1.$$

It is customary to write

$$W_0^{k,2}(D) = H_0^k(D).$$

Remark 4.7 The norm $\|\cdot\|_{W^{k,p}(D)}$ is actually a norm. Indeed

1. $\|v\|_{W^{k,p}(D)} = \Big(\sum_{|a|\leq k} \underbrace{\|\mathcal{D}^\alpha v\|^p_{L^p(D)}}_{\geq 0}\Big)^{1/p} \geq 0$.
2. If $v = 0$ then trivially $\|v\|_{W^{k,p}(D)} = 0$. On the other hand, if $\|v\|_{W^{k,p}(D)} = 0$ we have $\Big(\sum_{|a|\leq k} \|\mathcal{D}^\alpha v\|^p_{L^p(D)}\Big)^{1/p} = 0$, thus in particular $\|v\|_{L^p(D)} = 0$ which implies $v = 0$ a.e. in D.

3. Take $\lambda \in \mathbb{R}$: then

$$\|\lambda v\|_{W^{k,p}(D)} = \Big(\sum_{|a|\le k} \|\mathcal{D}^\alpha(\lambda v)\|^p_{L^p(D)} \Big)^{1/p}$$
$$= |\lambda| \Big(\sum_{|a|\le k} \|\mathcal{D}^\alpha v\|^p_{L^p(D)} \Big)^{1/p} = |\lambda|\, \|v\|_{W^{k,p}(D)} \,.$$

4. We have finally to verify that the triangular inequality $\|w + v\|_{W^{k,p}(D)} \le \|w\|_{W^{k,p}(D)} + \|v\|_{W^{k,p}(D)}$ holds true. Indeed, if $1 \le p < +\infty$, the discrete Minkowski's inequality implies

$$\|w+v\|_{W^{k,p}(D)} = \Big(\sum_{|a|\le k} \|\mathcal{D}^\alpha(w+v)\|^p_{L^p(D)} \Big)^{1/p}$$
$$= \Big(\sum_{|a|\le k} \|\mathcal{D}^\alpha w + \mathcal{D}^\alpha v\|^p_{L^p(D)} \Big)^{1/p}$$
$$\le \Big(\sum_{|a|\le k} (\|\mathcal{D}^\alpha w\|_{L^p(D)} + \|\mathcal{D}^\alpha v\|_{L^p(D)})^p \Big)^{1/p}$$
$$\le \Big(\sum_{|a|\le k} \|\mathcal{D}^\alpha w\|^p_{L^p(D)} \Big)^{1/p} + \Big(\sum_{|a|\le k} \|\mathcal{D}^\alpha v\|^p_{L^p(D)} \Big)^{1/p}$$
$$= \|w\|_{W^{k,p}(D)} + \|v\|_{W^{k,p}(D)} \,.$$

The case $p = +\infty$ is trivial.

Theorem 4.1 *The space $W^{k,p}(D)$ is a Banach space.*

Proof We have already proved that $W^{k,p}(D)$ is a normed space. It remains to prove that each Cauchy sequence $\{v_n\}_{n=1}^\infty$ is convergent in $W^{k,p}(D)$. Assume that for each $\varepsilon > 0$ it exists $M_\varepsilon \in \mathbb{N}$ such that for all $n, m > M_\varepsilon$

$$\|v_n - v_m\|_{W^{k,p}(D)} = \Big(\sum_{|\alpha|\le k} \|\mathcal{D}^\alpha(v_n - v_m)\|^p_{L^p(D)} \Big)^{1/p} \le \varepsilon$$

for $1 \le p < +\infty$ or

$$\|v_n - v_m\|_{W^{k,\infty}(D)} = \max_{|\alpha|\le k} \|D^\alpha(v_n - v_m)\|_{L^\infty(D)} \le \varepsilon \,.$$

In particular we have that for all α with $|\alpha| \le k$

$$\|\mathcal{D}^\alpha(v_n - v_m)\|_{L^p(D)} \le \varepsilon \,,$$

i.e., $\{\mathcal{D}^\alpha v_n\}_{n=1}^\infty$ is a Cauchy sequence in $L^p(D)$. Since $L^p(D)$ is a Banach space, for any α with $|\alpha| \leq k$ there exits $v_\alpha \in L^p(D)$, such that

$$\mathcal{D}^\alpha v_n \stackrel{L^p}{\rightarrow} v_\alpha \text{ as } n \rightarrow \infty \,.$$

In particular with $\alpha = (0, \ldots, 0)$ we have that $v_n \stackrel{L^p}{\rightarrow} v_{(0,\ldots,0)}$ (which we denote by v_0). We now claim that

$$v_0 \in W^{k,p}(D) \text{ and } \mathcal{D}^\alpha v_0 = v_\alpha \,.$$

To verify this assertion, fix $\varphi \in C_0^\infty(D)$. Then

$$\begin{aligned}\int_D v_0 \mathcal{D}^\alpha \varphi dx &= \lim_{n\rightarrow\infty} \int_D v_n \mathcal{D}^\alpha \varphi dx = \\ &= \lim_{n\rightarrow\infty} (-1)^{|\alpha|} \int_D \mathcal{D}^\alpha v_n \varphi dx = \\ &= (-1)^{|\alpha|} \int_D v_\alpha \varphi dx \,.\end{aligned}$$

Thus we have $\mathcal{D}^\alpha v_0 = v_\alpha$ and consequently $\mathcal{D}^\alpha v_n \stackrel{L^p}{\rightarrow} \mathcal{D}^\alpha v_0$ for all $|\alpha| \leq k$, which means $v_n \rightarrow v_0$ in $W^{k,p}(D)$, as required. □

Remark 4.8 The Sobolev space $W^{k,2}(D) = H^k(D)$ is a Hilbert space. In fact, it is easy to prove that the norm

$$\|v\|^2_{H^k(D)} = \sum_{|\alpha|\leq k} \int_D |\mathcal{D}^\alpha v|^2 \, dx = \sum_{|\alpha|\leq k} \|\mathcal{D}^\alpha v\|^2_{L^2(D)}$$

is induced by the scalar product

$$(w, v)_{H^k(D)} = \sum_{|\alpha|\leq k} \int_D \mathcal{D}^\alpha w \mathcal{D}^\alpha v \, dx \,.$$

In particular, if $k = 1$ we have that

$$(w, v)_{H^1(D)} = \int_D wv \, dx + \int_D \nabla w \cdot \nabla v \, dx$$

and therefore

$$\|v\|_{H^1(D)} = \left(\int_D v^2 dx + \int_D |\nabla v|^2 dx \right)^{1/2} .$$

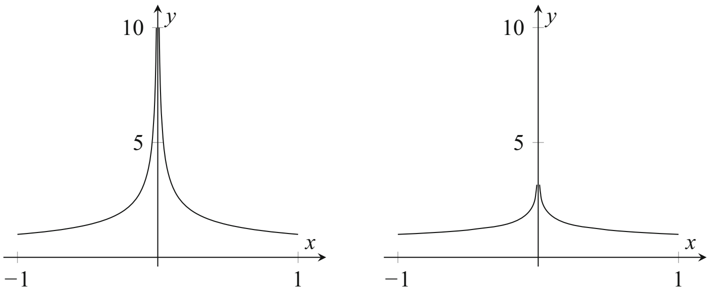


Fig. 4.4 The graph of the function $|x|^{-\alpha}$ for $\alpha = 1/2$ (left) and $\alpha = 1/4$ (right). (The graph is drawn for $0.01 \leq |x| \leq 1$)

Remark 4.9 It is proved that $W^{k,p}(D)$ is a reflexive Banach space when $1 < p < +\infty$ and is a separable Banach space when $1 \leq p < +\infty$ (see Adams [1, Theorem 3.5]).

Example 4.3 Take $D = B_1$, the open unit ball in $\mathbb{R}^n$ centered at 0, and

$$u(x) = |x|^{-\alpha} \quad (x \in D, x \neq 0)$$

(see Fig. 4.4). We notice that $u \notin L^\infty(D)$ and we want to find for which $\alpha > 0$, $p \in [1, +\infty)$, $n \geq 1$ the function u belongs to $W^{1,p}(D)$.

To answer, note first that u is smooth away from 0, i.e., for x with $|x| > 0$ we have that $x \mapsto u(x) \in C^\infty$; thus in this set we can compute the derivatives in the classical sense. We have

$$\begin{aligned} \mathcal{D}_i u &= (-\alpha)\, |x|^{-\alpha-1}\, \mathcal{D}_i(|x|) = (-\alpha)\, |x|^{-\alpha-1}\, \mathcal{D}_i\Big(\big(\sum_{j=1}^n x_j^2\big)^{1/2}\Big) \\ &= (-\alpha)\, |x|^{-\alpha-1}\, \frac{1}{2}\, \frac{1}{|x|}\, 2x_i = \frac{-\alpha x_i}{|x|^{\alpha+2}}\,; \end{aligned}$$

therefore for $x \neq 0$ it holds

$$|\nabla u(x)| = \frac{|-\alpha|\, |x|}{|x|^{\alpha+2}} = \frac{\alpha}{|x|^{\alpha+1}}\,.$$

For $i = 1, \ldots, n$ let us define

$$\omega_i(x) = \begin{cases} \mathcal{D}_i u(x) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0\,. \end{cases}$$

Let us determine for which values of α we have $u \in L^p(D)$ and $\omega_i \in L^p(D)$. We can employ polar coordinates (in dimension n), and we find

$$\int_D |u|^p dx = \kappa_n \int_0^1 \rho^{-\alpha p} \rho^{n-1} d\rho = \kappa_n \int_0^1 \rho^{-\alpha p+n-1} d\rho \,,$$

where κ_n is the $(n-1)$-measure of the set $\{x \in \mathbb{R}^n \,|\, |x| = 1\}$. Thus $u \in L^p(D)$ if and only if $\alpha p < n$ (in particular, $u \in L^1(D)$ if and only if $\alpha < n$). A similar calculation shows that

$$\int_D \Big(\sum_{1=1}^n \omega_i^2 \Big)^{p/2} dx = \alpha^p \kappa_n \int_0^1 \rho^{-(\alpha+1)p+n-1} d\rho \,,$$

thus $\omega_i \in L^p(D)$ if and only if $(\alpha + 1)p < n$ (and $\omega_i \in L^1(D)$ if and only if $\alpha + 1 < n$).

Assume therefore $n \geq 2$ and $\alpha < n - 1$, so that $u, \omega_i \in L^1(D)$ and we are allowed to consider weak derivatives of u. We want to show that the weak derivative $\mathcal{D}_i u$ is equal to ω_i. Let $\varphi \in C_0^\infty(D)$ and fix $\varepsilon > 0$. Then, denoting by B_ε the ball centered at 0 with radius $\varepsilon > 0$,

$$\begin{aligned} \textstyle\int_{D\setminus B_\varepsilon} u \mathcal{D}_i \varphi dx &= - \textstyle\int_{D\setminus B_\varepsilon} \mathcal{D}_i u \varphi dx - \int_{\partial B_\varepsilon} u \varphi n_i dS_x \\ &= - \textstyle\int_{D\setminus B_\varepsilon} \omega_i \varphi dx - \int_{\partial B_\varepsilon} u \varphi n_i dS_x \,, \end{aligned}$$

where n denotes the unit normal on ∂B_ε, external to B_ε. It holds

$$\left| \int_{\partial B_\varepsilon} u \varphi n_i dS_x \right| \leq \|\varphi\|_{L^\infty(D)} \int_{\partial B_\varepsilon} \varepsilon^{-\alpha} dS_x = C_{n,\varphi} \varepsilon^{n-1-\alpha} \to 0 \,,$$

as $\alpha < n - 1$. Thus passing to the limit as $\varepsilon \to 0^+$ and taking into account that $u\mathcal{D}_i\varphi \in L^1(D)$ and $\omega_i \varphi \in L^1(D)$ one finds

$$\int_D u \mathcal{D}_i \varphi dx = - \int_D \omega_i \varphi dx$$

for all $\varphi \in C_0^\infty(D)$. We have thus proved that $\mathcal{D}_i u = \omega_i$, and in conclusion $u \in W^{1,p}(D)$ if and only if $\alpha < (n - p)/p$; in particular $u \notin W^{1,p}(D)$ for each $p \geq n$.

This example seems to show that unbounded functions are not allowed to belong to $W^{1,p}(D)$ when $p \geq n$: we will see later on that this in fact true, but for the stronger restriction $p > n$.

Exercise 4.4 Let $1 \leq p \leq +\infty$, $u \in W^{1,p}(D)$, $\varphi \in C_0^\infty(D)$. Then $u\varphi \in W^{1,p}(D)$ and $\mathcal{D}_i(u\varphi) = \varphi \mathcal{D}_i u + u \mathcal{D}_i \varphi$.

Exercise 4.5 Let $u \in H^1_0(D)$ and $v \in H^1(D)$ (or viceversa). Then

$$\int_D v\mathcal{D}_i u dx = -\int_D u\mathcal{D}_i v dx\,.$$

4.3 Exercises

Exercise 4.1 Find a function $u : (-1,1) \mapsto \mathbb{R}$ which is differentiable and whose classical derivative u' does not belong to $L^1_{\rm loc}(-1,1)$ (therefore u' is not the weak derivative of u).

Solution Take

$$u(x) = \begin{cases} x^2\cos(x^{-2}) & \text{for } x \in (-1,1), x \neq 0 \\ 0 & \text{for } x = 0\,. \end{cases}$$

Clearly we have $u'(0) = 0$ and $u'(x) = 2x\cos(x^{-2}) + 2x^{-1}\sin(x^{-2})$ for $x \neq 0$. Since the first term in $u'(x)$ can be extended to a continuous function in $[-1,1]$, we focus on $x^{-1}\sin(x^{-2})$. Consider the interval $(-\frac{1}{2},\frac{1}{2}) \subset\subset (-1,1)$: it holds

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |x^{-1}|\,|\sin(x^{-2})|dx = 2\int_0^{\frac{1}{2}} x^{-1}|\sin(x^{-2})|dx \underbrace{=}_{x=\frac{1}{\sqrt{t}}} \int_4^{+\infty} \frac{|\sin t|}{t}dt = +\infty\,.$$

Exercise 4.2 Set $X_\alpha = \{v \in L^2(D)\,|\,\mathcal{D}^\alpha v \in L^2(D)\}$, where α is a multi-index. The operator $\mathcal{D}^\alpha : u \mapsto \mathcal{D}^\alpha u$ defined in X_α is a closed operator from $L^2(D)$ to $L^2(D)$, namely, if for $u_m \in X_\alpha$ one has $u_m \to u$ in $L^2(D)$ and $\mathcal{D}^\alpha u_m \to w_\alpha$ in $L^2(D)$ then it follows $w_\alpha = \mathcal{D}^\alpha u$.

Solution The definition of $\mathcal{D}^\alpha u_m$ reads

$$\int_D u_m\,\mathcal{D}^\alpha \varphi dx = (-1)^{|\alpha|}\int_D \mathcal{D}^\alpha u_m\,\varphi dx$$

for each $\varphi \in C^\infty_0(D)$. Then passing to the limit in this equality we find

$$\int_D u\,\mathcal{D}^\alpha \varphi dx = (-1)^{|\alpha|}\int_D w_\alpha\,\varphi dx\,,$$

hence $w_\alpha = \mathcal{D}^\alpha u$.

Exercise 4.3 Let φ_m as in (4.9) and set $\psi_m(x) = I_m^{-1}\varphi_m(x)$, $x \in (0,2)$, where $I_m = \int_0^2 \varphi_m dx$. Show that $\int_0^2 \psi_m \varphi dx \to \varphi(1)$ for each $\varphi \in C_0^\infty(0,2)$. Repeat the proof for each $\varphi \in C^0(0,2)$.

Solution Since $\int_0^2 \psi_m(x)dx = 1$, we have

$$\begin{aligned}
\left|\int_0^2 \psi_m(x)\varphi(x)dx - \varphi(1)\right| &= \left|\int_0^2 \psi_m(x)\big(\varphi(x)-\varphi(1)\big)dx\right| \\
&= \left|\int_{1-\frac{1}{2m}}^{1+\frac{1}{2m}} \psi_m(x)\big(\varphi(x)-\varphi(1)\big)dx\right| \\
&\le \max_{|x-1|\le\frac{1}{2m}} |\varphi(x)-\varphi(1)| \left|\int_{1-\frac{1}{2m}}^{1+\frac{1}{2m}} \psi_m(x)dx\right| \\
&= \max_{|x-1|\le\frac{1}{2m}} |\varphi(x)-\varphi(1)| .
\end{aligned}$$

Since in both cases $\varphi \in C_0^\infty(0,2)$ and $\varphi \in C^0(0,2)$ we have that φ is uniformly continuous in each compact subset K of $(0,2)$, the thesis follows with the same argument.

Exercise 4.4 Let $1 \le p \le +\infty$, $u \in W^{1,p}(D)$, $\varphi \in C_0^\infty(D)$. Then $u\varphi \in W^{1,p}(D)$ and $\mathcal{D}_i(u\varphi) = \varphi\mathcal{D}_i u + u\mathcal{D}_i\varphi$.

Solution Clearly $u\varphi, \varphi\mathcal{D}_i u, u\mathcal{D}_i\varphi \in L^p(D)$ (u and $\mathcal{D}_i u$ belong to $L^p(D)$, and φ is smooth…). Thus it is enough to show that $\mathcal{D}_i(u\varphi) = \varphi\mathcal{D}_i u + u\mathcal{D}_i\varphi$. We have, for $\psi \in C_0^\infty(D)$

$$\begin{aligned}
\int_D u\varphi\mathcal{D}_i\psi dx &= \int_D u\mathcal{D}_i(\varphi\psi)dx - \int_D u\psi\mathcal{D}_i\varphi dx \\
&= -\int_D (\mathcal{D}_i u)\varphi\psi dx - \int_D u\mathcal{D}_i\varphi\psi dx \\
&= -\int_D [\varphi\mathcal{D}_i u + u\mathcal{D}_i\varphi]\,\psi dx ,
\end{aligned}$$

as $\varphi\psi \in C_0^\infty(D)$.

Exercise 4.5 Let $u \in H_0^1(D)$ and $v \in H^1(D)$ (or viceversa). Then

$$\int_D v\mathcal{D}_i u dx = -\int_D u\mathcal{D}_i v dx .$$

Solution Take $u_k \to u$ in $H^1(D)$ with $u_k \in C_0^\infty(D)$. The result is true for u_k, v and then we just pass to the limit to conclude the proof.

Chapter 5
Weak Formulation of Elliptic PDEs

In this chapter we want to derive and analyze the weak formulation of the boundary value problems associated to the (uniformly) elliptic operator

$$Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w) + \sum_{i=1}^{n} b_i \mathcal{D}_i w + a_0 w \,, \tag{5.1}$$

where, as done in Sects. 2.1 and 2.2, we assume that $D \subset \mathbb{R}^n$ is a bounded, connected, open set, $a_{ij} \in L^\infty(D)$ for $i, j = 1, \ldots, n$, $b_i \in L^\infty(D)$ for $i = 1, \ldots, n$, $a_0 \in L^\infty(D)$. When considering the Robin problem, the assumptions on the coefficient are $\kappa \in L^\infty(\partial D)$, $\kappa \geq 0$ a.e. on ∂D and $\int_{\partial D} \kappa dS_x \neq 0$. On the data we assume that $f \in L^2(D)$ and $g \in L^2(\partial D)$ (Neumann and Robin problems), $g \in L^2(\Gamma_N)$ (mixed problem).

5.1 Weak Formulation of Boundary Value Problems

We have seen in Chap. 2 that a standard way for rewriting the boundary value problem

$$\begin{cases} Lu = f & \text{in } D \\ \text{BC} & \text{on } \partial D \end{cases}$$

is:

1. multiply the equation by a test function;
2. integrate in D;

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_5

3. reduce the problem to a more suitable form (we could say: a more balanced form) by integrating by parts the term stemming from the principal part (using in this computation the information given by the boundary condition).

   This typically leads to a problem of the form

$$u \in V \ : \ B(u, v) = F(v) \quad \forall\, v \in V$$

(see (2.23); see also (2.21), which has been specifically obtained taking into account the homogeneous Dirichlet boundary condition). In order to analyze this problem by means of tools from functional analysis, we have also clarified in Chap. 2 that the infinite dimensional vector space V must be a Hilbert space.

Our aim now is to make precise this procedure for all the boundary value problems we are interested in: Dirichlet (homogeneous case), Neumann, mixed (homogeneous case on Γ_D), Robin.

Dirichlet BC In this case the problem is

$$\begin{cases} Lu = f & \text{in } D \\ u = 0 & \text{on } \partial D\,. \end{cases} \tag{5.2}$$

For the ease of the reader, we repeat here the procedure presented in Chap. 2. This procedure is formal, namely, we are implicitly assuming that all the terms we are going to write have a meaning. We start choosing a function $v \in C_0^\infty(D)$, thus satisfying $v_{|\partial D} = 0$, and we multiply the equation by v. Integrating over D we obtain

$$-\int_D \sum_{i,j=1}^n \mathcal{D}_i(a_{ij}\mathcal{D}_j u) v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx = \int_D f v dx\,.$$

Integrating by parts, we obtain

$$\begin{aligned} -\int_D \sum_{i,j=1}^n \mathcal{D}_i(a_{ij}\mathcal{D}_j u) v dx &= \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx \\ &\qquad \underbrace{-\int_{\partial D} \sum_{i,j=1}^n n_i a_{ij} \mathcal{D}_j u v_{|\partial D} dS_x}_{=0,\ \text{as}\ v_{|\partial D}=0} \\ &= \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx\,. \end{aligned}$$

Thus we are left with

$$\int_D \sum_{i,j=1}^{n} a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx = \int_D f v dx \,.$$

Up to here, as we said, this is just a formal procedure; the aim now is to check for which choice of the space V this equation has a meaning for $u, v \in V$.
If $u \in H^1(D)$ (thus the derivatives appearing in the equation above have to be considered as weak derivatives) all the terms are well-defined. Moreover, since the space of test functions $C_0^\infty(D)$ is dense in the Sobolev space $H_0^1(D)$, it is easy to check that by continuity we can extend this equation to test functions $v \in H_0^1(D)$. Finally, a reasonable interpretation of the boundary condition $u_{|\partial D} = 0$ is that u can be approximated by functions vanishing near the boundary: thus we can require $u \in H_0^1(D)$. Our last step now is clear: the Hilbert space we choose is $V = H_0^1(D)$.
We observe that the original problem (5.2) has been transformed into a set of infinitely many integral equations, or, equivalently, into an equation in the infinite dimensional vector space $V = H_0^1(D)$.
We recall the definitions of the bilinear form

$$B_L(w, v) = \int_D \sum_{i,j=1}^{n} a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx$$

and the linear functional

$$F_D(v) = \int_D f v dx$$

(see (2.19) and (2.20)). Problem (5.2) has been therefore rewritten in the weak form:

$$\text{find } u \in V \ : \ B(u, v) = F(v) \quad \forall\, v \in V \,, \tag{5.3}$$

where

$$B(w, v) = B_L(w, v) \ , \ F(v) = \int_D f v dx \ , \ V = H_0^1(D) \,. \tag{5.4}$$

Neumann BC In this case the problem is

$$\begin{cases} Lu = f & \text{in } D \\ \displaystyle\sum_{i,j=1}^{n} n_i a_{ij}\mathcal{D}_j u = g & \text{on } \partial D \,. \end{cases} \tag{5.5}$$

Besides conditions (2.16) on the coefficients and (2.17) on the right hand side of the equation, as already told here we also assume $g \in L^2(\partial D)$.
In this case the structure of the boundary condition is qualitatively different from that of the Dirichlet problem. In particular, there is no longer reason to impose to the test function v to vanish on ∂D. Thus we choose $v \in C^\infty(\overline{D})$ and we multiply the differential equation by v. Proceeding formally, we integrate over D and obtain

$$\int_D - \sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j u) v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx = \int_D f v dx \, .$$

Integrating by parts the first term, the following boundary integral appears:

$$- \int_{\partial D} \sum_{i,j=1}^{n} n_i a_{ij} \mathcal{D}_j u v_{|\partial D} dS_x \, . \tag{5.6}$$

Using the Neumann condition it can be rewritten as $-\int_{\partial D} g v_{|\partial D} dS_x$; thus we have finally obtained

$$\int_D \sum_{i,j=1}^{n} a_{ij} \mathcal{D}_j u \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx$$
$$= \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \, .$$

Proceeding similarly to the Dirichlet case, we can choose V equal to the closure of $C^\infty(\overline{D})$ with respect to the $H^1(D)$-norm. We will see in Theorem 6.3 that, if D has a Lipschitz continuous boundary ∂D, the subspace $C^\infty(\overline{D})$ is dense in $H^1(D)$. Thus we choose $V = H^1(D)$, and assume that the boundary ∂D is Lipschitz continuous (see Appendix B for a precise definition of this regularity assumption).
Let us now give a look at the equation we have obtained. Four of its terms were also present in the Dirichlet case, thus we already know that they have a meaning for $u \in H^1(D)$. The new one is $\int_{\partial D} g v_{|\partial D} dS_x$: this needs some additional attention. In fact, first of all we have to show that it is possible to give a meaning to $v_{|\partial D}$ for $v \in H^1(D)$ (remember that ∂D is a set whose measure is equal to zero...), and moreover show that it belongs to $L^2(\partial D)$; secondly, if we want that the right hand side of the equation above is bounded for $v \in H^1(D)$, we need that the following inequality holds true:

$$\int_{\partial D} v_{|\partial D}^2 dS_x \leq C_* \int_D (v^2 + |\nabla v|^2) dx \quad \forall \, v \in H^1(D) \tag{5.7}$$

for a suitable $C_* > 0$. We will see in Theorem 6.5 that, for $v \in H^1(D)$, both these issues have a positive answer: the value $v_{|\partial D}$ will be called the *trace* of v and (5.7) will be called the *trace inequality*.

Problem (5.5) has been therefore rewritten in the weak form:

$$\text{find } u \in V \ : \ B(u, v) = F(v) \quad \forall\, v \in V \,, \tag{5.8}$$

where

$$B(w, v) = B_L(w, v) \ , \quad F(v) = \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \ , \quad V = H^1(D) \,. \tag{5.9}$$

Remark 5.1 The "thumb rule" for identifying which are the Dirichlet boundary condition and the Neumann boundary condition associated to a general second order partial differential operator $\mathcal{L}$ (not necessarily the elliptic operator L in (5.1)) is the following. Multiply $\mathcal{L}u$ by v, integrate in D and integrate by parts the principal (namely, second order) terms. Some terms given by integrals on the boundary ∂D will appear (for the operator L they are shown in (5.6)): they can be canceled either by putting to 0 the first order terms related to u or by putting to 0 the zero order terms related to v. The Neumann boundary condition is expressed by the first order terms related to u, the Dirichlet boundary condition is expressed by the zero order terms related to v. For the homogeneous Dirichlet boundary value problem the boundary condition is inserted as a constraint in the definition of the variational space V, whereas for the (non-homogeneous) Neumann boundary value problem the boundary condition is used to give a boundary contribution to the linear and bounded functional $F(\cdot)$ at the right hand side of the variational problem.

Mixed BC In this case the problem is

$$\begin{cases} Lu = f & \text{in } D \\ u = 0 & \text{on } \Gamma_D \\ \displaystyle\sum_{i,j=1}^{n} n_i a_{ij} \mathcal{D}_j u = g & \text{on } \Gamma_N \,, \end{cases} \tag{5.10}$$

where $\partial D = \overline{\Gamma_D} \cup \overline{\Gamma_N}$, $\Gamma_D \cap \Gamma_N = \emptyset$, and, besides (2.16) and (2.17), we assume $g \in L^2(\Gamma_N)$.

Choose as space of test functions

$$C^\infty_{\Gamma_D}(\overline{D}) = \{v \in C^\infty(\overline{D}) \mid v = 0 \text{ in a neighborhood of } \Gamma_D\} \,.$$

Multiplying the differential equation by $v \in C^\infty_{\Gamma_D}(\overline{D})$ and integrating over D we obtain

$$\int_D -\sum_{i,j=1}^n \mathcal{D}_i(a_{ij}\mathcal{D}_j u)v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx = \int_D f v dx \,.$$

By proceeding as in the previous cases, integrating by parts the first term and using the boundary conditions we obtain

$$\int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx$$
$$= \int_D f v dx + \int_{\Gamma_N} g v_{|\Gamma_N} dS_x \,.$$

We take the space V equal to the closure in $H^1(D)$ of $C^\infty_{\Gamma_D}(\overline{D})$. It will be shown that, if ∂D (here Γ_D would be enough…) is a Lipschitz continuous boundary, this closed subspace is $H^1_{\Gamma_D}(D)$ (see Sect. 6.5 for a precise definition and further details). Moreover, it will be also possible to define the trace of v on Γ_D and on Γ_N, to show that $v_{|\Gamma_D} = 0$, that $v_{|\Gamma_N} \in L^2(\Gamma_N)$ and finally that the map from $v \in H^1_{\Gamma_D}(D)$ to its trace $v_{|\Gamma_N} \in L^2(\Gamma_N)$ is continuous, namely, that the following *trace inequality* holds:

$$\int_{\Gamma_N} v^2_{|\Gamma_N} dS_x \le C_* \int_D (v^2 + |\nabla v|^2) dx \quad \forall\, v \in H^1_{\Gamma_D}(D) \tag{5.11}$$

for a suitable $C_* > 0$ (see Remark 6.7).
Problem (5.10) has been therefore rewritten in the weak form:

$$\text{find } u \in V \;:\; B(u,v) = F(v) \quad \forall\, v \in V \,, \tag{5.12}$$

where

$$B(w,v) = B_L(w,v) \;\;,\;\; F(v) = \int_D f v dx + \int_{\Gamma_N} g v_{|\Gamma_N} dS_x \;\;,\;\; V = H^1_{\Gamma_D}(D) \,. \tag{5.13}$$

Robin BC In this case the problem is

$$\begin{cases} Lu = f & \text{in } D \\ \displaystyle\sum_{i,j=1}^n n_i a_{ij} \mathcal{D}_j u + \kappa u = g & \text{on } \partial D \,, \end{cases} \tag{5.14}$$

where, besides (2.16) and (2.17), we also assume $g \in L^2(\partial D)$, $\kappa \in L^\infty(\partial D)$, $\kappa \geq 0$ a.e. on ∂D and $\int_{\partial D} \kappa dS_x \neq 0$.
We choose $C^\infty(\overline{D})$ as space of test functions. Multiplying the differential equation by $v \in C^\infty(\overline{D})$ and integrating over D we obtain

$$\int_D -\sum_{i,j=1}^n \mathcal{D}_i(a_{ij}\mathcal{D}_j u) v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx = \int_D f v dx \, .$$

Integrating by parts the first term, the following boundary integral appears:

$$-\int_{\partial D} \sum_{i,j=1}^n n_i a_{ij} \mathcal{D}_j u v_{|\partial D} dS_x \, .$$

Using the Robin condition it can be written as

$$-\int_{\partial D} (g - \kappa u_{|\partial D}) v_{|\partial D} dS_x \, .$$

Thus we have obtained

$$\int_D \sum_{i,j=1}^n a_{ij} \mathcal{D}_j u \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i u v dx + \int_D a_0 u v dx + \int_{\partial D} \kappa u_{|\partial D} v_{|\partial D} dS_x$$
$$= \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \, .$$

The results that have been used for giving a meaning to the Neumann problem are employed also here: thus we assume that ∂D is a Lipschitz continuous boundary, so that the trace $v_{|\partial D}$ of $v \in H^1(D)$ is defined in $L^2(\partial D)$ and depends continuously on v.
Problem (5.14) has been therefore rewritten in the weak form:

$$\text{find } u \in V \ : \ B(u, v) = F(v) \quad \forall\, v \in V \, , \tag{5.15}$$

where

$$\begin{aligned} B(w, v) &= B_L(w, v) + \textstyle\int_{\partial D} \kappa w_{|\partial D} v_{|\partial D} dS_x \\ F(v) &= \textstyle\int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \quad , \quad V = H^1(D) \, . \end{aligned} \tag{5.16}$$

5.2 Boundedness of the Bilinear Form $B(\cdot,\cdot)$ and the linear functional $F(\cdot)$

For the analysis of the boundary value problems we have derived in the previous section we want to apply the Lax–Milgram theorem 2.1. Thus, as a first step, we have to verify that $B(\cdot,\cdot)$ and $F(\cdot)$ are bounded in $H^1(D)$. Let us remind the assumptions on the coefficients and the right hand side: $a_{ij} \in L^\infty(D)$ for $i,j = 1,\dots,n$, $b_i \in L^\infty(D)$ for $i = 1,\dots,n$, $a_0 \in L^\infty(D)$, $f \in L^2(D)$ for all the problems, then $g \in L^2(\partial D)$ (for the Neumann and Robin problems) or $g \in L^2(\Gamma_N)$ (for the mixed problem), and finally $\kappa \in L^\infty(\partial D)$, $\kappa \geq 0$ a.e. on ∂D and $\int_{\partial D} \kappa dS_x \neq 0$ (for the Robin problem). Finally, we have assumed that D has a Lipschitz continuous boundary ∂D.

Let us denote by $A = \{a_{ij}\}_{i,j=1}^n$ the coefficient matrix of the principal part, by $\|A\| = \sqrt{\sum_{i,j=1}^n a_{ij}^2}$ its norm, and by $b = \{b_i\}_{i=1}^n$ the vector field describing the first order part of the operator L. We readily check, using the Cauchy–Schwarz inequality in $L^2(D)$,

$$
\begin{aligned}
|B_L(w,v)| &= \left| \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx \right| \\
&\leq \sup_D \|A\| \int_D |\nabla w||\nabla v| dx + \sup_D |b| \int_D |\nabla w||v| dx \\
&\qquad + \sup_D |a_0| \int_D |w||v| dx \\
&\leq \gamma \|w\|_{H^1(D)} \|v\|_{H^1(D)}
\end{aligned}
$$

for a suitable constant $\gamma > 0$ depending on the L^∞-norms of A, b and a_0. Moreover,

$$
\left| \int_{\partial D} \kappa w_{|\partial D} v_{|\partial D} dS_x \right| \leq \|\kappa\|_{L^\infty(\partial D)} \|w_{|\partial D}\|_{L^2(\partial D)} \|v_{|\partial D}\|_{L^2(\partial D)} ,
$$

by the Cauchy–Schwarz inequality in $L^2(\partial D)$. The trace inequality (5.7) permits to estimate $\|w_{|\partial D}\|_{L^2(\partial D)}$ and $\|v_{|\partial D}\|_{L^2(\partial D)}$ in terms of $\|w\|_{H^1(D)}$ and $\|v\|_{H^1(D)}$, respectively, and the boundedness of $B(\cdot,\cdot)$ is therefore proved.

Remark 5.2 Other conditions assuring boundedness of the bilinear form $B_L(\cdot,\cdot)$ can be found in Exercise 7.16, (i).

Let us come to the boundedness of the linear functional F. We have, again by the Cauchy–Schwarz inequality,

$$
\begin{aligned}
\left| \int_D f v dx \right| &\leq \|f\|_{L^2(D)} \|v\|_{L^2(D)} , \\
\left| \int_{\partial D} g v_{|\partial D} dS_x \right| &\leq \|g\|_{L^2(\partial D)} \|v_{|\partial D}\|_{L^2(\partial D)} \\
\left| \int_{\Gamma_N} g v_{|\Gamma_N} dS_x \right| &\leq \|g\|_{L^2(\Gamma_N)} \|v_{|\Gamma_N}\|_{L^2(\Gamma_N)} .
\end{aligned}
$$

The trace inequalities (5.7) and (5.11) give an estimate of $\|v_{|\partial D}\|_{L^2(\partial D)}$ and $\|v_{|\Gamma_N}\|_{L^2(\Gamma_N)}$ in terms of $\|v\|_{H^1(D)}$, and the boundedness of $F(\cdot)$ thus follows at once.

5.3 Weak Coerciveness of the Bilinear Form $B(\cdot, \cdot)$

First of all we need a new definition. Assume that $V \subset H^1(D)$ is a Hilbert space with respect to the $H^1(D)$-scalar product.

Definition 5.1 A bilinear form $B(\cdot, \cdot) : V \times V \mapsto \mathbb{R}$ is said to be *weakly* coercive in V if there exist two constants $\alpha > 0$ and $\sigma \geq 0$ such that

$$B(v, v) + \sigma \|v\|^2_{L^2(D)} \geq \alpha \|v\|^2_{H^1(D)} \quad \forall\, v \in V\,.$$

Remark 5.3 It is clearly seen that, if it possible to choose $\sigma = 0$ in this definition, then the bilinear form $B(\cdot, \cdot)$ is coercive in $H^1(D)$.

We consider the bilinear forms $B_L(\cdot, \cdot)$ and $B(\cdot, \cdot)$ defined as

$$B_L(w, v) = \int_D \sum_{i,j=1}^n a_{ij} \mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx$$

and

$$B(w, v) = \begin{cases} B_L(w, v) & \text{for the Dirichlet, Neumann, mixed problems} \\ B_L(w, v) + \displaystyle\int_{\partial D} \kappa w_{|\partial D} v_{|\partial D} dS_x & \text{for the Robin problem} \end{cases} ,$$

under the same assumptions of Sect. 5.2. Having assumed $\kappa \geq 0$ it follows $\int_{\partial D} \kappa v^2_{|\partial D} dS_x \geq 0$, thus we can limit our analysis to $B_L(v, v)$. We have

$$B_L(v, v) = \underbrace{\int_D \sum_{i,j=1}^n a_{ij} \mathcal{D}_j v \mathcal{D}_i v dx}_{[1]} + \underbrace{\int_D \sum_{i=1}^n b_i \mathcal{D}_i v v dx}_{[2]} + \underbrace{\int_D a_0 v^2 dx}_{[3]}\,.$$

[1] By ellipticity, for almost all $x \in D$ and for all $\eta \in \mathbb{R}^n$ we have that

$$\sum_{i,j=1}^n a_{ij}(x) \eta_j \eta_i \geq \alpha_0 |\eta|^2 \quad \text{for some } \alpha_0 > 0\,.$$

Thus, setting $\eta = \nabla v(x)$ and integrating in D it follows that

$$\int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j v \mathcal{D}_i v dx \geq \alpha_0 \int_D |\nabla v|^2 dx \,.$$

[2] Using the Cauchy–Schwarz inequality we find that

$$\begin{aligned} \left| \int_D \sum_{i=1}^n b_i \mathcal{D}_i v v dx \right| &\leq \int_D \sum_{i=1}^n |b_i|\,|\mathcal{D}_i v|\,|v| dx \leq \|b\|_{L^\infty(D)} \int_D |\nabla v||v| dx \\ &\leq \|b\|_{L^\infty(D)} \left(\int_D |\nabla v|^2 dx \right)^{1/2} \left(\int_D v^2 dx \right)^{1/2} \\ &= \left(\int_D |\nabla v|^2 dx \right)^{1/2} \left(\|b\|^2_{L^\infty(D)} \int_D v^2 dx \right)^{1/2} . \end{aligned}$$

Consider now the elementary inequality $|2AB| \leq A^2 + B^2$: from this, replacing A by $\sqrt{\varepsilon}A$ and B by $B/\sqrt{\varepsilon}$, where $\varepsilon > 0$, we can easily derive the following inequality

$$|AB| \leq \frac{\varepsilon}{2} A^2 + \frac{B^2}{2\varepsilon} \,.$$

Applying this we obtain

$$\left| \int_D \sum_{i=1}^n b_i \mathcal{D}_i v v dx \right| \leq \frac{\varepsilon}{2} \int_D |\nabla v|^2 dx + \frac{1}{2\varepsilon} \|b\|^2_{L^\infty(D)} \int_D v^2 dx$$

and so

$$\int_D \sum_{i=1}^n b_i \mathcal{D}_i v v dx \geq -\frac{\varepsilon}{2} \int_D |\nabla v|^2 dx - \frac{1}{2\varepsilon} \|b\|^2_{L^\infty(D)} \int_D v^2 dx \,.$$

[3] We have that

$$\int_D a_0 v^2 dx \geq \inf_D a_0 \int_D v^2 dx \,.$$

Putting everything together and choosing $\varepsilon = \alpha_0$ we have

$$B_L(v, v) \geq \frac{\alpha_0}{2} \int_D |\nabla v|^2 dx + \left(\inf_D a_0 - \frac{1}{2\alpha_0} \|b\|^2_{L^\infty(D)} \right) \int_D v^2 dx \,.$$

Therefore the following inequality holds

$$B_L(v,v)+\sigma\int_D v^2dx \geq \frac{\alpha_0}{2}\int_D |\nabla v|^2dx+\left(\sigma+\inf_D a_0-\frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D)}\right)\int_D v^2dx\,.$$

Set $\mu=\inf_D a_0-\frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D)}$. Choosing σ as follows:

$$\begin{cases}\sigma=0 & \text{if } \mu>0\\ \sigma>-\mu\geq 0 & \text{if } \mu\leq 0\,,\end{cases}$$

and denoting by $\rho=\sigma+\mu>0$ we find the desired result:

$$\begin{aligned}B_L(v,v)+\sigma\int_D v^2dx &\geq \frac{\alpha_0}{2}\int_D|\nabla v|^2dx+\rho\int_D v^2dx\\ &\geq \min\left(\frac{\alpha_0}{2},\rho\right)\int_D\left(|\nabla v|^2+v^2\right)dx\,.\end{aligned} \tag{5.17}$$

Remark 5.4 Weak coerciveness with $\sigma>0$ is not enough to apply the Lax–Milgram theorem 2.1. Therefore, in this respect the result just proved is satisfactory only when we can choose $\sigma=0$, namely, when $\mu=\inf_D a_0-\frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D)}>0$. This requires $\inf_D a_0>0$ and $\|b\|^2_{L^\infty(D)}$ small enough. The following example shows that for the "queen" of our operator, the Laplace operator $-\Delta$, this is not satisfied.

Example 5.1 Consider the (homogeneous) Dirichlet boundary value problem

$$\begin{cases}-\Delta u=f & \text{in } D\\ u=0 & \text{on } \partial D\,.\end{cases}$$

In this case we have $b=0$ and $a_0=0$, thus the condition $\inf_D a_0-\frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D)}>0$ is not satisfied. Since

$$B(v,v)=\int_D\nabla v\cdot\nabla v dx=\int_D|\nabla v|^2dx\,,$$

to prove coerciveness we have to find a constant α satisfying $0<\alpha<1$ such that

$$B(v,v)=\int_D|\nabla v|^2dx\geq\alpha\int_D\left(|\nabla v|^2+v^2\right)dx\quad\forall\, v\in H^1_0(D)$$

or, equivalently, we have to prove that

$$\text{there exists a constant } C_D > 0 \ : \quad \int_D v^2 dx \le C_D \int_D |\nabla v|^2 dx \quad \forall\, v \in H_0^1(D))\,. \tag{5.18}$$

Assuming that such a constant exists, we observe that

$$\begin{aligned} B(v, v) &= \int_D |\nabla v|^2 dx = \frac{1}{2} \int_D |\nabla v|^2 dx + \frac{1}{2} \int_D |\nabla v|^2 dx \\ &\ge \frac{1}{2} \int_D |\nabla v|^2 dx + \frac{1}{2C_D} \int_D v^2 dx \\ &\ge \min\left(\frac{1}{2}, \frac{1}{2C_D}\right) \int_D (v^2 + |\nabla v|^2) dx\,. \end{aligned}$$

Inequality (5.18) is called Poincaré inequality in $H_0^1(D)$: we will present its proof in Sect. 6.2. For the moment, let us note that this inequality is surely false if we can select as function v a non-zero constant. The fact that the only constant in $H_0^1(D)$ is 0 opens the possibility of showing that (5.18) is indeed true.

5.4 Coerciveness of the Bilinear Form $B(\cdot, \cdot)$

Assuming more regularity on the vector field b and some other qualitative relations, we want now to show that the bilinear form $B(\cdot, \cdot)$ is coercive for all the boundary value problems we have presented.

The starting point for this analysis is the remark that in some cases we succeed in proving the Poincaré inequality

$$\int_D v^2 dx \le C_* \int_D |\nabla v|^2 dx\,;$$

this tells us that the principal part of the bilinear form can be bounded from below by $\|v\|^2_{H^1(D)}$, namely, it is coercive. Thus we have only to be careful that the other terms, coming from b and a_0, do not destroy this property.

Let us consider the term coming from the vector field b. Assume that $b \in W^{1,\infty}(D)$ so that by the Sobolev immersion theorem 7.15 we also have $b_{|\partial D} \in L^\infty(\partial D)$ for the Neumann and Robin problems or $b_{|\Gamma_N} \in L^\infty(\Gamma_N)$ for the mixed problem (it is possible to require less restrictive assumptions, but the proof would become more technical). We proceed by analyzing each boundary condition.

Dirichlet BC. The choice of the Hilbert space is $V = H_0^1(D)$, and in this case Poincaré inequality holds (see Theorem 6.4). Since $C_0^\infty(D)$ is dense in $H_0^1(D)$

we can first suppose that $v \in C_0^\infty(D)$. We have, by integrating by parts (see Exercise 4.5)

$$\int_D \sum_{i=1}^n b_i \mathcal{D}_i v\, v dx = \sum_{i=1}^n \int_D b_i \mathcal{D}_i \frac{v^2}{2} dx$$
$$= -\sum_{i=1} \int_D \mathcal{D}_i b_i \frac{v^2}{2} dx \; = -\int_D \frac{1}{2} \operatorname{div} b\, v^2 dx\,.$$

By a density argument we see that this relation is also true for $v \in H_0^1(D)$. Hence we have

$$B(v,v) = \int_D \sum_{i,j=1}^n a_{ij} \mathcal{D}_j v\, D_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i v\, v dx + \int_D a_0 v^2 dx$$
$$\geq \alpha_0 \int_D |\nabla v|^2 dx + \int_D \Big(a_0 - \frac{1}{2} \operatorname{div} b\Big) v^2 dx$$

and coerciveness in $H_0^1(D)$ is guaranteed by the Poincaré inequality and assuming

$$\boxed{a_0 - \frac{1}{2} \operatorname{div} b \geq 0 \quad \text{in } D\,.}$$

Neumann BC The Hilbert space in this case is $V = H^1(D)$. Since in this space Poincaré inequality doesn't hold (e.g., consider $v = 1$), we could be led to modify this choice. Let us start, as before, by looking at the term coming from the first order part of the operator. We want to perform an integration by parts, which will show up an integral on ∂D involving the *trace* $v_{|\partial D}$ of v on ∂D. To give a meaning at this term we assume that ∂D is a Lipschitz continuous boundary, thus the space $C^\infty(\overline{D})$ is dense in $H^1(D)$ and the trace is defined (see Theorem 6.5). We can first assume that $v \in C^\infty(\overline{D})$. By integration by parts (see Exercise 6.7) we have

$$\int_D \sum_{i=1}^n b_i \mathcal{D}_i v\, v dx = \sum_{i=1}^n \int_D b_i \mathcal{D}_i \frac{v^2}{2} dx$$
$$= -\sum_{i=1} \int_D \mathcal{D}_i b_i \frac{v^2}{2} dx + \sum_{i=1}^n \int_{\partial D} b_{i|\partial D} n_i \frac{1}{2} v_{|\partial D}^2 dS_x$$
$$= -\int_D \frac{1}{2} \operatorname{div} b\, v^2 dx + \int_{\partial D} \frac{1}{2}\, b_{|\partial D} \cdot n v_{|\partial D}^2 dS_x\,.$$

By a density argument this relation is true also for $v \in H^1(D)$. In conclusion, we easily see that sufficient conditions for coerciveness are

$$a_0 - \frac{1}{2}\operatorname{div} b \geq \delta > 0 \text{ in } D \ , \quad b_{|\partial D} \cdot n \geq 0 \text{ on } \partial D \, .$$

However, these conditions are not satisfactory, as, for instance, the Laplace operator $-\Delta$ does not satisfy them. On the other hand this is not a surprise, as for the Neumann problem associated to the Laplace operator we cannot have a unique solution, as, if u is a solution, also $u + c$ with $c \in \mathbb{R}$ is a solution, and therefore the assumptions in the Lax–Milgram theorem 2.1 cannot be satisfied. (Remember that Lax–Milgram theorem guarantees the existence and uniqueness of the solution.)

In order to devise a weak problem for which the associated bilinear form is coercive, the idea is to define a new Hilbert space that doesn't contain constants different from 0. A space with this property is given by

$$H^1_*(D) = \left\{ v \in H^1(D) \, \Big| \, \int_D v dx = 0 \right\} \, . \tag{5.19}$$

This is a closed subspace of $H^1(D)$ (indeed if $v_k \to v$ in $H^1(D)$ and $\int_D v_k dx = 0$, then $\int_D v dx = 0$: the quantity $|\int_D (v_k - v) dx|$ is estimated by $\|v_k - v\|_{L^2(D)}$ by the Cauchy–Schwarz inequality), therefore it is a Hilbert space with respect to the same scalar product. In this space the Poincaré inequality holds (see Theorem 6.10) and therefore we can prove the coerciveness of $B(\cdot, \cdot)$ in $H^1_*(D)$ by following the same procedure we have employed in the case of the Dirichlet boundary condition. More precisely, sufficient conditions that guarantee the coerciveness of $B(\cdot, \cdot)$ are

$$\boxed{a_0 - \frac{1}{2}\operatorname{div} b \geq 0 \text{ in } D \ , \quad b_{|\partial D} \cdot n \geq 0 \text{ on } \partial D \, .}$$

Mixed BC The Hilbert space in this case is $V = H^1_{\Gamma_D}(D)$, and we will see that in this space the Poincaré inequality holds (see Theorem 6.11), provided ∂D is a Lipschitz continuous boundary. Therefore we can proceed exactly as in the case of the Neumann condition with the space $H^1_*(D)$ and we conclude that sufficient conditions that guarantee the coerciveness of $B(\cdot, \cdot)$ are

$$\boxed{a_0 - \frac{1}{2}\operatorname{div} b \geq 0 \text{ in } D \ , \quad b_{|\Gamma_N} \cdot n \geq 0 \text{ on } \Gamma_N \, .}$$

Robin BC The Hilbert space in this case is $V = H^1(D)$, and the bilinear form is given by

$$B(w,v) = \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx + \int_{\partial D} \kappa w_{|\partial D} v_{|\partial D} dS_x ,$$

where κ is a non-negative function defined on ∂D. By assuming that ∂D is Lipschitz continuous and performing an integration by parts in the first order term as in the Neumann case we have that

$$\begin{aligned} B(v,v) &\geq \alpha_0 \int_D |\nabla v|^2 dx + \int_D \left(a_0 - \frac{1}{2}\operatorname{div} b\right) v^2 dx \\ &\quad + \int_{\partial D} \left(\frac{1}{2} b_{|\partial D}\cdot n + \kappa\right) v^2_{|\partial D} dS_x \\ &= \alpha_0 \left(\int_D |\nabla v|^2 dx + \int_{\partial D} \alpha_0^{-1}\kappa v^2_{|\partial D} dS_x\right) + \int_D \left(a_0 - \frac{1}{2}\operatorname{div} b\right) v^2 dx \\ &\quad + \int_{\partial D} \frac{1}{2} b_{|\partial D}\cdot n v^2_{|\partial D} dS_x . \end{aligned}$$

We assume that

$$\boxed{a_0 - \frac{1}{2}\operatorname{div} b \geq 0 \text{ in } D \ , \ \ b_{|\partial D}\cdot n \geq 0 \text{ on } \partial D ,}$$

and we note that the function $q = \alpha_0^{-1}\kappa$ satisfies $q \geq 0$ on ∂D and $\int_{\partial D} q dS_x \neq 0$, thus we can apply the Poincaré-type inequality (see Theorem 6.12). In conclusion we are left with

$$\begin{aligned} B(v,v) &\geq \alpha_0 \left(\int_D |\nabla v|^2 dx + \int_{\partial D} \alpha_0^{-1}\kappa v^2_{|\partial D} dS_x\right) \\ &= \frac{\alpha_0}{2}\left(\int_D |\nabla v|^2 dx + \int_{\partial D} \alpha_0^{-1}\kappa v^2_{|\partial D} dS_x\right) \\ &\quad + \frac{\alpha_0}{2}\left(\int_D |\nabla v|^2 dx + \int_{\partial D} \alpha_0^{-1}\kappa v^2_{|\partial D} dS_x\right) \\ &\geq \frac{\alpha_0}{2}\left(\int_D |\nabla v|^2 dx + \int_{\partial D} \alpha_0^{-1}\kappa v^2_{|\partial D} dS_x\right) + \frac{\alpha_0}{2C_*}\int_D v^2 dx \\ &\geq \frac{\alpha_0}{2}\int_D |\nabla v|^2 dx + \frac{\alpha_0}{2C_*}\int_D v^2 dx \\ &\geq \min\left(\frac{\alpha_0}{2}, \frac{\alpha_0}{2C_*}\right)\left(\int_D v^2 dx + \int_D |\nabla v|^2 dx\right) . \end{aligned}$$

Exercise 5.1 Show that in all cases coerciveness is satisfied even if the assumption $a_0 - \frac{1}{2}\operatorname{div} b \geq 0$ in D is weakened to $a_0 - \frac{1}{2}\operatorname{div} b \geq -\nu$ in D for a constant $\nu > 0$ small enough.

Remark 5.5 Other conditions assuring coerciveness of the bilinear form $B_L(\cdot, \cdot)$ can be found in Exercise 7.16, (ii).

5.5 Interpretation of the Weak Problems

We want to clarify which is the "strong" interpretation of the weak problems we have presented up to now. To this aim, we first need a definition.

Definition 5.2 If we have $q_i \in L^1_{\mathrm{loc}}(D)$, $i = 1, \ldots, n$, we say that $w \in L^1_{\mathrm{loc}}(D)$ is the weak divergence of $q = (q_1, \ldots, q_n)$ if

$$\int_D \sum_{i=1}^n q_i \mathcal{D}_i \varphi dx = -\int_D w \varphi dx \qquad \forall\, \varphi \in C^\infty_0(D)\,.$$

Remark 5.6 If we know that the weak derivatives $\mathcal{D}_i q_i$ exist, for each $i = 1, \ldots, n$, then clearly $w = \sum_{i=1}^n \mathcal{D}_i q_i$.

Let us start our discussion from a simple example.

Example 5.2 Suppose we have found the solution $u \in H^1_0(D)$ of

$$\int_D \nabla u \cdot \nabla v dx = \int_D f v dx \qquad \forall\, v \in H^1_0(D)\,,$$

where $f \in L^2(D)$. What have we solved?

We can take $\varphi \in C^\infty_0(D) \subset H^1_0(D)$ and we get

$$\int_D \sum_{i=1}^n \mathcal{D}_i u \mathcal{D}_i \varphi dx = \int_D f \varphi dx\,,$$

thus from the definition above, with $q_i = \mathcal{D}_i u \in L^2(D)$, we obtain that

$$-\operatorname{div}\nabla u = f \quad \text{ in } D\,,$$

where div is the weak divergence and ∇ is the weak gradient. Thus, in this weak sense, $-\Delta u = f$ in D, where Δ is the weak Laplace operator.

This interpretation is based on the fact that $C_0^\infty(D) \subset V = H_0^1(D)$, the variational space where we have solved the problem. When considering the mixed problem, we have $V = H^1_{\Gamma_D}(D)$, and again $C_0^\infty(D) \subset V$. For the Robin problem, we have $V = H^1(D)$, and $C_0^\infty(D) \subset V$.

A difference comes for the Neumann problem for the Laplace operator, for the weak formulation in which we have chosen

$$V = H^1_*(D) = \left\{ v \in H^1(D) \,\Big|\, \int_D v dx = 0 \right\} ,$$

with the aim of obtaining the Poincaré inequality in this space.

This time $C_0^\infty(D) \not\subset V$, thus the interpretation in this case needs some care. Let us write the weak problem:

$$u \in H^1_*(D) \ : \ \int_D \nabla u \cdot \nabla v dx = \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \qquad \forall\, v \in H^1_*(D) \, .$$

Take a test function $w \in H^1(D)$, namely, without the restriction $\int_D w dx = 0$. Then we define

$$v = w - w_D \, , \qquad w_D = \frac{1}{\text{meas}(D)} \int_D w dx \, .$$

Then $v \in H^1_*(D)$, and we can use it as a test function. We have $\nabla w = \nabla v$, thus for each $w \in H^1(D)$ we have

$$\begin{aligned}
\int_D \nabla u \cdot \nabla w dx &= \int_D \nabla u \cdot \nabla v dx = \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \\
&= \int_D f (w - w_D) dx + \int_{\partial D} g (w_{|\partial D} - w_D) dS_x \\
&= \int_D f w dx - w_D \int_D f dx + \int_{\partial D} g w_{|\partial D} dS_x - w_D \int_{\partial D} g dS_x \\
&= \int_D f w dx + \int_{\partial D} g w_{|\partial D} dS_x \\
&\quad - \left(\int_D f dx + \int_{\partial D} g dS_x \right) \frac{1}{\text{meas}(D)} \int_D w dx \\
&= \int_D \left[f - \frac{1}{\text{meas}(D)} \left(\int_D f dx + \int_{\partial D} g dS_x \right) \right] w dx \\
&\quad + \int_{\partial D} g w_{|\partial D} dS_x \, .
\end{aligned} \tag{5.20}$$

Taking in particular $w \in C_0^\infty(D)$, it follows

$$-\Delta u = f - \frac{1}{\text{meas}(D)}\left(\int_D f dx + \int_{\partial D} g dS_x\right) \qquad \text{in } D\,.$$

If we have $p \in H^1(D)$, $q \in H^1(D)$ with $-\Delta q \in L^2(D)$, by approximation we have the integration by parts formula

$$\int_D \nabla q \cdot \nabla p dx = -\int_D \Delta q \; p dx + \int_{\partial D} (\nabla q \cdot n) p_{|\partial D} dS_x\,.$$

The last term should be clarified, indeed it is not obvious that there is a trace for $\nabla q \cdot n$. However, we do not deal here with this question, and we go on somehow formally. Let us come back now to the choice of a generic $w \in H^1(D)$: taking $p = w$ and $q = u$ in (5.20) we thus find

$$\begin{aligned} &\textstyle\int_{\partial D} \nabla u \cdot n \, w_{|\partial D} dS_x + \int_D \cancel{(-\Delta u)} w dx = \int_D \nabla u \cdot \nabla w dx \\ &= \textstyle\int_D \cancel{\left[f - \frac{1}{\text{meas}(D)} \left(\int_D f dx + \int_{\partial D} g dS_x\right)\right]} w dx + \int_{\partial D} g w_{|\partial D} dS_x\,. \end{aligned}$$

As a consequence

$$\int_{\partial D} (\nabla u \cdot n - g) w_{|\partial D} dS_x = 0 \qquad \forall\, w \in H^1(D)\,,$$

which is a weak form of $\nabla u \cdot n = g$ on ∂D. In conclusion, the "strong" form of the weak problem we have solved reads

$$\begin{cases} -\Delta u = f - \frac{1}{\text{meas}(D)}\left(\int_D f dx + \int_{\partial D} g dS_x\right) & \text{in } D \\ \nabla u \cdot n = g & \text{on } \partial D\,. \end{cases} \tag{5.21}$$

This problem has been solved for any $f \in L^2(D)$ and $g \in L^2(\partial D)$; but it is not the Neumann problem we had in mind, namely

$$\begin{cases} -\Delta u = f & \text{in } D \\ \nabla u \cdot n = g & \text{on } \partial D\,. \end{cases} \tag{5.22}$$

On the other hand, we know by the divergence theorem that this last problem cannot be solved unless the following compatibility condition is satisfied:

$$\int_D f dx + \int_{\partial D} g dS_x = 0\,.$$

In fact

$$\int_D f dx = -\int_D \Delta u dx = -\int_D \mathrm{div}\nabla u dx = -\int_{\partial D} \nabla u \cdot n dS_x = -\int_{\partial D} g dS_x \, .$$

In conclusion, if $\int_D f dx + \int_{\partial D} g dS_x = 0$ problem (5.21) becomes our original problem, and we have found a unique solution in $H^1_*(D)$, namely, with $\int_D u dx = 0$.

Remark 5.7 Why is problem (5.21) always solvable? It is a Neumann problem, therefore the compatibility condition on the data at the right hand side must be satisfied. The new right hand side in D is

$$\tilde{f} = f - \frac{1}{\mathrm{meas}(D)} \left(\int_D f dx + \int_{\partial D} g dS_x \right) .$$

Take its integral in D: it holds

$$\begin{aligned} &\int_D \left[f - \frac{1}{\mathrm{meas}(D)} \left(\int_D f dx + \int_{\partial D} g dS_x \right) \right] dx \\ &\quad = \int_D f dx - \left[\int_D f dx + \int_{\partial D} g dS_x \right] = -\int_{\partial D} g dS_x \, . \end{aligned}$$

Thus

$$\int_D \tilde{f} dx + \int_{\partial D} g dS_x = 0 \, ,$$

and the compatibility condition for the Neumann problem (5.21) is satisfied.

Exercise 5.2 Taking hint from the definition of the weak divergence in Definition 5.2, give the definition of the weak curl of a vector field $q \in (L^1_{\mathrm{loc}}(D))^3$, $D \subset \mathbb{R}^3$.

Exercise 5.3

(i) Show that there exists a unique solution of the weak problem

$$\begin{aligned} \text{find } u \in H^1_*(D) \ : \int_D \nabla u \cdot \nabla v dx + \int_{\partial D} u_{|\partial D} v_{|\partial D} dS_x \\ = \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \quad \forall \, v \in H^1_*(D) \, , \end{aligned}$$

where $H^1_*(D)$ is defined in (5.19).

(ii) Devise the "strong" interpretation of the weak problem above.

5.6 A Higher Order Example: The Biharmonic Operator

The biharmonic operator is $\Delta^2 = \Delta\Delta$, and the associated biharmonic equation is

$$\Delta^2 u = f \quad \text{in } D\,.$$

In this section we want to devise and then analyze the variational formulation of some "reasonable" boundary value problems associated to this equation. Assuming $f \in L^2(D)$, multiply this equation by a test function v and integrate in D:

$$\int_D (\Delta^2 u)\, v\, dx = \int_D f\, v\, dx\,.$$

Taking into account that $\Delta = \sum_{i=1}^n \mathcal{D}_i \mathcal{D}_i$, integrating by parts at the left hand side gives

$$\begin{aligned}\int_D (\Delta^2 u)\, v\, dx &= \int_D (\textstyle\sum_{i=1}^n \mathcal{D}_i \mathcal{D}_i \Delta u)\, v\, dx \\ &= -\int_D \textstyle\sum_{i=1}^n (\mathcal{D}_i \Delta u)\, \mathcal{D}_i v\, dx + \int_{\partial D} \textstyle\sum_{i=1}^n (n_i \mathcal{D}_i \Delta u)\, v\, dS_x\,.\end{aligned}$$

We still have a third order operator acting on u and a first order operator acting on v. Therefore we proceed with another integration by parts and we find

$$\begin{aligned}\int_D f\, v\, dx &= \int_D (\Delta^2 u)\, v\, dx \\ &= \int_D \Delta u\ \textstyle\sum_{i=1}^n \mathcal{D}_i \mathrm{d}_i v\, dx \\ &\qquad - \int_{\partial D} \Delta u \textstyle\sum_{i=1}^n n_i\, \mathcal{D}_i v\, dS_x + \int_{\partial D} \textstyle\sum_{i=1}^n (n_i \mathcal{D}_i \Delta u)\, v\, dS_x \\ &= \int_D \Delta u\, \Delta v\, dx - \int_{\partial D} \Delta u\, \nabla v \cdot n\, dS_x + \int_{\partial D} \nabla \Delta u \cdot n\, v\, dS_x\,. \end{aligned} \tag{5.23}$$

Looking at the boundary integrals that appear above it can be asserted that the "reasonable" boundary conditions associated to the biharmonic operator stem from the choice of a couple of the following ones:

$$u_{|\partial D} = 0 \quad \text{on } \partial D \tag{5.24}$$

$$(\nabla u \cdot n)_{|\partial D} = 0 \quad \text{on } \partial D \tag{5.25}$$

$$(\Delta u)_{|\partial D} = g \quad \text{on } \partial D \tag{5.26}$$

$$(\nabla \Delta u \cdot n)_{|\partial D} = h \quad \text{on } \partial D\,, \tag{5.27}$$

where it is assumed that $g \in L^2(\partial D)$ and $h \in L^2(\partial D)$.

When considering the Dirichlet boundary conditions (5.24) and (5.25) we assume that also the test function v satisfies the same conditions, namely, $v_{|\partial D} = 0$ and $(\nabla v \cdot n)_{|\partial D} = 0$. Therefore the integrals on the boundary disappear and we are left with the bilinear form

$$B(w, v) = \int_D \Delta w \, \Delta v \, dx \, .$$

The variational formulation associated to (5.24) and (5.25) is thus identified with the choices:

$$\begin{aligned} &V = \{v \in H^2(D) \mid v_{|\partial D} = 0, (\nabla v \cdot n)_{|\partial D} = 0\} \\ &B(w, v) = \textstyle\int_D \Delta w \, \Delta v \, dx \\ &F(v) = \textstyle\int_D f \, v \, dx \, . \end{aligned}$$

Since $v \in H^2(D)$ we have in particular $v \in H^1(D)$ and $\nabla v \in (H^1(D))^n$; therefore their values on ∂D have a meaning (by Theorem 6.5 we have that $v_{|\partial D}$ and $(\nabla v)_{|\partial D}$ belong to $(L^2(\partial D))^n$).

When considering the so-called Navier boundary conditions (5.24) and (5.26) we assume that also the test function v satisfies $v_{|\partial D} = 0$ and therefore the two integrals on the boundary become

$$-\int_{\partial D} \Delta u \, \nabla v \cdot n \, dS_x + \int_{\partial D} \nabla \Delta u \cdot n \, v \, dS_x = -\int_{\partial D} g \, \nabla v \cdot n \, dS_x \, .$$

The variational formulation associated to (5.24) and (5.26) is thus identified with the choices:

$$\begin{aligned} &V = H^2(D) \cap H^1_0(D) = \{v \in H^2(D) \mid v_{|\partial D} = 0\} \\ &B(w, v) = \textstyle\int_D \Delta w \, \Delta v \, dx \\ &F(v) = \textstyle\int_D f \, v \, dx + \int_{\partial D} g \, \nabla v \cdot n \, dS_x \, . \end{aligned}$$

When considering the boundary conditions (5.24) and (5.27) we assume that also the test function v satisfies $v_{|\partial D} = 0$ and therefore one of the two integrals on the boundary vanishes:

$$-\int_{\partial D} \Delta u \, \nabla v \cdot n \, dS_x + \int_{\partial D} \nabla \Delta u \cdot n \, v \, dS_x = -\int_{\partial D} \Delta u \, \nabla v \cdot n \, dS_x \, .$$

But the other boundary integral is not treatable, as in this boundary value problem we are not assigning $(\Delta u)_{|\partial D}$ nor $(\nabla u \cdot n)_{|\partial D} = 0$ (which would have allowed us to impose $(\nabla v \cdot n)_{|\partial D} = 0$). Moreover, Δu is only a $L^2(D)$-function, thus it has not a well-defined value on ∂D. Hence the boundary conditions (5.24) and (5.27) do not seem to lead to a good variational problem.

When considering the boundary conditions (5.25) and (5.26) we assume that also the test function v satisfies $(\nabla v \cdot n)_{|\partial D} = 0$ and therefore one of the two integrals on the boundary vanishes:

$$-\int_{\partial D} \Delta u \, \nabla v \cdot n \, dS_x + \int_{\partial D} \nabla \Delta u \cdot n \, v \, dS_x = \int_{\partial D} \nabla \Delta u \cdot n \, v \, dS_x \, .$$

But, similarly to the previous case, the other boundary integral is not treatable, as in this boundary value problem we are not assigning $(\nabla \Delta u \cdot n)_{|\partial D}$ nor $u_{|\partial D} = 0$ (which would have allowed us to impose $v_{|\partial D} = 0$). Moreover, $\nabla \Delta u \cdot n$ has not even a well-defined value on ∂D. Hence the boundary conditions (5.25) and (5.26) do not seem to lead to a good variational problem.

When considering the boundary conditions (5.25) and (5.27) we assume that also the test function v satisfies $(\nabla v \cdot n)_{|\partial D} = 0$ and therefore the two integrals on the boundary become

$$-\int_{\partial D} \Delta u \, \nabla v \cdot n \, dS_x + \int_{\partial D} \nabla \Delta u \cdot n \, v \, dS_x = \int_{\partial D} h \, v \, dS_x \, .$$

The variational formulation associated to (5.25) and (5.27) (sometimes called the Riquier–Neumann boundary conditions) is thus identified with the choices:

$$\begin{aligned}
&V = \{v \in H^2(D) \,|\, (\nabla v \cdot n)_{|\partial D} = 0\} \\
&B(w, v) = \textstyle\int_D \Delta w \, \Delta v \, dx \\
&F(v) = \textstyle\int_D f \, v \, dx - \int_{\partial D} h \, v \, dS_x \, .
\end{aligned}$$

However, it is clear that using this variational formulation the problem cannot be well-posed (adding a constant to a solution one still finds a solution). Therefore the space V should be replaced by a closed subspace of it which does not contain non-zero constants, say,

$$V_\star = \left\{ v \in H^2(D) \,\Big|\, (\nabla v \cdot n)_{|\partial D} = 0, \int_D v dx = 0 \right\} .$$

Finally, when considering the Neumann boundary conditions (5.26) and (5.27) the two integrals on the boundary become

$$-\int_{\partial D} \Delta u \, \nabla v \cdot n \, dS_x + \int_{\partial D} \nabla \Delta u \cdot n \, v \, dS_x = -\int_{\partial D} g \, \nabla v \cdot n \, dS_x + \int_{\partial D} h \, v \, dS_x \, .$$

Since both boundary conditions are imposed in a weak way (in the sense that they give a contribution to the linear functional at the right hand side, while do not appear in the definition of the variational space), the solution and the test functions are possibly expected to be less regular than in the preceding cases, namely, in general they are not foreseen to belong to $H^2(D)$. Therefore the variational formulation associated to (5.26) and (5.27) is at first identified with the choices:

$$\begin{aligned}&V = L^2(\Delta; D) = \{v \in L^2(D) \mid \Delta v \in L^2(D)\}\\&B(w, v) = \int_D \Delta w\, \Delta v\, dx\\&F(v) = \int_D f\, v\, dx + \int_{\partial D} g\, \nabla v \cdot n\, dS_x - \int_{\partial D} h\, v\, dS_x\,.\end{aligned}$$

Here the Laplace operator is intended in the following weak sense: for $v \in L^1_{\mathrm{loc}}(D)$ we say that a function $q \in L^1_{\mathrm{loc}}(D)$ is the weak Laplacian of v if

$$\int_D q\, \varphi\, dx = \int_D v\, \Delta\varphi\, dx \quad \forall\, \varphi \in C_0^\infty(D)\,.$$

By repeating the proof of Theorem 4.1 it is easily seen that $L^2(\Delta; D)$ is a Hilbert space with respect to the natural scalar product $\int_D (w\, v + \Delta w\, \Delta v)\, dx$. However, also in this case it is clear that by this choice of the variational space V the problem cannot be well-posed (adding to a solution a harmonic function belonging to $L^2(D)$ gives another solution). Therefore the space V should be replaced by a closed subspace of it which does not contain non-zero harmonic functions, say,

$$V_\sharp = \left\{ v \in L^2(\Delta; D) \;\Big|\; \int_D v\eta dx = 0 \text{ for each } \eta \in L^2(D) \text{ with } \Delta\eta = 0 \text{ in } D \right\}$$

(another example will be proposed below). Note also here that the definition of the linear operator $F(\cdot)$ is not completely clear: having only assumed $g \in L^2(\partial D)$ and $h \in L^2(\partial D)$, the boundary integrals would require $v_{|\partial D} \in L^2(\partial D)$ and $(\nabla v \cdot n)_{|\partial D} \in L^2(\partial D)$, and for a function v belonging to $L^2(\Delta; D)$ this is not always the case. Higher regularity of g and h could be the cure, but anyway one should also develop a suitable trace theory for $v_{|\partial D}$ and $(\nabla v \cdot n)_{|\partial D}$; this can be done, but we will not insist on it here. Clearly, a drastic answer to this issue is to take $g = 0$ and $h = 0$; but we will see later that a regularity assumption on the boundary ∂D (say, $\partial D \in C^4$) will be enough for giving a meaning to these boundary integrals (see Lemma 5.1).

Let us come now to the analysis of the first three boundary value problems we have described; we use the regularity results presented in Sect. 7.4. Assume that $D \subset R^n$ is a bounded, connected and open set, and that its boundary ∂D is of class C^2. In all the considered cases the bilinear form $B(\cdot, \cdot)$ is clearly bounded in $H^2(D)$, and the same holds for the linear functional $F(\cdot)$, provided we assume $f \in L^2(D)$, $g \in L^2(\partial D)$, $h \in L^2(\partial D)$ (the estimates for the boundary integrals come from the trace theorem 6.5). Moreover, for the Dirichlet boundary conditions (5.24) and (5.25) and for the Navier boundary conditions (5.24) and (5.26) from Theorem 7.12 and Exercise 7.13 we have that

$$\|u\|_{H^2(D)} \le C \|\Delta u\|_{L^2(D)}\,.$$

The same is true for the boundary conditions (5.25) and (5.27), just replacing in the argument above $L^2(D)$ and $H^1(D)$ with $L^2_*(D)$ and $H^1_*(D)$, namely, the corresponding spaces with the additional constraint $\int_D v\,dx = 0$. The conclusion is that in all these three cases the bilinear form $B(\cdot,\cdot)$ is bounded and coercive in $H^2(D)$, and the Lax–Milgram theorem 2.1 can be applied.

Since for the boundary conditions (5.25) and (5.27) the variational space $V_\star$ does not contain $C_0^\infty(D)$, it is appropriate to verify which problem we have indeed solved. Following what we have done in Sect. 5.5 we can easily see that we have found a solution of

$$\Delta^2 u = f - \frac{1}{\mathrm{meas}(D)}\Big(\int_D f\,dx - \int_{\partial D} h\,dS_x\Big)\,.$$

Since the condition $\int_D f\,dx - \int_{\partial D} h\,dS_x = 0$ is necessary to solve the problem (take $v = 1$ in (5.23)), the solution of the variational problem is the one we are looking for.

Note also that the solution of the boundary value problems associated to the Navier conditions (5.24) and (5.26) and to the Riquier–Neumann conditions (5.25) and (5.27) can be also obtained by a two step procedure only involving the Laplace operator $-\Delta$. In fact, proceeding formally without taking into account the regularity of the boundary data, for the Navier conditions we can solve

$$\begin{cases} -\Delta w = f & \text{in } D \\ w = -g & \text{on } \partial D \end{cases} \quad \text{and then} \quad \begin{cases} -\Delta u = w & \text{in } D \\ u = 0 & \text{on } \partial D\,, \end{cases}$$

while for the Riquier–Neumann conditions we can solve

$$\begin{cases} -\Delta w = f & \text{in } D \\ \nabla w \cdot n = -h & \text{on } \partial D \\ \int_D w\,dx = 0 \end{cases} \quad \text{and then} \quad \begin{cases} -\Delta u = w & \text{in } D \\ \nabla u \cdot n = 0 & \text{on } \partial D \\ \int_D u\,dx = 0\,. \end{cases}$$

5.6.1 The Analysis of the Neumann Boundary Value Problem

The analysis for the Neumann boundary conditions (5.26) and (5.27) is more complicated. In order to use the Lax–Milgram theorem 2.1 and thus obtaining well-posedness we need a Poincaré-type inequality like

$$\|v\|_{L^2(D)} \le C\|\Delta v\|_{L^2(D)}$$

for all $v \in V_\sharp$. This inequality would follow if the immersion $L^2(\Delta; D) \hookrightarrow L^2(D)$ was compact, as we could repeat line by line the proof of Theorem 6.10. But from Exercise 6.11 we know that this immersion is not compact! Then we could try to show that the immersion of $V_\sharp \hookrightarrow L^2(D)$ is compact. However, this result is more elusive, and for the moment we leave it apart (but see Remark 5.9).

A different attempt can be done by changing the variational space; it will be educational to follow this path as we will see again how a variational space that does not contain $C_0^\infty(D)$ can be a delicate choice. We introduce

$$X = \{\omega = \Delta r \mid r \in H^4(D) \cap H_0^2(D)\}$$

(see Definition 4.4). We verify at once that $X \subset H^2(D)$; moreover, assuming that the boundary ∂D is smooth, say, of class C^4, we see that X is closed in $H^2(D)$ with respect to the $H^2(D)$-norm (thus X is a Hilbert space with the $H^2(D)$-scalar product). In fact, if $\omega_k = \Delta r_k \to \omega$ in $H^2(D)$, we take the solution $r \in H_0^2(D)$ of the biharmonic problem $\Delta^2 r = \Delta\omega \in L^2(D)$ with the homogeneous Dirichlet boundary conditions (5.24) and (5.25), namely, $r_{|\partial D} = 0$ and $(\nabla r \cdot n)_{|\partial D} = 0$ (the existence and uniqueness of this solution has been proved above in this section). From the regularity results for higher order elliptic equations we obtain

$$\|r_k - r\|_{H^4(D)} \le C\|\Delta\omega_k - \Delta\omega\|_{L^2(D)}$$

(see Gazzola et al. [10, Corollary 2.21]). Since $\omega_k \to \omega$ in $H^2(D)$, it follows $r_k \to r$ in $H^4(D)$, and consequently $\Delta r_k = \omega_k \to \Delta r$ in $H^2(D)$, thus $\Delta r = \omega$ and X is closed.

On the other hand, the estimate above also says that for each $\omega \in X$ it holds

$$\|\omega\|_{H^2(D)} = \|\Delta r\|_{H^2(D)} \le C\|r\|_{H^4(D)} \le C\|\Delta\omega\|_{L^2(D)},$$

therefore the bilinear form $B(\upsilon, \omega) = \int_D \Delta\upsilon\, \Delta\omega\, dx$ is coercive in X, and we find a unique solution of the variational problem

$$\begin{aligned} \text{find } u \in X \ : \ \textstyle\int_D \Delta u\, \Delta\omega\, dx = \int_D f\, \omega\, dx + \int_{\partial D} g\, \nabla\omega \cdot n\, dS_x \\ - \textstyle\int_{\partial D} h\, \omega\, dS_x \quad \forall\, \omega \in X\,. \end{aligned} \tag{5.28}$$

Let us also remark that for $\omega \in X$ we have

$$\int_D \omega\eta\, dx = \int_D \Delta r\, \eta\, dx = \int_D r\, \Delta\eta\, dx = 0$$

for each $\eta \in L^2(D)$ with $\Delta\eta = 0$ in D, the integration by parts being justified by a density argument as $r \in H_0^2(D)$. Therefore $X \subset V_\sharp$.

We have thus proved the existence and uniqueness of a weak solution $u \in X \subset V_\sharp$ of a variational problem related to the same bilinear form and the same linear functional which describe the Neumann boundary value problem for the biharmonic operator. Moreover, the variational space X does not contain any directly imposed boundary condition, as it is usual for "natural" boundary conditions like those of Neumann type.

However, the proof that the solution we have found is the solution of the Neumann boundary value problem for the biharmonic operator needs some work: as we already underlined, difficulties come from the fact that the variational space X does not contain $C_0^\infty(D)$ and it is not straightforward how to trace back it to a space containing $C_0^\infty(D)$, thus problems arise when we try to interpret the meaning of the weak solution.

Let us see: we start trying to determine which equation satisfies u when we use test functions belonging to $H^2(D)$ (instead of X; note that $C_0^\infty(D)$ is contained in $H^2(D)$). Set $\mathcal{H} = \{\eta \in L^2(D) \mid \Delta\eta = 0 \text{ in } D\}$. This is a closed subspace of $L^2(D)$, thus we can use the $L^2(D)$-orthogonal projection $P_{\mathcal{H}}$ on $\mathcal{H}$. Take $q \in H^2(D)$ and set $\omega = q - P_{\mathcal{H}}q$: clearly, $\Delta\omega = \Delta q$, thus $\int_D \Delta u \Delta q \, dx = \int_D \Delta u \Delta\omega \, dx$. Moreover, $\omega \in L^2(\Delta; D)$ and

$$\int_D \omega\eta \, dx = \int_D (q - P_{\mathcal{H}}q)\eta \, dx = 0 \,,$$

for each $\eta \in \mathcal{H}$, hence $\omega \in V_\sharp$.

It is now useful the following lemma:

Lemma 5.1 *Assume that $D \subset \mathbb{R}^n$ is a bounded, connected, open set, with a smooth boundary ∂D, say, $\partial D \in C^4$. Then $X = V_\sharp$, and the norms $\| \cdot \|_X$ and $\| \cdot \|_{V_\sharp}$ are equivalent.*

Proof Here above we have verified that $X \subset V_\sharp$. Let us prove the opposite inclusion. Take $\omega \in V_\sharp$ and solve $\Delta^2 r = \Delta\omega \in L^2(D)$ with $r_{|\partial D} = 0$ and $(\nabla r \cdot n)_{|\partial D} = 0$ (namely, find $r \in H_0^2(D)$). We have already seen that this is possible and that, by the regularity results for the biharmonic operator and provided that $\partial D \in C^4$, we obtain a unique solution $r \in H^4(D) \cap H_0^2(D)$ with the estimate $\|r\|_{H^4(D)} \leq C\|\Delta\omega\|_{L^2(D)}$. Thus we have $(\Delta r - \omega) \in \mathcal{H}$. Moreover, for each $\eta \in \mathcal{H}$,

$$\int_D \Delta r \, \eta \, dx = \int_D r \, \Delta\eta \, dx = 0 \,,$$

due to the boundary conditions $r_{|\partial D} = 0$ and $(\nabla r \cdot n)_{|\partial D} = 0$; hence $\Delta r \in \mathcal{H}^\perp$ and also $(\Delta r - \omega) \in \mathcal{H}^\perp$. Having already seen $(\Delta r - \omega) \in \mathcal{H}$, we obtain $\omega = \Delta r$, $r \in H^4(D) \cap H_0^2(D)$, therefore $\omega \in X$.

Finally, we have

$$\|\omega\|_X = \|\omega\|_{H^2(D)} = \|\Delta r\|_{H^2(D)} \leq C\|r\|_{H^4(D)} \leq C\|\Delta\omega\|_{L^2(D)} \,,$$

and also

$$\|\omega\|_{V_\sharp}^2 = \|\omega\|_{L^2(D)}^2 + \|\Delta\omega\|_{L^2(D)}^2 \leq C\|\omega\|_{H^2(D)}^2 \,.$$

which ends the proof. □

Thus we can go on with the interpretation of the weak result, as now we know that $\omega = (q - P_{\mathcal{H}}q) \in X$. Moreover, we also see that $P_{\mathcal{H}}q \in H^2(D)$, as $q \in H^2(D)$ and $\omega \in X \subset H^2(D)$. By inserting ω in the variational problem (5.28) we easily find

$$\begin{aligned}\int_D \Delta u \Delta q\, dx = \int_D \Delta u \Delta \omega\, dx = \int_D f \omega\, dx + \int_{\partial D} g\, \nabla \omega \cdot n\, dS_x - \int_{\partial D} h\, \omega\, dS_x \\ = \int_D f q\, dx + \int_{\partial D} g\, \nabla q \cdot n\, dS_x - \int_{\partial D} h\, q\, dS_x \\ - \int_D f\, P_{\mathcal{H}}q\, dx - \int_{\partial D} g\, \nabla P_{\mathcal{H}}q \cdot n\, dS_x + \int_{\partial D} h\, P_{\mathcal{H}}q\, dS_x\,.\end{aligned}$$

Since a necessary condition for determining a solution of the Neumann problem is

$$\int_D f\, \eta\, dx + \int_{\partial D} g\, \nabla \eta \cdot n\, dS_x - \int_{\partial D} h\, \eta\, dS_x = 0$$

for each $\eta \in \mathcal{H} \cap H^2(D)$ (take $\Delta v = 0$ in (5.23)), by taking $\eta = P_{\mathcal{H}}q$ we conclude that the solution $u \in X$ we have found also solves (5.28) for each $q \in H^2(D)$.

Now selecting $q \in C_0^\infty(D)$ we first obtain $\Delta^2 u = f$ in D (in the weak sense). Then we take $q \in H^2(D)$ and we integrate by parts (assuming that u is smooth enough to give a meaning to the computations):

$$\begin{aligned}\int_D \Delta u\, \Delta q\, dx &= -\int_D \nabla \Delta u \cdot \nabla q\, dx + \int_{\partial D} \Delta u\, \nabla q \cdot n\, dS_x \\ &= \int_D \Delta^2 u\, q\, dx - \int_{\partial D} \nabla \Delta u \cdot n\, q\, dS_x + \int_{\partial D} \Delta u\, \nabla q \cdot n\, dS_x\,.\end{aligned}$$

Taking into account that $\Delta^2 u = f$ in D, it follows

$$-\int_{\partial D} (\nabla \Delta u \cdot n - h)\, q\, dS_x + \int_{\partial D} (\Delta u - g)\, \nabla q \cdot n\, dS_x \quad \forall\, q \in H^2(D)\,.$$

We must now select $q \in H^2(D)$ in a suitable way; precisely, it will be the solution $\rho \in H^2(D)$ of the Dirichlet boundary value problem $\Delta^2 \rho = 0$ in D with $\rho_{|\partial D} = p_1$, $(\nabla \rho \cdot n)_{|\partial D} = p_2$, with arbitrary p_1 and p_2 in suitable trace spaces. This solution ρ exists and is unique (see Gazzola et al. [10, Theorem 2.16]). Choosing $p_2 = 0$ we obtain $(\nabla \Delta u \cdot n)_{|\partial D} = h$; choosing $p_1 = 0$ it follows $(\Delta u)_{|\partial D} = g$.

Remark 5.8 We have thus realized that for the Neumann problem the situation is much more delicate than in the other three cases. An additional remark is in order: first, let us recall that for the biharmonic operator Δ^2 the three boundary conditions (5.24) and (5.25), (5.24) and (5.26), (5.25) and (5.27) satisfy the so-called Lopatinskiĭ–Šapiro condition (see Wloka [27, Sect. 11, Example 11.9]) or, equivalently, the Agmon–Douglis–Nirenberg complementing condition (see Gazzola et al. [10, Definition 2.9]). These conditions are notoriously a crucial tool for obtaining a priori estimates for classical solutions, and are often described as necessary conditions for well-posedness.

On the contrary, the Neumann boundary conditions (5.26) and (5.27) do not satisfy the Lopatinskiĭ–Šapiro condition (see Wloka [27, Sect. 11, Example 11.9]) or the Agmon–Douglis–Nirenberg complementing condition (see Gazzola et al. [10, Section 2.3]). Rather surprisingly, in spite of this fact we have proved existence and uniqueness of the weak solution for the Neumann problem associated to the biharmonic operator.

Remark 5.9 Looking at the proof just presented, we see that, under the assumption $\partial D \in C^4$, we have also proved that the continuous immersion

$$V_\sharp = L^2(\Delta; D) \cap \mathcal{H}^\perp \hookrightarrow H^2(D)$$

holds true. In fact, we have shown $V_\sharp = X$, with equivalence of the norms, and X is a closed subspace of $H^2(D)$. Therefore, by the Rellich theorem 6.9, the immersion $L^2(\Delta; D) \cap \mathcal{H}^\perp \hookrightarrow L^2(D)$ is compact; let us note again that for $L^2(\Delta; D)$ or for $L^2(\Delta; D) \cap \mathcal{H}$ this is not true (see Exercise 6.11).

5.7 Exercises

Exercise 5.1 Show that in all cases coerciveness is satisfied even if the assumption $a_0 - \frac{1}{2} \operatorname{div} b \geq 0$ in D is weakened to $a_0 - \frac{1}{2} \operatorname{div} b \geq -\nu$ in D for a constant $\nu > 0$ small enough.

Solution Let us consider the case of the Dirichlet boundary condition. We have, by using the Poincaré inequality 5.18 and proceeding as before,

$$\begin{aligned} B(v, v) &\geq \alpha_0 \int_D |\nabla v|^2 dx + \int_D (a_0 - \frac{1}{2} \operatorname{div} b) v^2 dx \\ &\geq \frac{\alpha_0}{2} \int_D |\nabla v|^2 dx + \frac{\alpha_0}{2C_D} \int_D v^2 dx - \nu \int_D v^2 dx \\ &= \frac{\alpha_0}{2} \int_D |\nabla v|^2 dx + \left(\frac{\alpha_0}{2C_D} - \nu \right) \int_D v^2 dx \,, \end{aligned}$$

therefore coerciveness holds provided that $\nu < \frac{\alpha_0}{2C_D}$. The proof in the other cases is similar, using the result provided by the Poincaré inequality in Theorem 6.10 (Neumann problem) or in Theorem 6.11 (mixed problem), or the Poincaré-type inequality in Theorem 6.12 (Robin problem).

Exercise 5.2 Taking hint from the definition of the weak divergence in Definition 5.2, give the definition of the weak curl of a vector field $q \in (L^1_{\mathrm{loc}}(D))^3$, $D \subset \mathbb{R}^3$.

Solution Having in mind the integration-by-parts formula (see Theorem C.7)

$$\int_D \operatorname{curl} q \cdot v dx = \int_D q \cdot \operatorname{curl} v dx$$

valid for $q \in (C^1(\overline{D}))^3$, $v \in (C_0^\infty(D))^3$, the weak curl of q is a vector field $\omega \in (L^1_{\text{loc}}(D))^3$ such that

$$\int_D \omega \cdot v dx = \int_D w \cdot \operatorname{curl} v dx$$

for each $v \in (C_0^\infty(D))^3$.

Exercise 5.3

(i) Show that there exists a unique solution of the weak problem

$$\text{find } u \in H^1_*(D) \ : \int_D \nabla u \cdot \nabla v dx + \int_{\partial D} u_{|\partial D} v_{|\partial D} dS_x = \int_D f v dx + \int_{\partial D} g v_{|\partial D} dS_x \quad \forall\, v \in H^1_*(D)\,,$$

where $H^1_*(D)$ is defined in (5.19).

(ii) Devise the "strong" interpretation of the weak problem above.

Solution

(i) The bilinear form

$$\int_D \nabla w \cdot \nabla v dx$$

is coercive in $H^1_*(D)$ (see Theorem 6.10), and $\int_{\partial D} v^2_{|\partial D} dS_x \geq 0$. Thus Lax–Milgram theorem 2.1 guarantees existence and uniqueness of the weak solution.

(ii) As in Sect. 5.5, take a test function $w \in H^1(D)$ and define $v = w - w_D$, where $w_D = \frac{1}{\text{meas}(D)} \int_D w dx$. Then $v \in H^1_*(D)$, and we can use it as a test function, obtaining

$$\int_D \nabla u \cdot \nabla w dx + \int_{\partial D} u_{|\partial D}(w_{|\partial D} - w_D) dS_x = \int_D f(w - w_D) dx + \int_{\partial D} g(w_{|\partial D} - w_D) dS_x\,,$$

which can be rewritten as

$$\int_D \nabla u \cdot \nabla w dx - \frac{1}{\text{meas}(D)} \int_D \left(\int_{\partial D} u_{|\partial D} dS_x \right) w dx + \int_{\partial D} u_{|\partial D} w_{|\partial D} dS_x$$
$$= \int_D f w dx + \int_{\partial D} g w_{|\partial D} dS_x$$
$$- \frac{1}{\text{meas}(D)} \int_D \left(\int_D f dx + \int_{\partial D} g dS_x \right) w dx \, .$$

Thus, following the procedure in Sect. 5.5, we obtain the equation

$$-\Delta u - \frac{1}{\text{meas}(D)} \int_{\partial D} u_{|\partial D} dS_x = f - \frac{1}{\text{meas}(D)} \left(\int_D f dx + \int_{\partial D} g dS_x \right) \quad \text{in } D \, ,$$

and the boundary condition

$$\frac{\partial u}{\partial n} + u_{|\partial D} = g \quad \text{on } \partial D \, ;$$

clearly, the solution u also satisfies the constraint $\int_D u dx = 0$.

Exercise 5.4

(i) Find $\omega \in N^\perp$, $\omega \neq 0$, where $N \subset V = L^2(D)$ is defined as in (3.2) and $\perp$ means orthogonality with respect to the scalar product in $(w, v)_V = \int_D wv dx$. Compare with Example 3.6.
(ii) Find $\omega \in N^\perp$, $\omega \neq 0$, where $N \subset V = H^1(D)$ is defined as in (3.2) and $\perp$ means orthogonality with respect to the scalar product in $((w, v))_V = \int_D (wv + \nabla w \cdot \nabla v) dx$. Compare with Example 3.6.

Solution

(i) We simply take $\omega = 1$. From an abstract point of view, it is the solution $\omega \in L^2(D)$ of the problem

$$(\omega, v)_V = \int_D v dx \quad \forall \, v \in L^2(D) \, ,$$

whose existence is assured by the Riesz representation theorem . The difference with Example 3.6 is that now we are working in the Hilbert space $L^2(D)$, so that the Riesz representation theorem holds.
(ii) Similarly to what done in (i), we take the solution $\omega \in H^1(D)$ of the problem

$$((\omega, v))_V = \int_D v dx \quad \forall \, v \in H^1(D) \, .$$

The well-posedness follows from the Riesz representation theorem 3.1, and $\omega \in N^{\perp}$. Again, the difference with Example 3.6 is that $H^1(D)$ is a Hilbert space, thus the Riesz representation theorem holds.

Exercise 5.5

(i) Devise a variational formulation for the homogeneous Dirichlet boundary value problem associated to the operator $Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w) + \sum_{i=1}^{n} b_i \mathcal{D}_i w + \sum_{i=1}^{n} \mathcal{D}_i(c_i w) + a_0 w$, where $c_i \in L^{\infty}(D)$, $i = 1, \ldots, n$.

(ii) Determine a sufficient condition on the coefficients c_i ensuring existence and uniqueness of the solution.

Solution

(i) Assuming $w, v \in H_0^1(D)$, a formal integration by parts yields the bilinear form

$$\hat{B}_L(w, v) = \int_D \sum_{i,j=1}^{n} a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i w \, v dx - \int_D \sum_{i=1}^{n} c_i w \mathcal{D}_i v dx + \int_D a_0 w v dx \,,$$

that is defined and bounded in $H_0^1(D) \times H_0^1(D)$ under the sole assumption $c_i \in L^{\infty}(D)$, $i = 1, \ldots, n$. The variational formulation is thus

$$u \in H_0^1(D) \; : \; \hat{B}_L(u, v) = \int_D f v dx \quad \forall \, v \in H_0^1(D) \,.$$

(ii) Taking $w = v$, the two terms coming from the first order terms of the operator become

$$\int_D \sum_{i=1}^{n} b_i \mathcal{D}_i v \, v dx - \int_D \sum_{i=1}^{n} c_i v \mathcal{D}_i v dx = \int_D \sum_{i=1}^{n} (b_i - c_i)\mathcal{D}_i v \, v dx \,.$$

Therefore, proceeding as in Sect. 5.4, coerciveness is achieved provided that $a_0 - \frac{1}{2}\text{div}(b - c) \geq 0$ in D.

Exercise 5.6 The physical conservation principles used to derive the time-independent linear Stokes system lead to the problem

$$\begin{cases} -\nu \sum_{i=1}^{n} \mathcal{D}_i(\mathcal{D}_i u_j + \mathcal{D}_j u_i) + \mathcal{D}_j p = f_j & \text{in } D \\ \text{div}\, u = 0 & \text{in } D \,, \end{cases} \tag{5.29}$$

for $\nu > 0$ (viscosity).

(i) Show that for a smooth solution u this problem can be rewritten as

$$\begin{cases} -\nu\Delta u + \nabla p = f & \text{in } D \\ \operatorname{div} u = 0 & \text{in } D\,. \end{cases} \tag{5.30}$$

(ii) Devise a variational formulation for the homogeneous Dirichlet boundary value problems associated to (5.29) and associated to (5.30), and show that these two variational formulations are equivalent.
(iii) Devise the variational formulation for the Neumann boundary value problem associated to (5.29), and determine the strong form of the Neumann boundary condition.
(iv) Devise the variational formulation for the Neumann boundary value problem associated to (5.30), and determine the strong form of the Neumann boundary condition.
(v) Compare the two Neumann boundary conditions in (iii) and (iv), and show that they are not equivalent.

Solution

(i) From the relation $\mathcal{D}_i\mathcal{D}_j u_i = \mathcal{D}_j\mathcal{D}_i u_i$ (that is valid for smooth functions) it follows

$$-\nu\sum_{i=1}^n \mathcal{D}_i(\mathcal{D}_i u_j + \mathcal{D}_j u_i) = -\nu\Delta u_j - \nu\mathcal{D}_j \operatorname{div} u\,,$$

thus using the second equation in (5.29) the result follows.
(ii) Taking the scalar product of (5.29) by a vector field v, integrating in D and integrating by parts we readily find

$$\begin{aligned} \int_D \sum_{j=1}^n f_j v_j dx &= \int_D \sum_{j=1}^n \Big[-\nu\sum_{i=1}^n \mathcal{D}_i(\mathcal{D}_i u_j + \mathcal{D}_j u_i) + \mathcal{D}_j p\Big] v_j dx \\ &= \nu\int_D \sum_{i,j=1}^n (\mathcal{D}_i u_j + \mathcal{D}_j u_i)\mathcal{D}_i v_j dx - \int_D p \operatorname{div} v dx \\ &\quad -\nu\int_{\partial D}\sum_{i,j=1}^n (\mathcal{D}_i u_j + \mathcal{D}_j u_i) n_i v_{j|\partial D} dS_x + \int_{\partial D} p\, v_{|\partial D}\cdot n dS_x\,. \end{aligned} \tag{5.31}$$

For the homogeneous Dirichlet boundary value problem we assume $v \in V = \{v \in H_0^1(D))^n \mid \operatorname{div} v = 0 \text{ in } D\}$, thus the term $\int_D p \operatorname{div} v dx$ and the boundary terms disappear and we are left with

$$u \in V\ :\ \nu\int_D \sum_{i,j=1}^n (\mathcal{D}_i u_j + \mathcal{D}_j u_i)\mathcal{D}_i v_j dx = \int_D \sum_{j=1}^n f_j v_j dx \quad \forall\, v \in V\,.$$

Repeating the same procedure for problem (5.30) we find

$$\begin{aligned}\int_D \sum_{j=1}^n f_j v_j dx &= \int_D \sum_{j=1}^n \Big[-\nu \sum_{i=1}^n \mathcal{D}_i \mathcal{D}_i u_j + \mathcal{D}_j p \Big] v_j dx \\ &= \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i u_j \mathcal{D}_i v_j dx - \int_D p \operatorname{div} v dx \\ &\quad -\nu \int_{\partial D} \sum_{i,j=1}^n \mathcal{D}_i u_j n_i v_{j|\partial D} dS_x + \int_{\partial D} p\, v_{|\partial D} \cdot n dS_x \,,\end{aligned} \tag{5.32}$$

and the variational formulation

$$u \in V \ : \ \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i u_j \mathcal{D}_i v_j dx = \int_D \sum_{j=1}^n f_j v_j dx \quad \forall\, v \in V \,.$$

The two formulations are equivalent as $\int_D \sum_{i,j=1}^n \mathcal{D}_j u_i \mathcal{D}_i v_j dx = \int_D \operatorname{div} u \operatorname{div} v dx$. In fact, by a density argument we can suppose $v \in C_0^\infty(D)$: thus

$$\begin{aligned}\int_D \sum_{i,j=1}^n \mathcal{D}_j u_i \mathcal{D}_i v_j dx &= -\int_D \sum_{i,j=1}^n u_i \mathcal{D}_j \mathcal{D}_i v_j dx \\ &= -\int_D \sum_{i,j=1}^n u_i \mathcal{D}_i \mathcal{D}_j v_j dx = \int_D \sum_{i,j=1}^n \mathcal{D}_i u_i \mathcal{D}_j v_j dx \,.\end{aligned}$$

(iii) Proceeding as in (ii) we obtain (5.31). The boundary terms

$$-\nu \int_{\partial D} \sum_{i,j=1}^n (\mathcal{D}_i u_j + \mathcal{D}_j u_i) n_i v_{j|\partial D} dS_x + \int_{\partial D} p\, v_{|\partial D} \cdot n dS_x$$

can be rewritten as

$$\int_{\partial D} \sum_{i,j=1}^n [-\nu(\mathcal{D}_i u_j + \mathcal{D}_j u_i) + p\delta_{ij}] n_i v_{j|\partial D} dS_x \,,$$

where δ_{ij} is the Kronecker symbol, and imposing the condition

$$\nu \sum_{i=1}^n (\mathcal{D}_i u_j + \mathcal{D}_j u_i) n_i - p n_j = g_j \ \ , \ \ j = 1, \ldots, n \,, \tag{5.33}$$

leads to the variational formulation

$$u \in W \ : \nu \int_D \sum_{i,j=1}^n (\mathcal{D}_i u_j + \mathcal{D}_j u_i)\mathcal{D}_i v_j dx$$
$$= \int_D \sum_{j=1}^n f_j v_j dx + \int_{\partial D} \sum_{j=1}^n g_j v_{j|\partial D} dS_x \quad \forall\, v \in W,$$

where $W = \{v \in (H^1(D))^n \mid \operatorname{div} v = 0 \text{ in } D\}$.

The strong form of the Neumann boundary condition (see Remark 5.1) is thus given by (5.33).

(iv) Proceeding as in (ii) we obtain (5.32). The boundary terms

$$-\nu \int_{\partial D} \sum_{i,j=1}^n \mathcal{D}_i u_j n_i v_{j|\partial D} dS_x + \int_{\partial D} p\, v_{|\partial D} \cdot n dS_x$$

can be rewritten as

$$\int_{\partial D} \sum_{i,j=1}^n (-\nu \mathcal{D}_i u_j + p\delta_{ij}) n_i v_{j|\partial D} dS_x \,,$$

where δ_{ij} is the Kronecker symbol, and imposing the condition

$$\nu \sum_{i=1}^n \mathcal{D}_i u_j n_i - p n_j = g_j \ \ , \quad j = 1, \ldots, n \,. \tag{5.34}$$

leads to the variational formulation

$$u \in W \ : \ \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i u_j \mathcal{D}_i v_j dx$$
$$= \int_D \sum_{j=1}^n f_j v_j dx + \int_{\partial D} \sum_{j=1}^n g_j v_{j|\partial D} dS_x \quad \forall\, v \in W \,,$$

where $W = \{v \in (H^1(D))^n \mid \operatorname{div} v = 0 \text{ in } D\}$.

The strong form of the Neumann boundary condition (see Remark 5.1) is thus given by (5.34).

(v) The two Neumann boundary conditions are different due to the term $\sum_{i=1}^n \mathcal{D}_j u_i n_i$, which is not present in (5.34). Anyway, there are divergence free vector fields for which this term is not vanishing, as, for instance,

$v(x_1, x_2) = (x_1, -x_2)$ on the flat boundary $\{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_2 = 0\}$. In this case we have $n = (0, 1)$ and

$$\sum_{i=1}^{n} \mathcal{D}_1 u_i n_i = 0 \ , \ \sum_{i=1}^{n} \mathcal{D}_2 u_i n_i = -1 \, .$$

Exercise 5.7

(i) Devise a variational formulation for the homogeneous Dirichlet boundary value problem associated to the linear elasticity operator $-\mu\Delta - \nu\nabla\text{div}$, $\mu > 0$, $\nu > 0$ (Lamé coefficients).

(ii) Show its well-posedness.

Solution

(i) In components, the equation $-\mu\Delta u - \nu\nabla\text{div}\, u = f$ can be rewritten as

$$-\mu \sum_{i=1}^{n} \mathcal{D}_i \mathcal{D}_i u_j - \nu \mathcal{D}_j \text{div}\, u = f_j \ , \ j = 1, \ldots, n \, ;$$

thus, multiplying by $v_j \in H_0^1(D)$, adding over $j = 1, \ldots, n$, integrating in D and integrating by parts we find:

$$\int_D \sum_{j=1}^{n} f_j v_j dx = \int_D \sum_{j=1}^{n} \Big(-\mu \sum_{i=1}^{n} \mathcal{D}_i \mathcal{D}_i u_j - \nu \mathcal{D}_j \text{div}\, u \Big) v_j dx$$
$$= \int_D \Big(\mu \sum_{i,j=1}^{n} \mathcal{D}_i u_j \mathcal{D}_i v_j + \nu \, \text{div}\, u \, \text{div}\, v \Big) dx \, ,$$

which leads to the variational formulation

$$u \in (H_0^1(D))^n \ : \int_D \Big(\mu \sum_{i,j=1}^{n} \mathcal{D}_i u_j \mathcal{D}_i v_j + \nu \, \text{div}\, u \, \text{div}\, v \Big) dx$$
$$= \int_D \sum_{j=1}^{n} f_j v_j dx \quad \forall \, v \in (H_0^1(D))^n \, .$$

(ii) Since $\int_D \nu (\text{div}\, v)^2 dx \geq 0$, well-posedness follows at once by the Poincaré inequality in $H_0^1(D)$ (see Theorem 6.4)) and Lax–Milgram theorem 2.1.

Exercise 5.8

(i) Devise a variational formulation for the homogeneous Dirichlet boundary value problem and for the Neumann boundary value problem associated to the operator $\text{curl}\,\text{curl} + \alpha I$.

(ii) Show their well-posedness.

Solution

(i) Take the scalar product of the equation $\operatorname{curl}\operatorname{curl} u + \alpha u = f$ by v, integrate in D and integrate by parts: taking into account Theorem C.8 we find

$$\int_D f \cdot v dx = \int_D (\operatorname{curl}\operatorname{curl} u + \alpha u) \cdot v dx$$
$$= \int_D (\operatorname{curl} u \cdot \operatorname{curl} v + \alpha u \cdot v) dx + \int_{\partial D} n \times \operatorname{curl} u \cdot v dS_x \,.$$

Since on the boundary it holds $v = (v \cdot n)n + n \times v \times n$, the boundary term $\int_{\partial D} n \times \operatorname{curl} u \cdot v dS_x$ can be rewritten as $\int_{\partial D} n \times \operatorname{curl} u \cdot (n \times v \times n) dS_x$. As explained in Remark 5.1, the Neumann boundary condition is thus given by $\operatorname{curl} u \times n = g$, with $g \cdot n = 0$, while the homogeneous Dirichlet boundary condition is given by $n \times v \times n = 0$, or, equivalently, $v \times n = 0$.
The variational formulations are the following: for the Neumann problem

$$u \in H(\operatorname{curl}; D) \ : \int_D (\operatorname{curl} u \cdot \operatorname{curl} v + \alpha u \cdot v) dx$$
$$= \int_D f \cdot v dx + \int_{\partial D} g \cdot (n \times v \times n) dS_x \ \ \forall\, v \in H(\operatorname{curl}; D) \,,$$

where $H(\operatorname{curl}; D) = \{v \in (L^2(D))^3 \,|\, \operatorname{curl} v \in (L^2(D))^3\}$, endowed with the scalar product

$$(w, v)_{\operatorname{curl}} = \int_D (\operatorname{curl} w \cdot \operatorname{curl} v + w \cdot v) dx$$

(the curl being intended in the weak sense), and for the homogeneous Dirichlet problem

$$u \in H_0(\operatorname{curl}; D) \ : \int_D (\operatorname{curl} u \cdot \operatorname{curl} v + \alpha u \cdot v) dx = \int_D f \cdot v dx \ \ \forall\, v \in H_0(\operatorname{curl}; D) \,,$$

where $H_0(\operatorname{curl}; D) = \{v \in H(\operatorname{curl}; D) \,|\, v \times n = 0 \text{ on } \partial D\}$.

(ii) The well-posedness of the two problems is easily proved, as the bilinear form $\int_D (\operatorname{curl} u \cdot \operatorname{curl} v + \alpha u \cdot v) dx$ defines a scalar product which is equivalent to $(w, v)_{\operatorname{curl}}$. Thus it is enough to apply the Riesz representation theorem 3.1.

[Indeed, here we are putting under the carpet some technical problems (that have a similar structure with those we had to face for the elliptic operator L):

- Is $H(\text{curl}; D)$, endowed with the scalar product $(\cdot, \cdot)_{\text{curl}}$, a Hilbert space? (The positive answer to this question is in Exercise 10.1.)
- Have the tangential component $n \times v \times n$ and the tangential trace $v \times n$ a meaning on ∂D for $v \in H(\text{curl}; D)$?
- Is the linear map $v \mapsto v \times n$ bounded from $H(\text{curl}; D)$ to a suitable tangential trace space (so that $H_0(\text{curl}; D)$ is a closed subspace of $H(\text{curl}; D)$, therefore a Hilbert space)?
- What is the real meaning of the term $\int_{\partial D} g \cdot (n \times v \times n) dS_x$? Namely, is it an integral?
- Which is the required regularity of the Neumann datum g?

We know all the answers (and for the first three questions they are positive), but it is not completely straightforward to obtain them… for these issues, see, e.g., Monk [22, Chapters 3 and 5].]

Exercise 5.9

(i) Devise a variational formulation for the homogeneous Dirichlet boundary value problem and for the Neumann boundary value problem associated to the operator $-\nabla \operatorname{div} + \alpha I$.
(ii) Show their well-posedness.

Solution

(i) As in the previous exercise, take the scalar product of the equation $-\nabla \operatorname{div} u + \alpha u = f$ by v, integrate in D and integrate by parts: taking into account Theorem C.6 we find

$$\begin{aligned} \int_D f \cdot v dx &= \int_D (-\nabla \operatorname{div} u + \alpha u) \cdot v dx \\ &= \int_D (\operatorname{div} u \operatorname{div} v + \alpha u \cdot v) dx - \int_{\partial D} \operatorname{div} u \, n \cdot v dS_x \,. \end{aligned}$$

As explained in Remark 5.1, the Neumann boundary condition is thus given by $\operatorname{div} u = g$, while the homogeneous Dirichlet boundary condition is given by $v \cdot n = 0$.
The variational formulations are the following: for the Neumann problem

$$\begin{aligned} u \in H(\text{div}; D) \ : \int_D (\operatorname{div} u \operatorname{div} v + \alpha u \cdot v) dx \\ = \int_D f \cdot v dx + \int_{\partial D} g \, v \cdot n dS_x \quad \forall \, v \in H(\text{div}; D) \,, \end{aligned}$$

where $H(\text{div}; D) = \{v \in (L^2(D))^n \,|\, \text{div}\, v \in L^2(D)\}$, endowed with the scalar product

$$(w, v)_{\text{div}} = \int_D (\text{div}\, w \,\text{div}\, v + w \cdot v) dx$$

(the divergence being intended in the weak sense), and for the homogeneous Dirichlet problem

$$u \in H_0(\text{div}; D) \,:\, \int_D (\text{div}\, u \,\text{div}\, v + \alpha u \cdot v) dx = \int_D f \cdot v dx \quad \forall\, v \in H_0(\text{div}; D)\,,$$

where $H_0(\text{div}; D) = \{v \in H(\text{div}; D) \,|\, v \cdot n = 0 \text{ on } \partial D\}$.

(ii) The well-posedness of the two problems is trivial, as the bilinear form $\int_D (\text{div}\, u \,\text{div}\, v + \alpha u \cdot v) dx$ defines a scalar product which is equivalent to $(w, v)_{\text{div}}$. Thus it is enough to apply the Riesz representation theorem 3.1.

[As in the previous exercise, here there are some technical problems:

- Is $H(\text{div}; D)$, endowed with the scalar product $(\cdot, \cdot)_{\text{div}}$, a Hilbert space? (The positive answer to this question is in Exercise 8.5.)
- Has the normal component $v \cdot n$ a meaning on ∂D for $v \in H(\text{div}; D)$?
- Is the linear map $v \mapsto v \cdot n$ bounded from $H(\text{div}; D)$ to a suitable tangential trace space (so that $H_0(\text{div}; D)$ is a closed subspace of $H(\text{div}; D)$, therefore a Hilbert space)?
- What is the real meaning of the term $\int_{\partial D} g\, v \cdot n dS_x$? Namely, is it an integral?
- Which is the required regularity of the Neumann datum g?

Again, we know all the answers (and for the first three questions they are positive): see, e.g., Monk [22, Chapters 3 and 5].]

Chapter 6
Technical Results

This chapter contains some technical results that have been frequently used in the previous sections: strictly speaking, if we had followed a "chronological" presentation, we should have proved these results before. We preferred to adopt a description without lateral interruptions, though it is quite clear that without these technical results the general ideas behind weak formulations would not have reached the desired end.

The following sections are devoted to approximation in Sobolev spaces, to the Poincaré and trace inequalities, to compactness results in $H^1(D)$ (the Rellich theorem), and to the du Bois-Reymond lemma. An "obvious" result assuring that if in a connected open set D the weak gradient of a function f vanishes then f is constant is also presented.

6.1 Approximation Results

Theorem 6.1 *Let $D \subset \mathbb{R}^n$ be an open set and define*

$$D_\varepsilon = \{x \in D \,|\, dist(x, \partial D) > \varepsilon\} .$$

Take $u \in W^{k,p}(D)$, where k is a non-negative integer and $1 \leq p < +\infty$. Then there exists a sequence $u_\varepsilon \in C^\infty(D_\varepsilon)$ with $u_\varepsilon \to u$ in $W^{k,p}_{loc}(D)$ as $\varepsilon \to 0$.

Proof We use the so-called *mollifiers*, introduced and named by Kurt O. Friedrichs[1] (earlier versions of them can be found in some seminal papers by Jean Leray[2] and

[1] Friedrichs [9].

[2] Leray [18].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_6

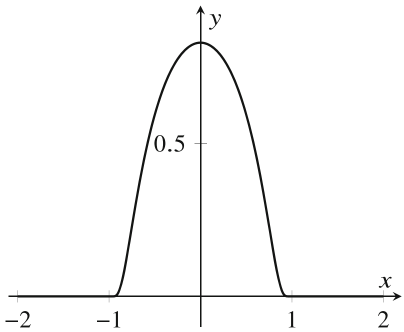


Fig. 6.1 The graph of the function η in (6.1)

Sergei L. Sobolev[3]). To define them let us consider the function

$$\eta(x) = \begin{cases} c_0 \exp\left(-\frac{1}{1-|x|^2}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1 \,, \end{cases} \tag{6.1}$$

where c_0 is such that $\int_{\mathbb{R}^n} \eta dx = 1$. In the one-dimensional case the graph of η is drawn in Fig. 6.1.

For every $\varepsilon > 0$ set

$$\eta_\varepsilon(x) = \frac{1}{\varepsilon^n}\eta\left(\frac{x}{\varepsilon}\right) .$$

This is called a ε-mollifier. It is known that if $u \in L^p_{\text{loc}}(D)$ then the ε-mollified version u_ε defined in D_ε as

$$u_\varepsilon(x) = (\eta_\varepsilon * u)(x) = \int_D \eta_\varepsilon(x-y)u(y)dy$$

belongs to $C^\infty(D_\varepsilon)$ and converges to u in $L^p_{\text{loc}}(D)$ (see, e.g., Evans [8, Theorem 6, pp. 630–631]). We need to prove that $\mathcal{D}^\alpha u_\varepsilon \to \mathcal{D}^\alpha u$ in $L^p_{\text{loc}}(D)$. To this aim, it is sufficient to show that

$$\mathcal{D}^\alpha u_\varepsilon = \eta_\varepsilon * \mathcal{D}^\alpha u \,,$$

that is, the ordinary α^{th}-partial derivative of the smooth function u_ε is the ε-mollified version of the α^{th}-weak partial derivative of u. To confirm this, we compute for $x \in D_\varepsilon$

$$\mathcal{D}^\alpha u_\varepsilon(x) = \int_D \mathcal{D}^\alpha_x \eta_\varepsilon(x-y)u(y)dy = (-1)^{|\alpha|}\int_D \mathcal{D}^\alpha_y \eta_\varepsilon(x-y)u(y)dy \,,$$

[3] Sobolev [25].

where the results comes from the fact that any derivative with respect to x is the opposite of the correspondent derivative with respect to y. For fixed $x \in D_\varepsilon$, the function $\phi(y) = \eta_\varepsilon(x-y)$ belongs to $C_0^\infty(D)$, because its support is given by $\{y \in \mathbb{R}^n \mid |y-x| \le \varepsilon\}$. Consequently, the definition of the α^{th}-weak partial derivative implies:

$$\int_D \mathcal{D}_y^\alpha \eta_\varepsilon(x-y)u(y)dy = (-1)^{|\alpha|} \int_D \eta_\varepsilon(x-y)\mathcal{D}^\alpha u(y)dy\,.$$

The proof is thus complete, as $(-1)^{|\alpha|}(-1)^{|\alpha|} = 1$. □

Exercise 6.1 Prove that $H_0^1(\mathbb{R}^n) = H^1(\mathbb{R}^n)$.

We now ask when it is possible to approximate a given function $u \in W^{k,p}(D)$ by functions belonging to $C^\infty(\overline{D})$. Such an approximation requires some conditions on the regularity of the boundary ∂D. We start with an extension result.

Theorem 6.2 (Extension Result in $W^{k,p}(D)$**)** *Let D be a bounded, connected, open subset of $\mathbb{R}^n$ with Lipschitz continuous boundary ∂D. Let $1 \le p \le +\infty$ and $k \ge 1$, and let Q a bounded, connected, open subset with $D \subset\subset Q$. Then there exists a linear and bounded operator*

$$E : W^{k,p}(D) \mapsto W^{k,p}(\mathbb{R}^n)$$

such that

(i) $Eu_{|D} = u$ a.e. in D;
(ii) $supp(Eu) \subset\subset Q$.

Proof We only present an idea of the proof, in the case $k = 1$. As a first step we consider a flat boundary. Set $B_{R,+} = \{\xi \in \mathbb{R}^n \mid |\xi| < R, \xi_n > 0\}$ and $B_{R,-} = \{\xi \in \mathbb{R}^n \mid |\xi| < R, \xi_n < 0\}$, and consider $w \in W^{1,p}(B_{R,+})$. We set, by reflection,

$$E_- w(x', x_n) = \begin{cases} w(x', x_n) & \text{if } x \in B_{R,+} \\ w(x', -x_n) & \text{if } x \in B_{R,-}\,, \end{cases}$$

having set $x = (x', x_n)$, $x' = (x_1, \dots, x_{n-1})$. As shown in Exercise 6.8, we see that $E_- w \in W^{1,p}(B_R)$.

Let us consider now a general domain D and $u \in W^{1,p}(D)$. As in Theorem 6.7 we can cover the domain $\overline{D}$ by a finite union of open balls $B_s \subset\subset Q$, $s = 1, \dots, M$, each one centered at a point $x_s \in \partial D$, plus an internal open set B_0 (say, D_{ϵ_0} for a suitable $\epsilon_0 > 0$; the covering is finite as ∂D is a closed and bounded set, therefore a compact set in $\mathbb{R}^n$). Consider a partition of unity ζ_s associated to the covering B_s of $\overline{D}$ (in particular, the support of ζ_s is a compact set in B_s: see Appendix A). The assumption on the regularity of the boundary tells us that there is a finite set of local charts ψ_s, $s = 1, \dots, M$, bijective Lipschitz continuous maps from B_s onto $B_R =$

$\{\xi \in \mathbb{R}^n \mid |\xi| < R\}$, with the inverse map ψ_s^{-1} that is Lipschitz continuous, and such that $B_s \cap D$ is mapped onto $B_{R,+}$. The functions $(\zeta_s u) \circ \psi_s^{-1}$ belong $W^{1,p}(B_{R,+})$ (see, e.g., Ziemer [29, Theor. 2.2.2, p. 52]) and has a compact support in $\overline{B_{R,+} \cap B_R}$. We can thus apply the reflection result obtained above, and we construct the function $E_-((\zeta_s u) \circ \psi_s^{-1})$ belonging to $W^{1,p}(B_R)$ (with compact support in B_R). Then we have to go back to the domain D by defining in B_s the extension $u_s = E_-((\zeta_s u) \circ \psi_s^{-1}) \circ \psi_s$; since it has a compact support in B_s, we can extend it by 0 outside B_s, obtaining $E_s u_s \in W^{1,p}(\mathbb{R}^n)$. It can be noted that $(E_s u_s)_{|D} = (\zeta_s u)_{|D}$. We finally set $Eu = \sum_{s=0}^{M} E_s u_s$ (having simply set $E_0 u_0$ the extension by 0 outside B_0 of $\zeta_0 u$). Now it is not difficult to check that Eu has the property listed in the statement of the theorem.

For more details on this proof see, e.g., Salsa [24, Section 7.8.2]. A similar proof for the general case $k \geq 1$ would need the introduction of higher order "reflections" and, due to the use of local charts, a C^k-regularity of the boundary ∂D. The result for a Lipschitz continuous boundary is proved in Stein [26, Section VI.3], by means of a different approach. □

Remark 6.1 It is also easily checked that the "extension-by-reflection" Eu constructed in the proof of the theorem satisfies $Eu \in W^{1,p}(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$ if $u \in W^{1,p}(D) \cap C^0(\overline{D})$.

The following approximation result is now an easy consequence.

Theorem 6.3 *Let D be a bounded, connected, open subset of $\mathbb{R}^n$ with Lipschitz continuous boundary ∂D. Let $u \in W^{k,p}(D)$, $1 \leq p < +\infty$. Then there exists a sequence $u_\varepsilon \in C^\infty(\overline{D})$ with $u_\varepsilon \to u$ in $W^{k,p}(D)$.*

Proof We consider the extension $Eu \in W^{k,p}(\mathbb{R}^n)$ of u, with $\text{supp}(Eu) \subset\subset Q$. Then, by Theorem 6.1 we can construct a sequence of ε-mollified versions $\widetilde{u}_\varepsilon \in C^\infty(Q_\varepsilon)$ with $\widetilde{u}_\varepsilon \to Eu$ in $W^{k,p}_{\text{loc}}(Q)$ as $\varepsilon \to 0^+$. Taking $u_\varepsilon = \widetilde{u}_{\varepsilon|\overline{D}}$ we have the desired result. □

Remark 6.2 We also obtain that, if $u \in W^{1,p}(D) \cap C^0(\overline{D})$, then the sequence $u_\varepsilon \in C^\infty(\overline{D})$ constructed in Theorem 6.3 converges to u not only in $W^{1,p}(D)$ but also in $C^0(\overline{D})$. In fact, from Remark 6.1 we know that in this case the extension $Eu \in C^0(\overline{Q})$, and it is well-known that the ε-mollified versions of a continuous function in $\overline{Q}$ converge uniformly on compact subsets of Q as $\varepsilon \to 0^+$, thus converge uniformly on $\overline{D} \subset Q$.

Exercise 6.2 Let $1 \leq p \leq +\infty$ and let p' be given by $\frac{1}{p} + \frac{1}{p'} = 1$ (with $p' = +\infty$ for $p = 1$ and viceversa). If $f_k \to f$ in $L^p(D)$ and $g_k \to g$ in $L^{p'}(D)$, then $\int_D f_k g_k dx \to \int_D fg dx$.

Exercise 6.3

(i) Let $u \in H^1(D)$, $v \in H^1(D)$. Then $uv \in W^{1,1}(D)$ and

$$\mathcal{D}_i(uv) = (\mathcal{D}_i u)v + u(\mathcal{D}_i v) \, .$$

(ii) The same result holds for $u \in W^{1,p}(D)$, $v \in W^{1,p'}(D)$, $1 < p < +\infty$, $\frac{1}{p} + \frac{1}{p'} = 1$.

6.2 Poincaré Inequality in $H_0^1(D)$

Theorem 6.4 (Poincaré Inequality in $H_0^1(D)$**)** *Let D be a bounded, connected, open subset of $\mathbb{R}^n$. Then there exists a constant $C_D > 0$ such that*

$$\int_D v^2 dx \leq C_D \int_D |\nabla v|^2 dx \quad \forall\, v \in H_0^1(D)\,.$$

Proof (1st Way) Since $H_0^1(D)$ is the closure of $C_0^\infty(D)$, we can proceed by approximation. Indeed, if we assume that the inequality holds in $C_0^\infty(D)$ it can be easily extended to $H_0^1(D)$ by the following continuity procedure: consider $v \in H_0^1(D)$, then there exists a sequence $\{v_k\}$ in $C_0^\infty(D)$ such that $v_k \to v$ in $H^1(D)$; in particular we have that

$$\int_D v_k^2 dx \to \int_D v^2 dx \;\;,\quad \int_D |\nabla v_k|^2 dx \to \int_D |\nabla v|^2 dx$$

(see Exercise 6.4), and therefore the inequality holds for v by passing to the limit in

$$\int_D v_k^2 dx \leq C_D \int_D |\nabla v_k|^2 dx\,.$$

We thus need now to prove the inequality in $C_0^\infty(D)$; let $v \in C_0^\infty(D)$, and choose a ball large enough to contain the bounded set D, say $D \subset B(x_0, R)$ with $x_0 \in D$. Note that $\operatorname{div}(x - x_0) = n$, then integrating by parts and using the Cauchy–Schwarz inequality

$$\begin{aligned}
\int_D v^2 dx &= n^{-1} \int_D n v^2 dx = n^{-1} \int_D \operatorname{div}(x - x_0) v^2 dx \\
&= -n^{-1} \int_D (x - x_0) \cdot \nabla(v^2) dx = -n^{-1} \int_D (x - x_0) \cdot 2v \nabla v dx \\
&\leq 2n^{-1} \underbrace{\sup_{x \in D} |x - x_0|}_{\leq R} \left(\int_D v^2 dx \right)^{1/2} \left(\int_D |\nabla v|^2 dx \right)^{1/2} .
\end{aligned}$$

We simplify $\left(\int_D v^2 dx\right)^{1/2}$ and defining

$$C_D = \left(\frac{2R}{n}\right)^2 = \frac{4R^2}{n^2}$$

we obtain the estimate. □

Exercise 6.4 Prove that if $v_k \to v$ in $H^1(D)$ then

$$\int_D v_k^2 dx \to \int_D v^2 dx \;\;,\;\; \int_D |\nabla v_k|^2 dx \to \int_D |\nabla v|^2 dx\,.$$

Exercise 6.5 Using an approach similar to the one presented in the first proof of Theorem 6.4, prove the Poincaré inequality for D bounded in one direction, with constant S^2 (S being the dimension of the strip containing D).

Proof of the Poincaré Inequality, 2nd Way We have already noted that, since $H_0^1(D)$ is the closure of $C_0^\infty(D)$, we can proceed by approximation. Take $v \in C_0^\infty(D)$ and extend it by 0 outside D. Since D is bounded, it is bounded in all directions; let us say that, having set $x = (x', x_n)$, $x' = (x_1, \ldots, x_{n-1})$, for each $x \in D$ we have $a \le x_n \le b$. Thus we have $v(x', a) = 0$ for all x' such that $(x', x_n) \in D$ and therefore

$$v(x', x_n) = \int_a^{x_n} \mathcal{D}_n v(x', \xi)d\xi + \underbrace{v(x', a)}_{=0} = \int_a^{x_n} \mathcal{D}_n v(x', \xi)d\xi\,.$$

Consequently,

$$\begin{aligned} v^2(x', x_n) &= \left(\int_a^{x_n} 1 \cdot \mathcal{D}_n v(x', \xi)d\xi\right)^2 \\ &\le \left(\left(\int_a^{x_n} 1^2 d\xi\right)^{\frac{1}{2}} \left(\int_a^{x_n} (\mathcal{D}_n v(x', \xi))^2 d\xi\right)^{\frac{1}{2}}\right)^2 \\ &\le (x_n - a)\int_a^{x_n} (\mathcal{D}_n v(x', \xi))^2 d\xi\,. \end{aligned}$$

Integrating in dx' we obtain

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} v^2(x', x_n)dx' &\le (x_n - a)\int_{\mathbb{R}^{n-1}}\int_a^{x_n} (\mathcal{D}_n v(x', \xi))^2 d\xi dx' \\ &\le (x_n - a)\int_{\mathbb{R}^n} (\mathcal{D}_n v(x', \xi))^2 d\xi dx'\,. \end{aligned}$$

Thus

$$\int_a^b \int_{\mathbb{R}^{n-1}} v^2(x', x_n)dx'dx_n \le \int_a^b (x_n - a)\left[\int_{\mathbb{R}^n} (\mathcal{D}_n v(x', \xi))^2 d\xi dx'\right] dx_n$$
$$= \frac{1}{2}(b-a)^2 \int_{\mathbb{R}^n} (\mathcal{D}_n v(x))^2 dx$$
$$= \frac{1}{2}(b-a)^2 \int_D (\mathcal{D}_n v(x))^2 dx \qquad (v = 0 \text{ outside } D)$$

and

$$\int_a^b \int_{\mathbb{R}^{n-1}} v^2(x', x_n)dx'dx_n = \int_{\mathbb{R}^n} v(x)^2 dx \qquad (v = 0 \text{ for } x_n \notin (a, b))$$
$$= \int_D v(x)^2 dx \qquad (v = 0 \text{ outside } D)\,.$$

In conclusion

$$\int_D v^2 dx \le \frac{1}{2}(b-a)^2 \int_D (\mathcal{D}_n v)^2 dx \le \frac{1}{2}(b-a)^2 \int_D |\nabla v|^2 dx\,,$$

thus the stated estimate with $C_D = \frac{1}{2}(b-a)^2$. □

Exercise 6.6 The Poincaré inequality still holds in $W_0^{1,p}(D)$, $1 \le p < +\infty$: there exists a constant $C_D > 0$ such that

$$\int_D |v|^p dx \le C_D \int_D |\nabla v|^p dx \quad \forall\, v \in W_0^{1,p}(D)\,.$$

6.3 Trace Inequality

Next we discuss the possibility of assigning "boundary values" on ∂D to a function $v \in H^1(D)$, assuming that ∂D is Lipschitz continuous. When we deal with $v \in C(\overline{D})$, clearly it has values on ∂D in the usual sense. The problem is that a typical function $v \in H^1(D)$ is not in general continuous and, even worse, is only defined almost everywhere in D. Since ∂D can have n-dimensional Lebesgue measure equal to zero, it seems that we cannot give a clear meaning to the expression "v restricted to ∂D". The notion of a *trace on the boundary* solves this problem.

Theorem 6.5 (Trace on ∂D and Trace Inequality) *Let D be a bounded, connected, open set with a Lipschitz continuous boundary ∂D. Then for $v \in H^1(D)$ there is a way to determine a function $\gamma_0 v \in L^2(\partial D)$ such that*

$$\gamma_0 v = v_{|\partial D} \quad \text{for } v \in C^\infty(\overline{D})$$

and

$$\int_{\partial D} (\gamma_0 v)^2 dx \le C_* \int_D (v^2 + |\nabla v|^2) dx$$

for a suitable $C_* > 0$ (independent of v). Moreover, the map $v \to \gamma_0 v$ is linear and, from the inequality above, continuous from $H^1(D)$ to $L^2(\partial D)$.

Definition 6.1 We call $\gamma_0 v$ the *trace* of v on ∂D, and, even if this can lead to some confusion, very often in the sequel we will continue to write $v_{|\partial D}$ instead of $\gamma_0 v$.

The proof of this theorem needs some steps. We start by proving it for smooth functions defined in a half-space. To clarify this point, we need some notation. Suppose we have $v \in C^1(\overline{\mathbb{R}^n_+})$, where $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \,|\, x_n > 0\}$, with $v = 0$ out of

$$\overline{B_{R,+}} = \{x \in \mathbb{R}^n \,|\, x_n \ge 0\,,\ |x| \le R\}\,.$$

Then we have

Theorem 6.6 (Trace Inequality in $\mathbb{R}^n_+$ for C^1-Functions) *For any $v \in C^1(\overline{\mathbb{R}^n_+})$ vanishing outside $\overline{B_{R,+}}$ it holds*

$$\int_{\mathbb{R}^{n-1}} v^2(x', 0) dx' \le R \int_{\mathbb{R}^n_+} (\mathcal{D}_n v)^2 dx\,.$$

Proof For $(x', 0) \in \overline{B_{R,+}}$ we have

$$v(x', 0) = -\int_0^R \mathcal{D}_n v(x', \xi) d\xi + \underbrace{v(x', R)}_{=0} = -\int_0^R \mathcal{D}_n v(x', \xi) d\xi\,.$$

Thus, as in the second proof of the Poincaré inequality:

$$\int_{\mathbb{R}^{n-1}} v^2(x', 0) dx' \le R \int_{\mathbb{R}^{n-1}} \left(\int_0^R (\mathcal{D}_n v(x', \xi))^2 d\xi \right) dx' = R \int_{\mathbb{R}^n_+} (\mathcal{D}_n v)^2 dx\,,$$

where the last equality is justified since $v = 0$ outside $\overline{B_{R,+}}$. □

Now we can obtain the following theorem:

Theorem 6.7 (Trace Inequality in D for C^1-Functions) *Let D be a bounded, connected, open set with a Lipschitz continuous boundary ∂D. There exists a constant $C_* > 0$ such that*

$$\int_{\partial D} v^2_{|\partial D} dS_x \leq C_* \int_D (v^2 + |\nabla v|^2) dx \qquad \forall\, v \in C^1(\overline{D}) .$$

Proof The proof is rather technical and we will only enlighten some essential ideas. To simplify a little the procedure, let us also suppose that the regularity of the boundary is C^1; the proof for the Lipschitz case is just a little bit more complicate, as in that case we have to deal with almost everywhere differentiable functions with bounded derivatives (this is the case of Lipschitz functions, by the Rademacher theorem: see, e.g., Ziemer [29, Theor. 2.2.1, p. 50]).

We can cover the boundary ∂D by a finite union of open balls B_s, $s = 1, \dots, M$, each one centered at a point $x_s \in \partial D$ (the covering is finite as ∂D is a closed and bounded set, therefore a compact set in $\mathbb{R}^n$). Consider a sub-covering $\widehat{B}_s$, $s = 1, \dots, M$, with $\widehat{B}_s \subset\subset B_s$, and a set of cut-off functions ζ_s such that $\zeta_s \in C_0^\infty(B_s)$, $0 \leq \zeta_s(x) \leq 1$ for $x \in \mathbb{R}^n$ and $\zeta_s(x) = 1$ for $x \in \widehat{B}_s$ (in particular, the support of ζ_s is a compact set in B_s; see Corollary A.1). The assumption on the regularity of the boundary tells us that there is a finite set of local charts ψ_s, bijective C^1-maps from B_s onto $B_R = \{\xi \in \mathbb{R}^n \,|\, |\xi| < R\}$, with the inverse map ψ_s^{-1} that is C^1, and such that $B_s \cap D$ is mapped onto $B_{R,+} = \{\xi \in \mathbb{R}^n \,|\, |\xi| < R, \xi_n > 0\}$. The functions $(\zeta_s v) \circ \psi_s^{-1}$ are C^1-functions in $\overline{\mathbb{R}^n_+}$, vanishing outside $\overline{B_{R,+}}$. Therefore we can apply to each of them the result of Theorem 6.6, and we get

$$\int_{\mathbb{R}^{n-1}} ((\zeta_s v) \circ \psi_s^{-1})^2 (x', 0) dx' \leq R \int_{\mathbb{R}^n_+} (\mathcal{D}_n((\zeta_s v) \circ \psi_s^{-1}))^2 dx .$$

Transforming these integrals into integrals in $B_s \cap \partial D$ and $B_s \cap D$ we find, by the chain rule and some straightforward estimates,

$$\int_{B_s \cap \partial D} (\zeta_s v)^2 dS_x \leq C \int_{B_s \cap D} (|\nabla v|^2 + v^2) dx .$$

Now we can add for $s = 1, \dots, M$, and using the fact that ζ_s is equal to 1 on $\widehat{B}_s$ we obtain the final result. □

We can now give the proof of the trace theorem (Theorem 6.5).

Proof of Theorem 6.5 We proceed by approximation. Consider $v_k \in C^\infty(\overline{D})$ such that $v_k \to v$ in $H^1(D)$. By the trace theorem for C^1-functions we have that

$$\int_{\partial D} v^2_{k|\partial D} dS_x \leq C_* \int_D (v_k^2 + |\nabla v_k|^2) dx \quad \forall\, k \geq 1 \tag{6.2}$$

and

$$\int_{\partial D} (v_{k|\partial D} - v_{s|\partial D})^2 dS_x \leq C_* \int_D [(v_k - v_s)^2 + |\nabla(v_k - v_s)|^2] dx \quad \forall\, k, s \geq 1\,. \tag{6.3}$$

Since v_k is convergent, it is a Cauchy sequence in $H^1(D)$. Therefore

$$\int_{\partial D} (v_{k|\partial D} - v_{s|\partial D})^2 dS_x \leq C_* \int_D [(v_k - v_s)^2 + |\nabla(v_k - v_s)|^2] dx \leq C_* \epsilon$$

for k, s large enough, and thus we see that $v_{k|\partial D}$ is a Cauchy sequence in $L^2(\partial D)$. Since $L^2(\partial D)$ is a Hilbert space, we find $q \in L^2(\partial D)$ such that $v_{k|\partial D} \to q$ in $L^2(\partial D)$. Taking the limit in (6.2) we have

$$\int_{\partial D} q^2 dS_x \leq C_* \int_D (v^2 + |\nabla v|^2) dx\,.$$

This value q does not depend on the approximating sequence v_k, but only on v. In fact, if w_k is another approximating sequence of v, and p is the limit in $L^2(\partial D)$ of $w_{k|\partial D}$, it follows

$$\begin{aligned}
\int_{\partial D} |q - p|^2 dS_x &= \int_{\partial D} |q - v_{k|\partial D} + v_{k|\partial D} - w_{k|\partial D} + w_{k|\partial D} - p|^2 dS_x \\
&\leq 3\Big[\int_{\partial D} (q - v_{k|\partial D})^2 dS_x + \int_{\partial D} (p - w_{k|\partial D})^2 dS_x \\
&\quad + \int_{\partial D} (v_{k|\partial D} - w_{k|\partial D})^2 dS_x\Big] \\
&\leq 3\Big[\int_{\partial D} (q - v_{k|\partial D})^2 dS_x + \int_{\partial D} (p - w_{k|\partial D})^2 dS_x \\
&\quad + C_* \int_D \Big[(v_k - w_k)^2 + |\nabla(v_k - w_k)|^2\Big] dx\Big]\,,
\end{aligned}$$

and all the terms go to 0, as $v_k \to v$ in $H^1(D)$ and $w_k \to v$ in $H^1(D)$.

In conclusion, we define the trace $\gamma_0 v$ as the unique value $q \in L^2(\partial D)$ obtained with the above procedure. Clearly the map $v \mapsto q$ is linear; moreover, if $v \in C^\infty(\overline{D})$ we can choose $v_k = v$ for all $k \geq 1$, therefore

$$v_{k|\partial D} = v_{|\partial D} \to \gamma_0 v\,,$$

showing that the trace of a smooth function v (the limit of $v_{k|\partial D} \ldots$) is coincident with its restriction on the boundary. □

Remark 6.3 As we have seen the proof of the trace inequality is based on an elementary argument that we have already met many times. Indeed, if we consider a continuous function $f : \mathbb{Q} \mapsto \mathbb{R}$ and we want to extend this function to all $\mathbb{R}$, how can we do? Let x be an irrational number; since $\mathbb{Q}$ is dense in $\mathbb{R}$, we can take a sequence $\{r_k\} \subset \mathbb{Q}$ such that $r_k \to x$. Then the natural step is to define $f(x)$ as the limit of $f(r_k)$. To led this argument to its end we have to verify that the limit exists, proving for example that $\{f(r_k)\}$ is a Cauchy sequence, and that its limit does not depend on the sequence $\{r_k\}$ we have chosen.

Remark 6.4 If $v \in H^1(D) \cap C^0(\overline{D})$, we know from Remark 6.2 we can find a sequence $v_k \in C^\infty(\overline{D})$ that converges to v in $H^1(D)$ and in $C^0(\overline{D})$ (namely, uniformly in $\overline{D}$). Then on one side

$$v_{k|\partial D} \to \gamma_0 v \quad \text{in } L^2(\partial D) \quad \text{(definition of the trace } \gamma_0 v)$$

and on the other side

$$v_{k|\partial D} \to v_{|\partial D} \quad \text{in } C^0(\partial D) \quad \text{(uniform convergence in } \overline{D}) \,,$$

in particular $v_{k|\partial D} \to v_{|\partial D}$ in $L^2(\partial D)$. Thus the trace $\gamma_0 v$ on ∂D is equal to the restriction $v_{|\partial D}$ on ∂D for all functions $v \in H^1(D) \cap C^0(\overline{D})$.

Remark 6.5 It can be proved that $H^1_0(D)$ is equal to the space $\{v \in H^1(D) \,|\, v_{|\partial D} = 0 \text{ on } \partial D\}$. The proof of the inclusion $H^1_0(D) \subset \{v \in H^1(D) \,|\, v_{|\partial D} = 0 \text{ on } \partial D\}$ is easy. In fact, an element $v \in H^1_0(D)$ can be approximated by a sequence $v_k \in C^\infty_0(D)$; since $v_{k|\partial D} = 0$, it follows that the trace $v_{|\partial D}$ satisfies $v_{|\partial D} = 0$. The opposite inclusion is also true, but the proof is a little bit technical, therefore we do not present it here (see Evans [8, Theorem 2, pp. 259–261]).

Remark 6.6 Let us note that the trace inequality still holds in $W^{1,p}(D)$ $(1 \leq p < +\infty)$. The proof of the basic estimate for smooth functions in Theorem 6.6 is essentially the same of the similar estimate for the Poincaré inequality (see Exercise 6.6).

Remark 6.7 A result similar to that presented in Theorem 6.5 can be proved for the trace on Γ, a (non-empty) open and Lipschitz continuous subset of ∂D.

Having defined the trace, we can prove an integration by parts formula. We state it as an exercise.

Exercise 6.7 Let D a bounded, connected, open set with a Lipschitz continuous boundary ∂D, and take $u \in H^1(D)$, $v \in H^1(D)$. Then the integration by parts formula

$$\int_D (\mathcal{D}_i u) v dx = -\int_D u \mathcal{D}_i v dx + \int_{\partial D} n_i u_{|\partial D} v_{|\partial D} dS_x$$

holds.

Another couple of exercises are the following:

Exercise 6.8 Let us assume that D is a bounded, connected, open set with a Lipschitz continuous boundary ∂D, and that $\overline{D} = \overline{D_1} \cup \overline{D_2}$, $D_1 \cap D_2 = \emptyset$, where D_1 and D_2 are (non-empty) open sets with a Lipschitz continuous boundary. Set $\Gamma = \partial D_1 \cap \partial D_2$ and take $v \in L^p(D)$, $1 \leq p < +\infty$. Then $v \in W^{1,p}(D)$ if and only if $v_{|D_1} \in W^{1,p}(D_1)$, $v_{|D_2} \in W^{1,p}(D_2)$ and the trace of $v_{|D_1}$ and $v_{|D_2}$ on Γ is the same.

Exercise 6.9 Let D a bounded, connected, open set with a Lipschitz continuous boundary ∂D. The statement "there exists a constant $C > 0$ such that

$$\int_{\partial D} |v|^p dS_x \leq C \int_D |v|^p dx \quad \forall\, v \in C^0(\overline{D})\text{"}$$

is *false* for $1 \leq p < +\infty$.

6.4 Compactness and Rellich Theorem

First of all, we see a compactness criterion (similar to Ascoli-Arzelà theorem, and due to Kolmogorov and M. Riesz).

Theorem 6.8 (Precompactness) *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set. Consider $1 \leq p < +\infty$ and $X \subset L^p(D)$. Then X is precompact if and only if*

(i) there exists $M > 0$ such that

$$\|v\|_{L^p(D)} \leq M \quad \forall\, v \in X\,;$$

(ii) extending v by 0 outside D, it holds

$$\lim_{h \to 0} \|v(\cdot + h) - v(\cdot)\|_{L^p(D)} = 0\,,$$

uniformly with respect to $v \in X$.

Remark 6.8 Remember that a subset X of a Banach space Y is said to be precompact if its closure is compact, i.e., from any sequence in X we can extract a subsequence convergent in Y to an element that does not necessarily belong to X.

The principal compactness result in Sobolev spaces is the following:

Theorem 6.9 (Rellich Theorem) *Let D a bounded, connected, open subset of $\mathbb{R}^n$, with a Lipschitz continuous boundary ∂D, and let $1 \leq p < +\infty$. Then $W^{1,p}(D)$ is compactly immersed in $L^p(D)$: from any bounded sequence $v_k \in W^{1,p}(D)$ it is possible to extract a subsequence v_{k_s} that converges in $L^p(D)$ to a limit $v \in L^p(D)$.*

Proof We use the precompactness theorem, and we limit ourselves to the case $p = 2$. Let us start with an estimate that is valid for smooth functions. Taking $v \in C_0^\infty(\mathbb{R}^n)$ it follows

$$v(x+h) - v(x) = \int_0^1 \frac{d}{dt}[v(x+th)]dt = \int_0^1 \nabla v(x+th) \cdot h dt \,,$$

hence

$$|v(x+h) - v(x)|^2 = \left| \int_0^1 \nabla v(x+th) \cdot h dt \right|^2 \le |h|^2 \left| \int_0^1 \nabla v(x+th) dt \right|^2$$
$$\le |h|^2 \int_0^1 |\nabla v(x+th)|^2 dt$$

by the Cauchy–Schwarz inequality. Integrating in $\mathbb{R}^n$

$$\begin{aligned} \int_{\mathbb{R}^n} |v(x+h) - v(x)|^2 dx &\le |h|^2 \int_{\mathbb{R}^n} \left(\int_0^1 |\nabla v(x+th)|^2 dt \right) dx \\ &= |h|^2 \int_0^1 \left(\int_{\mathbb{R}^n} |\nabla v(x+th)|^2 dx \right) dt \\ &= |h|^2 \int_{\mathbb{R}^n} |\nabla v|^2 dx \,, \end{aligned}$$

having performed the change of variable $x + th = y$ (and then replaced dy with $dx \ldots$). By approximation, since $C_0^\infty(\mathbb{R}^n)$ is dense in $H_0^1(\mathbb{R}^n)$, we have that this inequality is true for $v \in H_0^1(\mathbb{R}^n)$:

$$\int_{\mathbb{R}^n} |v(x+h) - v(x)|^2 dx \le |h|^2 \int_{\mathbb{R}^n} |\nabla v|^2 dx \,. \tag{6.4}$$

Now we want to prove that a bounded set $X \subset H^1(D)$ is precompact in $L^2(D)$. Consider

$$X \subset \{v \in H^1(D) \mid \|v\|_{H^1(D)} \le M\} \,.$$

By the extension theorem (Theorem 6.2) we know that, for $v \in X$, $Ev \in H_0^1(\mathbb{R}^n)$, $\mathrm{supp}(Ev) \subset\subset Q$. Thus $Ev \in H_0^1(Q)$ and is vanishing outside Q; moreover, from the continuity of the extension operator we have

$$\|Ev\|_{H^1(Q)} = \|Ev\|_{H^1(\mathbb{R}^n)} \le C_* \|v\|_{H^1(D)} \le C_* M \qquad \forall\, v \in X \,.$$

Let us denote by EX the set of the extensions of elements of X

$$EX = \{w \in H_0^1(Q) \,|\, \exists\, v \in X \text{ such that } w = Ev\}\,.$$

We have just shown that EX is bounded in $L^2(Q)$. Furthermore we know that (6.4) is satisfied for all $w \in EX$, thus

$$\int_Q |(Ev)(x+h) - (Ev)(x)|^2 dx \le \int_{\mathbb{R}^n} |(Ev)(x+h) - (Ev)(x)|^2 dx$$

$$\underset{(6.4)}{\le} |h|^2 \int_{\mathbb{R}^n} |\nabla Ev|^2 dx \le C_*^2 M^2 |h|^2 \qquad \forall\, v \in X\,.$$

Applying Theorem 6.7 we obtain that EX is precompact in $L^2(Q)$. Take now a sequence $v_k \in X$: since EX is precompact in $L^2(Q)$, we can select a subsequence Ev_{k_s} convergent to w_0 in $L^2(Q)$. Then $v_{k_s} = Ev_{k_s|D}$ converges to $w_{0|D}$ in $L^2(D)$, and the proof is complete. □

Exercise 6.10 Let D a bounded, connected, open set with a Lipschitz continuous boundary ∂D. Let v_k be a bounded sequence in $W^{1,p}(D)$, $1 < p < +\infty$, and consider a subsequence v_{k_s} which converges to v in $L^p(D)$ by the Rellich theorem. Prove that the limit v indeed belongs to $W^{1,p}(D)$.

6.5 Other Poincaré Inequalities

We are now in a condition to prove other Poincaré inequalities that are useful in the proof of the coerciveness of the bilinear form $B_L(\cdot,\cdot)$ introduced in (2.19) (see Sect. 5.4 for these coerciveness results).

Theorem 6.10 *Let D be bounded, connected, open subset of $\mathbb{R}^n$ with a Lipschitz continuous boundary ∂D. Denote by*

$$H_*^1(D) = \left\{ v \in H^1(D) \,\middle|\, \int_D v dx = 0 \right\}\,.$$

Then there exists $C_* > 0$ such that

$$\int_D v^2 dx \le C_* \int_D |\nabla v|^2 dx \quad \forall\, v \in H_*^1(D)\,.$$

Proof Assume, by contradiction, that for each $k \in \mathbb{N}$, $k \neq 0$, we can find $v_k \in H_*^1(D)$ such that

$$\int_D v_k^2 dx > k \int_D |\nabla v_k|^2 dx\,.$$

Thus $\int_D v_k^2 dx > 0$, and we can consider

$$w_k = \frac{v_k}{\left(\int_D v_k^2 dx\right)^{1/2}} \in H^1_*(D)\,,$$

which satisfies $\int_D w_k^2 dx = 1$. We clearly have that

$$1 = \int_D w_k^2 dx > k \int_D |\nabla w_k|^2 dx \implies \int_D |\nabla w_k|^2 dx < \frac{1}{k}\,, \tag{6.5}$$

in particular

$$\|w_k\|_{H^1(D)} = \left(\int_D w_k^2 dx + \int_D |\nabla w_k|^2 dx\right)^{1/2} \le \sqrt{2}\,.$$

From Rellich theorem we can extract a subsequence w_{k_s} which converges to w_0 in $L^2(D)$, therefore

$$\int_D w_0^2 dx = \lim_{s\to\infty} \int_D w_{k_s}^2 dx = 1\,.$$

From (6.5) we have $\nabla w_{k_s} \to 0$ in $(L^2(D))^n$; therefore for each $\varphi \in C_0^\infty(D)$ and for each $i = 1, \dots, n$ it holds

$$\int_D w_0 \mathcal{D}_i \varphi dx = \lim_{s\to\infty} \int_D w_{k_s} \mathcal{D}_i \varphi dx = -\lim_{s\to\infty} \int_D (\mathcal{D}_i w_{k_s}) \varphi dx = 0\,.$$

As a consequence $\nabla w_0 = 0$ and $w_0 \in H^1(D)$. From $w_{k_s} \to w_0$ in $L^2(D)$ we also have that

$$\int_D w_0 dx = \lim_{s\to\infty} \int_D w_{k_s} dx = 0\,,$$

thus $w_0 \in H^1_*(D)$. From $\mathcal{D}_i w_0 = 0$ for each $i = 1, \dots, n$ we can infer $w_0 = \text{const}$ (see Sect. 6.7) and thus we have a contradiction, as the only constant belonging to $H^1_*(D)$ is the null constant, but then $\int_D w_0^2 dx = 1$ is impossible. □

Let us continue by presenting other similar results. We start with this remark:

Remark 6.9 Let D be bounded, connected, open subset of $\mathbb{R}^n$ with a Lipschitz continuous boundary ∂D, and let $\Gamma_D \subset \partial D$ be a non-empty, open Lipschitz continuous subset. It can be proved that $H^1_{\Gamma_D}(D)$, the closure of $C^\infty_{\Gamma_D}(\overline{D})$ in $H^1(D)$, is equal to the space $\{v \in H^1(D) \,|\, v_{|\Gamma_D} = 0\}$, where $v_{|\Gamma_D}$ is the trace on Γ_D (see Remark 6.7). As already seen in Remark 6.5, the easy part is the inclusion $H^1_{\Gamma_D}(D) \subset \{v \in H^1(D) \,|\, v_{|\Gamma_D} = 0\}$; the inverse inclusion is more technical.

Theorem 6.11 *Let D be bounded, connected, open subset of $\mathbb{R}^n$ with a Lipschitz continuous boundary ∂D. Denote by*

$$H^1_{\Gamma_D}(D) = \left\{ v \in H^1(D) \mid v_{|\Gamma_D} = 0 \right\} ,$$

where $\Gamma_D \subset \partial D$ is a non-empty, open Lipschitz continuous subset. Then there exists $C_* > 0$ such that

$$\int_D v^2 dx \le C_* \int_D |\nabla v|^2 dx \quad \forall\, v \in H^1_{\Gamma_D}(D) .$$

Proof It is essentially the same as before. The only change is a consequence of the remark that, having found $w_{k_s} \to w_0$ in $L^2(D)$ with $\nabla w_{k_s} \to 0 = \nabla w_0$ in $(L^2(D))^n$, we have indeed obtained $w_{k_s} \to w_0$ in $H^1(D)$. Thus by the continuity of the trace operator we find $0 = w_{k_s|\Gamma_D} \to w_{0|\Gamma_D}$ in $L^2(\Gamma_D)$, hence $w_0 \in H^1_{\Gamma_D}(D)$. Since we also know that $w_0 =$ const and that the only constant belonging to $H^1_{\Gamma_D}(D)$ is the null constant, again we obtain a contradiction from $\int_D w_0^2 dx = 1$. □

For the Robin problem this Poincaré-type inequality is important.

Theorem 6.12 *Let D be bounded, connected, open subset of $\mathbb{R}^n$ with a Lipschitz continuous boundary ∂D. Let $q : \partial D \mapsto \mathbb{R}$ be a non-negative and bounded function, not identically vanishing, namely, such that $\int_{\partial D} q dS_x > 0$. Then there exists $C_* > 0$ such that*

$$\int_D v^2 dx \le C_* \left(\int_D |\nabla v|^2 dx + \int_{\partial D} q v^2 dS_x \right) \quad \forall\, v \in H^1(D) . \tag{6.6}$$

Proof The result is proved as before. We arrive at $w_{k_s} \to w_0$ in $H^1(D)$, with $\int_D w_0^2 dx = 1$ and $w_0 =$ const. By the continuity of the trace operator we obtain that $w_{k_s|\partial D} \to w_{0|\partial D}$ in $L^2(\partial D)$, thus also $\sqrt{q} w_{k_s|\partial D} \to \sqrt{q} w_{0|\partial D}$ in $L^2(\partial D)$. As a consequence,

$$\int_{\partial D} q w_{k_s}^2 dS_x \to \int_{\partial D} q w_0^2 dS_x ,$$

by applying Exercise 6.2 in $L^2(\partial D)$. On the other hand, from the assumption that inequality (6.6) does not hold we have

$$\int_D |\nabla w_{k_s}|^2 dx + \int_{\partial D} q w_{k_s}^2 dS_x < \frac{1}{k_s} ,$$

hence $\int_{\partial D} q w_{k_s}^2 dS_x \to 0$. The contradiction comes from the fact that $\int_{\partial D} q w_0^2 dS_x = 0$ implies $w_0 = 0$, as w_0 is constant and $\int_{\partial D} q dS_x > 0$. □

We conclude with the following theorem:

Theorem 6.13 *Let D be bounded, connected, open subset of $\mathbb{R}^n$ with a Lipschitz continuous boundary ∂D. Then there exists $C_* > 0$ such that*

$$\int_D (v - v_D)^2 dx \le C_* \int_D |\nabla v|^2 dx \quad \forall\, v \in H^1(D)\,,$$

where $v_D = \frac{1}{meas(D)} \int_D v dx$.

Proof The proof is trivial. Indeed it is sufficient to consider $w = v - v_D$, which is average free and satisfies $\nabla w = \nabla v$. Thus we can apply Theorem 6.10. □

6.6 du Bois-Reymond Lemma

Lemma 6.1 *Let D be an open set in $\mathbb{R}^n$. If $f \in L^1_{loc}(D)$ satisfies*

$$\int_D f\varphi dx = 0 \qquad \forall\, \varphi \in C_0^\infty(D) \tag{6.7}$$

then $f = 0$ a.e. in D.

Proof For $r > 0$ and $\varepsilon > 0$ denote by $B_r = \{x \in \mathbb{R}^n \mid |x| < r\}$ and by $D_\varepsilon = \{x \in D \mid \text{dist}(x, \partial D) > \varepsilon\}$. Take k_0 large enough to have $D_{1/k_0} \cap B_{k_0} \neq \emptyset$. For a fixed $k \in \mathbb{N}$, $k \ge k_0$ and for $0 < \delta < 1/k$ consider the δ-mollified version $f_\delta = \eta_\delta * f$ defined in $D_\delta \supset D_{1/k}$.

For any fixed $x \in D_{1/k}$ the map $y \mapsto \eta_\delta(x - y) \in C_0^\infty(D)$, thus by (6.7) we obtain

$$f_\delta(x) = \int_D f(y)\eta_\delta(x - y)dy = 0\,.$$

We also know that $f_\delta \to f$ in $L^1_{\text{loc}}(D)$, in particular $f_\delta \to f$ in $L^1(D_{1/k} \cap B_k)$. Therefore, for a suitable subsequence we find $f_{\delta_s} \to f$ a.e. in $D_{1/k} \cap B_k$.

Putting together the two results it follows $f(x) = 0$ a.e. in $D_{1/k} \cap B_k$. Since $D = \cup_{k=k_0}^\infty (D_{1/k} \cap B_k)$, the thesis is proved. □

6.7 $\nabla f = 0$ implies $f = \text{const}$

Proposition 6.1 *Let D be an open and connected set in $\mathbb{R}^n$. Suppose that $f \in L^1_{loc}(D)$ satisfies $\mathcal{D}_i f = 0$ for each $i = 1, \ldots, n$. Then $f = const$ a.e. in D.*

Proof It is enough to prove that there exists $c_0 \in \mathbb{R}$ such that

$$\int_D f\varphi dx = c_0 \int_D \varphi dx \qquad \forall\, \varphi \in C_0^\infty(D)\,.$$

In fact, from this it follows $\int_D (f - c_0)\varphi dx = 0$ for each $\varphi \in C_0^\infty(D)$, thus from du Bois-Reymond Lemma 6.1 we obtain $f = c_0$ a.e. in D. Consider now $\varphi \in C_0^\infty(D)$: the assumption says that the weak gradient of f is vanishing, namely,

$$0 = \int_D f\mathcal{D}_i\varphi dx \qquad \text{for each } i = 1, \ldots, n\,.$$

Take $Q \subset\subset D$, Q open and connected. Consider the ε-mollified version $f_\varepsilon = \eta_\varepsilon * f$, defined in Q for $\varepsilon < \varepsilon_Q$. We already know that

$$\mathcal{D}_i f_\varepsilon = \eta_\varepsilon * \mathcal{D}_i f$$

(see the proof of Theorem 6.1), thus

$$\mathcal{D}_i f_\varepsilon = 0 \quad \text{in } Q\,.$$

Therefore we have

$$f_\varepsilon = c_{\varepsilon,Q} \quad \text{in } Q\,,$$

and for any $\varphi \in C_0^\infty(Q)$ it follows, for $\varepsilon < \varepsilon_Q$,

$$\int_Q f_\varepsilon\varphi dx = c_{\varepsilon,Q}\int_Q \varphi dx\,. \tag{6.8}$$

Selecting $\hat\varphi_Q \in C_0^\infty(Q)$ such that $\int_Q \hat\varphi_Q dx \neq 0$, for $\varepsilon < \varepsilon_Q$ we have from (6.8)

$$c_{\varepsilon,Q} = \frac{\int_Q f_\varepsilon\hat\varphi_Q dx}{\int_Q \hat\varphi_Q dx}\,.$$

Since $f_\varepsilon \to f$ in $L^1_{\text{loc}}(D)$, we get $\int_Q f_\varepsilon\hat\varphi_Q dx \to \int_Q f\hat\varphi_Q dx$, hence

$$c_{\varepsilon,Q} \to \frac{\int_Q f\hat\varphi_Q dx}{\int_Q \hat\varphi_Q dx} = c_{0,Q}\,.$$

On the other hand, we also have $\int_Q f_\varepsilon \varphi dx \to \int_Q f \varphi dx$ for any $\varphi \in C_0^\infty(Q)$, thus from (6.8) we obtain

$$\int_Q f\varphi dx = c_{0,Q} \int_Q \varphi dx \qquad \forall\, \varphi \in C_0^\infty(Q)\,.$$

In conclusion, we have $f = c_{0,Q}$ a.e. in Q. Since when $Q_1 \cap Q_2 \neq \emptyset$ it follows $c_{0,Q_1} = c_{0,Q_2}$, the proof is completed by "invading" D by a sequence of open and connected sets $Q_m \subset\subset D$. □

6.8 Exercises

Exercise 6.1 Prove that $H_0^1(\mathbb{R}^n) = H^1(\mathbb{R}^n)$.

Solution We only need to show that a function $v \in H^1(\mathbb{R}^n)$ can be approximated in $H^1(\mathbb{R}^n)$ by functions belonging to $C_0^\infty(\mathbb{R}^n)$. For this aim, the keywords are: "truncate" and "mollify". In fact, adapting the proof of Theorem 6.1, one sees that the ε-mollified versions $v_\varepsilon \in C^\infty(\mathbb{R}^n)$ converge to v in $H^1(\mathbb{R}^n)$, but v_ε have not a compact support, unless v itself has a compact support.

Then let us first suppose that $v \in H^1(\mathbb{R}^n)$ and has a compact support. We take $v_\varepsilon = \eta_\varepsilon * v$, where η_ε is the Friedrichs ε-mollifier introduced in the proof of Theorem 6.1. It is known that $v_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ (here it is used that v has a compact support) and that $v_\varepsilon \to v$ in $L^2(\mathbb{R}^n)$ (here it is used that $v \in L^2(\mathbb{R}^n)$). Moreover, adapting the proof of Theorem 6.1 to the whole space $\mathbb{R}^n$, we see that $\mathcal{D}_i v_\varepsilon = (\mathcal{D}_i v)_\varepsilon$ in $\mathbb{R}^n$, thus $\mathcal{D}_i v_\varepsilon \to \mathcal{D}_i v$ in $L^2(\mathbb{R}^n)$.

Now we have to show that each function $v \in H^1(\mathbb{R}^n)$ can be approximated by a function belonging to $H^1(\mathbb{R}^n)$ with compact support. It is enough to "truncate" v out of a compact set. Precisely, we take a function $\zeta \in C_0^\infty(\mathbb{R}^n)$ such that $0 \leq \zeta(x) \leq 1$, $\zeta(x) = 1$ for $|x| \leq 1$ and $\zeta(x) = 0$ for $|x| \geq 2$ and for $t > 0$ we define $v_t(x) = v(x)\zeta(x/t)$. Clearly $v_t \in H^1(\mathbb{R}^n)$ and has a compact support. Then

$$\nabla v_t(x) = \nabla v(x)\zeta(x/t) + \frac{1}{t} v(x)\nabla\zeta(x/t)\,.$$

We have

$$\int_{\mathbb{R}^n} (v(x) - v_t(x))^2 dx = \int_{\mathbb{R}^n} v^2(x)(1 - \zeta(x/t))^2 dx \leq \int_{|x|\geq t} v^2(x) dx$$

and

$$\begin{aligned}\int_{\mathbb{R}^n} |\nabla v(x) - \nabla v_t(x)|^2 dx &= \int_{\mathbb{R}^n} \left|\nabla v(x)(1-\zeta(x/t)) - \tfrac{1}{t} v(x)\nabla\zeta(x/t)\right|^2 dx \\ &\le 2\int_{\mathbb{R}^n} |\nabla v(x)|^2 (1-\zeta(x/t))^2 dx + \tfrac{2}{t^2}\int_{\mathbb{R}^n} v^2(x)|\nabla\zeta(x/t)|^2 dx \\ &\le 2\int_{|x|\ge t} |\nabla v(x)|^2 dx + \tfrac{2M^2}{t^2}\int_{\mathbb{R}^n} v^2(x) dx\,,\end{aligned}$$

where $M = \sup_{x\in\mathbb{R}^n} |\nabla\zeta(x)|$. Taking the limit for $t \to +\infty$ we obtain the result.

Exercise 6.2 Let $1 \le p \le +\infty$ and let p' be given by $\frac{1}{p} + \frac{1}{p'} = 1$ (with $p' = +\infty$ for $p = 1$ and viceversa). If $f_k \to f$ in $L^p(D)$ and $g_k \to g$ in $L^{p'}(D)$, then $\int_D f_k g_k dx \to \int_D fg dx$.

Solution Indeed, by Hölder inequality,

$$\begin{aligned}\left|\int_D (f_k g_k - fg)dx\right| &= \left|\int_D (f_k g_k - f_k g + f_k g - fg)dx\right| \\ &\le \left|\int_D f_k(g_k - g)dx\right| + \left|\int_D g(f_k - f)dx\right| \\ &\le \|f_k\|_{L^p(D)}\|g_k - g\|_{L^{p'}(D)} + \|g\|_{L^{p'}(D)}\|f_k - f\|_{L^p(D)} \to 0\,,\end{aligned}$$

as $\|f_k\|_{L^p(D)} \to \|f\|_{L^p(D)}$ (by the triangular inequality).

Exercise 6.3

(i) Let $u \in H^1(D)$, $v \in H^1(D)$. Then $uv \in W^{1,1}(D)$ and

$$\mathcal{D}_i(uv) = (\mathcal{D}_i u)v + u(\mathcal{D}_i v)\,.$$

(ii) The same result holds for $u \in W^{1,p}(D)$, $v \in W^{1,p'}(D)$, $1 < p < +\infty$, $\frac{1}{p} + \frac{1}{p'} = 1$.

Solution

(i) The proof is similar to that of Exercise 4.4. First of all, we know that $uv \in L^1(D)$. Moreover $(\mathcal{D}_i u)v$ and $u(\mathcal{D}_i v)$ belong to $L^1(D)$, as products of functions in $L^2(D)$. Thus it is enough to prove $\mathcal{D}_i(uv) = (\mathcal{D}_i u)v + u(\mathcal{D}_i v)$. We choose $\varphi \in C_0^\infty(D)$ and we set $Q = \mathrm{supp}(\varphi)$. Then we take an open set $\widehat{Q}$ such that $Q \subset\subset \widehat{Q} \subset\subset D$. By Theorem 6.1 we find $u_k \in C^\infty(\widehat{Q})$, $v_k \in C^\infty(\widehat{Q})$ such that $u_k \to u$ in $H^1(\widehat{Q})$, $v_k \to v$ in $H^1(\widehat{Q})$. Since $\varphi \in C_0^\infty(\widehat{Q})$ we have

$$\begin{aligned}\int_{\widehat{Q}} u_k v_k \mathcal{D}_i\varphi dx &= -\int_{\widehat{Q}} \mathcal{D}_i(u_k v_k)\varphi dx \\ &= -\int_{\widehat{Q}} [(\mathcal{D}_i u_k)v_k + u_k(\mathcal{D}_i v_k)]\,\varphi dx\,.\end{aligned}$$

Taking into account Exercise 6.2, the result follows passing to the limit for $k \to \infty$, as we obtain

$$\begin{aligned}\int_D uv\mathcal{D}_i\varphi dx &= \int_{\widehat{Q}} uv\mathcal{D}_i\varphi dx \\ &= -\int_{\widehat{Q}} [(\mathcal{D}_i u)v + u(\mathcal{D}_i v)]\varphi\, dx \\ &= -\int_D [(\mathcal{D}_i u)v + u(\mathcal{D}_i v)]\varphi\, dx\,.\end{aligned}$$

(ii) The proof is the same, just noting that uv, $(\mathcal{D}_i u)v$ and $u(\mathcal{D}_i v)$ belong to $L^1(D)$, as products of functions in $L^p(D)$ and $L^{p'}(D)$, and using the approximation results given by Theorem 6.1 for functions belonging to $W^{1,p}(D)$ and $W^{1,p'}(D)$.

Exercise 6.4 Prove that if $v_k \to v$ in $H^1(D)$ then

$$\int_D v_k^2 dx \to \int_D v^2 dx \;\;,\;\; \int_D |\nabla v_k|^2 dx \to \int_D |\nabla v|^2 dx\,.$$

Solution It is enough to apply Exercise 6.2, since from $v_k \to v$ in $H^1(D)$ we have in particular that $v_k \to v$ in $L^2(D)$ and $\nabla v_k \to \nabla v$ in $L^2(D)$. An alternative proof is simply based on the triangular inequality:

$$\left|\left(\int_D v_k^2 dx\right)^{1/2} - \left(\int_D v^2 dx\right)^{1/2}\right| = \big|\|v_k\|_{L^2(D)} - \|v\|_{L^2(D)}\big| \le \|v_k - v\|_{L^2(D)} \to 0\,.$$

Similarly we prove that $\int_D |\nabla v_k|^2 dx \to \int_D |\nabla v|^2 dx$.

Exercise 6.5 Using an approach similar to the one presented in the first proof of Theorem 6.4, prove the Poincaré inequality for D bounded in one direction, with constant S^2 (S being the dimension of the strip containing D).

Solution By proceeding as in Theorem 6.4 it is enough to prove the inequality for $v \in C_0^\infty(D)$. Suppose that D is contained in the strip $\{x \in \mathbb{R}^n \,|\, |x_n - x_n^0| \le S/2\}$. Since $\mathcal{D}_n x_n = 1$, we have

$$\begin{aligned}\int_D v^2 dx &= \int_D \mathcal{D}_n(x_n - x_n^0)v^2 dx \\ &= -\int_D (x_n - x_n^0)\,\mathcal{D}_n(v^2) dx = -\int_D (x_n - x_n^0)\, 2v\mathcal{D}_n v dx \\ &\le 2\,\frac{S}{2}\left(\int_D v^2 dx\right)^{1/2}\left(\int_D |\mathcal{D}_n v|^2 dx\right)^{1/2},\end{aligned}$$

thus the Poincaré inequality holds with $C_D = S^2$.

Exercise 6.6 The Poincaré inequality still holds in $W_0^{1,p}(D)$, $1 \le p < +\infty$: there exists a constant $C_D > 0$ such that

$$\int_D |v|^p dx \le C_D \int_D |\nabla v|^p dx \quad \forall\, v \in W_0^{1,p}(D)\,.$$

Solution As in the proof of Theorem 6.4, 2nd way, we assume that $a \le x_n \le b$ and we start writing, for $v \in C_0^\infty(D)$,

$$v(x', x_n) = \int_a^{x_n} \mathcal{D}_n v(x', \xi) d\xi\,.$$

For $1 \le p < +\infty$ it follows

$$|v(x', x_n)|^p = \left|\int_a^{x_n} \mathcal{D}_n v(x', \xi) d\xi\right|^p \le \left(\int_a^{x_n} 1 \cdot |\mathcal{D}_n v(x', \xi)| d\xi\right)^p\,.$$

By Hölder inequality, for $1 < p < +\infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$ (for $p = 1$ you do not even need the Hölder inequality...) it follows

$$\begin{aligned}\left(\int_a^{x_n} 1 \cdot |\mathcal{D}_n v(x', \xi)| d\xi\right)^p &\le \left(\left(\int_a^{x_n} 1^{p'} d\xi\right)^{\frac{1}{p'}} \left(\int_a^{x_n} |\mathcal{D}_n v(x', \xi)|^p d\xi\right)^{\frac{1}{p}}\right)^p \\ &\le (x_n - a)^{p/p'} \int_a^{x_n} |\mathcal{D}_n v(x', \xi)|^p d\xi\,.\end{aligned}$$

Since $\frac{p}{p'} = p - 1$, integrating in dx' we obtain

$$\begin{aligned}\int_{\mathbb{R}^{n-1}} |v(x', x_n)|^p dx' &\le (x_n - a)^{p-1} \int_{\mathbb{R}^{n-1}} \int_a^{x_n} |\mathcal{D}_n v(x', \xi)|^p d\xi dx' \\ &\le (x_n - a)^{p-1} \int_{\mathbb{R}^n} |\mathcal{D}_n v(x', \xi)|^p d\xi dx'\,.\end{aligned}$$

Thus

$$\begin{aligned}\int_a^b \int_{\mathbb{R}^{n-1}} |v(x', x_n)|^p dx' dx_n &\le \int_a^b (x_n - a)^{p-1} \left[\int_{\mathbb{R}^n} |\mathcal{D}_n v(x', \xi)|^p d\xi dx'\right] dx_n \\ &= \frac{1}{p}(b - a)^p \int_{\mathbb{R}^n} |\mathcal{D}_n v(x)|^p dx\,.\end{aligned}$$

In conclusion, taking into account that $v = 0$ outside D,

$$\int_D |v|^p dx \le \frac{1}{p}(b - a)^p \int_D |\nabla v|^p dx\,.$$

Exercise 6.7 Let D a bounded, connected, open set with a Lipschitz continuous boundary ∂D, and take $u \in H^1(D)$, $v \in H^1(D)$. Then the integration by parts formula

$$\int_D (\mathcal{D}_i u) v dx = -\int_D u \mathcal{D}_i v dx + \int_{\partial D} n_i u_{|\partial D} v_{|\partial D} dS_x$$

holds.

Solution We proceed by approximation. We have $u_k \in C^\infty(\overline{D})$, $u_k \to u$ in $H^1(D)$, $v_k \in C^\infty(\overline{D})$, $v_k \to v$ in $H^1(D)$. Then

$$\underbrace{\int_D (\mathcal{D}_i u_k) v_k dx}_{[1]} = -\underbrace{\int_D u_k \mathcal{D}_i v_k dx}_{[2]} + \underbrace{\int_{\partial D} n_i u_{k|\partial D} v_{k|\partial D} dS_x}_{[3]} .$$

By Exercise 6.2 we have these first two results:

[1] As $k \to \infty$ we have

$$\int_D (\mathcal{D}_i u_k) v_k dx \to \int_D (\mathcal{D}_i u) v dx .$$

[2] As $k \to \infty$ we have

$$\int_D u_k \mathcal{D}_i v_k dx \to \int_D u \mathcal{D}_i v dx .$$

[3] The final step is to check that

$$\int_{\partial D} n_i u_k v_k dS_x \to \int_{\partial D} n_i u_{|\partial D} v_{|\partial D} dS_x .$$

We know that the map $v \to v_{|\partial D}$ is continuous from $H^1(D)$ to $L^2(\partial D)$, thus

$$u_{k|\partial D} \to u_{|\partial D} \quad \text{in } L^2(\partial D)$$

and

$$v_{k|\partial D} \to v_{|\partial D} \quad \text{in } L^2(\partial D) .$$

Since n is a bounded vector field,

$$n_i u_{k|\partial D} \to n_i u_{|\partial D} \quad \text{in } L^2(\partial D) ,$$

which ends the proof, applying the result of Exercise 6.2 in $L^2(\partial D)$.

Exercise 6.8 Let us assume that D is a bounded, connected, open set with a Lipschitz continuous boundary ∂D, and that $\overline{D} = \overline{D_1} \cup \overline{D_2}$, $D_1 \cap D_2 = \emptyset$, where D_1 and D_2 are (non-empty) open sets with a Lipschitz continuous boundary. Set $\Gamma = \partial D_1 \cap \partial D_2$ and take $v \in L^p(D)$, $1 \leq p < +\infty$. Then $v \in W^{1,p}(D)$ if and only if $v_{|D_1} \in W^{1,p}(D_1)$, $v_{|D_2} \in W^{1,p}(D_2)$ and the trace of $v_{|D_1}$ and $v_{|D_2}$ on Γ is the same.

Solution ($\Rightarrow$) The proof that $v_{|D_1} \in W^{1,p}(D_1)$ and $v_{|D_2} \in W^{1,p}(D_2)$ is straightforward. Then consider a sequence $v_k \in C^\infty(\overline{D})$ which converges to v in $W^{1,p}(D)$ (see Theorem 6.3); in particular, $w_{1,k} = v_{k|D_1} \in C^\infty(\overline{D_1})$ converges to $v_{|D_1}$ in $W^{1,p}(D_1)$ and $w_{2,k} = v_{k|D_2} \in C^\infty(\overline{D_2})$ converges to $v_{|D_2}$ in $W^{1,p}(D_2)$. Hence $w_{1,k|\Gamma}$ converges in $L^p(\Gamma)$ to the trace of $v_{|D_1}$ on Γ and $w_{2,k|\Gamma}$ converges in $L^p(\Gamma)$ to the trace of $v_{|D_2}$ on Γ. Since $w_{1,k|\Gamma} = w_{2,k|\Gamma}$, the thesis follows.

($\Leftarrow$) For the sake of simplicity, let us write v_1 and v_2 for $v_{|D_1}$ and $v_{|D_2}$. Take a test function $\varphi \in C_0^\infty(D)$ (and thus not necessarily vanishing on the interface Γ) and define $\omega_i \in L^p(D)$ by setting $\omega_{i|D_1} = \mathcal{D}_i v_1$ and $\omega_{i|D_2} = \mathcal{D}_i v_2$, $i = 1, \ldots, n$. We find, by integration by parts as in Exercise 6.7,

$$\begin{aligned}\int_D \omega_i \varphi dx &= \int_{D_1} \mathcal{D}_i v_1 \varphi dx + \int_{D_2} \mathcal{D}_i v_2 \varphi dx \\ &= -\int_{D_1} v_1 \mathcal{D}_i \varphi dx + \int_\Gamma n_{1,i} v_{1|\Gamma} \varphi_{|\Gamma} dS_x \\ &\quad - \int_{D_2} v_2 \mathcal{D}_i \varphi dx + \int_\Gamma n_{2,i} v_{2|\Gamma} \varphi_{|\Gamma} dS_x \,,\end{aligned}$$

where n_j is the unit normal vector on Γ directed outside D_j, $j = 1, 2$. Since $v_{1|\Gamma} = v_{2|\Gamma}$ and $n_{1,i} = -n_{2,i}$, it follows

$$\int_D \omega_i \varphi dx = -\int_{D_1} v_1 \mathcal{D}_i \varphi dx - \int_{D_2} v_2 \mathcal{D}_i \varphi dx = -\int_D v \mathcal{D}_i \varphi dx \,,$$

hence $\mathcal{D}_i v = \omega_i \in L^p(D)$.

Exercise 6.9 Let D a bounded, connected, open set with a Lipschitz continuous boundary ∂D. The statement "there exists a constant $C > 0$ such that

$$\int_{\partial D} |v|^p dS_x \leq C \int_D |v|^p dx \quad \forall\, v \in C^0(\overline{D})\text{"} \tag{6.9}$$

is *false* for $1 \leq p < +\infty$.

Solution Consider the sequence $v_k \in C^0(\overline{D})$ satisfying $0 \leq v_k(x) \leq 1$ and defined as follows:

$$v_k(x) = \begin{cases} 1 & \text{for } x \in \overline{D} \setminus D_{1/k} \\ \text{continuous} & \text{for } x \in \overline{D_{1/k}} \setminus D_{2/k} \\ 0 & \text{for } x \in \overline{D_{2/k}} \,, \end{cases}$$

where D_ϵ is as in Theorem 6.1. Then

$$\int_{\partial D} |v_k|^p dS_x = \text{meas}(\partial D) > 0$$

and

$$\int_D |v_k|^p dx \le \text{meas}(D \setminus D_{2/k}) \le C\frac{1}{k} \,,$$

thus (6.9) cannot hold.

Exercise 6.10 Let D a bounded, connected, open set with a Lipschitz continuous boundary ∂D. Let v_k be a bounded sequence in $W^{1,p}(D)$, $1 < p < +\infty$, and consider a subsequence v_{k_s} which converges to v in $L^p(D)$ by the Rellich theorem. Prove that the limit v indeed belongs to $W^{1,p}(D)$.

Solution Since $W^{1,p}(D)$ is a reflexive Banach space (see Remark 4.9), from the bounded sequence v_{k_s} we can extract a subsequence, still denoted by v_{k_s}, which converges weakly to $w \in W^{1,p}(D)$. In particular, v_{k_s} converges weakly to w in $L^p(D)$, and since it converges to v in $L^p(D)$, it follows $v = w$ by the uniqueness of the weak limit and thus $v \in W^{1,p}(D)$.

Exercise 6.11 Let $D = B_R \subset \mathbb{R}^2$ be the disc of radius R centered at 0 and consider the Hilbert space $L^2(\Delta; D) = \{v \in L^2(D) \,|\, \Delta v \in L^2(D)\}$, endowed with the natural scalar product $\int_D (w\, v + \Delta w\, \Delta v)\, dx$. Show that the immersion $L^2(\Delta; D) \hookrightarrow L^2(D)$ is not compact.

Solution In polar coordinates, for $k \ge 1$ take $v_k(r,\theta) = c_k \frac{r^k}{R^k} \sin(k\theta)$, where $c_k = \frac{\sqrt{2}}{R\sqrt{\pi}}\sqrt{k+1}$ is chosen so that $\|v_k\|_{L^2(D)} = 1$. Clearly we have $\Delta v_k = 0$ in D, therefore $\|v_k\|_{L^2(\Delta; D)} = 1$. If we had a subsequence v_{k_s} of v_k which converges to v in $L^2(D)$, then we would also have a subsequence of v_{k_s} which pointwise converges to v almost everywhere in D. Therefore we would obtain $v = 0$ almost everywhere (for $r < R$ we easily see that $v_k \to 0$ pointwise) and $\|v\|_{L^2(D)} = 1$, a contradiction.

Chapter 7
Additional Results

In this chapter a series of additional results are described and analyzed: the Fredholm alternative theory applied to second order elliptic problems; the spectral theory for an elliptic operator (in the general case and in the symmetric case); the maximum principle for weak subsolution of elliptic equations; some results concerning further regularity of weak solutions, together with higher summability or regularity results in the classical sense for functions belonging to Sobolev spaces; and finally the Galerkin approximation method.

7.1 Fredholm Alternative

We can employ the Fredholm theory for a compact perturbation of the identity operator to glean more detailed information regarding the solvability of second order elliptic PDE.

We start by briefly analyzing the finite dimensional case. Let A be a $n \times m$ matrix, associated to the linear map $v \mapsto Av$, $v \in \mathbb{R}^m$, $Av \in \mathbb{R}^n$. From linear algebra it is known that $\dim N(A) + \dim R(A) = m$, where $N(A) = \{v \in \mathbb{R}^m \mid Av = 0\}$ is the kernel of A and $R(A) = \{Av \in \mathbb{R}^n \mid v \in \mathbb{R}^m\}$ its range. Therefore, if $n = m$ it follows that $N(A) = \{0\}$ implies $R(A) = \mathbb{R}^n$ and viceversa: in other words, from uniqueness one obtains existence and viceversa.

Another interesting and well-known result is a characterization of the range of A, given by $R(A) = N(A^T)^\perp$ (see Exercise 7.2).

We want to understand if something of this type is also true in a Hilbert space V whose dimension is infinite. The answer is provided by the Fredholm alternative. Before stating the result, we need a definition.

Definition 7.1 A linear operator $K : X \mapsto Y$, X and Y Banach spaces, is said to be compact if it is bounded and it maps bounded sets into precompact sets (namely, sets whose topological closure is a compact set).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_7

The following result is the core of Fredholm theory (see, e.g., Evans [8, Theorem 5, pp. 641–643]).

Theorem 7.1 (Fredholm Alternative) *Let V be a Hilbert space and $K : V \mapsto V$ be a compact linear operator. Then:*

1. $N(I - K) = \{0\}$ *if and only if* $R(I - K) = V$;
2. $N(I - K)$ *is a finite dimensional subspace;*
3. $\dim N(I - K) = \dim N(I - K^T)$;
4. $R(I - K)$ *is closed and therefore* $R(I - K) = N(I - K^T)^\perp$ *(see Exercise 7.3).*

Let us recall that, if $A : X \mapsto Y$ is a bounded linear operator, X and Y being Hilbert spaces, its adjoint operator $A^T : Y \mapsto X$ is defined as

$$(A^T y, x)_X = (y, Ax)_Y \quad \forall\, y \in Y, x \in X\,.$$

Let us consider the elliptic operator

$$Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w) + \sum_{i=1}^{n} b_i \mathcal{D}_i w + a_0 w\,,$$

with $a_{ij} \in L^\infty(D)$ for $i, j = 1, \dots, n$, $b_i \in L^\infty(D)$ for $i = 1, \dots, n$, $a_0 \in L^\infty(D)$. The formal adjoint L^T is defined by

$$L^T w = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ji}\mathcal{D}_j w) - \sum_{i=1}^{n} \mathcal{D}_i(b_i w) + a_0 w\,.$$

The bilinear form $B_L(\cdot, \cdot)$ is defined as

$$B_L(w, v) = \int_D \sum_{i,j=1}^{n} a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx\,,$$

while the adjoint bilinear form is defined by

$$B_{L^T}(w, v) = \int_D \sum_{i,j=1}^{n} a_{ji}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i w \mathcal{D}_i v dx + \int_D a_0 w v dx\,,$$

where integration by parts has been applied not only to the second order term but also to $-\int_D \mathcal{D}_i(b_i w) v dx$. Consequently,

$$B_{L^T}(w, v) = B_L(v, w) \quad \forall\, v, w \in H^1(D)\,.$$

Let us start focusing on the homogeneous Dirichlet boundary condition. As usual, we will say that u is a weak solution of $Lu = f$ with homogeneous Dirichlet boundary value if $u \in H_0^1(D)$ is a solution of

$$B_L(u, v) = \int_D f v dx \quad \forall\, v \in H_0^1(D)\,.$$

Similarly, we will say that w is a weak solution of $L^T w = p$ with homogeneous Dirichlet boundary value if $w \in H_0^1(D)$ is a solution of

$$B_{L^T}(w, v) = \int_D p v dx \quad \forall\, v \in H_0^1(D)\,.$$

Theorem 7.2 (Existence and Uniqueness Theorem Based on Fredholm Alternative) *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set.*

(i) Precisely one of the following statements holds:

(α) either for each $f \in L^2(D)$ there exists a unique solution $u \in H_0^1(D)$ of

$$B_L(u, v) = \int_D f v dx \quad \forall\, v \in H_0^1(D)\,, \tag{7.1}$$

(β) or else there exists a solution $w \in H_0^1(D)$, $w \neq 0$, of

$$B_L(w, v) = 0 \quad \forall\, v \in H_0^1(D)\,. \tag{7.2}$$

The dichotomy (α), (β) is called the Fredholm alternative.

(ii) Furthermore, when assertion (β) holds, the dimension of $N(L)$, the space of the solutions of problem (7.2), is finite, and it is equal to the dimension of $N(L^T)$, the space of the solutions of the problem

$$B_{L^T}(w, v) = 0 \quad \forall\, v \in H_0^1(D)\,.$$

(iii) Finally, when assertion (β) holds, problem (7.1) has a solution if and only if

$$\int_D f v_* dx = 0 \quad \forall\, v_* \in N(L^T)\,.$$

Proof

(i) Choose $\tau > 0$ in a such a way that

$$B_\tau(w, v) = B_L(w, v) + \tau \int_D w v dx$$

is coercive in $H_0^1(D)$. We have seen in Sect. 5.3 that this is possible choosing

$$\tau > \max(0, -\mu)\,,$$

where $\mu = \inf_D a_0 - \frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D)}$. Then for each $q \in L^2(D)$ there exists a unique solution $u_\star \in H_0^1(D)$ of

$$B_\tau(u_\star, v) = \int_D q v dx \quad \forall\, v \in H_0^1(D)\,. \tag{7.3}$$

Let us write $u_\star = (L + \tau I)^{-1} q$ whenever (7.3) holds. Indeed (7.3) is the weak form of $Lu_\star + \tau u_\star = q$.

Now observe that $u \in H_0^1(D)$ is a solution of (7.1) if and only if

$$B_\tau(u, v) = \int_D (\tau u + f) v dx \quad \forall\, v \in H_0^1(D)\,,$$

namely, if and only if

$$u = (L + \tau I)^{-1}(\tau u + f) = \tau (L + \tau I)^{-1} u + (L + \tau I)^{-1} f\,.$$

Let us write this as

$$u - Ku = (L + \tau I)^{-1} f\,,$$

where $K = \tau(L + \tau I)^{-1}$. We have thus found that a solution $u \in H_0^1(D) \subset L^2(D)$ to (7.1) is a solution of $u - Ku = h$, with a right hand side $h = (L + \tau I)^{-1} f \in H_0^1(D) \subset L^2(D)$.

On the other hand, let us take a solution $\hat{u} \in L^2(D)$ of $\hat{u} - K\hat{u} = h$ with $h \in L^2(D)$, namely, we have

$$\hat{u} - \tau(L + \tau I)^{-1}\hat{u} = h\,.$$

If we know the additional information that $h \in H_0^1(D)$, then $\hat{u} = \tau(L + \tau I)^{-1}\hat{u} + h \in H_0^1(D)$. Moreover, we can rewrite the problem as $(L + \tau I)\hat{u} - \tau\hat{u} = (L + \tau I)h$ or simply $L\hat{u} = (L + \tau I)h$. Therefore, choosing $h = (L + \tau I)^{-1} f = \frac{1}{\tau} K f$ with $f \in L^2(D)$, the two problems $Lu = f$ with $u \in H_0^1(D)$ and $(I - K)u = h$ with $u \in L^2(D)$ are equivalent.

We claim that $K : L^2(D) \mapsto L^2(D)$ is a linear and compact operator. In fact, from the coerciveness of $B_\tau(\cdot,\cdot)$ for the solution $u_\star$ of (7.3) we have

$$\begin{aligned} \alpha\|u_\star\|^2_{H^1(D)} &\le B_\tau(u_\star, u_\star) = \textstyle\int_D q u_\star dx \\ &\le \|q\|_{L^2(D)}\|u_\star\|_{L^2(D)} \le \|q\|_{L^2(D)}\|u_\star\|_{H^1(D)}\,, \end{aligned}$$

hence, being $u_\star = (L + \tau I)^{-1} q$, $K = \tau(L + \tau I)^{-1}$ and $Kq = \tau u_\star$,

$$\|Kq\|_{H^1(D)} \le \frac{\tau}{\alpha} \|q\|_{L^2(D)} . \tag{7.4}$$

In particular we have

$$\|Kq\|_{L^2(D)} \le \|Kq\|_{H^1(D)} \le \frac{\tau}{\alpha} \|q\|_{L^2(D)} ,$$

that proves the boundedness of K. Moreover estimate (7.4) and Rellich Theorem 6.9, (that in $H_0^1(D)$ is valid without assumptions on ∂D, as we can freely use the trivial extension by 0 outside D) tell us that K is compact.

We now apply the Fredholm alternative that states that

$$N(I - K) = \{0\} \text{ if and only if } R(I - K) = L^2(D) .$$

In other words

(α) we always find $u \in L^2(D)$, solution of $u - Ku = h \in L^2(D)$, and u is unique
or
(β) $N(I - K)$ is not trivial and has finite positive dimension.

We have already seen that case (α) can be rephrased as follows: choosing $h = (L + \tau I)^{-1} f$, $f \in L^2(D)$, we always find $u \in H_0^1(D)$ solution of $Lu = f$.

In case (β) we have that there exists $w \in N(I - K)$, $w \ne 0$; this means $w = Kw$, namely,

$$w = \tau(L + \tau I)^{-1} w \iff (L + \tau I)w = \tau w \iff Lw = 0 ,$$

thus $w \in N(L)$.

(ii) In case (β) we know that $\dim N(I - K) = \dim N(I - K^T)$ and also that $\dim N(I - K)$ is finite; since we have just seen that $\dim N(I - K) = \dim N(L)$, we obtain that $\dim N(L)$ is finite. Moreover, it is easy to check that $K^T = \tau(L^T + \tau I)^{-1}$ (see Exercise 7.1). Thus, similarly to what proved for the operator L, we deduce that $v \in N(I - K^T)$ is equivalent to $v \in N(L^T)$, and consequently $\dim N(L^T) = \dim N(I - K^T) = \dim N(I - K) = \dim N(L)$.

(iii) Finally, we know that $R(I - K) = N(I - K^T)^\perp$. Thus $u - Ku = h$ has a solution if and only if $h \in N(I - K^T)^\perp$. Let us make explicit this condition: take $v_* \in N(I - K^T)$, i.e., $v_* = K^T v_*$, and remember that we are interested in solving the problem for $h = \frac{1}{\tau} K f$. Then we can solve the problem if and only if h satisfies

$$0 = \int_D h v_* dx = \int_D \frac{1}{\tau} K f \, v_* dx = \int_D \frac{1}{\tau} f K^T v_* dx = \int_D \frac{1}{\tau} f v_* dx .$$

Thus $h = \frac{1}{\tau} K f \in N(I - K^T)^\perp$ is equivalent to $f \in N(I - K^T)^\perp$, which means $f \in N(L^T)^\perp$ or, explicitly, $\int_D f v_* dx = 0$ for all $v_* \in N(L^T)$.

□

Exercise 7.1 Prove that in Theorem 7.2 one has $K^T = \tau(L^T + \tau I)^{-1}$.

Similar arguments can be used for other boundary value problems. Let us present how the result can be adapted to the Neumann problem for Laplace operator $-\Delta$. Let us restrict our attention to the homogeneous case $\nabla u \cdot n = 0$, namely, $g = 0$. The weak problem reads:

$$\text{find } u \in H^1(D) \ : \ \int_D \nabla u \cdot \nabla v dx = \int_D f v dx \quad \forall\, v \in H^1(D)\,. \tag{7.5}$$

Theorem 7.3 (Existence and Uniqueness Theory for the Neumann Problem) *Assume that $D \subset \mathbb{R}^n$ is a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. There exists a weak solution $w \in H^1(D)$, $w \neq 0$, of*

$$\int_D \nabla w \cdot \nabla v dx = 0 \quad \forall\, v \in H^1(D)\,. \tag{7.6}$$

The dimension of the space of such solutions is 1, and problem (7.5) has a solution if and only if

$$\int_D f dx = 0\,.$$

Proof We can repeat the procedure used for the homogeneous Dirichlet boundary value problem. We can introduce the operator $K = \tau(L + \tau I)^{-1}$, from $L^2(D)$ to $H^1(D)$, and prove that K is compact from $L^2(D)$ into itself (the regularity of the boundary ∂D assures that the Rellich theorem is valid in $H^1(D)$). Then Fredholm alternative can be applied, and in this case we see that there are non-trivial solutions of the homogeneous problem. In fact, a weak solution w of (7.6) must satisfy $\int_D |\nabla w|^2 dx = 0$, hence w is a constant. Note now that the bilinear form $\int_D \nabla w \cdot \nabla v dx$ is symmetric, thus the adjoint problem coincides with the given problem, and therefore the solutions of the homogeneous adjoint problem are only constants. Then from the Fredholm alternative theorem applied to this problem we have that (7.5) has a solution if and only if

$$\int_D f \omega dx = 0$$

for all the solutions ω of the homogeneous adjoint problem, thus for all the constants $\omega \in \mathbb{R}$. This is equivalent to $\int_D f dx = 0$. □

As final remark, let us note that for a weak solution w of (7.6) the conclusion $w = 0$ follows if we require $\int_D w dx = 0$; thus with this additional condition the solution of problem (7.5) is unique. We have already proved this result: if $\int_D f dx = 0$ there is a solution of (7.5) and it is unique in $H^1_*(D) = \{v \in H^1(D) \mid \int_D v dx = 0\}$ (see Sect. 5.4).

Exercise 7.2 Let A be a $n \times m$ matrix, associated to the linear map $v \mapsto Av$, $v \in \mathbb{R}^m$, $Av \in \mathbb{R}^n$. Prove that $R(A) = N(A^T)^\perp$.

Exercise 7.3 Let $A : X \mapsto Y$ be a linear and bounded operator, X and Y Hilbert spaces. Define the adjoint operator $A^T : Y \mapsto X$ as $(A^T y, x)_X = (y, Ax)_Y$ for all $y \in Y$, $x \in X$. Prove that

(i) $\overline{R(A)} = N(A^T)^\perp$
(ii) $R(A)^\perp = N(A^T)$.

7.2 Spectral Theory

Definition 7.2 Let V be a Banach space and $A : V \mapsto V$ a bounded linear operator.

(i) The resolvent set of A is

$$\rho(A) = \{\eta \in \mathbb{R} \mid A - \eta I \text{ is one-to-one and onto}\}\,.$$

(ii) The spectrum of A is

$$\sigma(A) = \mathbb{R} \setminus \rho(A)\,.$$

(iii) $\eta \in \sigma(A)$ is an eigenvalue of A if $N(A - \eta I) \neq \{0\}$.
(iv) If η is an eigenvalue of A, any $w \in V$, $w \neq 0$, satisfying

$$Aw = \eta w$$

is an associated eigenvector.

Theorem 7.4 (Spectrum of a Compact Operator) *Let V be a Hilbert space and assume that* $\dim V = +\infty$. *Let $K : V \mapsto V$ be a linear and compact operator. Then*

(i) $0 \in \sigma(K)$.
(ii) If $\eta \neq 0$ belongs to $\sigma(K)$, then η is an eigenvalue of K.
(iii) The eigenvalues $\eta \neq 0$ are either the empty set, or a finite set, or a sequence tending to 0.
(iv) If $\eta \neq 0$ is an eigenvalue, then $\dim N(K - \eta I) < +\infty$.

We now apply this general theorem to a boundary value problem. We focus on the elliptic operator L (with bounded coefficients) and the homogeneous Dirichlet boundary condition.

Theorem 7.5 *Let D be a bounded, connected and open set in $\mathbb{R}^n$. There exists an at most countable set $\Sigma \subset \mathbb{R}$ such that the problem*

$$u \in H_0^1(D) \ : \ B_L(u, v) = \lambda \int_D uvdx + \int_D fvdx \qquad \forall\, v \in H_0^1(D) \tag{7.7}$$

has a unique solution for each $f \in L^2(D)$ if and only if $\lambda \notin \Sigma$. Moreover, if Σ is infinite, then $\Sigma = \{\lambda_k\}_{k=1}^{\infty}$ with $\lambda_k \to +\infty$. In particular, λ_k can be reordered in a non-decreasing way, with $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$.

Proof Choose $\tau > 0$ in such a way that

$$B_\tau(w, v) = B_L(w, v) + \tau \int_D wvdx$$

is coercive in $H_0^1(D)$. We have seen in Sect. 5.3 that this is possible choosing

$$\tau > \max(0, -\mu)\,,$$

where $\mu = \inf_D a_0 - \frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D)}$ (let us note that there we wrote σ instead of τ, but now σ denotes the spectrum...).

For $\lambda = -\tau$ we know that (7.7) has a unique solution, as B_τ is coercive. Thus let us assume from now on that $\lambda \neq -\tau$. According to the Fredholm alternative (applied to the bilinear form $B_L(w, v) - \lambda \int_D wv \dots$), we know that problem (7.7) has a unique solution for each $f \in L^2(D)$ if and only if the only solution of

$$B_L(u, v) = \lambda \int_D uvdx \qquad \forall\, v \in H_0^1(D)$$

is $u = 0$ (see Theorem 7.2). This means that $u = 0$ is the only solution to

$$B_L(u, v) + \tau \int_D uvdx = (\tau + \lambda) \int_D uvdx \qquad \forall\, v \in H_0^1(D)\,.$$

We can rewrite this relation as

$$u = (L + \tau I)^{-1}(\tau + \lambda)u = \frac{\tau + \lambda}{\tau} Ku\,,$$

having set $K = \tau(L + \tau I)^{-1}$. We have already proved that $K : L^2(D) \mapsto L^2(D)$ is linear and compact. Thus its spectrum is given by 0 (in fact K is not onto form $L^2(D)$ to $L^2(D)$: $H_0^1(D)$ is a subspace of $L^2(D)$, strictly contained in it) and by eigenvalues. Thus $u = 0$ is the only solution to $Ku = \frac{\tau}{\tau+\lambda}u$ if and only if $\frac{\tau}{\tau+\lambda}$ is not an eigenvalue of K. The eigenvalues $\eta_k \neq 0$ of K are either the empty set, or a finite set or a sequence convergent to 0. In the last case, from

$$\eta_k = \frac{\tau}{\tau + \lambda_k}$$

we get

$$\lambda_k = \tau \frac{1-\eta_k}{\eta_k}\,.$$

We want now to show that $\eta_k > 0$. Being η_k an eigenvalue of K, we have

$$Kw_k = \eta_k w_k\ ,\ \ w_k \neq 0\,,$$

which is equivalent to

$$\tau(L+\tau I)^{-1}w_k = \eta_k w_k \Longleftrightarrow \tau w_k = \eta_k(L+\tau I)w_k \Longleftrightarrow (L+\tau I)w_k = \frac{\tau}{\eta_k}w_k\,.$$

This means

$$B_\tau(w_k, v) = \frac{\tau}{\eta_k}\int_D w_k v dx$$

and from the coerciveness of $B_\tau(\cdot\ ,\ \cdot)$ we get

$$\alpha\|w_k\|^2_{H^1(D)} \leq B_\tau(w_k, w_k) = \frac{\tau}{\eta_k}\int_D w_k^2 dx \leq \frac{\tau}{\eta_k}\|w_k\|^2_{H^1(D)}\,.$$

Thus $\frac{\tau}{\eta_k} \geq \alpha > 0$, hence $\eta_k > 0$ (and consequently $\lambda_k > -\tau$). In conclusion

$$\eta_k \to 0^+ \quad \text{and} \quad \lambda_k = \tau\frac{1-\eta_k}{\eta_k} \to +\infty\,,$$

which is the stated result. □

Exercise 7.4 Under the assumptions of Theorem 7.5, take $\lambda \notin \Sigma$ and for each $f \in L^2(D)$ let $u \in H_0^1(D)$ be the unique solution of (7.7). Prove that the solution operator $S_\lambda : f \mapsto u$ is a bounded operator in $L^2(D)$, namely, there exists a constant $C > 0$ such that

$$\|u\|_{L^2(D)} \leq C\|f\|_{L^2(D)}\,.$$

Exercise 7.5 Under the assumptions of Theorem 7.5, take $\lambda \notin \Sigma$ and for each $f \in L^2(D)$ let $u \in H_0^1(D)$ be the unique solution of (7.7). Prove that the solution operator $S_\lambda : f \mapsto u$ is a bounded operator from $L^2(D)$ to $H_0^1(D)$, namely, there exists a constant $C > 0$ such that

$$\|u\|_{H^1(D)} \leq C\|f\|_{L^2(D)} .$$

Another important result is the following.

Theorem 7.6 (Spectrum of a Compact and Self-Adjoint Operator) *Let V be a separable Hilbert space and let $K : V \mapsto V$ be a linear, compact and self-adjoint operator. Then there exists an (at most) countable orthonormal Hilbertian basis of V consisting of eigenvectors of K; in particular, if* $\dim V = +\infty$ *the eigenvectors of K are an infinite sequence, and if moreover* $\dim N(K) < +\infty$ *the eigenvalues of K are an infinite sequence.*

As a consequence, it holds:

Theorem 7.7 (Spectrum of a Symmetric Elliptic Operator) *Let D be a bounded, connected and open subset of $\mathbb{R}^n$. Let the coefficients of the operator L be bounded and satisfy $a_{ij} = a_{ji}$ for $i, j = 1, \ldots, n$, $b_i = 0$ for $i = 1, \ldots, n$. Then there exist an infinite sequence $\{\lambda_k\}_{k=1}^\infty$ of eigenvalues of L and a countable $L^2(D)$-orthonormal Hilbertian basis $\{w_k\}_{k=1}^\infty$ given by eigenvectors of L with homogeneous Dirichlet boundary condition, namely, solutions $w_k \in H_0^1(D)$ of*

$$B_L(w_k, v) = \lambda_k \int_D w_k v dx \qquad \forall\, v \in H_0^1(D) .$$

The eigenvectors

$$\omega_k = \frac{w_k}{\sqrt{\lambda_k + \tau}}$$

are an orthonormal Hilbertian basis of $H_0^1(D)$ with respect to the scalar product given by

$$B_\tau(w, v) = B_L(w, v) + \tau \int_D w v dx ,$$

where $\tau \geq 0$ is such that $B_\tau(w, v)$ is coercive in $H_0^1(D)$.

Proof Let us first consider the case $\tau > 0$ (namely, $B_L(\cdot, \cdot)$ is not coercive $H_0^1(D)$). We know that $L^2(D)$ is a separable Hilbert space; furthermore, we have already seen that the operator $K = \tau(L + \tau I)^{-1}$ is compact in $L^2(D)$, whose dimension is infinite, and we trivially see that $N(K) = \{0\}$. Moreover, from $a_{ij} = a_{ji}$ and $b_i = 0$

we see that K is also self-adjoint. Indeed we have that

$$L v = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j v) + \sum_{i=1}^{n} b_i \mathcal{D}_i v + a_0 v$$

$$L^T v = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ji}\mathcal{D}_j v) - \sum_{i=1}^{n} \mathcal{D}_i(b_i v) + a_0 v$$

and so $L = L^T$. Thus there exists a sequence of eigenvalues η_k and eigenfunctions w_k of K such that w_k are an orthonormal Hilbertian basis in $L^2(D)$. Let us see what is the meaning of this statement. We have $w_k \in L^2(D)$, $w_k \neq 0$, such that $K w_k = \eta_k w_k$; this is equivalent to

$$\tau(L + \tau I)^{-1} w_k = \eta_k w_k \iff \tau w_k = \eta_k (L + \tau I) w_k$$
$$\iff (L + \tau I) w_k = \frac{\tau}{\eta_k} w_k \iff L w_k = \tau \frac{1 - \eta_k}{\eta_k} w_k \, ,$$

thus w_k are the eigenvectors of L corresponding to the eigenvalues $\lambda_k = \tau \frac{1-\eta_k}{\eta_k} > -\tau$. Coming back to the bilinear forms, we see that

$$B_\tau(w_k, v) = B_L(w_k, v) + \tau \int_D w_k v dx = (\lambda_k + \tau) \int_D w_k v dx \quad \forall\, v \in H_0^1(D) \, .$$

Thus

$$B_\tau(w_k, w_j) = (\lambda_k + \tau) \int_D w_k w_j dx = (\lambda_k + \tau)\delta_{kj} \, .$$

In conclusion,

$$\omega_k = \frac{w_k}{\sqrt{\lambda_k + \tau}}$$

is an orthonormal system with respect to the scalar product $B_\tau(\cdot\,,\cdot)$ in $H_0^1(D)$. For verifying that it is a Hilbertian basis, it is sufficient to see that if $v \in H_0^1(D)$ satisfies $B_\tau(v, \omega_k) = 0$ for every $k \geq 1$, then it follows $v = 0$. This is true as

$$0 = B_\tau(v, \omega_k) = B_L(v, \omega_k) + \tau \int_D v\omega_k dx = (\lambda_k + \tau) \int_D v\omega_k dx \, ,$$

thus $\int_D v w_k dx = 0$ for every $k \geq 1$. Since w_k is an orthonormal Hilbertian basis in $L^2(D)$ it follows that $v = 0$.

The proof for the case $\tau = 0$ (namely, $B_L(\cdot, \cdot)$ is coercive $H_0^1(D)$) is essentially the same, just replacing the compact operator K by L^{-1}. This leads to a sequence of eigenvalues $\lambda_k = \frac{1}{\eta_k} > 0$, η_k being the eigenvalues of L^{-1}, of eigenvectors w_k orthonormal in $L^2(D)$, and of eigenvectors $\omega_k = \frac{w_k}{\sqrt{\lambda_k}}$ orthonormal in $H^1_0(D)$ with respect to the scalar product given by $B_L(w, v)$. □

Exercise 7.6 Prove that the minimum eigenvalue λ_1 of the Laplace operator $-\Delta$ associated to the homogeneous Dirichlet boundary condition is equal to $\frac{1}{C_D}$, where

$$C_D = \sup_{v \in H^1_0(D), v \neq 0} \frac{\int_D v^2 dx}{\int_D |\nabla v|^2 dx}$$

is the "best" Poincaré constant (see Sect. 6.2).

Exercise 7.7

(i) Consider the elliptic operator

$$Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w) + a_0 w \,,$$

with $a_{ij} = a_{ji}$ and $a_0 \geq 0$. If $\lambda_\star$ is an eigenvalue of L associated to anyone of the boundary conditions of Dirichlet, Neumann, mixed or Robin type, then $\lambda_\star \geq 0$.

(ii) The case $\lambda_\star = 0$ is possible if and only if the boundary condition is of Neumann type and $a_0 = 0$. In that case the corresponding eigenvector $w_\star$ is a constant (different from 0).

7.3 Maximum Principle

A peculiar property of a solution of an elliptic boundary value problem is that, under suitable assumptions, its values on the boundary ∂D are a bound for its values in the interior D. Just to propose a simple physical example, one can think to an elastic membrane fixed on the boundary: looking for the position u in the vertical direction, the simplest model is given by the solution of the Poisson equation $-\Delta u = f$, where f is the external force. When the membrane is charged by a load (thus $f \leq 0$), the values of u on the boundary are higher than its values inside (or viceversa, if you pushes it from below, with $f \geq 0$).

We start by underlying a clear fact: for a function $v \in H^1(D)$ the meaning of $v \geq 0$ on ∂D is that its trace $v_{|\partial D} \in L^2(\partial D)$ satisfies $v_{|\partial D} \geq 0$ (since we are considering the trace $v_{|\partial D}$, we have to assume that the boundary ∂D is Lipschitz continuous).

Another remark is that it is possible to see that $v^+ = \max(v, 0)$ and $v^- = \max(-v, 0)$ belong to $H^1(D)$ for $v \in H^1(D)$ (see Exercise 7.8); moreover, $v \leq v^+$

and $v \geq -v^-$ a.e. in D. A consequence is the fact that the statement $v \geq 0$ on ∂D can to be interpreted as $v^- \in H_0^1(D)$; similarly $v \leq 0$ on ∂D means $v^+ \in H_0^1(D)$.

Exercise 7.8 Let $D \subset \mathbb{R}^n$ an open set. Prove that $v^+ = \max(v, 0)$ and $v^- = \max(-v, 0)$ belong to $W^{1,p}(D)$ for $v \in W^{1,p}(D)$, $1 \leq p \leq +\infty$. More precisely, defining

$$w_i^+ = \begin{cases} \mathcal{D}_i v & \text{where } v > 0 \\ 0 & \text{where } v \leq 0 \end{cases}, \quad w_i^- = \begin{cases} -\mathcal{D}_i v & \text{where } v < 0 \\ 0 & \text{where } v \geq 0 \end{cases},$$

one has $\mathcal{D}_i v^+ = w_i^+$ and $\mathcal{D}_i v^- = w_i^-$, $i = 1, \ldots, n$.

We need now a definition. We say that $u \in H^1(D)$ satisfies $Lu \leq 0$ on D if

$$B_L(u, v) \leq 0 \qquad \forall\, v \in H_0^1(D)\,,\ v \geq 0 \text{ a.e. in } D\,.$$

Definition 7.3 If $u \in H^1(D)$ satisfies $Lu \leq 0$ in D, then it is called subsolution of L. A function $u \in H^1(D)$ is called supersolution of L if $-u$ is a subsolution of L (namely, if $B_L(u, v) \geq 0$ for all $v \in H_0^1(D)$, $v \geq 0$ a.e. in D).

Theorem 7.8 *Let $D \subset \mathbb{R}^n$ be a bounded, connected and open set with a Lipschitz continuous boundary ∂D. Let L be the elliptic operator*

$$Lv = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j v) + \sum_{i=1}^{n} b_i \mathcal{D}_i v + a_0 v\,,$$

with bounded coefficients a_{ij}, b_i and a_0. Assume that $a_0 \geq 0$ a.e. in D. Then:

(i) if u is a subsolution of L we have

$$\sup_D u \leq \sup_{\partial D} u^+\,;$$

in particular, if $u \leq 0$ on ∂D (thus $u^+ \in H_0^1(D)$) it follows $u \leq 0$ a.e. in D;

(ii) if u is a supersolution of L we have

$$\inf_D u \geq \inf_{\partial D}(-u^-)\,;$$

in particular, if $u \geq 0$ on ∂D (thus $u^- \in H_0^1(D)$) it follows $u \geq 0$ a.e. in D.

Proof Let us give the proof under the assumption that the weak divergence $\operatorname{div} b$ exists and satisfies $\operatorname{div} b \leq 0$ a.e. in D. The proof for the general case can be found in Gilbarg and Trudinger [11, Theorem 8.1, p. 168].

(i) Let u be a subsolution of L. Then

$$\int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx \le -\int_D \sum_{i=1}^n b_i \mathcal{D}_i u\, v dx - \int_D a_0 u v dx$$

for each $v \in H_0^1(D)$, $v \ge 0$ a.e. in D. Set $M = \sup_{\partial D} u^+$, which clearly is ≥ 0 (this is an important point in the proof). We can suppose $M < +\infty$, otherwise we would have $\sup_{\partial D} u^+ = +\infty$ and nothing has to be proved. Take $v = \max(u - M, 0)$; clearly $v \ge 0$ a.e. in D and from $u \le M$ on ∂D we have $v \in H_0^1(D)$. Moreover, note that in the set $\{u > M\}$ we have $v = u - M$, thus $\nabla v = \nabla u$; instead, where $\{u \le M\}$ one has $v = 0$ and $\nabla v = 0$. Then we have

$$\begin{aligned}\int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx &= \int_{\{u>M\}} \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx \\ &\quad + \int_{\{u\le M\}} \sum_{i,j=1}^n a_{ij}\mathcal{D}_j u \mathcal{D}_i v dx \\ &= \int_{\{u>M\}} \sum_{i,j=1}^n a_{ij}\mathcal{D}_j v \mathcal{D}_i v dx \\ &= \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j v \mathcal{D}_i v dx \\ &\ge \alpha_0 \int_D |\nabla v|^2 dx\,,\end{aligned}$$

where $\alpha_0 > 0$ is the ellipticity constant. Moreover

$$\begin{aligned}-\int_D \sum_{i=1}^n b_i \mathcal{D}_i u\, v dx &= -\int_{\{u>M\}} \sum_{i=1}^n b_i \mathcal{D}_i v\, v dx = -\int_D \sum_{i=1}^n \frac{1}{2} b_i\, \mathcal{D}_i(v^2) dx \\ &= \int_D \frac{1}{2} \underbrace{\mathrm{div}\, b}_{\le 0} v^2 dx \le 0\end{aligned}$$

and

$$\begin{aligned}-\int_D a_0 u v dx &= -\int_{\{u>M\}} a_0 u v dx - \int_{\{u\le M\}} a_0 u v dx = -\int_{\{u>M\}} a_0 u v dx \\ &= -\int_{\{u>M\}} \underbrace{a_0}_{\ge 0} \underbrace{u}_{\ge M \ge 0} \underbrace{(u-M)}_{\ge 0} dx \le 0\,.\end{aligned}$$

Thus

$$\int_D |\nabla v|^2 dx \leq 0\,,$$

hence $\nabla v = 0$ in D. Since $v \in H^1_0(D)$, it follows $v = 0$ a.e. in D, hence $u \leq M$ a.e. in D.

(ii) The proof in the case of u supersolution comes from the fact that $-u$ is a subsolution and $(-u)^+ = u^-$.

□

Exercise 7.9 Prove that

$$\sup_{\partial D} u^+ = \max(\sup_{\partial D} u, 0) \quad \text{and} \quad \inf_{\partial D}(-u^-) = \min(\inf_{\partial D} u, 0)$$

(so that the conclusion of Theorem 7.8 can be written as $\sup_D u \leq \max(\sup_{\partial D} u, 0)$ for a subsolution and $\inf_D u \geq \min(\inf_{\partial D} u, 0)$ for a supersolution).

Remark 7.1 Note that in the Theorem 7.8 we cannot substitute $\sup_{\partial D} u^+$ with $\sup_{\partial D} u$ or $\inf_{\partial D} u^+$ with $\inf_{\partial D} u$. The following example can clarify the point: consider the one dimensional elliptic problem

$$\begin{cases} -u'' + u = 0 \\ u(-1) = 1\,,\ u(1) = 1\,. \end{cases} \tag{7.8}$$

To find the solution, consider the associated polynomial $-r^2 + 1$, whose roots are $r = 1$, $r = -1$. The general solution of $-u'' + u = 0$ is thus given by

$$u(x) = c_1 e^x + c_2 e^{-x}\,.$$

Imposing the boundary conditions, it follows

$$c_1 e^{-1} + c_2 e = 1 \;\;,\;\; c_1 e + c_2 e^{-1} = 1\,,$$

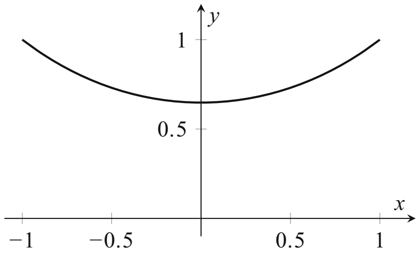
thus $c_1 = c_2 = \frac{1}{e+e^{-1}}$, and we finally obtain

$$u(x) = \frac{1}{e + e^{-1}}(e^x + e^{-x})$$

(see Fig. 7.1).

Taking the derivative we see that $u'(x) = \frac{1}{e+e^{-1}}(e^x - e^{-x})$, which satisfies $u' > 0$ for $x > 0$ and $u' < 0$ for $x < 0$, therefore u has its minimum for $x = 0$ (as it is also clear from Fig. 7.1). This minimum value is $\frac{2}{e+e^{-1}}$, which is larger than 0 and

Fig. 7.1 The graph of the solution $u(x) = \frac{1}{e+e^{-1}}(e^x + e^{-x})$ of problem (7.8)



smaller than 1. Thus

$$\inf_{(-1,1)} u = \frac{2}{e + e^{-1}} < 1 = \inf_{\partial(-1,1)} u \,,$$

but, as the theorem says,

$$\inf_{(-1,1)} u = \frac{2}{e + e^{-1}} > 0 = \inf_{\partial(-1,1)} (-u^-) \,.$$

One can revisit this example noting that the solution u satisfies $u \geq 0$. Therefore $-u'' = -u \leq 0$, and u is a subsolution of the elliptic operator $Lv = -v''$. Therefore the theorem assures that the (positive) maximum is on the boundary, as it is reasonable for a charged elastic membrane.

Remark 7.2 Instead, if $a_0 = 0$ we can substitute $\sup_{\partial D} u^+$ with $\sup_{\partial D} u$ and $\inf_{\partial D} u^+$ with $\inf_{\partial D} u$. In fact, in this case one can repeat the same proof (again, for simplicity, with $\operatorname{div} b \leq 0$), but now setting $M = \sup_{\partial D} u$ (which is no longer assured to be non-negative). Choosing $v = \max(u - M, 0)$, the assumptions that u is a subsolution, that $\operatorname{div} b \leq 0$ and that $a_0 = 0$ still yield

$$\int_D \sum_{i,j=1}^n a_{ij} \mathcal{D}_j u \mathcal{D}_i v dx \leq 0 \,,$$

and everything goes on as in the previous case.

An interesting consequence is the following result.

Theorem 7.9 (Existence Theorem via Fredholm Alternative) *Let $D \subset \mathbb{R}^n$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Let L be an elliptic operator with bounded coefficients a_{ij}, b_i, a_0. Assume that $a_0 \geq 0$ a.e. in D. Then there exists a unique solution $u \in H_0^1(D)$ of the homogeneous Dirichlet boundary value problem*

$$B_L(u, v) = \int_D f v dx \qquad \forall\, v \in H_0^1(D) \,.$$

Proof Let $w \in H_0^1(D)$ be a solution with $f = 0$. Then it is both a subsolution and a supersolution, thus

$$0 = \inf_{\partial D}(-w^-) \leq \inf_D w \leq \sup_D w \leq \sup_{\partial D} w^+ = 0\,,$$

hence $w = 0$ in D. Thus the thesis follows from the Fredholm alternative, see Theorem 7.2. □

Remark 7.3 The existence and uniqueness of a solution for the homogeneous Dirichlet boundary value problem has been proved, via coerciveness, if $b \in W^{1,\infty}(D)$ and $a_0 - \frac{1}{2}\mathrm{div}b \geq -\nu$, with $\nu > 0$ and small enough (precisely, such that $\alpha_0 - 2C_D\nu > 0$, with $\alpha_0 > 0$ the ellipticity constant and $C_D > 0$ the Poincaré constant; see Exercise 5.1). Therefore the two results are not comparable. In one case b is only assumed to be bounded, but one needs $a_0 \geq 0$ in D. In the other case b is assumed to belong to $W^{1,\infty}(D)$ and to satisfy $\mathrm{div}\, b \leq 2(a_0 + \nu)$, but no assumption on the sign of a_0 in D is required.

7.4 Regularity Issues and Sobolev Embedding Theorems

7.4.1 Regularity Issues

Let us look back at the existence theorems for the four boundary value problems we have considered. In all cases, we have found a weak solution $u \in V$ of

$$B(u, v) = \int_D f v dx \quad \forall\, v \in V\,,$$

where V is a infinite dimensional, closed subspace of $H^1(D)$.

Since this is the weak form of the second order elliptic equation

$$Lu = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j u) + \sum_{i=1}^{n} b_i \mathcal{D}_i u + a_0 u = f\,,$$

and the right hand side f belongs to $L^2(D)$, we could expect $u \in H^2(D)$.

Let us show with a formal example that this is reasonable. Suppose that u is a solution to $-\Delta u = f$ in D, and assume that $u \in C_0^\infty(D)$. Then we have

$$\int_D (-\Delta u)^2 dx = \int_D f^2 dx\,.$$

Integrating by parts we obtain

$$
\begin{aligned}
\int_D f^2 dx = \int_D (-\Delta u)^2 dx &= \int_D \sum_{i,j}^{n} \mathcal{D}_i\mathcal{D}_i u \mathcal{D}_j\mathcal{D}_j u dx \\
&= -\int_D \sum_{i,j=1}^{n} \overbrace{\mathcal{D}_j\mathcal{D}_i}^{=\mathcal{D}_i\mathcal{D}_j} \mathcal{D}_i u \mathcal{D}_j u dx = \int_D \sum_{i,j=1}^{n} \mathcal{D}_j\mathcal{D}_i u \mathcal{D}_i\mathcal{D}_j u dx \\
&= \sum_{i,j=1}^{n} \int_D (\mathcal{D}_i\mathcal{D}_j u)^2 dx \geq \int_D (\mathcal{D}_k\mathcal{D}_l u)^2 dx \,,
\end{aligned}
\tag{7.9}
$$

for any fixed couple of indices $k, l = 1, \ldots, n$. Hence the $L^2(D)$-norm of all the second order derivatives is bounded by the $L^2(D)$-norm of the right-hand side f.

For a general operator L it is necessary to take into account the regularity of the coefficients. Rewriting the second order term we have

$$
-\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j u) = -\sum_{i,j=1}^{n} a_{ij}\mathcal{D}_i\mathcal{D}_j u - \sum_{i,j=1}^{n} (\mathcal{D}_i a_{ij})\mathcal{D}_j u \,,
$$

thus

$$
-\sum_{i,j=1}^{n} a_{ij}\mathcal{D}_i\mathcal{D}_j u = \sum_{i,j=1}^{n} (\mathcal{D}_i a_{ij})\mathcal{D}_j u - \sum_{i=1}^{n} b_i\mathcal{D}_i u - a_0 u + f \,. \tag{7.10}
$$

Already knowing that $u \in H^1(D)$, this suggests that we have to assume

$$
a_{ij} \in C^1(\overline{D}) \text{ for } i, j = 1, \ldots, n
$$

(or simply $a_{ij} \in W^{1,\infty}(D)$). With this choice the right-hand side in (7.10) belongs to $L^2(D)$, because only products between $L^\infty(D)$-functions and $L^2(D)$-functions appear.

Theorem 7.10 (Interior Regularity) *Assume that $D \subset \mathbb{R}^n$ is a bounded, connected and open set. Let $u \in H^1(D)$ be a weak solution of $Lu = f$ in D, with $f \in L^2(D)$. Assume that $a_{ij} \in C^1(D)$, $b_i \in L^\infty(D)$, $a_0 \in L^\infty(D)$ for $i, j = 1, \ldots, n$. Then $u \in H^2_{loc}(D)$ and for each subset $Q \subset\subset D$ it holds*

$$
\|u\|_{H^2(Q)} \leq C(\|f\|_{L^2(D)} + \|u\|_{L^2(D)}) \,,
$$

where the constant $C > 0$ only depends on D, Q and a_{ij}, b_i, a_0.

Proof We only give a brief description of the ideas. There are some steps:

1. To localize the problem into Q use a cut-off function ζ, namely, a C^∞-function with $\zeta(x) = 1$ in Q, $\zeta(x) = 0$ on $\mathbb{R}^n \setminus T$, $0 \le \zeta(x) \le 1$ in D, where $Q \subset\subset T \subset\subset D$ (see Corollary A.1).
2. For $w \in L^2(D)$, $k = 1, \dots, n$ and $h \neq 0$ consider the difference quotients

$$\mathcal{D}_k^h w(x) = \frac{w(x + he_k) - w(x)}{h}, \tag{7.11}$$

   defined in $Q \subset\subset D$ for $0 < |h| < \mathrm{dist}(Q, \partial D)$.
3. Take as test function in the weak formulation

$$v = -\mathcal{D}_k^{-h}(\zeta^2 \mathcal{D}_k^h u)$$

   and proceed to estimate all the terms.
4. This leads to the estimate

$$\|u\|_{H^2(Q)} \le C(\|f\|_{L^2(D)} + \|u\|_{H^1(T)})\,;$$

   with a similar procedure one finds

$$\|u\|_{H^1(T)} \le C(\|f\|_{L^2(D)} + \|u\|_{L^2(D)})\,,$$

   thus the stated result.

Two important properties of difference quotients are used: see Exercises 7.10 and 7.11. □

Exercise 7.10 Take $v \in L^2(D)$, $\varphi \in L^2(D)$ with $\Phi = \mathrm{supp}\,\varphi \subset D$, and consider the difference quotients defined in (7.11). Then we have the integration by parts formula

$$\int_D v\, \mathcal{D}_k^h \varphi dx = -\int_D \mathcal{D}_k^{-h} v\, \varphi dx\,,$$

for all h with $0 < |h| < \mathrm{dist}(\Phi, \partial D)$, $k = 1, \dots, n$.

Exercise 7.11

(i) Take $v \in H^1(D)$ and consider $Q \subset\subset D$. Then the difference quotient $\mathcal{D}^h v = (\mathcal{D}_1^h v, \dots, \mathcal{D}_n^h v)$ defined in (7.11) satisfies

$$\|\mathcal{D}^h v\|_{L^2(Q)} \le \|\nabla v\|_{L^2(D)}$$

for each h with $0 < |h| < \mathrm{dist}(Q, \partial D)$.

(ii) Take k with $1 \leq k \leq n$, $v \in L^2(D)$ and $Q \subset\subset D$. Suppose that there exists a constant $C_* > 0$ such that

$$\|\mathcal{D}_k^h v\|_{L^2(Q)} \leq C_*$$

for each h with $0 < |h| < \mathrm{dist}(Q, \partial D)$. Then $\mathcal{D}_k v \in L^2(Q)$.

(iii) Take k with $1 \leq k \leq n$, $v \in L^2(D)$ and suppose there exists a constant $C_\sharp > 0$ such that

$$\|\mathcal{D}_k^h v\|_{L^2(D_{|h|})} \leq C_\sharp$$

for each $h \neq 0$, where $D_{|h|} = \{x \in D \,|\, \mathrm{dist}(x, \partial D) > |h|\}$. Then $\mathcal{D}_k v \in L^2(D)$ and $\|\mathcal{D}_k v\|_{L^2(D)} \leq C_\sharp$.

An inductive argument gives:

Theorem 7.11 (Higher Interior Regularity) *Assume that $D \subset \mathbb{R}^n$ is a bounded, connected and open set. Let $u \in H^1(D)$ be a weak solution of $Lu = f$ in D, with $f \in H^m(D)$, $m \geq 1$. Assume that $a_{ij} \in C^{m+1}(D)$, $b_i \in C^m(D)$, $a_0 \in C^m(D)$ for $i, j = 1, \ldots, n$. Then $u \in H^{m+2}_{loc}(D)$, and for each $Q \subset\subset D$ we have the estimate*

$$\|u\|_{H^{m+2}(Q)} \leq C \left(\|f\|_{H^m(D)} + \|u\|_{L^2(D)}\right) ,$$

where the constant $C > 0$ only depends on m, D, Q and a_{ij}, b_i, a_0.

These regularity results can be extended up to the boundary ∂D. For simplicity, let us focus on the homogeneous Dirichlet boundary value problem; however, the results are also true for the homogeneous Neumann and Robin problems.

Theorem 7.12 (Regularity Up to the Boundary) *Let the assumptions of the interior regularity Theorem 7.10 be satisfied. Assume moreover that $a_{ij} \in C^1(\overline{D})$ and that ∂D is of class C^2. Assume that $u \in H^1_0(D)$ is a weak solution of $Lu = f$, $u_{|\partial D} = 0$. Then $u \in H^2(D)$ and it holds*

$$\|u\|_{H^2(D)} \leq C \left(\|f\|_{L^2(D)} + \|u\|_{L^2(D)}\right) ,$$

where the constant $C > 0$ only depends on D and a_{ij}, b_i, a_0.

Proof As for the interior regularity result, there are some steps.

1. Since it is assumed that $u \in H^1_0(D)$, the lower order terms of the operator L belong to $L^2(D)$ and thus can be put at the right hand side, focusing only on the principal part of L.
2. Reduce the problem to a flat boundary by local charts (here the fact that the boundary ∂D is of class C^2 is used). Note that the transformed differential operator remains uniformly elliptic.

3. To localize the problem into $B_{R,+} = \{x \in \mathbb{R}^n \mid |x| < R, x_n > 0\}$ use a cut-off function $\zeta \in C_0^\infty(B_R)$, namely, a function with $\zeta(x) = 1$ in B_r, $\zeta(x) = 0$ on $\mathbb{R}^n \setminus B_\rho$, $0 \leq \zeta(x) \leq 1$ in $\mathbb{R}^n$, where $B_r \subset\subset B_\rho \subset\subset B_R$ (see Corollary A.1).
4. Rewrite the elliptic problem in the half-ball $B_{R,+}$ and use as test function v the difference quotient

$$v = -\mathcal{D}_k^{-h}(\zeta^2 \mathcal{D}_k^h u) \quad , \quad k = 1, \ldots, n-1 \, ,$$

namely, only acting in the directions tangential to the boundary $\{x_n = 0\}$ (this will give a control on all the second order derivatives in which at least one is tangential).
5. Use the ellipticity of the transformed operator for estimating the second order normal derivative $\mathcal{D}_n \mathcal{D}_n u$ in terms of the other derivatives (see also Exercise 7.12).
6. Use the fact that $\zeta = 1$ in B_r and put together all the estimates.
7. This gives the estimate

$$\|u\|_{H^2(D)} \leq C \left(\|f\|_{L^2(D)} + \|u\|_{H^1(D)} \right) \, .$$

By using the weak coerciveness of the bilinear form $B_L(\cdot, \cdot)$ (see Sect. 5.3) it follows at once

$$\|u\|_{H^1(D)} \leq C \left(\|f\|_{L^2(D)} + \|u\|_{L^2(D)} \right) \, ,$$

thus the stated result.

□

Exercise 7.12 Prove that all the terms $a_{ii}(x)$ on the diagonal of a uniformly positive definite matrix in D (namely, a matrix $\{a_{ij}(x)\}$ such that $\sum_{ij} a_{ij}(x)\eta_j\eta_i \geq \alpha_0|\eta|^2$ for all $\eta \in \mathbb{R}^n$ and almost every $x \in D$) satisfy $a_{ii}(x) \geq \alpha_0$ for almost every in $x \in D$.

Exercise 7.13 Under the assumptions of Theorem 7.12, the stronger estimate

$$\|u\|_{H^2(D)} \leq C\|f\|_{L^2(D)}$$

holds, provided that we know that for each $f \in L^2(D)$ there exists a unique weak solution $u \in H_0^1(D)$.

By induction, we obtain:

Theorem 7.13 (Higher Regularity Up to the Boundary) *Let the assumption of Theorem 7.11 be satisfied. Assume moreover that $a_{ij} \in C^{m+1}(\overline{D})$, $b_i \in C^m(\overline{D})$,*

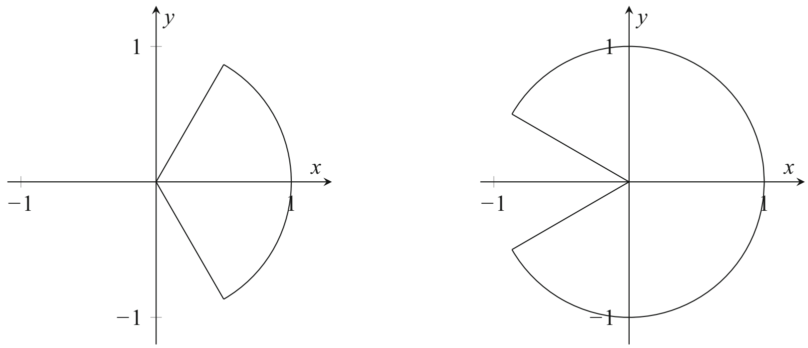


Fig. 7.2 The sectors S_α for $\alpha = \frac{2\pi}{3}$ (left) and $\alpha = \frac{5\pi}{3}$ (right)

$a_0 \in C^m(\overline{D})$ for $i,\ j = 1, \ldots, n$ and that ∂D is of class C^{m+2}. Assume that $u \in H_0^1(D)$ is a weak solution of $Lu = f$, $u_{|\partial D} = 0$. Then $u \in H^{m+2}(D)$ and it holds

$$\|u\|_{H^{m+2}(D)} \leq C \left(\|f\|_{H^m(D)} + \|u\|_{L^2(D)} \right) ,$$

where the constant $C > 0$ only depends on m, D and a_{ij}, b_i, a_0.

Remark 7.4 Similar results hold for the Neumann and Robin problems, having assumed a boundary datum $g = 0$. In the case $g \neq 0$ the trace theory for the derivatives of u and for higher order Sobolev spaces is needed.

As we have seen, the regularity results require some assumptions on the smoothness of the boundary and have been stated for Dirichlet, Neumann and Robin problems. It is interesting to give a couple of examples on the regularity of the solution in domains with corners and for the mixed problem.

Example 7.1 (Domains with Corners) Consider $S_\alpha = \{(r, \theta) \,|\, 0 < r < 1, -\alpha/2 < \theta < \alpha/2\}$ with $0 < \alpha < 2\pi$ and $\alpha \neq \pi$ (for $\alpha = \pi$ there are no corners; see Fig. 7.2 for the cases $\alpha = \frac{2\pi}{3}$ and $\alpha = \frac{5\pi}{3}$).

Consider

$$u(r, \theta) = r^{\frac{\pi}{\alpha}} \cos\left(\frac{\pi}{\alpha}\theta\right) .$$

Remember that the Laplace operator in polar coordinates is given by

$$\Delta = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2$$

and that the length of the gradient is given by

$$|\nabla v|^2 = (\partial_r v)^2 + \frac{1}{r^2}(\partial_\theta v)^2$$

(see Exercise 7.14). Thus it is easy to check that

$$\begin{cases} \Delta u = 0 & \text{in } S_\alpha \\ |\nabla u|^2 = \frac{\pi^2}{\alpha^2} r^{2(\frac{\pi}{\alpha}-1)} & \text{in } S_\alpha \,. \end{cases}$$

Moreover for $\theta = -\alpha/2$ and $\theta = \alpha/2$ we have $u = 0$, and for $r = 1$ we have $u = \cos(\frac{\pi}{\alpha}\theta)$. Thus u is the solution in S_α of a (non-homogeneous) Dirichlet boundary value problem for the Laplace operator, and the boundary datum is a continuous function on the boundary. Moreover,

$$\int_{S_\alpha} |\nabla u|^2 dx = \int_{-\alpha/2}^{\alpha/2} d\theta \int_0^1 \frac{\pi^2}{\alpha^2} r^{2(\frac{\pi}{\alpha}-1)} r dr = \alpha \, \frac{\pi^2}{\alpha^2} \, \frac{\alpha}{2\pi} = \frac{\pi}{2} \,,$$

thus $u \in H^1(S_\alpha)$. On the other hand $|\mathcal{D}^2 u| \sim r^{\frac{\pi}{\alpha}-2}$ as $r \sim 0$, therefore

$$\int_{S_\alpha} |\mathcal{D}^2 u|^2 dx \sim \int_0^1 r^{2(\frac{\pi}{\alpha}-2)} r dr = \int_0^1 r^{2\frac{\pi}{\alpha}-3} dr \,,$$

and this integral is convergent if and only if $3 - 2\pi/\alpha < 1$, namely if $\alpha < \pi$. In conclusion, if S_α is convex we have $u \in H^2(S_\alpha)$; if S_α is not convex we have $u \notin H^2(S_\alpha)$. Re-entrant corners are a threshold for regularity.

Exercise 7.14 Prove that the Laplace operator in polar is given by

$$\Delta = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2 \,,$$

and that the gradient is given by

$$\mathcal{D}_{x_1} = \cos\theta\,\partial_r - \frac{1}{r}\sin\theta\,\partial_\theta \quad , \quad \mathcal{D}_{x_2} = \sin\theta\,\partial_r + \frac{1}{r}\cos\theta\,\partial_\theta \,.$$

Example 7.2 (The Mixed Problem) Consider $u = r^{1/2}\sin(\theta/2)$ in $S = \{(r,\theta)\,|\,0 < r < 1, 0 < \theta < \pi\}$ (see Fig. 7.3). As before, we have $\Delta u = 0$ in S, $u_{|r=1} = \sin(\theta/2)$, $u_{|\theta=0} = 0$. We have seen in Exercise 7.14 that $\mathcal{D}_{x_2} u$ is given by $\mathcal{D}_{x_2} = \sin\theta\,\partial_r + \frac{1}{r}\cos\theta\,\partial_\theta$, thus

$$\mathcal{D}_{x_2} u = \sin\theta \frac{1}{2r^{1/2}} \sin\left(\frac{\theta}{2}\right) + \frac{1}{r}\cos\theta\, r^{1/2}\frac{1}{2}\cos\left(\frac{\theta}{2}\right) =$$
$$= \frac{1}{2r^{1/2}}\left(\sin\theta\sin\frac{\theta}{2} + \cos\theta\cos\frac{\theta}{2}\right) ,$$

which vanishes for $\theta = \pi$. Therefore u is the solution in S of the mixed problem for the Laplace operator, with homogeneous Neumann boundary datum on $\theta = \pi$,

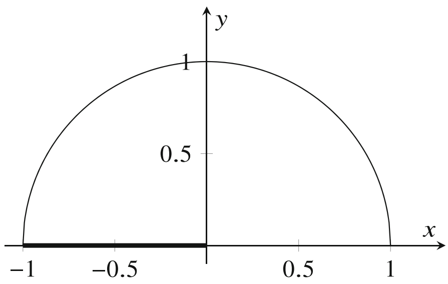


Fig. 7.3 The domain S: the homogeneous Neumann condition is imposed on the part of the boundary represented by a thicker line, while the Dirichlet condition is imposed on the remaining part of the boundary

homogeneous Dirichlet boundary datum on $\theta = 0$ and non-homogeneous Dirichlet boundary datum for $r = 1$ (note however that the Dirichlet boundary datum is continuous on the boundary).

We have $|\nabla u|^2 = (\partial_r u)^2 + 1/r^2(\partial_\theta u)^2 = \frac{1}{4r}$, thus

$$\int_S |\nabla u|^2 dx = \int_0^\pi d\theta \int_0^1 \frac{1}{4r} r dr = \frac{\pi}{4}$$

and $u \in H^1(S)$. On the other hand, we have

$$|\mathcal{D}^2 u| \sim r^{-3/2} \text{ as } r \sim 0\,,$$

thus

$$\int_S |\mathcal{D}^2 u|^2 dx \sim \int_0^1 r^{-3} r dr = \int_0^1 r^{-2} dr = +\infty$$

and $u \notin H^2(S)$.

In conclusion, the mixed boundary value problem can have solutions that are not regular. Note that the singularity has nothing to do with the corners at the points $(1, 0)$ and $(-1, 0)$. In fact, we can modify S in such a way that it becomes as smooth as we want at those points, and we can then reconsider this same example in that smooth domain.

7.4.2 *Sobolev Embedding Theorems*

An element in the Sobolev space $W^{1,p}(D)$ has additional "summability" or "regularity" properties. These properties are usually stated as "Sobolev embedding theorems". We will not present here the proofs (for that, see Evans [8, Section 5.6]), which are not so difficult but present some technicalities: we only underline that the idea is to prove suitable inequalities for smooth functions, and then use the fact that smooth functions are dense in $W^{1,p}(D)$. We divide the final statement in two cases: $1 \le p < n$ and $n < p < +\infty$.

Theorem 7.14 *Let $D \subset \mathbb{R}^n$ be a bounded, connected and open set. Suppose that ∂D is Lipschitz continuous. Assume $1 \le p < n$. Then if $u \in W^{1,p}(D)$ it follows $u \in L^{p^*}(D)$, where*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$$

and the estimate

$$\|u\|_{L^{p*}(D)} \le C\|u\|_{W^{1,p}(D)}$$

holds with a constant $C > 0$ only depending on p, n and D.

Note that $p^* > p$ and $p^* < +\infty$ (with $p^* \to +\infty$ for $p \to n^-$).

Example 7.3 Take $n = 2$ and $u \in W^{1,2}(D)$: then $u \in L^q(D)$ for all $q < +\infty$. Indeed $u \in W^{1,p}(D)$ for an arbitrary $p < 2 = n$, so that $u \in L^{p^*}(D)$ for p^* converging to $+\infty$ as $p \to 2^-$.

Example 7.4 Take $n = 3$ and $u \in W^{1,2}(D)$: then $u \in L^6(D)$, as

$$\frac{1}{p^*} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}\,.$$

Remark 7.5 We have already seen that $|x|^{-\alpha}$ belongs to $W^{1,p}(B_1)$ (B_1 being the ball centered at 0 with radius 1), provided that $p < n$ and $0 < \alpha < \frac{n-p}{p}$. Thus for $p < n$, unbounded functions are admitted in $W^{1,p}(D)$. This is also true for $p = n > 1$. Consider in fact $u(x) = (-\log|x|)^\alpha$ for $\alpha > 0$ and $B_{1/2} = \{x \in \mathbb{R}^n \mid |x| < 1/2\}$. We have, writing $|x| = r$:

$$|\nabla u| = \alpha(-\log r)^{\alpha-1}|\nabla \log r| = \alpha(-\log r)^{\alpha-1}\frac{1}{r}\,,$$

thus

$$\int_{B_{1/2}} |\nabla u|^n dx \sim \int_0^{1/2} \alpha^n(-\log r)^{(\alpha-1)n}\frac{1}{r^n}r^{n-1}dr$$
$$= \alpha^n \int_0^{1/2} (-\log r)^{(\alpha-1)n}\frac{1}{r}dr\,.$$

Changing variable with $t = -\log r$, $dt = -\frac{1}{r}dr$, we have

$$\int_{B_{1/2}} |\nabla u|^n dx \sim \alpha^n \int_{\log 2}^{+\infty} t^{(\alpha-1)n} dt \, ,$$

which is convergent for $(\alpha - 1)n < -1$, namely $0 < \alpha < \frac{n-1}{n}$. For these values of α the unbounded function $u(x) = (-\log |x|)^\alpha$ belongs to $W^{1,n}(B_{1/2})$.

Let us come now to the second result we want to present. We first introduce the Hölder space $C^{m,\lambda}(\overline{D})$, with $m \geq 0$, $0 < \lambda < 1$. This is given by the functions $u \in C^m(\overline{D})$ such that

$$\max_{|\alpha|=m} |\mathcal{D}^\alpha u(x_1) - \mathcal{D}^\alpha u(x_2)| \leq K|x_1 - x_2|^\lambda \qquad \forall\, x_1, x_2 \in \overline{D} \, ,$$

where the constant K does not depend on x_1 and x_2.

Theorem 7.15 *Let $D \subset \mathbb{R}^n$ be a bounded, connected and open set. Suppose that ∂D is Lipschitz continuous. Assume $n < p < +\infty$. Then if $u \in W^{1,p}(D)$, possibly modifying it on a set of measure equal to 0, we have $u \in C^{0,\lambda}(\overline{D})$ with $\lambda = 1 - \frac{n}{p}$ and the estimate*

$$\|u\|_{C^{0,\lambda}(\overline{D})} \leq C\|u\|_{W^{1,p}(D)}$$

holds with a constant $C > 0$ only depending on p, n and D.

The norm $\|u\|_{C^{m,\lambda}(\overline{D})}$ is given by the sum of $\|u\|_{C^m(\overline{D})}$ and

$$[u]_{C^{m,\lambda}(\overline{D})} = \sum_{|\alpha|=m} \sup_{x_1,x_2 \in \overline{D},\, x_1 \neq x_2} \frac{|\mathcal{D}^\alpha u(x_1) - \mathcal{D}^\alpha u(x_2)|}{|x_1 - x_2|^\lambda} \, .$$

Example 7.5 Take $n = 2$ and $u \in W^{1,3}(D)$: then $u \in C^{0,\lambda}(\overline{D})$ with $\lambda = 1 - \frac{2}{3} = \frac{1}{3}$.

Example 7.6 Take $n = 3$ and $u \in W^{1,6}(D)$: then $u \in C^{0,\lambda}(\overline{D})$ with $\lambda = 1 - \frac{3}{6} = \frac{1}{2}$.

The following characterization of Lipschitz continuous functions $\mathrm{Lip}(\overline{D})$ (see Appendix B, Definition B.1) is also interesting:

Theorem 7.16 *Let $D \subset \mathbb{R}^n$ be a bounded, connected and open set and suppose that ∂D is Lipschitz continuous. If $u \in W^{1,\infty}(D)$ then $u \in Lip(\overline{D})$ (possibly having modified it on a set of measure equal to 0). Vice versa, if $u \in Lip(\overline{D})$ then $u \in W^{1,\infty}(D)$.*

Proof This proof is essentially taken from Evans [8, Theor. 4, p. 279]. Let $u \in W^{1,\infty}(D)$; from Theorem 6.2 we know that there exists an extension Eu, in the following denoted $\overline{u}$, that has a compact support in $\mathbb{R}^n$ and belongs to $W^{1,\infty}(\mathbb{R}^n)$ (thus, possibly having modified it on a set of measure equal to 0, is continuous in

$\mathbb{R}^n$ by Theorem 7.15). Taking the ε-mollified version $\overline{u}_\varepsilon$ we know that $\overline{u}_\varepsilon$ uniformly converges to $\overline{u}$ and that

$$\|\nabla \overline{u}_\varepsilon\|_{L^\infty(\mathbb{R}^n)} = \|(\nabla \overline{u})_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \leq \|\nabla \overline{u}\|_{L^\infty(\mathbb{R}^n)}$$

(for the first equality see Theorem 6.1). Then, for $x, y \in \mathbb{R}^n$

$$\begin{aligned}|\overline{u}_\varepsilon(x) - \overline{u}_\varepsilon(y)| &= \left| \int_0^1 \tfrac{d}{dt}[\overline{u}_\varepsilon(y + t(x-y)] dt \right| \\ &= \left| \int_0^1 \nabla \overline{u}_\varepsilon(y + t(x-y)) \cdot (x-y)\, dt \right| \leq |x-y| \, \|\nabla \overline{u}_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \\ &\leq |x-y| \, \|\nabla \overline{u}\|_{L^\infty(\mathbb{R}^n)} .\end{aligned}$$

By passing to the limit for $\varepsilon \to 0^+$ we see that $\overline{u}$ is a Lipschitz function in $\mathbb{R}^n$, with Lipschitz constant $\|\nabla \overline{u}\|_{L^\infty(\mathbb{R}^n)}$, and therefore u a Lipschitz function in $\overline{D}$.

Assume now that $u \in \mathrm{Lip}(\overline{D})$, with L as Lipschitz constant; in particular, u is continuous in $\overline{D}$ and therefore bounded. The function

$$\widehat{u}(x) = \inf_{w \in \overline{D}} \big(u(w) + L|x-w|\big)$$

is defined for each $x \in \mathbb{R}^n$ (as $u(\cdot) + L|x - \cdot|$ is bounded from below) and is an extension of u. In fact, first of all for a fixed $x \in \overline{D}$ and for any $w \in \overline{D}$ we have

$$u(x) - u(w) \leq |u(x) - u(w)| \leq L|x-w| ,$$

thus $u(x) \leq u(w) + L|x-w|$ and $u(x) \leq \inf_{w \in \overline{D}} \big(u(w) + L|x-w|\big) = \widehat{u}(x)$. Secondly, taking into account that $x \in \overline{D}$,

$$\widehat{u}(x) = \inf_{w \in \overline{D}} \big(u(w) + L|x-w|\big) \leq u(x) + L|x-x| = u(x) .$$

Note now that the function $\widehat{u}$ belongs to $\mathrm{Lip}(\mathbb{R}^n)$ with Lipschitz constant L. In fact, for each fixed $x \in \mathbb{R}^n$ the function $w \to u(w) + L|x-w|$ is continuous in $\overline{D}$, thus $\widehat{u}(x) = \inf_{w \in \overline{D}} \big(u(w) + L|x-w|\big) = u(w_x) + L|x-w_x|$ for a some $w_x \in \overline{D}$. Take now $x, y \in \mathbb{R}^n$ and assume that $\widehat{u}(y) \geq \widehat{u}(x)$ (the opposite case is treated similarly); it follows

$$\begin{aligned}|\widehat{u}(x) - \widehat{u}(y)| = \widehat{u}(y) - \widehat{u}(x) &= \inf_{w \in \overline{D}} \big(u(w) + L|y-w|\big) - u(w_x) - L|x-w_x| \\ &\leq u(w_x) + L|y-w_x| - u(w_x) - L|x-w_x| \leq L|x-y| .\end{aligned}$$

For each fixed $i = 1, \ldots, n$ and each $h \in \mathbb{R}$, $h \neq 0$, the difference quotient $\mathcal{D}_i^{-h}\widehat{u}$ satisfies $\|\mathcal{D}_i^{-h}\widehat{u}\|_{L^\infty(\mathbb{R}^n)} \leq L$; since $L^\infty(\mathbb{R}^n)$ is the dual space of $L^1(\mathbb{R}^n)$, we can find a sequence $h_m \to 0$ and a function $\omega_i \in L^\infty(\mathbb{R}^n)$ such that $\mathcal{D}_i^{-h_m}\widehat{u}$ converges to ω_i with respect to the weak* convergence in $L^\infty(\mathbb{R}^n)$ (see, e.g., Yosida [28, Corollary to Theor. 1, p. 137]). Therefore, for each $\varphi \in C_0^\infty(\mathbb{R}^n)$ and taking into account Exercise 7.10 we have

$$\begin{aligned}\int_{\mathbb{R}^n} \widehat{u}\, \mathcal{D}_i\varphi\, dx &= \lim_{m\to\infty} \int_{\mathbb{R}^n} \widehat{u}\, \mathcal{D}_i^{h_m} \varphi\, dx \\ &= -\lim_{m\to\infty} \int_{\mathbb{R}^n} \varphi\, \mathcal{D}_i^{-h_m} \widehat{u}\, dx = -\int_{\mathbb{R}^n} \varphi\, \omega_i\, dx\,,\end{aligned}$$

thus, in the weak sense, $\mathcal{D}_i \widehat{u} = \omega_i \in L^\infty(\mathbb{R}^n)$ and $\mathcal{D}_i u = \omega_{i|D} \in L^\infty(D)$. □

Clearly, by a simple induction argument one can also obtain immersion theorems for higher order Sobolev spaces.

Theorem 7.17 *Let $D \subset \mathbb{R}^n$ be a bounded, connected and open set. Suppose that ∂D is Lipschitz continuous. Assume $u \in W^{k,p}(D)$, $k \geq 2$, $1 \leq p < +\infty$.*

1. *If $pk < n$, then $u \in L^q(D)$, where*

$$\frac{1}{q} = \frac{1}{p} - \frac{k}{n}$$

and

$$\|u\|_{L^q(D)} \leq C\|u\|_{W^{k,p}(D)}\,,$$

with a constant $C > 0$ only depending on k, p, n and D.

2. *If $pk > n$, then $u \in C^{k-[n/p]-1,\lambda}(\overline{D})$, where*

$$\lambda = \begin{cases} [n/p] + 1 - n/p & \text{if } n/p \text{ is not an integer} \\ \text{any positive number } < 1 & \text{if } n/p \text{ is an integer} \end{cases}$$

and

$$\|u\|_{C^{k-[n/p]-1,\lambda}(\overline{D})} \leq C\|u\|_{W^{k,p}(D)}\,,$$

with a constant $C > 0$ only depending on k, p, n and D.

Example 7.7 Take $n = 3$ and $u \in H^2(D) = W^{2,2}(D)$: then $u \in C^{0,\lambda}(\overline{D})$, with $\lambda = [3/2] + 1 - 3/2 = 1/2$.

Exercise 7.15 Let $D \subset \mathbb{R}^3$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Show that the immersion $W^{2,2}(D) \hookrightarrow C^{0,1/2}(\overline{D})$ holds, using Theorems 7.14 and 7.15.

Remark 7.6 (About Compactness)

(i) Let $p < n$. We have seen that

$$W^{1,p}(D) \hookrightarrow L^{p^*}(D)$$

for $p^* = \frac{np}{n-p}$; thus, since D is bounded, we also have

$$W^{1,p}(D) \hookrightarrow L^q(D)$$

for q satisfying $p \le q \le p^*$. It can be proved that this immersion is compact for $p \le q < p^*$ (note the strict inequality between q and p^*).

(ii) Let $p > n$. We have seen that

$$W^{1,p}(D) \hookrightarrow C^{0,\lambda}(\overline{D})$$

for $\lambda = 1 - n/p$; thus, since D is bounded, we also have

$$W^{1,p}(D) \hookrightarrow C^{0,\mu}(\overline{D})$$

for μ satisfying $0 < \mu \le \lambda$. It can be proved that this immersion is compact for $0 < \mu < \lambda$ (note the strict inequality between μ and λ).

Exercise 7.16

(i) Let $D \subset \mathbb{R}^3$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Show that the bilinear form

$$B_L(w, v) = \int_D \sum_{i,j=1}^{n} a_{ij} \mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx$$

is bounded provided that the coefficients satisfy $a_{ij} \in L^\infty(D)$, $b_i \in L^3(D)$ and $a_0 \in L^{3/2}(D)$.

(ii) Prove that $B_L(w, v)$ is coercive in $H^1_0(D)$, $H^1_*(D)$ and $H^1_{\Gamma_D}(D)$, provided that $\|b_i\|_{L^3(D)}$, $i = 1, \dots, n$, and $\|a_0\|_{L^{3/2}(D)}$ are small enough.

Exercise 7.17 Show that the solution u of the homogeneous Dirichlet boundary value problem

$$\begin{cases} -\Delta u = 1 \text{ in } D \\ u_{|\partial D} = 0 \ \text{ on } \partial D\,, \end{cases}$$

where $D = \{x \in \mathbb{R}^n \mid |x| < 1\}$, belongs to $C^\infty(\overline{D})$.

Exercise 7.18 Show that the eigenvectors w_k of the homogeneous Dirichlet boundary value problem

$$\begin{cases} -\Delta w_k = \lambda_k w_k \text{ in } D \\ w_{k|\partial D} = 0 \qquad \text{on } \partial D \,, \end{cases}$$

where $D = \{x \in \mathbb{R}^n \mid |x| < 1\}$, belong to $C^\infty(\overline{D})$.

7.5 Galerkin Numerical Approximation

The general form of the variational problem we have dealt with is:

$$\text{find } u \in V \ : \ B(u, v) = F(v) \quad \forall\, v \in V \,, \tag{7.12}$$

where V is an infinite dimensional Hilbert space.

This is a problem with infinitely many "degrees of freedom" (as we need infinitely many informations for determining a function in an infinite dimensional Hilbert space). Moreover, very often we have not an explicit formula for representing the solution. Therefore, in concrete applications it is important to devise an approximation method to compute a suitable approximate solution.

To this aim, a very popular and efficient idea is to discretize the problem by projecting it onto a finite dimensional subspace of V, say $V_N \subset V$, such that $\dim V_N = N < +\infty$. Notice that V_N is a Hilbert space because it is a finite dimensional subspace.

The approximate problem in V_N can be simply formulated as follows:

$$\text{find } u_N \in V_N \ : \ B(u_N, v_N) = F(v_N) \quad \forall\, v_N \in V_N \,. \tag{7.13}$$

Let us assume that $\psi_1, \ldots, \psi_N$ is basis of V_N: as a consequence of the linearity of $B(\cdot, \cdot)$ and $F(\cdot)$ this problem is equivalent to

$$\text{find } u_N \in V_N \ : \ B(u_N, \psi_j) = F(\psi_j) \quad \forall\, j = 1, \ldots, N \,.$$

This is the so-called Galerkin method. Note that it corresponds to the solution of the linear system

$$A\mathbf{U} = \mathbf{F} \,,$$

with $u_N = \sum_{j=1}^N U_j \psi_j$, $U_j \in \mathbb{R}$, $\mathbf{U} = (U_1, \ldots, U_N)$, $A = \{A_{jl}\}$ with $A_{jl} = B(\psi_l, \psi_j)$ and $\mathbf{F} = (F(\psi_1), \ldots, F(\psi_N))$.

The convergence analysis is very easy, and it is based on the following important result.

Theorem 7.18 (Céa Theorem) *Assume that bilinear form B and the linear functional F satisfy to hypotheses of the Lax-Milgram theorem, i.e., that the following conditions hold*

(i) $|B(w, v)| \leq \gamma \|w\|_V \|v\|_V$ *for* $\gamma > 0$ *[boundedness of* $B(\cdot, \cdot)$*]*
(ii) $B(v, v) \geq \alpha \|v\|_V^2$ *for* $\alpha > 0$ *[coerciveness of* $B(\cdot, \cdot)$*]*
(iii) $|F(v)| \leq M\|v\|_V$ *for* $M > 0$ *[boundedness of* $F(\cdot)$*].*

Then by Lax-Milgram theorem in V there exists a unique $u \in V$, solution of the infinite dimensional problem (7.12)*, and by Lax-Milgram theorem in V_N there exists a unique $u_N \in V_N$, solution of the approximated problem* (7.13)*. Moreover, the following error estimate holds*

$$\|u - u_N\|_V \leq \frac{\gamma}{\alpha} \inf_{v_N \in V_N} \|u - v_N\|_V = \frac{\gamma}{\alpha} \operatorname{dist}(u, V_N) .$$

Therefore, the convergence of the Galerkin method follows at once, provided that for all $w \in V$ we have that $\operatorname{dist}(w, V_N) \to 0$ *as* $N \to \infty$*.*

Proof Since $B(u, v) = F(v)$ for all $v \in V$, in particular we have that $B(u, v_N) = F(v_N)$ for all $v_N \in V_N \subset V$. Moreover $B(u_N, v_N) = F(v_N)$ for all $v_N \in V_N$. Therefore $B(u-u_N, v_N) = 0$ for all $v_N \in V_N$. Employing this consistency property, we easily have that

$$\begin{aligned}\alpha\|u - u_N\|_V^2 &\leq B(u - u_N, u - u_N) = \overbrace{B(u - u_N, u)}^{\text{as } B(u-u_N, u_N)=0} \\ &= \overbrace{B(u - u_N, u - v_N)}^{\text{as } B(u-u_N, v_N)=0} \leq \gamma\|u - u_N\|_V \|u - v_N\|_V \quad \forall\, v_N \in V_N ,\end{aligned}$$

and so we have obtained that

$$\|u - u_N\|_V \leq \frac{\gamma}{\alpha} \inf_{v_N \in V_N} \|u - v_N\|_V ,$$

the desired estimate. □

Exercise 7.19 Let $D \subset \mathbb{R}^3$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Let V be a closed subspace of $H^1(D)$, and let the assumptions of Theorem 7.18 be satisfied. Suppose moreover that for each $w \in C^0(\overline{D})$ one can find $\pi_N(w) \in V_N$ such that $\|w - \pi_N(w)\|_V \to 0$ as $N \to \infty$. Then show that the Galerkin method is convergent.

Remark 7.7 One of the most important examples of Galerkin approximation is that based on finite elements. For the variational problems described in Chap. 5 the finite dimensional subspace V_N is given by piecewise-polynomial and globally continuous functions (see Exercise 6.8 for the proof that this is indeed a subspace of $H^1(D)$). Here it is assumed that the domain D is the union of (non-overlapping) subsets of simple shape T, the elements: say, for $n = 3$, tetrahedra or hexahedra. Denoting by h the maximum diameter of the elements, let N_h be the dimension of the space

$$V_{N_h} = \{v : \overline{D} \mapsto \mathbb{R} \mid v \in C^0(\overline{D}), v_{|T} \in \mathbb{P}_r \ \forall\, T\},$$

where $\mathbb{P}_r$ is the space of polynomials of degree less than or equal to r, $r \geq 1$. Thus when $h \to 0$ the number of elements T goes to infinity, and therefore one has $N_h \to +\infty$.

For this type of finite elements one has an error estimate between the exact solution u and the approximate solution u_h that satisfies $\|u - u_h\|_{H^1(D)} = O(h^r)$ (having assumed that the hypotheses of Theorem 7.18 are satisfied and provided that the solution u is smooth enough).

7.6 Exercises

Exercise 7.1 Prove that in Theorem 7.2 one has $K^T = \tau(L^T + \tau I)^{-1}$.

Solution Let us first observe that this result is clearly reasonable, as this would be the case for a matrix $K = \tau(L + \tau I)^{-1}$.

Let us write for simplicity $(\cdot,\cdot)$ instead of $(\cdot,\cdot)_{L^2(D)}$, and for $w, v \in L^2(D)$ compute (Kw, v): defining by $q \in H^1_0(D)$ the solution of $(L + \tau I)q = w$ (in the weak sense, $B_\tau(q, \psi) = (w, \psi)$ for each $\psi \in H^1_0(D)$), we have

$$(Kw, v) = (\tau(L + \tau I)^{-1}w, v) = (\tau q, v) = \tau(q, v)\,.$$

Then define by $p \in H^1_0(D)$ the solution of $(L^T + \tau I)p = v$ (namely, $B_{L^T}(p, \psi) + \tau(p, \psi) = (v, \psi)$ for each $\psi \in H^1_0(D)$) and compute $(\tau(L^T + \tau I)^{-1}v, w)$: it holds

$$(\tau(L^T + \tau I)^{-1}v, w) = (\tau p, w) = \tau(p, w)\,.$$

Thus we must prove that $(q, v) = (p, w)$. We have

$$(q, v) = \overbrace{(q, (L^T + \tau I)p)}^{B_{L^T}(p,q)+\tau(p,q)} = \tau(q, p) + \overbrace{(q, L^T p)}^{B_{L^T}(p,q)} = \tau(p, q) + \overbrace{(p, Lq)}^{B(q,p)}$$

and

$$(p, w) = \underbrace{(p, (L + \tau I)q)}_{B_\tau(q,p)} = \tau(p, q) + \underbrace{(p, Lq)}_{B(q,p)}\,,$$

thus the result

$$(Kw, v) = (\tau(L^T + \tau I)^{-1}v, w)$$

is proved.

Exercise 7.2 Let A be a $n \times m$ matrix, associated to the linear map $v \mapsto Av$, $v \in \mathbb{R}^m$, $Av \in \mathbb{R}^n$. Prove that $R(A) = N(A^T)^\perp$.

Solution

($\subset$) $y \in R(A)$ means that exists $x \in \mathbb{R}^m$ such that $Ax = y$. Taking now $w \in N(A^T)$, namely, $A^T w = 0$, it is easily checked that $(y, w) = (Ax, w) = (x, A^T w) = 0$.

($\supset$) $y \in N(A^T)^\perp$ can be written (as any vector in $\mathbb{R}^n$) as

$$y = \hat{y} + Ax, \quad \hat{y} \in R(A)^\perp, \quad x \in \mathbb{R}^m .$$

Since we already know that $Ax \in N(A^T)^\perp$, it follows at once $\hat{y} \in N(A^T)$. Also

$$(A^T \hat{y}, x) = (\hat{y}, Ax) = 0 \quad \forall\, x \in \mathbb{R}^m \implies A^T \hat{y} = 0 \implies \hat{y} \in N(A^T) .$$

Since $\hat{y} \in N(A^T) \cap N(A^T)^\perp$, it follows $\hat{y} = 0$ and $y = Ax \in R(A)$.

Exercise 7.3 Let $A : X \mapsto Y$ be a linear and bounded operator, X and Y Hilbert spaces. Define the adjoint operator $A^T : Y \mapsto X$ as $(A^T y, x)_X = (y, Ax)_Y$ for all $y \in Y$, $x \in X$. Prove that

(i) $\overline{R(A)} = N(A^T)^\perp$
(ii) $R(A)^\perp = N(A^T)$.

Solution

(i) The proof that $R(A) \subset N(A^T)^\perp$ is as in Exercise 7.2; since $N(A^T)^\perp$ is closed, we have $\overline{R(A)} \subset N(A^T)^\perp$. On the other hand, let us first verify that for a subspace $W \subset Y$ it holds $W^\perp = \overline{W}^\perp$. In fact, a vector v orthogonal to all the elements of $\overline{W}$ is clearly orthogonal to all the elements of W; viceversa, suppose we have $(v, w)_Y = 0$ for all $w \in W$ and take $w_* \in \overline{W}$: then $w_* = \lim_k w_k$, $w_k \in W$, and therefore $(v, w_*)_Y = \lim_k (v, w_k)_Y = 0$. As a second step, consider the orthogonal decomposition given by $Y = \overline{R(A)} \oplus \overline{R(A)}^\perp$ and take $y \in N(A^T)^\perp$. We can write $y = \hat{y} + q$, where $\hat{y} \in \overline{R(A)}^\perp = R(A)^\perp$ and $q \in \overline{R(A)}$. Now the proof is as that of Exercise 7.2: since we already know that $q \in \overline{R(A)} \subset N(A^T)^\perp$, it follows $\hat{y} \in N(A^T)^\perp$; moreover

$$(A^T \hat{y}, x)_X = (\hat{y}, Ax)_Y = 0 \quad \forall\, x \in X ,$$

thus $\hat{y} \in N(A^T)$ and therefore $\hat{y} = 0$. In conclusion, $y = q \in \overline{R(A)}$.

(ii) Follows at once from (i) by passing to the orthogonal.

Exercise 7.4 Under the assumptions of Theorem 7.5, take $\lambda \notin \Sigma$ and for each $f \in L^2(D)$ let $u \in H_0^1(D)$ be the unique solution of (7.7). Prove that the solution operator $S_\lambda : f \mapsto u$ is a bounded operator in $L^2(D)$, namely, there exists a constant $C > 0$ such that

$$\|u\|_{L^2(D)} \le C\|f\|_{L^2(D)} .$$

Solution We prove that the operator S_λ is closed, thus, being defined on the whole space $L^2(D)$, it is bounded as a consequence of the closed graph theorem (see Yosida [28, Theorem 1, p. 79]). Take $f_k \to f$ in $L^2(D)$ and $u_k = S_\lambda f_k \to q$ in $L^2(D)$. For a suitable $\tau > 0$ we know that u_k is the solution of the coercive problem

$$B_L(u_k, v) + \tau \int_D u_k v dx = (\tau + \lambda) \int_D u_k v dx + \int_D f_k v dx \quad \forall\, v \in H_0^1(D) . \tag{7.14}$$

Thus by Lax–Milgram theorem we have the estimate

$$\|u_k\|_{H^1(D)} \le C(\|u_k\|_{L^2(D)} + \|f_k\|_{L^2(D)}) .$$

Therefore u_k is bounded in $H^1(D)$, and since $H^1(D)$ is a Hilbert space we can extract a subsequence u_{k_s} which is weakly convergent to $w \in H^1(D)$ (see Yosida [28, Theorem 1, p. 126, and Theorem of Eberlein–Shmulyan, p. 141]), in particular is weakly convergent to w in $L^2(D)$. As a consequence of the uniqueness of the weak limit we obtain $q = w$, and passing to the limit in (7.14) we find

$$B_L(q, v) + \tau \int_D q v dx = (\tau + \lambda) \int_D q v dx + \int_D f v dx \quad \forall\, v \in H_0^1(D) .$$

This shows that $q = S_\lambda f$, thus S_λ is closed.

Exercise 7.5 Under the assumptions of Theorem 7.5, take $\lambda \notin \Sigma$ and for each $f \in L^2(D)$ let $u \in H_0^1(D)$ be the unique solution of (7.7). Prove that the solution operator $S_\lambda : f \mapsto u$ is a bounded operator from $L^2(D)$ to $H_0^1(D)$, namely, there exists a constant $C > 0$ such that

$$\|u\|_{H^1(D)} \le C\|f\|_{L^2(D)} .$$

Solution In Exercise 7.4 we have seen that u is the solution of the coercive problem

$$B_L(u, v) + \tau \int_D u v dx = (\tau + \lambda) \int_D u v dx + \int_D f v dx \quad \forall\, v \in H_0^1(D) ,$$

$\tau > 0$ being a suitable constant, and that by Lax–Milgram theorem u satisfies the estimate

$$\|u\|_{H^1(D)} \leq C(\|u\|_{L^2(D)} + \|f\|_{L^2(D)})\,.$$

Thus the result follows from Exercise 7.4.

Exercise 7.6 Prove that the minimum eigenvalue λ_1 of the Laplace operator $-\Delta$ associated to the homogeneous Dirichlet boundary condition is equal to $\frac{1}{C_D}$, where

$$C_D = \sup_{v \in H_0^1(D), v \neq 0} \frac{\int_D v^2 dx}{\int_D |\nabla v|^2 dx}$$

is the "best" Poincaré constant (see Sect. 6.2).

Solution The eigenvalues λ_k and their related eigenvectors $w_k \in H_0^1(D)$, $w_k \neq 0$, $k = 1, 2, \ldots$, satisfy

$$\int_D \nabla w_k \cdot \nabla v dx = \lambda_k \int_D w_k v dx \quad \forall\, v \in H_0^1(D)\,, \tag{7.15}$$

thus λ_1 can be represented by the Rayleigh quotient

$$\lambda_1 = \frac{\int_D |\nabla w_1|^2 dx}{\int_D w_1^2 dx}$$

and we have at once

$$\lambda_1 \geq \inf_{v \in H_0^1(D), v \neq 0} \frac{\int_D |\nabla v|^2 dx}{\int_D v^2 dx} = \frac{1}{C_D}\,.$$

On the other hand, knowing that the sequence of eigenvectors w_k is an $L^2(D)$-orthonormal Hilbertian basis (see Theorem 7.7), we can write $v = \sum_{k=1}^{\infty} v_k w_k$, where $v_k = \int_D v w_k dx$, so that

$$\int_D v^2 dx = \int_D \Big(\sum_{k=1}^{\infty} v_k w_k\Big) \Big(\sum_{j=1}^{\infty} v_j w_j\Big) dx = \sum_{k=1}^{\infty} v_k^2$$

and, using (7.15),

$$\begin{aligned}\int_D |\nabla v|^2 dx &= \int_D \Big(\sum_{k=1}^{\infty} v_k \nabla w_k\Big) \cdot \Big(\sum_{j=1}^{\infty} v_j \nabla w_j\Big) dx \\ &= \sum_{k,j=1}^{\infty} v_k v_j \int_D \nabla w_k \cdot \nabla w_j dx \\ &= \sum_{k,j=1}^{\infty} v_k v_j \lambda_k \int_D w_k\, w_j dx = \sum_{k=1}^{\infty} v_k^2 \lambda_k \\ &\geq \lambda_1 \sum_{k=1}^{\infty} v_k^2 \, .\end{aligned}$$

In conclusion, for any $v \in H_0^1(D)$, $v \neq 0$,

$$\frac{\displaystyle\int_D |\nabla v|^2 dx}{\displaystyle\int_D v^2 dx} \geq \frac{\lambda_1 \displaystyle\sum_{k=1}^{\infty} v_k^2}{\displaystyle\sum_{k=1}^{\infty} v_k^2} = \lambda_1$$

thus

$$\frac{1}{C_D} = \inf_{v \in H_0^1(D), v \neq 0} \frac{\int_D |\nabla v|^2 dx}{\int_D v^2 dx} \geq \lambda_1 \, ,$$

and the thesis is proved.

Exercise 7.7

(i) Consider the elliptic operator

$$Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij} \mathcal{D}_j w) + a_0 w \, ,$$

with $a_{ij} = a_{ji}$ and $a_0 \geq 0$. If $\lambda_\star$ is an eigenvalue of L associated to anyone of the boundary conditions of Dirichlet, Neumann, mixed or Robin type, then $\lambda_\star \geq 0$.

(ii) The case $\lambda_\star = 0$ is possible if and only if the boundary condition is of Neumann type and $a_0 = 0$. In that case the corresponding eigenvector $w_\star$ is a constant (different from 0).

Solution

(i) The eigenvalue $\lambda_\star$ and the correspondent eigenvector $w_\star \in V$, $w_\star \neq 0$, satisfy

$$B(w_\star, v) = \lambda_\star \int_D w_\star v dx \quad \forall\, v \in V\,,$$

where V and $B(\cdot,\cdot)$ are the Hilbert space and the bilinear form associated to the different boundary value problems (see Sect. 5.1). In particular, we have

$$\lambda_\star = \frac{B(w_\star, w_\star)}{\int_D w_\star^2 dx}$$

and, by the ellipticity assumption (and the assumption that the coefficient κ for the Robin problem is non-negative) we obtain

$$\begin{aligned} B(w_\star, w_\star) \geq B_L(w_\star, w_\star) &= \textstyle\int_D \sum_{i,j=1}^n a_{ij} \mathcal{D}_j w_\star \mathcal{D}_i w_\star dx + \int_D a_0 w_\star^2 dx \\ &\geq \alpha_0 \textstyle\int_D |\nabla w_\star|^2 dx + \int_D a_0 w_\star^2 dx \geq 0\,. \end{aligned}$$

(ii) When we have $\lambda_\star = 0$, from the arguments in (i) we deduce $B(w_\star, w_\star) = 0$. Therefore coerciveness and the assumption $a_0 \geq 0$ imply $w_\star =$ const. For the Dirichlet, mixed and Robin boundary value problems this would give $w_\star = 0$, a contradiction. (Note that for the Robin problem this follows from the fact that

$$0 = B(w_\star, w_\star) = B_L(w_\star, w_\star) + \int_{\partial D} \kappa w_\star^2 dS_x\,,$$

thus $\int_{\partial D} \kappa w_\star^2 dS_x = w_\star^2 \int_{\partial D} \kappa dS_x = 0$, only possible for $w_\star = 0$.) For the Neumann boundary condition knowing that $w_\star =$ const has as a consequence $\int_D a_0 dx = 0$, which gives $a_0 = 0$. Finally, it is trivial to show that the Neumann problem with $a_0 = 0$ has a vanishing eigenvalue correspondent to a constant eigenvector (different from 0).

Exercise 7.8 Let $D \subset \mathbb{R}^n$ an open set. Prove that $v^+ = \max(v, 0)$ and $v^- = \max(-v, 0)$ belong to $W^{1,p}(D)$ for $v \in W^{1,p}(D)$, $1 \leq p \leq +\infty$. More precisely, defining

$$w_i^+ = \begin{cases} \mathcal{D}_i v & \text{where } v > 0 \\ 0 & \text{where } v \leq 0 \end{cases}\,, \quad w_i^- = \begin{cases} -\mathcal{D}_i v & \text{where } v < 0 \\ 0 & \text{where } v \geq 0 \end{cases}\,,$$

one has $\mathcal{D}_i v^+ = w_i^+$ and $\mathcal{D}_i v^- = w_i^-$, $i = 1, \ldots, n$.

Solution Since $|v^+| \leq |v|$, $|v^-| \leq |v|$, $|w_i^+| \leq |\mathcal{D}_i v|$ and $|w_i^-| \leq |\mathcal{D}_i v|$ it is clear that v^+, v^-, w_i^+ and w_i^- belong to $L^p(D)$; therefore the only things to be proved are the differentiation formulas. Let us first make some naive considerations. Nothing has to be proved if either $v > 0$ a.e. in D or $v \leq 0$ a.e. in D. Thus we can consider the case in which both sets $\{v > 0\}$ and $\{v \leq 0\}$ have positive measure. Let us focus on v^+. Defining w_i^+ as above and taking $\varphi \in C_0^\infty(D)$ we formally have, by integration by parts,

$$\int_D w_i^+ \varphi dx = \int_{\{v>0\}} w_i^+ \varphi dx + \int_{\{v\leq 0\}} w_i^+ \varphi dx = \int_{\{v>0\}} \mathcal{D}_i v \varphi dx$$
$$= -\int_{\{v>0\}} v \mathcal{D}_i \varphi dx + \int_{\partial\{v>0\}} n_i v \varphi dS_x = -\int_D v^+ \mathcal{D}_i \varphi dx \,,$$

where we have deduced that $\int_{\partial\{v>0\}} n_i v \varphi dS_x = 0$ as $\partial\{v > 0\} = (\partial\{v > 0\} \cap D) \cup (\partial\{v > 0\} \cap \partial D)$, $\varphi = 0$ on ∂D and we expect that $v = 0$ on $\partial\{v > 0\} \cap D$. However, this formal proof is not rigorous, as when $v \in W^{1,p}(D)$ is not smooth the set $\{v > 0\}$ is only a measurable set, and an integration by parts formula like the one here above is not necessarily valid. Moreover, even the additional information that v is smooth would not solve the problem, as in that situation it would be true that the set $\{v > 0\}$ is an open set and that $v = 0$ on $\partial\{v > 0\}$, but still this boundary $\partial\{v > 0\}$ could be as wild as you (do not) like.

Thus we have to adopt a different strategy, that we essentially borrow from Gilbarg and Trudinger [11, Lemma 7.6, p. 145]. First of all let us prove the following "chain rule": if $v \in L^1_{\mathrm{loc}}(D)$ with $\mathcal{D}_i v \in L^1_{\mathrm{loc}}(D)$ and $F \in C^1(\mathbb{R})$ with $F' \in L^\infty(\mathbb{R})$, then $\mathcal{D}_i[F(v)] = F'(v)\,\mathcal{D}_i v$ in D, $i = 1, \ldots, n$. In fact, take $\varphi \in C_0^\infty(D)$, set $\Phi = \operatorname{supp}\varphi$ and take an open set Q with a Lipschitz continuous boundary ∂Q and such that $\Phi \subset Q \subset\subset D$. From Theorem 6.1 there exists a sequence $v_m \in C^\infty(\overline{Q})$ such that $v_m \to v$ and $\mathcal{D}_i v_m \to \mathcal{D}_i v$ in $L^1(Q)$; for these smooth functions (and knowing that ∂Q is Lipschitz continuous) by integration by parts we clearly have

$$\int_Q F(v_m)\,\mathcal{D}_i\varphi = -\int_Q F'(v_m)\mathcal{D}_i v_m\,\varphi \,.$$

The "chain rule" $\int_D F(v)\,\mathcal{D}_i\varphi = -\int_D F'(v)\mathcal{D}_i v\,\varphi$ thus follows as

$$\int_Q |F(v_m) - F(v)| \leq \sup|F'| \int_Q |v_m - v| \to 0$$

and

$$\begin{aligned}
&\textstyle\int_Q |F'(v_m)\mathcal{D}_i v_m - F'(v)\mathcal{D}_i v| \\
&\quad \leq \textstyle\int_Q |F'(v_m)||\mathcal{D}_i v_m - \mathcal{D}_i v| + \int_Q |F'(v_m) - F'(v)||\mathcal{D}_i v| \\
&\quad \leq \sup|F'| \textstyle\int_Q |\mathcal{D}_i v_m - \mathcal{D}_i v| + \int_Q |F'(v_m) - F'(v)||\mathcal{D}_i v| \to 0\,.
\end{aligned}$$

This last result holds true since (for a subsequence…) $v_m \to v$ a.e. in Q, $F'(v_m) \to F'(v)$ a.e. in Q (here the continuity of F' has been used) and thus

$$\int_Q |F'(v_m) - F'(v)||\mathcal{D}_i v| \to 0$$

by the Lebesgue dominated convergence theorem.

Take now $v \in W^{1,p}(D)$, $1 \le p \le +\infty$, and consider the approximation of the function $s \mapsto \max(s, 0)$ given by

$$F_\epsilon(s) = \begin{cases} 0 & \text{for } s \le 0 \\ \frac{1}{2\epsilon}s^2 & \text{for } 0 < s < \epsilon \\ s - \frac{\epsilon}{2} & \text{for } s \ge \epsilon\,, \end{cases}$$

which clearly satisfies $F_\epsilon \in C^1(\mathbb{R})$, $F'_\epsilon \in L^\infty(\mathbb{R})$, $F_\epsilon(s) \to \max(s, 0)$ and $F'_\epsilon(s) \to \chi^+(s)$ for $s \in \mathbb{R}$, where $\chi^+(s)$ is the characteristic function of $\{s > 0\}$. Thus we have

$$\int_D F_\epsilon(v)\,\mathcal{D}_i\varphi = -\int_D F'_\epsilon(v)\mathcal{D}_i v\,\varphi\,,$$

and by the Lebesgue dominated convergence theorem we find

$$\int_D v^+\,\mathcal{D}_i\varphi = -\int_D \chi^+(v)\mathcal{D}_i v\,\varphi = -\int_{\{v>0\}} \mathcal{D}_i v\,\varphi = -\int_D w_i^+\,\varphi\,.$$

For another proof of this exercise and other related results we refer to the classical book by Kinderlehrer and Stampacchia [15, Theorem A.1, p. 50].

As a final remark, take into account that it is not even trivial to prove the following "trivial" result: for $v \in W^{1,p}(D)$ it holds $\nabla v = 0$ a.e. in $E = \{x \in D \,|\, v(x) = 0\}$. Its proof is indeed a consequence of the results provided by this exercise, as $v = v^+ - v^-$.

Exercise 7.9 Prove that

$$\sup_{\partial D} u^+ = \max(\sup_{\partial D} u, 0) \quad \text{and} \quad \inf_{\partial D}(-u^-) = \min(\inf_{\partial D} u, 0)$$

(so that the conclusion of Theorem 7.8 can be written as $\sup_D u \le \max(\sup_{\partial D} u, 0)$ for a subsolution and $\inf_D u \ge \min(\inf_{\partial D} u, 0)$ for a supersolution).

Solution For the sake of simplicity let us write $B = \sup_{\partial D} u^+$ and $A = \max(\sup_{\partial D} u, 0)$. Suppose that $\sup_{\partial D} u > 0$ and define $Q = \{x \in \partial D \,|\, u(x) > 0\}$: we have $u^+ = u$ in Q and $u^+ = 0$ in $\partial D \setminus Q$, thus $B = \sup_{\partial D} u^+ = \sup_Q u^+ = \sup_Q u = \sup_{\partial D} u = A$. On the other hand, if $\sup_{\partial D} u \le 0$ we have $A = 0$

and $u \le 0$ on ∂D, thus $u^+ = 0$ on ∂D and finally $B = 0 = A$. The proof of $\inf_{\partial D}(-u^-) = \min(\inf_{\partial D} u, 0)$ is similar.

Exercise 7.10 Take $v \in L^2(D)$, $\varphi \in L^2(D)$ with $\Phi = \operatorname{supp}\varphi \subset D$, and consider the difference quotients defined in (7.11). Then we have the integration by parts formula

$$\int_D v\, \mathcal{D}_k^h \varphi dx = -\int_D \mathcal{D}_k^{-h} v\, \varphi dx\,,$$

for each h with $0 < |h| < \operatorname{dist}(\Phi, \partial D)$, $k = 1, \dots, n$.

Solution Set $\Phi = \operatorname{supp}\varphi$ and define $\Phi_h^k = \{y \in D \mid y = x - he_k, x \in \Phi\}$. Then we have

$$\begin{aligned}\int_D v(x - he_k)\varphi(x)dx &= \int_\Phi v(x - he_k)\varphi(x)dx \\ &= \int_{\Phi_h^k} v(y)\varphi(y + he_k)dy = \int_D v(y)\varphi(y + he_k)dy\,,\end{aligned}$$

having used the change of variable $y = x - he_k$. Then it easily follows

$$\int_D v(x)\frac{\varphi(x + he_k) - \varphi(x)}{h}dx = -\int_D \frac{v(x - he_k) - v(x)}{-h}\varphi(x)dx\,,$$

which is the stated result.

Exercise 7.11

(i) Take $v \in H^1(D)$ and consider $Q \subset\subset D$. Then the difference quotient $\mathcal{D}^h v = (\mathcal{D}_1^h v, \dots, \mathcal{D}_n^h v)$ defined in (7.11) satisfies

$$\|\mathcal{D}^h v\|_{L^2(Q)} \le \|\nabla v\|_{L^2(D)}$$

for each h with $0 < |h| < \operatorname{dist}(Q, \partial D)$.

(ii) Take k with $1 \le k \le n$, $v \in L^2(D)$ and $Q \subset\subset D$. Suppose that there exists a constant $C_* > 0$ such that

$$\|\mathcal{D}_k^h v\|_{L^2(Q)} \le C_*$$

for each h with $0 < |h| < \operatorname{dist}(Q, \partial D)$. Then $\mathcal{D}_k v \in L^2(Q)$.

(iii) Take k with $1 \le k \le n$, $v \in L^2(D)$ and suppose there exists a constant $C_\sharp > 0$ such that

$$\|\mathcal{D}_k^h v\|_{L^2(D_{|h|})} \le C_\sharp$$

for each $h \neq 0$, where $D_{|h|} = \{x \in D \,|\, \mathrm{dist}(x, \partial D) > |h|\}$. Then $\mathcal{D}_k v \in L^2(D)$ and $\|\mathcal{D}_k v\|_{L^2(D)} \le C_\sharp$.

Solution

(i) By approximation, we can assume that v is smooth. Take $x \in Q$ and let e_k the unit vector in the k-th direction. Since

$$\frac{d}{dt} v(x + the_k) = \sum_{j=1}^{n} (\mathcal{D}_j v)(x + the_k) \frac{d}{dt}(x_j + th\delta_{kj}) = h\,(\mathcal{D}_k v)(x + the_k)\ ,$$

we have

$$v(x + he_k) - v(x) = h \int_0^1 (\mathcal{D}_k v)(x + the_k) dt$$

and consequently

$$\begin{aligned}
\int_Q (\mathcal{D}_k^h v)^2(x) dx &= \int_Q \frac{|v(x + he_k) - v(x)|^2}{h^2} dx \\
&= \int_Q \left(\int_0^1 (\mathcal{D}_k v)(x + the_k) dt \right)^2 dx \\
&\le \int_Q \left(\int_0^1 (\mathcal{D}_k v)^2(x + the_k) dt \right) dx \\
&= \int_0^1 \left(\int_Q (\mathcal{D}_k v)^2(x + the_k) dx \right) dt \\
&\le \int_0^1 \left(\int_D (\mathcal{D}_k v)^2(y) dy \right) dt = \int_D (\mathcal{D}_k v)^2(x) dx\ ,
\end{aligned}$$

having used the change of variable $x + the_k = y$.

(ii) The idea is to pass to the limit in the integration by parts formula in Exercise 7.10:

$$\int_Q \mathcal{D}_k^{-1/m} v\, \varphi dx = - \int_Q v\, \mathcal{D}_k^{1/m} \varphi dx\ , \tag{7.16}$$

where $\varphi \in C_0^\infty(Q)$ and m is such that $1/m < \mathrm{dist}(\mathrm{supp}\,\varphi, \partial Q)$. Since $L^2(Q)$ is a Hilbert space, the estimate $\|\mathcal{D}^h v\|_{L^2(Q)} \le C_*$ for $h = -1/m$ (and m large enough to have $1/m < \mathrm{dist}(Q, \partial D)$) has as a consequence that from the sequence $\mathcal{D}_k^{-1/m} v$ we can extract a subsequence, still denote by $\mathcal{D}_k^{-1/m} v$, which converges weakly to w_k in $L^2(Q)$ (see Yosida [28, Theorem 1, p. 126, and Theorem of Eberlein–Shmulyan, p. 141]). On the other hand, it is easily seen that $\mathcal{D}_k^{1/m}\varphi$ converges to $\mathcal{D}_k\varphi$ in $L^2(Q)$: in fact, by Taylor expansion

$$\frac{\varphi(x+he_k)-\varphi(x)}{h}-\mathcal{D}_k\varphi(x)=\frac{h}{2}\mathcal{D}_k^2\varphi(\hat{x})\,,$$

where $\hat{x}$ is between x and $x+he_k$. Thus

$$\int_Q\left|\frac{\varphi(x+he_k)-\varphi(x)}{h}-\mathcal{D}_k\varphi(x)\right|^2dx\leq(\max_D|\mathcal{D}_k^2\varphi|)^2\,\mathrm{meas}(Q)\,\frac{h^2}{4}\,.$$

Passing to the limit in (7.16) we obtain

$$\int_Q w_k\,\varphi dx=-\int_Q v\,\mathcal{D}_k\varphi dx\,,$$

namely, $\mathcal{D}_k v=w_k\in L^2(Q)$.

(iii) From part (ii) we know that the weak derivative $\mathcal{D}_k v$ exists in each subset Q with $Q\subset\subset D$ and that $\mathcal{D}_k^{-1/m}v$ converges weakly to $\mathcal{D}_k v$ in $L^2(Q)$. Since the weak derivatives are unique, by the arbitrariness of Q we deduce that the weak derivative $\mathcal{D}_k v$ exists in D and moreover it satisfies

$$\|\mathcal{D}_k v\|_{L^2(Q)}\leq\liminf_{m\to+\infty}\|\mathcal{D}_k^{-1/m}v\|_{L^2(Q)}\leq C_\sharp\,,$$

(see Yosida [28, Theorem 1, p. 120]). If we define

$$q_{k,m}=\begin{cases}(\mathcal{D}_k v)_{|D_{1/m}} & \text{in } D_{1/m}\\ 0 & \text{in } D\setminus D_{1/m}\,,\end{cases}$$

we readily see that $q_{k,m}^2\to(\mathcal{D}_k v)^2$ pointwise in D as m goes to $+\infty$ and $q_{k,m}^2$ is an increasing sequence with respect to m. Then by the Beppo Levi monotone convergence theorem it follows that $\int_{D_{1/m}}(\mathcal{D}_k v)^2dx=\int_D q_{k,m}^2dx\to\int_D(\mathcal{D}_k v)^2dx$, thus $\mathcal{D}_k v\in L^2(D)$ and $\|\mathcal{D}_k v\|_{L^2(D)}\leq C_\sharp$.

Exercise 7.12 Prove that all the terms $a_{ii}(x)$ on the diagonal of a uniformly positive definite matrix in D (namely, a matrix $\{a_{ij}(x)\}$ such that $\sum_{ij}a_{ij}(x)\eta_j\eta_i\geq\alpha_0|\eta|^2$ for all $\eta\in\mathbb{R}^n$ and almost every $x\in D$) satisfy $a_{ii}(x)\geq\alpha_0$ for almost every in $x\in D$.

Solution Take $\eta=e^{(k)}$, the k-th element of the Euclidean basis, $k=1,\ldots,n$. Then

$$\alpha_0=\alpha_0|e^{(k)}|^2\leq\sum_{ij}a_{ij}(x)e_j^{(k)}e_i^{(k)}=a_{kk}(x)\,.$$

Exercise 7.13 Under the assumptions of Theorem 7.12, the stronger estimate

$$\|u\|_{H^2(D)} \leq C\|f\|_{L^2(D)}$$

holds, provided that we know that for each $f \in L^2(D)$ there exists a unique weak solution $u \in H_0^1(D)$.

Solution Knowing that for each $f \in L^2(D)$ there exists a unique weak solution $u \in H_0^1(D)$ means that the solution operator $S_0 : f \mapsto u$ is well-defined and thus 0 is not an eigenvalue. Then, looking at Exercise 7.4, we know that $\|u\|_{L^2(D)} \leq C\|f\|_{L^2(D)}$ and therefore from Theorem 7.12 we find

$$\|u\|_{H^2(D)} \leq C\|f\|_{L^2(D)} .$$

Exercise 7.14 Prove that the Laplace operator in polar coordinates is given by

$$\Delta = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2 ,$$

and that the gradient is given by

$$\mathcal{D}_{x_1} = \cos\theta\,\partial_r - \frac{1}{r}\sin\theta\,\partial_\theta \quad , \quad \mathcal{D}_{x_2} = \sin\theta\,\partial_r + \frac{1}{r}\cos\theta\,\partial_\theta .$$

Solution Polar coordinates are given by $x_1 = r\cos\theta$, $x_2 = r\sin\theta$. Setting $\widehat{f}(r,\theta) = f(r\cos\theta, r\sin\theta)$, we have

$$\begin{aligned} \frac{\partial \widehat{f}}{\partial r} &= \frac{\partial f}{\partial x_1}\cos\theta + \frac{\partial f}{\partial x_2}\sin\theta \\ \frac{1}{r}\frac{\partial \widehat{f}}{\partial \theta} &= -\frac{\partial f}{\partial x_1}\sin\theta + \frac{\partial f}{\partial x_2}\cos\theta \end{aligned}$$

(here and in the sequel, for the sake of simplicity and with abuse of notation, we are not writing that the derivatives of f have to be computed at $(x, y) = (r\cos\theta, r\sin\theta)$). For determining $\frac{\partial f}{\partial x_1}$, multiply the first equation by $\cos\theta$ and the second one by $-\sin\theta$, and add the equations; for determining $\frac{\partial f}{\partial x_2}$, multiply the first equation by $\sin\theta$ and the second one by $\cos\theta$, and add the equations. The final result is

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= \cos\theta\,\frac{\partial \widehat{f}}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial \widehat{f}}{\partial \theta} \\ \frac{\partial f}{\partial x_2} &= \sin\theta\,\frac{\partial \widehat{f}}{\partial r} + \frac{\cos\theta}{r}\,\frac{\partial \widehat{f}}{\partial \theta} , \end{aligned}$$

hence $\mathcal{D}_1 = \cos\theta\,\partial_r - \frac{\sin\theta}{r}\,\partial_\theta$ and $\mathcal{D}_2 = \sin\theta\,\partial_r + \frac{\cos\theta}{r}\,\partial_\theta$. This permits to compute the second order derivatives, yielding

$$\frac{\partial^2 f}{\partial x_1^2} = \cos\theta\frac{\partial}{\partial r}\left(\cos\theta\,\frac{\partial \widehat{f}}{\partial r} - \frac{\sin\theta}{r}\frac{\partial \widehat{f}}{\partial \theta}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\cos\theta\,\frac{\partial \widehat{f}}{\partial r} - \frac{\sin\theta}{r}\frac{\partial \widehat{f}}{\partial \theta}\right)$$
$$\frac{\partial^2 f}{\partial x_2^2} = \sin\theta\frac{\partial}{\partial r}\left(\sin\theta\,\frac{\partial \widehat{f}}{\partial r} + \frac{\cos\theta}{r}\frac{\partial \widehat{f}}{\partial \theta}\right) + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\left(\sin\theta\,\frac{\partial \widehat{f}}{\partial r} + \frac{\cos\theta}{r}\frac{\partial \widehat{f}}{\partial \theta}\right).$$

By straightforward computations we obtain the representation of the Laplace operator in polar coordinates:

$$\Delta f = \frac{\partial^2 \widehat{f}}{\partial r^2} + \frac{1}{r}\frac{\partial \widehat{f}}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \widehat{f}}{\partial \theta^2}.$$

Exercise 7.15 Let $D \subset \mathbb{R}^3$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Show that the immersion $W^{2,2}(D) \hookrightarrow C^{0,1/2}(\overline{D})$ holds, using Theorems 7.14 and 7.15.

Solution We have that $\nabla u \in W^{1,2}(D)$, thus, by Theorem 7.14, $\nabla u \in L^6(D)$. The same holds for u, therefore we have $u \in W^{1,6}(D)$. Since $p = 6 > 3 = n$, from Theorem 7.15 it follows that the Hölder exponent is $\lambda = 1 - \frac{3}{6} = \frac{1}{2}$, thus $u \in C^{0,1/2}(\overline{D})$.

Exercise 7.16

(i) Let $D \subset \mathbb{R}^3$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Show that the bilinear form

$$B_L(w,v) = \int_D \sum_{i,j=1}^n a_{ij}\mathcal{D}_j w \mathcal{D}_i v\,dx + \int_D \sum_{i=1}^n b_i \mathcal{D}_i w v\,dx + \int_D a_0 w v\,dx$$

is bounded provided that the coefficients satisfy $a_{ij} \in L^\infty(D)$, $b_i \in L^3(D)$ and $a_0 \in L^{3/2}(D)$.

(ii) Prove that $B_L(w,v)$ is coercive in $H_0^1(D)$, $H_*^1(D)$ and $H_{\Gamma_D}^1(D)$, provided that $\|b_i\|_{L^3(D)}$, $i = 1, \dots, n$, and $\|a_0\|_{L^{3/2}(D)}$ are small enough.

Solution

(i) We have, using Hölder inequality,

$$\left|\int_D \sum_{i=1}^n b_i \mathcal{D}_i w v\,dx\right| \le \sum_{i=1}^n \int_D |b_i||\mathcal{D}_i w||v|\,dx$$
$$\le \sum_{i=1}^n \|b_i\|_{L^3(D)}\|\mathcal{D}_i w\|_{L^2(D)}\|v\|_{L^6(D)}$$

and

$$\left|\int_D a_0 w v dx\right| \le \int_D |a_0||w||v| dx \le \|a_0\|_{L^{3/2}(D)} \|w\|_{L^6(D)} \|v\|_{L^6(D)} .$$

The result follows from the Sobolev embedding Theorem 7.14.

(ii) From the Sobolev embedding Theorem 7.14 we have $\|v\|_{L^6(D)} \le C\|v\|_{H^1(D)}$; from the Poincaré inequality, that holds in all the spaces $H^1_0(D)$, $H^1_*(D)$ and $H^1_{\Gamma_D}(D)$, we have $\|v\|_{L^2(D)} \le \sqrt{C_D}\|\nabla v\|_{L^2(D)}$. Therefore it holds $\|v\|_{L^6(D)} \le C_*\|\nabla v\|_{L^2(D)}$. Then we have found

$$B_L(v,v) \ge \left[\alpha_0 - C_*\Big(\sum_{i=1}^n \|b_i\|^2_{L^3(D)}\Big)^{1/2} - C_*^2\|a_0\|_{L^{3/2}(D)}\right]\|\nabla v\|^2_{L^2(D)} ,$$

and the result follows.

Exercise 7.17 Show that the solution u of the homogeneous Dirichlet boundary value problem

$$\begin{cases} -\Delta u = 1 \text{ in } D \\ u_{|\partial D} = 0 \ \text{ on } \partial D , \end{cases}$$

where $D = \{x \in \mathbb{R}^n \mid |x| < 1\}$, belongs to $C^\infty(\overline{D})$.

Solution The coefficients of the operator and the right hand side are constant and the boundary is a C^∞-manifold, thus by the regularity result in Theorem 7.13 we see that $u \in H^{m+2}(D)$ for any $m \ge 0$. Therefore by the Sobolev embedding Theorem 7.17 we deduce $u \in C^{m+1-[n/2]}(\overline{D})$ for any $m \ge [n/2] - 1$, hence $u \in C^\infty(\overline{D})$.

Exercise 7.18 Show that the eigenvectors w_k of the homogeneous Dirichlet boundary value problem

$$\begin{cases} -\Delta w_k = \lambda_k w_k \text{ in } D \\ w_{k|\partial D} = 0 \qquad \text{on } \partial D , \end{cases}$$

where $D = \{x \in \mathbb{R}^n \mid |x| < 1\}$, belong to $C^\infty(\overline{D})$.

Solution The coefficients of the operator are constant and the boundary is a C^∞-manifold; moreover, we can consider $\lambda_k w_k$ as a right hand side for the Laplace operator. Since the variational solution w_k belongs to $H^1(D)$, by the regularity result in Theorem 7.13 we see that $w_k \in H^3(D)$. Now we can apply a bootstrap argument: from what we have just proved, the right hand side $\lambda_k w_k$ belongs to $H^3(D)$. Therefore we apply once again Theorem 7.13 and we find $w_k \in H^5(D)$. By iterating this procedure, we see that $w_k \in H^m(D)$ for any $m \ge 0$. Therefore by the Sobolev embedding Theorem 7.17 we deduce $u \in C^{m-1-[n/2]}(\overline{D})$ for any $m \geq [n/2]+1$, hence $u \in C^\infty(\overline{D})$.

Exercise 7.19 Let $D \subset \mathbb{R}^3$ be a bounded, connected and open set, with a Lipschitz continuous boundary ∂D. Let V be a closed subspace of $H^1(D)$, and let the assumptions of Theorem 7.18 be satisfied. Suppose moreover that for each $w \in C^0(\overline{D})$ one can find $\pi_N(w) \in V_N$ such that $\|w - \pi_N(w)\|_V \to 0$ as $N \to \infty$. Then show that the Galerkin method is convergent.

Solution Let $u \in V$ be the exact solution of the problem. By the approximation Theorem 6.3 for each $\epsilon > 0$ we can find $u_* \in C^\infty(\overline{D})$ such that $\|u - u_*\|_V \leq \epsilon$. Thus, using Theorem 7.18, we have

$$\begin{aligned}\|u - u_N\|_V &\leq \frac{\gamma}{\alpha} \inf_{v_N \in V_N} \|u - v_N\|_V \leq \frac{\gamma}{\alpha}\|u - \pi_N(u_*)\|_V \\ &\leq \frac{\gamma}{\alpha}(\|u - u_*\|_V + \|u_* - \pi_N(u_*)\|_V) \leq 2\,\frac{\gamma}{\alpha}\,\epsilon\end{aligned}$$

for N large enough.

Exercise 7.20 Let X be a Hilbert space with scalar product $(\cdot,\cdot)_X$ and let $\{\varphi_m\}$, $m \geq 1$, be an orthonormal Hilbertian basis of X. Prove that $\varphi_m \to 0$ weakly in X, thus furnishing an example of a sequence which is weakly convergent in X but not convergent in X.

Solution Since $\{\varphi_m\}$ is an orthonormal Hilbertian basis of X, for each $v \in X$ we have the Fourier expansion

$$v = \sum_{m=1}^{\infty} (v, \varphi_m)_X \varphi_m \quad , \quad \|v\|_X^2 = \sum_{m=1}^{\infty} (v, \varphi_m)_X^2 \, .$$

Being the series $\sum_{m=1}^{\infty} (v, \varphi_m)_X^2$ convergent, we have at once $(v, \varphi_m)_X \to 0$.

Exercise 7.21 Let X be a Hilbert space with scalar product $(\cdot,\cdot)_X$. Prove that $v_m \to v$ in X if and only if $v_m \to v$ weakly in X and $\|v_m\|_X \to \|v\|_X$.

Solution Suppose that $v_m \to v$ in X: then for any $w \in X$ we have $(v_m, w)_X \to (v, w)_X$ by the Cauchy–Schwarz inequality: moreover, $\|v_m\|_X \to \|v\|_X$ by the triangular inequality.

Vice versa, it holds

$$\|v_m - v\|_X^2 = (v_m - v, v_m - v)_X = \|v_m\|_X^2 - 2(v_m, v)_X + \|v\|_X^2 \, ;$$

since $(v_m, w)_X \to (v, w)_X$ for each $w \in X$, it follows $(v_m, v)_X \to \|v\|_X^2$ and thus $\|v_m - v\|_X^2 \to 0$.

Exercise 7.22 Let X and Y be Hilbert spaces and $K : X \mapsto Y$ a linear and compact operator. Prove that if $u_j \to u$ weakly in X then $Ku_j \to Ku$ in Y.

Solution Being weakly convergent, the sequence u_j is bounded in X (see, e.g., Yosida [28, Theorem 1, p. 120]). Therefore, due to the compactness of K, from each subsequence u_{j_s} of u_j we can extract another subsequence, denoted by $u_{j_{s_m}}$, such that $Ku_{j_{s_m}}$ converges to an element $\omega_\star \in Y$. Then for each $v \in Y$ we have

$$(\omega_\star - Ku, v)_Y = \lim_m (Ku_{j_{s_m}} - Ku, v)_Y = \lim_m (u_{j_{s_m}} - u, K^T v)_X = 0\,,$$

where K^T is the adjoint operator of K, and thus $Ku = \omega_\star$. Hence from any subsequence Ku_{j_s} we have extracted another subsequence $Ku_{j_{s_m}}$ which converges to Ku, and this limit is the same for all the possible choices of the subsequence Ku_{j_s}. This implies that the whole sequence Ku_j converges to Ku in Y.

Chapter 8
Saddle Points Problems

This chapter is devoted to the solution of saddle point problems that can be written in the abstract form

$$\begin{cases} Au + B^T \lambda = F \\ Bu = G \end{cases}$$

for some linear operators A and B, λ having the role of a Lagrangian multiplier associated to the constraint $Bu = G$.

The first section, concerned with constrained minimization, is divided into two parts: the finite dimensional case and the infinite dimensional case. Then we describe and analyze the Galerkin approximation method for saddle point problems, and finally we present some issues of the Galerkin method based on finite elements.

8.1 Constrained Minimization

This section is divided into two parts, regarding the finite dimensional and the infinite dimensional case, respectively. We chose this approach as we believe that the leading ideas are more easily caught when dealing with vectors. In this way we hope that the process of extending known results of finite dimensional linear algebra to the infinite dimensional case can become an easier task.

8.1.1 The Finite Dimensional Case

Let us start from a problem in $\mathbb{R}^n$. We have a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ and we want to minimize it subject to a set of constraints, expressed by $g(x) = 0$, with $g : \mathbb{R}^n \mapsto$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_8

$\mathbb{R}^m$, with $m < n$. If $m = 1$ and $\nabla g \neq 0$ on $\{g(x) = 0\}$, we know that at a minimum point $\hat{x}$ we must have

$$\nabla f(\hat{x}) = \lambda \nabla g(\hat{x}) ,$$

where $\lambda \in \mathbb{R}$ is a Lagrange multiplier. If $1 < m < n$ and ∇g_k are linearly independent on $\{g(x) = 0\}$, we know that at a minimum point $\hat{x}$ we must have

$$\nabla f(\hat{x}) = \sum_{k=1}^{m} \lambda_k \nabla g_k(\hat{x}) ,$$

where $\lambda_k \in \mathbb{R}$, $k = 1, \dots, m$, are Lagrange multipliers.

In other words, we can look for the stationary points (i.e., the points where the gradient vanishes) of the Lagrangian

$$\mathcal{L}(w, \mu) = f(w) + \sum_{k=1}^{m} \mu_k g_k(w) ;$$

clearly, we mean stationary points related to derivatives with respect to all the components of w and μ.

Suppose now we have a quadratic function

$$f(w) = \frac{1}{2}(Aw, w) - (F, w) ,$$

where A is a $n \times n$ matrix and $F \in \mathbb{R}^n$ and we denote by $(\cdot, \cdot)$ the scalar product in $\mathbb{R}^n$. Let us also consider linear (indeed, affine) constraints

$$g(w) = Bw - G ,$$

where B is an $m \times n$ full-rank matrix and $G \in \mathbb{R}^m$. Assuming that A is symmetric, it is well-known that the problem

$$\min_{w \in \mathbb{R}^n,\, g(w)=0} f(w) \tag{8.1}$$

can be interpreted in a suitable matrix form.

Theorem 8.1 *Suppose that A is a symmetric matrix. Let $u \in \mathbb{R}^n$ be a solution of problem (8.1). Then there exists $\lambda \in \mathbb{R}^m$ such that the couple (u, λ) is a solution to*

$$\begin{cases} Au - F + B^T \lambda = 0 \\ Bu - G = 0 . \end{cases} \tag{8.2}$$

Proof As explained above, it is enough to take the derivatives of the Lagrangian

$$\mathcal{L}(w,\mu) = \frac{1}{2}(Aw, w) - (F, w) + \sum_{k=1}^{m} \mu_k (Bw - G)_k\,.$$

Taking the derivative with respect to w_j we obtain

$$\begin{aligned}
\frac{\partial}{\partial w_j}\left[\frac{1}{2}\sum_{i,s=1}^{n} A_{is} w_s w_i - \sum_{s=1}^{n} F_s w_s\right] &= \frac{1}{2}\sum_{i,s=1}^{n} A_{is}\frac{\partial}{\partial w_j}(w_s w_i) - \sum_{s=1}^{n} F_s \frac{\partial w_s}{\partial w_j} \\
&= \frac{1}{2}\sum_{i,s=1}^{n} A_{is}(\delta_{sj} w_i + w_s \delta_{ij}) - \sum_{s=1}^{n} F_s \delta_{sj} \\
&= \frac{1}{2}\Big(\sum_{i=1}^{n} \underbrace{A_{ij}}_{=A^T_{ji}} w_i + \sum_{s=1}^{n} A_{js} w_s\Big) - F_j = \Big(\frac{A^T + A}{2} w - F\Big)_j
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial w_j}\left[\sum_{k=1}^{m} \mu_k (Bw - G)_k\right] &= \frac{\partial}{\partial w_j}\left[\sum_{k=1}^{m} \mu_k \left(\sum_{s=1}^{n} B_{ks} w_s - G_k\right)\right] \\
&= \sum_{k=1}^{m} \mu_k \sum_{s=1}^{n} B_{ks} \frac{\partial w_s}{\partial w_j} = \sum_{k=1}^{m} \mu_k B_{kj} = (B^T \mu)_j\,.
\end{aligned}$$

Differentiating with respect to μ_l, $l = 1, \dots, m$, it easily follows

$$\frac{\partial \mathcal{L}}{\partial \mu_l}(w,\mu) = \frac{\partial}{\partial \mu_l}\left(\sum_{k=1}^{m} \mu_k (Bw - G)_k\right) = (Bw - G)_l\,.$$

Therefore the Euler equations of the Lagrangian $\mathcal{L}$ are

$$\begin{cases} \frac{A^T + A}{2} w - F + B^T \mu = 0 \\ Bw - G = 0\,, \end{cases} \tag{8.3}$$

and, having assumed that the matrix A is symmetric, a stationary point (u, λ) of $\mathcal{L}$ satisfies problem (8.2). □

We can also show that problems (8.2) and (8.1) are indeed equivalent (provided that A is not only symmetric but also non-negative definite). In fact, it holds:

Theorem 8.2 *Suppose that A is a symmetric and non-negative definite matrix. A solution (u, λ) to (8.2) furnishes a solution u of the minimization problem (8.1).*

Proof Take v such that $g(v) = 0$, namely $Bv = G$. Then it can be written as $v = u + w$, with $Bw = 0$. We have

$$
\begin{aligned}
\frac{1}{2}(Av, v) - (F, v) &= \frac{1}{2}(A(u+w), u+w) - (F, u+w) \\
&= \frac{1}{2}(Au, u) + (Au, w) + \frac{1}{2}(Aw, w) - (F, u) - (F, w) \quad (A \text{ is symmetric}) \\
&= \frac{1}{2}(Au, u) - (F, u) - (\underbrace{B^T\lambda}_{=F-Au}, w) + \frac{1}{2}(Aw, w) \\
&= \frac{1}{2}(Au, u) - (F, u) - (\lambda, \underbrace{Bw}_{=0}) + \frac{1}{2}\underbrace{(Aw, w)}_{\geq 0} \geq \frac{1}{2}(Au, u) - (F, u),
\end{aligned}
$$

thus u solves the minimization problem (8.1). □

We can give some additional information on the stationary point (u, λ) of the Lagrangian $\mathcal{L}$. In fact we have:

Proposition 8.1 *Suppose that A is a symmetric and non-negative definite matrix. A solution (u, λ) of (8.2) is a saddle point of the Lagrangian*

$$
\mathcal{L}(w, \mu) = \frac{1}{2}(Aw, w) - (F, w) + \sum_{k=1}^{m} \mu_k (Bw - G)_k ,
$$

i.e., it satisfies

$$
\mathcal{L}(u, \eta) \leq \mathcal{L}(u, \lambda) \leq \mathcal{L}(v, \lambda) \tag{8.4}
$$

for each $v \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^m$.

Proof Writing $v = u + w$, we have for each $w \in \mathbb{R}^n$,

$$
\begin{aligned}
\mathcal{L}(u+w, \lambda) &= \frac{1}{2}(A(u+w), u+w) - (F, u+w) + \sum_{k=1}^{m} \lambda_k (B(u+w) - G)_k \\
&= \underbrace{\frac{1}{2}(Au, u)}_{\star} + (Au, w) + \frac{1}{2}(Aw, w) - \underbrace{(F, u)}_{\star} - (F, w) \\
&\qquad + \underbrace{\sum_{k=1}^{m} \lambda_k (Bu - G)_k}_{\star} + \sum_{k=1}^{m} \lambda_k (Bw)_k \qquad (A \text{ is symmetric}) \\
&= \underbrace{\mathcal{L}(u, \lambda)}_{\star} + (Au - F, w) + \frac{1}{2}(Aw, w) + \sum_{k=1}^{m} \lambda_k \sum_{s=1}^{n} B_{ks} w_s
\end{aligned}
$$

$$= \mathcal{L}(u,\lambda) + (Au - F, w) + \frac{1}{2}(Aw, w) + \sum_{s=1}^{n} w_s \sum_{k=1}^{m} \underbrace{B_{ks}}_{=B^T_{sk}} \lambda_k$$

$$= \mathcal{L}(u,\lambda) + (\underbrace{Au - F + B^T\lambda}_{=0}, w) + \frac{1}{2}\underbrace{(Aw, w)}_{\geq 0} \geq \mathcal{L}(u,\lambda)\,.$$

Moreover, for each $\eta \in \mathbb{R}^m$

$$\mathcal{L}(u,\eta) = \frac{1}{2}(Au, u) - (F, u) + \sum_{k=1}^{m} \eta_k \underbrace{(Bu - G)_k}_{=0}$$

$$= \frac{1}{2}(Au, u) - (F, u) = \mathcal{L}(u,\lambda)\,,$$

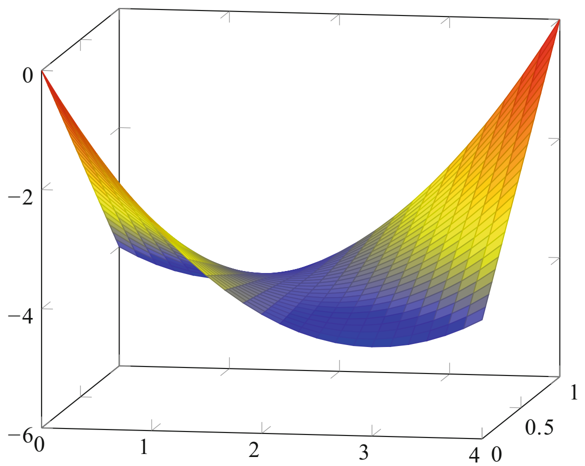
and (8.4) is completely proved. □

Example 8.1 In order to show, by means of a figure, the saddle point structure of a constrained minimization problem like those we are considering, let us take $n = 1$, $m = 1$, $A = 1$, $B = 2$, $F = 3$ and $G = 4$. This leads to the Lagrangian $\mathcal{L}(w,\mu) = \frac{1}{2}w^2 - 3w + \mu(2w - 4)$. The graph of this function is drawn in Fig. 8.1, where it can be possible to recognize that $(2, \frac{1}{2})$ is a saddle point, and that $w \to \mathcal{L}(w, \frac{1}{2})$ has a minimum at $w = 2$, while $\mu \to \mathcal{L}(2,\mu)$ is constant.

We are now in a position to prove the well-posedness of problem (8.2).

Theorem 8.3 *Suppose that A is a positive definite matrix and that $N(B^T) = \{0\}$. Then (8.2) has a unique solution.*

Fig. 8.1 The graph of the Lagrangian $\mathcal{L}(w,\mu) = \frac{1}{2}w^2 - 3w + \mu(2w - 4)$



Proof For a finite dimensional linear problem existence and uniqueness are equivalent. Let us prove the uniqueness, namely, let us show that if $F = 0$ and $G = 0$ in (8.2) we obtain $u = 0$ and $\lambda = 0$. Take the scalar product of the first equation by u:

$$
\begin{aligned}
0 &= (Au, u) + (B^T\lambda, u) = (Au, u) + (\lambda, \underbrace{Bu}_{=0}) \\
&= (Au, u) \Longrightarrow u = 0 \quad \text{(as } A \text{ is positive definite)}.
\end{aligned}
$$

Since $u = 0$, we have $B^T\lambda = 0$, then the assumption $N(B^T) = \{0\}$ gives $\lambda = 0$. □

Remark 8.1 The condition $N(B^T) = \{0\}$ is necessary for uniqueness. If we had $B^T\eta_* = 0$ for $\eta_* \neq 0$, from a solution (u, λ) of (8.2) we could construct another solution $(u, \lambda + \eta_*)$.

Remark 8.2 The symmetry of A is not needed in this theorem. On the other hand, it has been used to show that the solution of the minimization problem (8.1) is a solution to (8.2) and viceversa (see Theorems 8.1 and 8.2).

Remark 8.3 Giving a deeper look at the proof, we see that it is possible to weaken a little bit the assumption on A. In fact, the proof of the theorem also works if we only assume that

$$
(Aw, w) = 0 \text{ for } w \text{ with } Bw = 0 \text{ implies } w = 0 \,.
$$

8.1.2 *The Infinite Dimensional Case*

Before entering the problem of how we can extend Theorem 8.3 to Hilbert spaces having infinite dimension, let pose the following question: in the infinite dimensional case, do we encounter problems with a structure like (8.2)?

Example 8.2 Consider the Stokes problem

$$
\begin{cases}
-\nu\Delta u + \nabla p = f & \text{in } D \\
\operatorname{div} u = 0 & \text{in } D \\
u = 0 & \text{on } \partial D\,,
\end{cases}
\tag{8.5}
$$

where u is the velocity of a fluid, p is the pressure (indeed, the pressure divided by the density), $\nu > 0$ a constant (the kinematic viscosity) and f is the acceleration of the external forces. The constraint $\operatorname{div} u = 0$ represents the incompressibility of the fluid.

We know that formally ∇ is the adjoint operator of $-\text{div}$:

$$\int_D \nabla\varphi \cdot v dx = -\int_D \varphi \,\text{div}\, v dx \qquad \text{for } \varphi \in C_0^\infty(D)\,,\ v \in C_0^\infty(D)\,.$$

Then if we call $A = -\nu\Delta$ (Δ being the Laplace operator acting on vector functions, associated with the homogeneous Dirichlet boundary condition) and $B = -\text{div}$ (so that $B^T = \nabla$), we rewrite the Stokes problem as

$$\begin{cases} Au + B^T p = f \\ Bu = 0\,. \end{cases}$$

Example 8.3 Consider the elliptic operator (without the first order and zero order terms)

$$L\varphi = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j\varphi)$$

and define

$$q_i = -\sum_{j=1}^{n} a_{ij}\mathcal{D}_j\varphi\ ,\ i = 1,\ldots,n\,.$$

Then the problem

$$\begin{cases} L\varphi = g & \text{in } D \\ \varphi = 0 & \text{on } \partial D \end{cases}$$

can be rewritten

$$\begin{cases} q_i + \sum_{j=1}^n a_{ij}\mathcal{D}_j\varphi = 0 & \text{in } D\ ,\ i = 1,\ldots,n \\ \sum_{i=1}^n \mathcal{D}_i q_i = g & \text{in } D \\ \varphi = 0 & \text{on } \partial D\,. \end{cases}$$

Due to the ellipticity assumption we know that the matrix $\{a_{ij}\}$ is (uniformly) positive definite, hence non-singular. If we define $Z = \{z_{ij}\}$ its inverse matrix, which is also positive definite, we have, since $\sum_{j=1}^n z_{ij}a_{js} = \delta_{is}$,

$$\sum_{j=1}^{n} z_{ij}q_j + \mathcal{D}_i\varphi = 0 \quad \text{in } D\ ,\ i = 1,\ldots,n\,.$$

Thus we have finally rewritten the problem as a first order elliptic system:

$$\begin{cases} Z\,q + \nabla\varphi = 0 & \text{in } D \\ -\mathrm{div}\, q = -g & \text{in } D \\ \varphi = 0 & \text{on } \partial D\,. \end{cases} \tag{8.6}$$

In this case the operator A is not a differential operator, but simply $Aq = Zq$, where the matrix Z has entries $\{z_{ij}\}$. Instead, as before, the operator B is $-\mathrm{div}$ and $B^T = \nabla$.

We want to extend to infinite dimensional Hilbert spaces the results in Theorem 8.3; in particular we want to devise which sufficient conditions will take the place of those appearing there.

Let us present the abstract theory that covers both cases (8.5) and (8.6). It can be described in two equivalent ways. In the first one we are given with two bounded bilinear forms $a : V \times V \mapsto \mathbb{R}$ and $b : V \times M \mapsto \mathbb{R}$, where V and M are two Hilbert spaces. Clearly, these two forms define two linear and bounded operators $A : V \mapsto V'$, $B : V \mapsto M'$, where V' and M' are the dual spaces of V and M, respectively, namely, the space of linear and bounded operators from V to $\mathbb{R}$ and from M to $\mathbb{R}$, respectively. This is done as follows: for each $w \in V$ we define

$$\begin{aligned} &Aw \ \text{ is the map } \ v \mapsto a(w,v) \quad \forall\, v \in V \\ &Bw \ \text{ is the map } \ \psi \mapsto b(w,\psi) \quad \forall\, \psi \in M\,; \end{aligned}$$

in this way $B^T : M \mapsto V'$ is defined by saying that, for each $\mu \in M$, $B^T\mu$ is the map $v \mapsto b(v,\mu)$ for all $v \in V$.

The other way around is described by starting from two linear and bounded operators $A : V \mapsto V'$ and $B : V \mapsto M'$, and introducing two bilinear and bounded forms $a : V \times V \mapsto \mathbb{R}$ and $b : V \times M \mapsto \mathbb{R}$ by setting

$$\begin{aligned} &a(w,v) = \langle Aw, v\rangle \qquad \forall\, w, v \in V \\ &b(w,\psi) = \langle Bw, \psi\rangle \qquad \forall\, w \in V,\ \psi \in M\,, \end{aligned}$$

where $\langle\cdot,\cdot\rangle$ are the duality pairings between V and V' and M and M' (we use the same notation for both of them, and the specific context will permit to identify which duality pairing is considered). As a consequence, one can also see that $B^T : M \mapsto V'$ is defined as

$$\langle B^T\mu, v\rangle = b(v,\mu) = \langle Bv, \mu\rangle \qquad \forall\, \mu \in M,\ v \in V\,.$$

We will present and analyze the problem in terms of the operators A, B and B^T.

Before going on, a clearer picture of the situation in the infinite dimensional case can come from a more direct proof of the existence of a solution to problem (8.2). We can devise a procedure that have three steps, as described here below.

1. Find a solution $u_G \in \mathbb{R}^n$ of $Bu_G = G$: this requires that the range of B, namely, the space $R(B) = \{\mu \in \mathbb{R}^m \,|\, \exists\, v \in \mathbb{R}^n$ such that $\mu = Bv\}$, satisfies $R(B) = \mathbb{R}^m$.
2. Find $\hat{u} \in \mathbb{R}^n$ solution to

$$\begin{cases} A\hat{u} = -B^T\lambda + F - Au_G \\ B\hat{u} = 0 \, . \end{cases}$$

This would require the knowledge of λ. However, if we project the first equation on the kernel $N(B)$ we find that

$$\hat{u} \in N(B) \;:\; (A\hat{u}, v) = \underbrace{-(B^T\lambda, v)}_{=-(\lambda, Bv)=0} + (F - Au_G, v) \qquad \forall\, v \in N(B)\,,$$

a problem where λ is no longer present. For solvability, here a sufficient assumption is that A is positive definite on $N(B)$.
3. Find a solution $\lambda \in \mathbb{R}^m$ to

$$B^T\lambda = F - Au_G - A\hat{u}\,.$$

Here we have, by the second step, $(F - Au_G - A\hat{u}, v) = 0$ for all $v \in N(B)$, therefore the needed property is that $R(B^T) = N(B)^\perp$.

In the finite dimensional case we know that the property $R(B^T) = N(B)^\perp$ is always satisfied, as well as $R(B) = N(B^T)^\perp$ (see Exercise 7.2). Thus the existence of a solution to problem (8.2) follows by assuming that A is positive definite on $N(B)$ and that $N(B^T) = \{0\}$, so that $R(B) = N(B^T)^\perp = \mathbb{R}^m$.

In this respect, the situation at the infinite dimensional level is somehow different. First, for a linear and bounded operator $K : X \mapsto Y$, X and Y Hilbert spaces, it is no longer true that $R(K) = N(K^T)^\perp$, as in general the range $R(K)$ is not a *closed* subspace in Y (see Sect. 3.1, item 5, and Exercise 7.3; in particular, in the latter it is proved that $R(K)^\perp = N(K^T)$ and $R(K) \subset \overline{R(K)} = (R(K)^\perp)^\perp = N(K^T)^\perp$, thus the equality in this last relation is true if and only if $R(K)$ is closed in Y). Moreover, here we have to deal with operators $B : V \mapsto M'$ and $B^T : M \mapsto V'$, V' and M' being the *dual* spaces of V and M, respectively, and it is more suitable to focus in a more precise way on this specific situation.

Thus we start with a definition.

Definition 8.1 The polar set of $N(B)$ is

$$N(B)_\sharp = \{g \in V' \,|\, \langle g, v\rangle = 0 \quad \forall\, v \in N(B)\}\,.$$

As seen in Exercise 8.1, $N(B)_\sharp$ can be identified with a suitable dual space.

Exercise 8.1 $N(B)_\sharp$ can be isometrically identified with the dual of $N(B)^\perp$.

We are now in a position to "translate" conditions 1, 2 and 3 for the infinite dimensional case. With respect to condition 2, when considering the Lax–Milgram theorem 2.1 we have already seen that a natural extension of the assumption that the matrix A is positive definite is that the operator $A : V \mapsto V'$ is coercive, namely, there exists $\alpha > 0$ such that $\langle Av, v\rangle \geq \alpha\|v\|_V^2$ for all $v \in V$. However, we have seen in Remark 8.3 that in the present case it could be sufficient to assume that coerciveness is satisfied only in the kernel of B, namely, it holds $\langle Av, v\rangle \geq \alpha\|v\|_V^2$ for all $v \in N(B) = \{v \in V \mid Bv = 0\}$.

A remark is in order about condition 3: since the operator A takes values in the dual space V', the relation $R(B^T) = N(B)^\perp$ clearly has to be replaced by $R(B^T) = N(B)_\sharp$.

Conditions 1 and 3 are strictly related. In fact, by a suitable version of the closed range theorem (see Yosida [28, Theorem 1, p. 205]) we know that

Theorem 8.4 (Closed Range) *Let $B : V \mapsto M'$ be a linear and bounded operator, where V and M are Hilbert spaces and M' is the dual space of M. Denote by $B^T : M \mapsto V'$ the adjoint operator of B, V' being the dual space of V. Then*

(i) The range $R(B)$ is closed in M' if and only if the range $R(B^T)$ is closed in V'.
(ii) The range $R(B)$ is closed in M' if and only if $R(B) = N(B^T)_\sharp$.
(iii) The range $R(B^T)$ is closed in V' if and only if $R(B^T) = N(B)_\sharp$.

It is now easy to see that, for repeating the finite dimensional existence procedure, it is sufficient to assume that A is coercive on $N(B)$, $N(B^T) = \{0\}$ and $R(B^T)$ is closed in V'. In fact, in this case from (i) we have that $R(B)$ is closed in M', hence from (ii) we see that $R(B) = N(B^T)_\sharp = M'$ and finally from (iii) we obtain $R(B^T) = N(B)_\sharp$. Moreover, from the coerciveness of A in $N(B)$ and $N(B^T) = \{0\}$ it follows that the solution is unique.

To this end, the key point is the following result.

Proposition 8.2 *Suppose that there exists $\beta > 0$ such that*

$$\forall\, \mu \in M\ \exists\, v_\mu \in V, v_\mu \neq 0\ :\ \langle B^T\mu, v_\mu\rangle \geq \beta\|\mu\|_M\|v_\mu\|_V\,. \tag{8.7}$$

Then $N(B^T) = \{0\}$ and $R(B^T)$ is closed in V'.

Proof Condition (8.7) clearly says that $N(B^T) = \{0\}$. Moreover, in Theorem 2.1 we have already presented an argument that shows that $R(B^T)$ is closed in V'. Let us repeat it here for the ease of the reader. From (8.7) we see that for all $\mu \in M$ it holds

$$\|B^T\mu\|_{V'} = \sup_{v\in V, v\neq 0} \frac{\langle B^T\mu, v\rangle}{\|v\|_V} \geq \frac{\langle B^T\mu, v_\mu\rangle}{\|v_\mu\|_V} \geq \beta\|\mu\|_M\,. \tag{8.8}$$

Suppose that $B^T\mu_k \to \varphi$ in V', thus $B^T\mu_k$ is a Cauchy sequence in V' and by condition (8.8) μ_k is a Cauchy sequence in M. Since M is a Hilbert space we find $\mu_k \to \mu_0$ in M and by the continuity of B^T it follows $B^T\mu_k \to B^T\mu_0$, hence $\varphi = B^T\mu_0$. □

Remark 8.4 Condition (8.7) is called *inf–sup condition* since it can be rewritten as

$$\inf_{\mu\in M,\mu\neq 0}\left(\frac{1}{\|\mu\|_M}\sup_{v\in V,v\neq 0}\frac{\langle B^T\mu, v\rangle}{\|v\|_V}\right) = \inf_{\mu\in M,\mu\neq 0}\sup_{v\in V,v\neq 0}\frac{\langle B^T\mu, v\rangle}{\|\mu\|_M\|v\|_V} \geq \beta > 0\,.$$

Exercise 8.2 The inf–sup condition (8.7) is equivalent to each one of the following conditions:

(a) The operator B^T is an isomorphism from M onto $N(B)_\sharp$ and

$$\exists\, \beta > 0 \,:\, \|B^T\mu\|_{V'} \geq \beta\|\mu\|_M \qquad \forall\, \mu \in M\,.$$

(b) The operator B is an isomorphism from $N(B)^\perp$ onto M' and

$$\exists\, \beta > 0 \,:\, \|Bv\|_{M'} \geq \beta\|v\|_V \qquad \forall\, v \in N(B)^\perp\,.$$

For the solution of Exercise 8.2 it is useful to use the following result:

Exercise 8.3 Let V be a Hilbert space and $F \in V'$. Show that the norm $\|F\|_{V'}$ defined as

$$\|F\|_{V'} = \sup_{v\in V,v\neq 0}\frac{\langle F, v\rangle}{\|v\|_V}$$

is indeed equal to

$$\|F\|_{V'} = \max_{v\in V,v\neq 0}\frac{\langle F, v\rangle}{\|v\|_V}\,,$$

namely, there is $v_F \in V$, $v_F \neq 0$, such that

$$\|F\|_{V'} = \frac{\langle F, v_F\rangle}{\|v_F\|_V}\,.$$

We are now in a position to prove the existence and uniqueness theorem we are interested in. The problem reads: for each $F \in V'$, $G \in M'$, find a unique solution $(u, \varphi) \in V \times M$ of

$$\begin{cases} Au + B^T\varphi = F \\ Bu = G\,. \end{cases} \tag{8.9}$$

Theorem 8.5 *Let A be a linear and bounded operator from V to V', with $\|A\| = \gamma$. Let B be a linear and bounded operator from V in M'. Assume that the operator A is coercive over the kernel of the operator B, namely,*

$$\exists\, \alpha > 0 \text{ such that } \langle Av, v\rangle \geq \alpha \|v\|_V^2 \quad \forall\, v \in N(B)\,, \tag{8.10}$$

and that the inf–sup condition (8.7) is satisfied, namely,

$$\exists\, \beta > 0 \text{ such that } \forall\, \mu \in M\ \exists\, v_\mu \in V,\ v_\mu \neq 0\ :\ \langle B^T \mu, v_\mu\rangle \geq \beta \|\mu\|_M \|v_\mu\|_V\,. \tag{8.11}$$

Then there exists a unique solution (u, φ) to (8.9). Moreover

$$\|u\|_V \leq \frac{1}{\alpha}\|F\|_{V'} + \frac{1}{\beta}\left(1 + \frac{\gamma}{\alpha}\right)\|G\|_{M'}$$

$$\|\varphi\|_M \leq \frac{1}{\beta}\left(1 + \frac{\gamma}{\alpha}\right)\|F\|_{V'} + \frac{\gamma}{\beta^2}\left(1 + \frac{\gamma}{\alpha}\right)\|G\|_{M'}\,.$$

Proof Uniqueness is easy: from $F = 0$ and $G = 0$ it follows $Bu = 0$ and from the first equation we get

$$0 = \langle Au, u\rangle + \langle B^T\varphi, u\rangle = \langle Au, u\rangle + \langle \varphi, \underbrace{Bu}_{=0}\rangle\,,$$

thus $u = 0$ from condition (8.10), as $u \in N(B)$. Hence it follows $B^T\varphi = 0$ and, taking $\mu = \varphi$ in condition (8.11), we obtain $\|\varphi\|_M \|v_\varphi\|_V = 0$ for $v_\varphi \neq 0$, thus $\varphi = 0$.

Now, from Proposition 8.2 and Theorem 8.4 we know that $R(B) = M'$, thus we find $u_G \in N(B)^\perp$ such that $Bu_G = G$ and moreover

$$\|u_G\|_V \leq \frac{1}{\beta}\|G\|_{M'}$$

(see Exercise 8.2 (b)). Then we rewrite problem (8.9) as

$$\begin{cases} A\hat{u} + B^T\varphi = F - Au_G \\ B\hat{u} = 0\,, \end{cases} \tag{8.12}$$

with $\hat{u} = u - u_G$. Taking the pairing with $v \in N(B)$, we can eliminate φ: we find

$$\langle F - Au_G, v\rangle = \langle A\hat{u}, v\rangle + \langle B^T\varphi, v\rangle = \langle A\hat{u}, v\rangle + \langle \varphi, \underbrace{Bv}_{=0}\rangle = \langle A\hat{u}, v\rangle\,.$$

Since we look for $\hat{u} \in N(B)$, we can apply the Lax–Milgram theorem 2.1 in $N(B)$, where A is coercive by condition (8.10). Then we have a unique solution $\hat{u} \in N(B)$ of

$$\langle A\hat{u} + Au_G - F, v\rangle = 0 \qquad \forall\, v \in N(B)\,,$$

satisfying

$$\|\hat{u}\|_V \le \frac{1}{\alpha}\|F - Au_G\|_{V'}\,.$$

Setting $u = \hat{u} + u_G$, we have that

$$\langle Au - F, v\rangle = 0 \qquad \forall\, v \in N(B)\,,$$

thus $(Au - F) \in N(B)_\sharp$. From Proposition 8.2 and Theorem 8.4 there exists a unique $\varphi \in M$ such that

$$B^T\varphi = F - Au\,,$$

and estimate (8.8) holds, i.e.,

$$\begin{aligned}\|\varphi\|_M &\le \frac{1}{\beta}\|B^T\varphi\|_{V'} = \frac{1}{\beta}\|Au - F\|_{V'} \le \frac{1}{\beta}\left(\|Au\|_{V'} + \|F\|_{V'}\right)\\ &\le \frac{\gamma}{\beta}\|u\|_V + \frac{1}{\beta}\|F\|_{V'}\,.\end{aligned}$$

Thus (u, φ) is a solution to problem (8.9). Moreover we have

$$\begin{aligned}\|u\|_V &\le \|\hat{u}\|_V + \|u_G\|_V \le \frac{1}{\alpha}\|F - Au_G\|_{V'} + \|u_G\|_V\\ &\le \frac{1}{\alpha}\|F\|_{V'} + \left(1 + \frac{\gamma}{\alpha}\right)\|u_G\|_V \le \frac{1}{\alpha}\|F\|_{V'} + \frac{1}{\beta}\left(1 + \frac{\gamma}{\alpha}\right)\|G\|_{M'}\,.\end{aligned}$$

Concerning φ, we easily obtain

$$\|\varphi\|_M \le \frac{1}{\beta}\|F\|_{V'} + \frac{\gamma}{\beta}\|u\|_V \le \frac{1}{\beta}\left(1 + \frac{\gamma}{\alpha}\right)\|F\|_{V'} + \frac{\gamma}{\beta^2}\left(1 + \frac{\gamma}{\alpha}\right)\|G\|_{M'}\,,$$

which ends the proof. □

Exercise 8.4 Give a proof of the Lax–Milgram theorem 2.1 based on the closed range theorem 8.4.

Let us come back now to our Examples 8.2 and 8.3. We want to show that they can be written in the general form we have described in Theorem 8.5. The first step is the identification of the variational spaces: in case (8.5) we take $u \in V = (H_0^1(D))^n$, so that each component of the velocity vector u belongs to $H_0^1(D)$, and $p \in M \subset L^2(D)$ (M yet to be determined). The reason of this choice is that integrating by parts we obtain

$$\int_D (-\nu\Delta u)\cdot v dx = \int_D -\nu \sum_{k,s=1}^n (\mathcal{D}_s\mathcal{D}_s u_k) v_k dx$$
$$= \int_D \nu \sum_{k,s=1}^n \mathcal{D}_s u_k \mathcal{D}_s v_k dx - \int_{\partial D} \nu \sum_{k,s=1}^n \mathcal{D}_s u_k n_s \underbrace{v_k}_{=0} dS_x \,,$$

and the last integral vanishes if $v \in (H_0^1(D))^n$. Moreover

$$\int_D \nabla p \cdot v dx = -\int_D p \,\mathrm{div}\, v dx + \int_{\partial D} pn \cdot \underbrace{v}_{=0} dS_x \,,$$

and again the last integral vanishes if $v \in (H_0^1(D))^n$, while the first integral has a meaning for $p \in L^2(D)$.

Concerning the second equation $\mathrm{div}\, u = 0$ in D, it is easily seen that it can be simply written in weak form as

$$\int_D (\mathrm{div}\, u)\, r dx = 0 \quad \text{for each } r \in L^2(D)\,.$$

However, here it is worthy to note that, by the divergence theorem C.3, $\int_D \mathrm{div}\, v dx = \int_{\partial D} v \cdot n dS_x = 0$ for each $v \in (H_0^1(D))^n$; namely, $\mathrm{div}\, v$ is orthogonal to the constants. Therefore, it is sufficient to require that the equation above is satisfied for each $r \in L^2_*(D) = \left\{ r \in L^2(D) \mid \int_D r dx = 0 \right\}$. In conclusion, the right choice of the pressure space is $M = L^2_*(D)$. Let us note that in (8.5) the pressure p is determined up to an additive constant: thus this choice permits to select a unique pressure.

Let us see now which are the variational spaces in case (8.6). Take the scalar product of the first equation in (8.6) by m: integrating in D and integrating by parts we obtain

$$0 = \int_D Zq \cdot m dx + \int_D \nabla\varphi \cdot m dx$$
$$= \int_D Zq \cdot m dx - \int_D \varphi \,\mathrm{div}\, m dx + \int_{\partial D} \underbrace{\varphi}_{=0} n \cdot m dS_x$$
$$= \int_D Zq \cdot m dx - \int_D \varphi \,\mathrm{div}\, m dx \,.$$

From the second equation in (8.6) we get, for any ψ,

$$\int_D (-\operatorname{div} q)\psi\, dx = -\int_D g\psi\, dx\,.$$

Thus we need $q, m \in (L^2(D))^n$ with $\operatorname{div} q$, $\operatorname{div} m \in L^2(D)$, and $\varphi, \psi \in L^2(D)$. Summing up, in this second case (8.6) we have

$$V = H(\text{div}; D) = \{m \in (L^2(D))^n \mid \operatorname{div} m \in L^2(D)\}$$

and $M = L^2(D)$. It is easy to see that $H(\text{div}; D)$ is a Hilbert space with respect to the scalar product

$$(q, m)_{H(\text{div};D)} = \int_D (q \cdot m + \operatorname{div} q \operatorname{div} m) dx\,. \tag{8.13}$$

Exercise 8.5 Prove that $H(\text{div}; D)$ is a Hilbert space with respect to the scalar product (8.13).

In order to apply Theorem 8.5, let us check if the operator A is coercive over the kernel of the operator B. In the first case (8.5) we have $V = (H_0^1(D))^n$ and

$$\begin{aligned}
\langle Av, v\rangle &= \nu \int_D \sum_{k=1}^n \nabla v_k \cdot \nabla v_k dx = \nu \sum_{k=1}^n \int_D |\nabla v_k|^2 dx \\
&= \frac{\nu}{2} \sum_{k=1}^n \int_D |\nabla v_k|^2 dx + \frac{\nu}{2} \sum_{k=1}^n \int_D |\nabla v_k|^2 dx \\
&\geq \frac{\nu}{2} \sum_{k=1}^n \int_D |\nabla v_k|^2 dx + \frac{\nu}{2C_D} \sum_{k=1}^n \int_D v_k^2 dx \\
&\qquad \text{(Poincaré inequality in } H_0^1(D)) \\
&\geq \alpha \|v\|^2_{H^1(D)}
\end{aligned}$$

where $\alpha = \min\left(\frac{\nu}{2}, \frac{\nu}{2C_D}\right)$, and C_D is the Poincaré constant in $H_0^1(D)$.

We have thus seen that for problem (8.5) the operator A is indeed coercive in V, and not only on the kernel of B. A natural question then arises: are there interesting cases for which the "strong" assumption

$$(Av, v) \geq \alpha \|v\|_V^2\,, \quad \alpha > 0$$

is not satisfied and we really need a weaker assumption? The answer is yes, as the second Example 8.3 shows.

In fact, in case (8.6) we have $V = H(\text{div}; D)$ and

$$\langle Am, m\rangle = \int_D Zm \cdot m dx \\ \geq \alpha \|m\|^2_{L^2(D)} \qquad (Z \text{ is positive definite, uniformly in } x \in D)\,, \tag{8.14}$$

but this is not enough as the control on $\int_D (\text{div}\, m)^2 dx$ is missing. However, we note that in this case $Bm = 0$ means

$$\int_D \text{div}\, m\, \psi\, dx = 0$$

for each $\psi \in L^2(D)$; thus it follows at once $\text{div}\, m = 0$ in D. Summing up, for m satisfying $\text{div}\, m = 0$ in D we can rewrite (8.14) as

$$\langle Am, m\rangle \geq \alpha \|m\|^2_V = \alpha \Big(\|m\|^2_{L^2(D)} + \underbrace{\|\text{div}\, m\|^2_{L^2(D)}}_{=0} \Big)\,,$$

and we have a control from below in terms of the norm of the space V, namely, coerciveness is restored in the closed subspace of V given by $N(B)$.

Let us now verify that the condition (8.11) is fulfilled for the Stokes problem (8.5) and the first order elliptic system (8.6). Let us start from problem (8.5). We have to check that for each $q \in L^2_*(D)$, $q \neq 0$, we can find $v_q \in (H^1_0(D))^n$, $v_q \neq 0$, such that

$$\langle B^T q, v_q\rangle = -\int_D q\, \text{div}\, v_q\, dx \geq \beta \|q\|_{L^2(D)} \|v_q\|_{H^1(D)}\,,$$

with a positive constant β not depending on q. Since q is average-free, i.e., $\int_D q dx = 0$, it is known that there exists $v_q \in (H^1_0(D))^n$ such that $\text{div}\, v_q = -q$ in D (with $v_q \neq 0$, as $q \neq 0$) and

$$\|v_q\|_{H^1(D)} \leq c_* \|q\|_{L^2(D)}$$

(see Remark 8.5 here below).

Remark 8.5 There are many ways to prove the result here above, and all of them require some work. Just to quote a classical result, it is possible to furnish an explicit formula, at least for a (connected) bounded open set that is star-shaped with respect to all the points of a ball $B^0 = B(x_0, r_0) \subset\subset D$, $x_0 \in D$, $r_0 > 0$. In this geometrical case, take $w \in C_0^\infty(B^0)$ with $\int_{B^0} w\,dx = 1$. For $q \in C_0^\infty(D)$ with $\int_D q\,dx = 0$, define for $i = 1, \dots, n$

$$(v_q)_i(x) = -\int_D q(y) \left[\frac{x_i - y_i}{|x-y|^n} \int_0^{+\infty} w\left(x + t\frac{x-y}{|x-y|}\right)(|x-y|+t)^{n-1}\,dt \right] dy\,.$$

In 1979 Mikhail E. Bogovskii[1] has proved that $v_q \in (H_0^1(D))^n$ and $\operatorname{div} v_q = -q$ in D, with $\|v_q\|_{H^1(D)} \le c_* \|q\|_{L^2(D)}$. Since a bounded, connected, open set D with Lipschitz continuous boundary ∂D is the finite union of domains that are star-shaped with respect to all the points of a ball, the result for this general geometrical situation is obtained by localization. Then by a density argument the result is also extended to all $q \in L^2(D)$ with $\int_D q\,dx = 0$.

Let us use the function v_q thus determined for checking condition (8.11). We have

$$-\int_D q \operatorname{div} v_q dx = \int_D q^2 dx = \|q\|_{L^2(D)}\|q\|_{L^2(D)} \ge \|q\|_{L^2(D)} \frac{1}{c_*}\|v_q\|_{H^1(D)}\,,$$

thus we get $\beta = 1/c_*$, independent of q.

Let us come now to problem (8.6). For any $q \in L^2(D)$, take the solution $\varphi_q \in H_0^1(D)$ of the weak form of the homogeneous Dirichlet problem

$$\begin{cases} -\Delta\varphi_q = q & \text{in } D \\ \varphi_q = 0 & \text{on } \partial D\,, \end{cases}$$

and set $v_q = \nabla\varphi_q$. We have

$$-\operatorname{div} v_q = -\Delta\varphi_q = q \quad \text{in } D$$

and

$$\|\varphi_q\|_{H^1(D)} \le c_* \|q\|_{L^2(D)}$$

by the Lax–Milgram theorem 2.1. Thus

$$\|v_q\|^2_{H(\operatorname{div};D)} = \|v_q\|^2_{L^2(D)} + \|\underbrace{\operatorname{div} v_q}_{-q}\|^2_{L^2(D)} \le c_*^2\|q\|^2_{L^2(D)} + \|q\|^2_{L^2(D)}\,.$$

[1] Bogovskii [3].

Hence

$$\|v_q\|_{H(\mathrm{div};D)} \leq \sqrt{c_*^2 + 1}\, \|q\|_{L^2(D)} ,$$

and the thesis now follows as in the previous case.

8.2 Galerkin Numerical Approximation

Let us now give a look at the Galerkin numerical approximation. In the present case we change the notation used in Sect. 7.5, and we take $V_h \subset V$, $M_h \subset M$, two finite dimensional subspaces of dimension N_h^V and N_h^M, respectively, where $h > 0$ is a parameter; for $h \to 0^+$ one has $N_h^V \to +\infty$ and $N_h^M \to +\infty$.

Writing the saddle point problem in terms of the bilinear forms, we want to solve the finite dimensional problem

$$u_h \in V_h \ , \ \varphi_h \in M_h \ : \ \begin{cases} a(u_h, v_h) + b(v_h, \varphi_h) = \langle F, v_h \rangle & \forall\, v_h \in V_h \\ b(u_h, \psi_h) = \langle G, \psi_h \rangle & \forall\, \psi_h \in M_h \, . \end{cases} \tag{8.15}$$

The assumptions assuring well-posedness are:

$$\exists\, \alpha_h > 0 \ : \ a(v_h, v_h) \geq \alpha_h \|v_h\|_V^2 \qquad \forall\, v_h \in N_h \tag{8.16}$$

where $N_h = \{v_h \in V_h \,|\, b(v_h, \psi_h) = 0 \ \forall\, \psi_h \in M_h\}$ (coerciveness of $a(\cdot, \cdot)$ on the discrete kernel of $b(\cdot, \cdot)$) and

$$\exists\, \beta_h > 0 \ : \ \forall\, \mu_h \in M_h, \ \exists\, \hat{v}_h \in V_h \ , \ \hat{v}_h \neq 0 \ : \ b(\hat{v}_h, \mu_h) \geq \beta_h \|\mu_h\|_M \|\hat{v}_h\|_V \tag{8.17}$$

(discrete inf–sup condition for $b(\cdot, \cdot)$). In this case, in fact, we can repeat the procedure that has led to determine the solution (u, φ) to problem (8.9).

Note that these two assumptions are not a consequence of conditions (8.10) and (8.11). Indeed in general $N_h \not\subset N(B)$ (as M_h is a proper closed subspace of M). Moreover, from condition (8.11) we know that for each $\mu_h \in M_h \subset M$ we can find $\hat{v} \in V$, $\hat{v} \neq 0$, satisfying the desired estimate, but not $\hat{v}_h \in V_h$, $\hat{v}_h \neq 0$.

8.2.1 Error Estimates

Under assumptions (8.16) and (8.17) it is possible to prove the convergence of the Galerkin approximation method. This can be done as follows. The first step is a consistency property: since $V_h \subset V$, we can take a test function $v_h \in V_h$ in (8.9). Thus the first equations in (8.15) and (8.9) give

$$a(u_h, v_h) + b(v_h, \varphi_h) = \langle F, v_h \rangle = a(u, v_h) + b(v_h, \varphi)\,. \tag{8.18}$$

Now we want to make appearing a difference between the approximate solution φ_h and a test function $\mu_h \in M_h$: subtracting from (8.18) $b(v_h, \mu_h)$ we find

$$a(u_h, v_h) + b(v_h, \varphi_h - \mu_h) = a(u, v_h) + b(v_h, \varphi - \mu_h)\,. \tag{8.19}$$

A similar procedure is in order for the approximate solution u_h: take $v_h^* \in V_h$ such that $b(v_h^*, \psi_h) = \langle G, \psi_h \rangle$ for each $\psi_h \in M_h$. Note that any element of the form $u_h + w_h$, $w_h \in N_h$, has this property. We will denote by N_h^G the affine subspace $\{\omega_h^* \in V_h \,|\, \omega_h^* = u_h + w_h, w_h \in N_h\}$: we have thus selected $v_h^* \in N_h^G$. Subtracting $a(v_h^*, v_h)$ we get

$$a(u_h - v_h^*, v_h) + b(v_h, \varphi_h - \mu_h) = a(u - v_h^*, v_h) + b(v_h, \varphi - \mu_h)\,. \tag{8.20}$$

Taking now $v_h = u_h - v_h^*$, it follows

$$\begin{aligned} \alpha_h \|u_h - v_h^*\|_V^2 &\le a(u_h - v_h^*, u_h - v_h^*) \\ &= -b(u_h - v_h^*, \varphi_h - \mu_h) + a(u - v_h^*, u_h - v_h^*) \\ &\quad + b(u_h - v_h^*, \varphi - \mu_h)\,. \end{aligned}$$

Since

$$b(u_h - v_h^*, \psi_h) = \langle G, \psi_h \rangle - \langle G, \psi_h \rangle = 0 \qquad \forall\, \psi_h \in M_h\,,$$

the term $b(u_h - v_h^*, \varphi_h - \mu_h)$ vanishes. Therefore we have found

$$\|u_h - v_h^*\|_V^{\cancel{2}} \le \frac{1}{\alpha_h} \left(\gamma \|u - v_h^*\|_V \cancel{\|u_h - v_h^*\|_V} + \|b\| \cancel{\|u_h - v_h^*\|_V} \|\varphi - \mu_h\|_M \right)\,.$$

Thus

$$\begin{aligned} \|u - u_h\|_V &\le \|u - v_h^*\|_V + \|u_h - v_h^*\|_V \\ &\le \|u - v_h^*\|_V + \frac{\gamma}{\alpha_h} \|u - v_h^*\|_V + \frac{\|b\|}{\alpha_h} \|\varphi - \mu_h\|_M \\ &\le \left(1 + \frac{\gamma}{\alpha_h}\right) \|u - v_h^*\|_V + \frac{\|b\|}{\alpha_h} \|\varphi - \mu_h\|_M\,, \end{aligned} \tag{8.21}$$

for each $v_h^* \in N_h^G$ and for each $\mu_h \in M_h$.

For a fixed $v_h \in V_h$ consider now the linear functional $\psi_h \mapsto b(u - v_h, \psi_h)$, $\psi_h \in M_h$. From condition (8.17) we know that there exists a unique $z_h \in N_h^\perp$ such that

$$b(z_h, \psi_h) = b(u - v_h, \psi_h) \qquad \forall\, \psi_h \in M_h\,,$$

with

$$\|z_h\|_V \le \frac{1}{\beta_h} \sup_{\psi_h \in M_h,\, \psi_h \neq 0} \frac{b(u - v_h, \psi_h)}{\|\psi_h\|_M} \le \frac{\|b\|}{\beta_h} \|u - v_h\|_V\,.$$

Setting $w_h^* = z_h + v_h$, we see that

$$b(w_h^*, \psi_h) = b(z_h + v_h, \psi_h) = b(u, \psi_h) = \langle G, \psi_h \rangle \quad \forall\, \psi_h \in M_h\,.$$

Thus $w_h^* \in N_h^G$ and

$$\begin{aligned} \inf_{\omega_h^* \in N_h^G} \|u - \omega_h^*\|_V &\le \|u - w_h^*\|_V \le \|u - v_h\|_V + \|z_h\|_V \\ &\le \left(1 + \frac{\|b\|}{\beta_h}\right) \|u - v_h\|_V \qquad \forall\, v_h \in V_h\,. \end{aligned}$$

In conclusion, inserting this estimate in (8.21) we have found the error estimate

$$\|u - u_h\|_V \le \left(1 + \frac{\gamma}{\alpha_h}\right)\left(1 + \frac{\|b\|}{\beta_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_V + \frac{\|b\|}{\alpha_h} \inf_{\mu_h \in M_h} \|\varphi - \mu_h\|_M\,. \tag{8.22}$$

The estimate of the error $\|\varphi - \varphi_h\|_M$ is obtained as follows: by condition (8.17), in correspondence with $\varphi_h - \mu_h$ we can find $v_h \in V_h$, $v_h \neq 0$, such that

$$b(v_h, \varphi_h - \mu_h) \ge \beta_h \|v_h\|_V \|\varphi_h - \mu_h\|_M, . \tag{8.23}$$

On the other hand from (8.18) we have $a(u - u_h, v_h) + b(v_h, \varphi - \varphi_h) = 0$ for each $v_h \in V_h$, hence

$$\begin{aligned} b(v_h, \varphi_h - \mu_h) &= b(v_h, \varphi_h - \varphi) + b(v_h, \varphi - \mu_h) \\ &= a(u - u_h, v_h) + b(v_h, \varphi - \mu_h)\,. \end{aligned}$$

Thuse from condition (8.23) we have

$$\begin{aligned}\|\varphi_h - \mu_h\|_M &\le \frac{1}{\beta_h} \frac{a(u-u_h, v_h) + b(v_h, \varphi - \mu_h)}{\|v_h\|_V} \\ &\le \frac{\gamma}{\beta_h}\|u-u_h\|_V + \frac{\|b\|}{\beta_h}\|\varphi - \mu_h\|_M .\end{aligned}$$

Finally, we have found

$$\begin{aligned}\|\varphi - \varphi_h\|_M &\le \|\varphi - \mu_h\|_M + \|\varphi_h - \mu_h\|_M \\ &\le \left(1 + \frac{\|b\|}{\beta_h}\right)\|\varphi - \mu_h\|_M + \frac{\gamma}{\beta_h}\|u-u_h\|_V \quad \forall\, \mu_h \in M_h ,\end{aligned}$$

hence

$$\|\varphi - \varphi_h\|_M \le \left(1 + \frac{\|b\|}{\beta_h}\right) \inf_{\mu_h \in M_h} \|\varphi - \mu_h\|_M + \frac{\gamma}{\beta_h}\|u-u_h\|_V , \tag{8.24}$$

which, together with (8.22), is the error estimate we wanted to prove.

Remark 8.6 It is evident that a speed of convergence that only depends on the approximation properties of V_h in V and of M_h in M is achieved if $\alpha_h \ge \alpha > 0$ and $\beta_h \ge \beta > 0$, uniformly with respect to the parameter h. Thus the art of the approximation here is to find finite dimensional subspaces V_h and M_h such that conditions (8.16) and (8.17) are satisfied uniformly with respect to h.

8.2.2 Finite Element Approximation

The uniform approximation of V and M by V_h and M_h is possible for many interesting cases, for instance for $V = (H^1_0(D))^n$ and $M = L^2_*(D)$ or $V = H(\mathrm{div}; D)$ and $M = L^2(D)$, the spaces related to Examples 8.2 and 8.3 that we have considered here. To illustrate this fact, let us focus on a very important type of Galerkin approximation: the finite element method.

As already noted in Remark 7.7, the main ingredients of a finite element approximation are the facts that the domain D is the union of a finite number of non-overlapping subsets T of simple shape (say, triangles or tetrahedra) and that the finite dimensional spaces V_h and M_h are given by functions whose restrictions to the elements T are polynomials. The parameter h represents the mesh size, namely, the maximum diameter of the elements T.

Let us show some examples of finite elements that satisfy the two conditions (8.16) and (8.17), focusing on the two-dimensional case. A first example for the Stokes problem described in Example 8.2 is the $\mathbb{P}_2$-$\mathbb{P}_0$ element, in which the two components of the velocity are piecewise-quadratic polynomials and the pressure

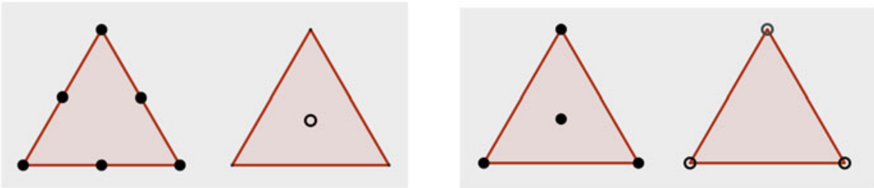


Fig. 8.2 The degrees of freedom of the $\mathbb{P}_2$-$\mathbb{P}_0$ element (left) and of the "mini-element" $(\mathbb{P}_1 \oplus B)$-$\mathbb{P}_1$ (right): point values for the velocity and for the pressure

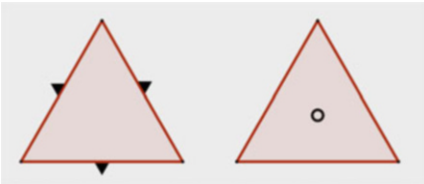


Fig. 8.3 The degrees of freedom of the Raviart–Thomas element: fluxes for the vector u_h and point values for scalar φ_h

is a piecewise-constant, therefore a discontinuous function; its degrees of freedom are point values, at the nodes drawn in Fig. 8.2, left. A second example is the "mini-element" $(\mathbb{P}_1 \oplus B)$-$\mathbb{P}_1$, in which the two components of the velocity u_h are linear combination of first order polynomials and of a fixed third order polynomial vanishing on the sides (this is called "a bubble"), and the pressure φ_h is a continuous piecewise-linear polynomial; its degrees of freedom are point values, at the nodes drawn in Fig. 8.2, right.

For the first order elliptic system presented in Example 8.3 a classical instance is the Raviart–Thomas element, for which in each element T the vector field u_h is of the form $a+bx$, with $a \in \mathbb{R}^2$ and $b \in \mathbb{R}$, and the scalar φ_h is a piecewise constant; its degrees of freedom are point values of the scalar φ_h, at the node drawn in Fig. 8.3, and fluxes of the vector u_h across the sides of T, i.e., integrals of $u_h \cdot n$ on the sides.

For all these elements it is proved that the convergence in $V \times M$ of the approximate solutions to the exact solution is linear with respect to the mesh size h.

8.3 Exercises

Exercise 8.1 $N(B)_\sharp$ can be isometrically identified with the dual of $N(B)^\perp$.

Solution Take $g \in (N(B)^\perp)'$, we define $\hat{g} \in V'$ by setting

$$\langle \hat{g}, v \rangle = \langle g, P_\perp v \rangle \quad \forall\, v \in V,$$

where $P_\perp v$ is the orthogonal projection on $N(B)^\perp$. Clearly $\hat{g} \in N(B)_\sharp$, as $P_\perp v = 0$ for $v \in N(B)$. The map $g \mapsto \hat{g}$ from $(N(B)^\perp)'$ to $N(B)_\sharp$ is clearly one-to-one, as $\hat{g} = g$ on $N(B)^\perp$. It is also onto: in fact, taking $\tilde{g} \in N(B)_\sharp$, we need to verify that there exists $g_* \in (N(B)^\perp)'$ such that $\widehat{g_*} = \tilde{g}$. Let us define $g_* \in (N(B)^\perp)'$ by

$$\langle g_*, w\rangle = \langle \tilde{g}, w\rangle \qquad \forall\, w \in N(B)^\perp\,.$$

Thus we have $g_* = \tilde{g}$ on $N(B)^\perp$, and also $\widehat{g_*} = g_*$ on $N(B)^\perp$, thus $\widehat{g_*} = \tilde{g}$ on $N(B)^\perp$. On the other hand, $\widehat{g_*} = 0$ and $\tilde{g} = 0$ on $N(B)$, as both of them belong to $N(B)_\sharp$, thus $\widehat{g_*} = \tilde{g}$ on V.

Finally, for each $v \in V$, $v \neq 0$, one has $\langle \hat{g}, v\rangle = 0$ if $v \in N(B)$, while for $v \in N(B)^\perp$

$$\frac{\langle \hat{g}, v\rangle}{\|v\|} = \frac{\langle g, v\rangle}{\|v\|} \leq \sup_{w\in N(B)^\perp, w\neq 0} \frac{\langle g, w\rangle}{\|w\|} = \|g\|\,,$$

thus $\|\hat{g}\| \leq \|g\|$. Moreover, for $w \in N(B)^\perp$, $w \neq 0$, it holds

$$\frac{\langle g, w\rangle}{\|w\|} = \frac{\langle \hat{g}, w\rangle}{\|w\|} \leq \sup_{v\in V, v\neq 0} \frac{\langle \hat{g}, v\rangle}{\|v\|} = \|\hat{g}\|\,.$$

Exercise 8.2 The inf–sup condition (8.7) is equivalent to each one of the following conditions:

(a) The operator B^T is an isomorphism from M onto $N(B)_\sharp$ and

$$\exists\, \beta > 0 \,:\, \|B^T\mu\|_{V'} \geq \beta\|\mu\|_M \qquad \forall\, \mu \in M\,.$$

(b) The operator B is an isomorphism from $N(B)^\perp$ onto M' and

$$\exists\, \beta > 0 \,:\, \|Bv\|_{M'} \geq \beta\|v\|_V \qquad \forall\, v \in N(B)^\perp\,.$$

Solution (b) $\Rightarrow$ (a). From (b) we know that $R(B) = M'$ is closed, so that by the closed range theorem 8.4 $R(B^T)$ is closed in V' and $R(B^T) = N(B)_\sharp$, $R(B) = N(B^T)_\sharp = M'$, thus $N(B^T) = \{0\}$. In conclusion, B^T is an isomorphism from M onto $N(B)_\sharp$. The estimate in (b) says that $\|B^{-1}\|_{\mathcal{L}(M';N(B)^\perp)} \leq 1/\beta$, while the estimate in (a) says that $\|(B^T)^{-1}\|_{\mathcal{L}(N(B)_\sharp;M)} \leq 1/\beta$. Thus they are equivalent, since

$$\|B^{-1}\|_{\mathcal{L}(M';N(B)^\perp)} = \|(B^T)^{-1}\|_{\mathcal{L}(N(B)_\sharp;M)}\,,$$

as it can be easily verified by looking at the definition of adjoint operator and taking into account that $(B^{-1})^T = (B^T)^{-1}$ and the identification $N(B)_\sharp = (N(B)^\perp)'$.

(a) $\Rightarrow$ (8.7). It is enough to note that

$$\|B^T\mu\|_{V'} = \max_{v\in V, v\neq 0} \frac{\langle B^T\mu, v\rangle}{\|v\|_V}$$

(see Exercise 8.3).

(8.7) $\Rightarrow$ (b). By Proposition 8.2 we know that (8.8) is satisfied, $R(B^T)$ is closed in V' and $N(B^T) = \{0\}$, so that, by the closed range theorem 8.4, $R(B) = M'$. By decomposing V into the two ortoghonal subspaces $N(B)$ and $N(B)^\perp$, it is easy to check that also the restriction of B to $N(B)^\perp$ is onto M'. Therefore B is an isomorphism from $N(B)^\perp$ onto M'. Finally, (8.8) is equivalent to the estimate in (a), which, as already seen, is equivalent to the estimate in (b).

Exercise 8.3 Let V be a Hilbert space and $F \in V'$. Show that the norm $\|F\|_{V'}$ defined as

$$\|F\|_{V'} = \sup_{v\in V, v\neq 0} \frac{\langle F, v\rangle}{\|v\|_V}$$

is indeed equal to

$$\|F\|_{V'} = \max_{v\in V, v\neq 0} \frac{\langle F, v\rangle}{\|v\|_V} ,$$

namely, there is $v_F \in V$, $v_F \neq 0$, such that

$$\|F\|_{V'} = \frac{\langle F, v_F\rangle}{\|v_F\|_V} .$$

Solution We can assume that $F \neq 0$, otherwise the result is trivial. By the Riesz representation theorem 3.1 we know that there exists a unique $v_F \in V$ such that $\langle F, v\rangle = (v_F, v)_V$ for any $v \in V$. Moreover, $\|F\|_{V'} = \|v_F\|_V$: in fact

$$\langle F, v\rangle = (v_F, v)_V \leq \|v_F\|_V \|v\|_V \quad \forall\, v \in V ,$$

which implies $\|F\|_{V'} \leq \|v_F\|_V$. On the other hand

$$\frac{\langle F, v_F\rangle}{\|v_F\|_V} = \frac{\|v_F\|_V^2}{\|v_F\|_V} = \|v_F\|_V \leq \|F\|_{V'} .$$

Thus $\|F\|_{V'} = \|v_F\|_V = \frac{\langle F, v_F\rangle}{\|v_F\|_V}$.

Exercise 8.4 Give a proof of the Lax–Milgram theorem 2.1 based on the closed range theorem 8.4.

Solution For any $w \in V$ define the linear and bounded operator $Q : V \mapsto V'$ as

$$\langle Qw, v \rangle = B(w, v) \qquad \forall\, v \in V\,.$$

As a consequence we have that $Q^T : V \mapsto V'$ is defined as $\langle Q^T w, v \rangle = B(v, w)$ for all $v \in V$. The existence and uniqueness result in the Lax–Milgram theorem 2.1 has been thus transformed into the existence of a unique $u \in V$ such that $Qu = F \in V'$, namely, in showing that Q is one-to-one and onto, or, equivalently, in showing that $N(Q) = \{0\}$ and $R(Q) = V'$. The coerciveness assumption on the bilinear form $B(\cdot, \cdot)$ straightforwardly shows that $N(Q) = \{0\}$ and $N(Q^T) = \{0\}$. Moreover, by proceeding as in step 4 of the proof of the Lax–Milgram theorem 2.1, we obtain that the range of Q is closed. Therefore the closed range theorem 8.4 gives that $R(Q) = N(Q^T)_\sharp = V'$ and the proof is completed.

Exercise 8.5 Prove that $H(\mathrm{div}; D)$ is a Hilbert space with respect to the scalar product (8.13).

Solution Take a Cauchy sequence q_k in $H(\mathrm{div}; D)$: in particular q_k and $\mathrm{div}\, q_k$ are Cauchy sequences in $(L^2(D))^n$ and $L^2(D)$, respectively, thus we have that $q_k \to q$ and $\mathrm{div}\, q_k \to w$ in $(L^2(D))^n$ and in $L^2(D)$, respectively. From the definition of weak divergence we know that $\mathrm{div}\, q_k$ satisfies

$$\int_D \mathrm{div}\, q_k\, \varphi dx = -\int_D q_k \cdot \nabla\varphi dx \quad \forall\, \varphi \in C_0^\infty(D)\,.$$

Passing to the limit we find

$$\int_D w\, \varphi dx = -\int_D q \cdot \nabla\varphi dx \quad \forall\, \varphi \in C_0^\infty(D)\,,$$

which means that $w \in L^2(D)$ is the weak divergence of q. As a consequence we have proved that the sequence q_k converges to q in $H(\mathrm{div}; D)$.

Chapter 9
Parabolic PDEs

Parabolic equations are equations of the form

$$\frac{\partial u}{\partial t} + Lu = f \quad \text{in } D \times (0, T)\,,$$

where L is an elliptic operator, whose coefficients can depend on t. The "prototype" is the *heat equation*

$$\frac{\partial u}{\partial t} - \Delta u = f \quad \text{in } D \times (0, T)\,.$$

Since with respect to the space derivative the operator $\frac{\partial}{\partial t} + L$ is associated to an elliptic operator, it is necessary to add boundary conditions (for instance, one of the four types we have considered before: Dirichlet, Neumann, mixed, Robin). Since with respect to the time derivative the operator $\frac{\partial}{\partial t} + L$ is a first order operator, it is necessary to add one initial condition on u, the value of u in D at $t = 0$.

In the first two sections of this chapter we present the abstract variational theory related to parabolic equations and its application to various examples of initial-boundary value problems, including linear Navier–Stokes equations. The last section is devoted to an important property of the solutions: the maximum principle.

9.1 Variational Theory

Before considering some specific problems, let us present an abstract theory for first order evolution equations in Hilbert spaces. First of all we need to clarify some theoretical results concerning functions with values in an infinite dimensional Hilbert space. We will not enter in depth this topic, limiting ourselves to give some

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_9

general ideas. A complete description of the functional analysis framework can be found in Dautray and Lions [6, Chapter XVIII, §1].

We start with some definitions. Let X be an Hilbert space: we set, for $1 \leq p \leq +\infty$,

$$L^p(0,T;X) = \{v : (0,T) \mapsto X \mid t \mapsto v(t) \text{ is measurable in } (0,T) \text{ and } t \mapsto \|v(t)\|_X \text{ is a } L^p\text{- function in } (0,T)\}$$

$$C^0([0,T];X) = \{v : [0,T] \mapsto X \mid t \mapsto v(t) \;\text{ is a } C^0\text{- function in } [0,T]\}\,.$$

The norms in these spaces are, respectively,

$$\|v\|_{L^p(0,T;X)} = \Big(\int_0^T \|v(t)\|_X^p dt\Big)^{1/p} \;, \quad \|v\|_{C^0([0,T];X)} = \max_{t\in[0,T]} \|v(t)\|_X\,.$$

For $p = 2$ the scalar product of the Hilbert space $L^2(0,T;X)$ is defined as

$$(v,w)_{L^2(0,T;X)} = \int_0^T (v(t),w(t))_X dt\,,$$

having as usual denoted by $(\cdot,\cdot)_X$ the scalar product in X.

Then we define the weak derivative with respect to $t \in [0,T]$.

Definition 9.1 We say that $q \in L^1_{\text{loc}}(0,T;X)$ is the weak derivative of $u \in L^1_{\text{loc}}(0,T;X)$ if, as elements of the space X,

$$\int_0^T \Phi(t)q(t)dt = -\int_0^T \Phi'(t)u(t)dt$$

for each $\Phi \in C_0^\infty(0,T)$, or, equivalently, if

$$\int_0^T \Phi(t)(q(t),v)_X dt = -\int_0^T \Phi'(t)(u(t),v)_X dt$$

for each $v \in X$ and $\Phi \in C_0^\infty(0,T)$. In this case we write $u' = q$, as an element of $L^1_{\text{loc}}(0,T;X)$.

Now it is a standard task to define the Sobolev spaces $W^{1,p}(0,T;X)$. We write, as usual, $H^1(0,T;X) = W^{1,2}(0,T;X)$.

An important theorem is the following:

Theorem 9.1 *If $u \in H^1(0,T;X)$, then $u \in C^0([0,T];X)$ and*

$$\|u\|_{C^0([0,T];X)} \leq C_T\|u\|_{H^1(0,T;X)}\,.$$

This is not enough for our needs, and we are going to present a similar theorem which is even more important. Before giving its statement, we need some preliminary considerations. First of all the following result holds:

Exercise 9.1 Suppose that V and H are two Hilbert spaces, that V is immersed in H with continuity and that V is dense in H. Then H', the dual space of H, is immersed with continuity in V', the dual space of V. Moreover, H' is dense in V'.

Identifying H with H' we can thus write

$$V \hookrightarrow H \approx H' \hookrightarrow V' ,$$

or simply $V \subset H \subset V'$.

We can now furnish a definition of the derivative of u with respect to t which is weaker than that given in 9.1. Suppose that $u \in L^1_{\mathrm{loc}}(0, T; H)$; we say that there exists the derivative $u' \in L^1_{\mathrm{loc}}(0, T; V')$ if there exists $q \in L^1_{\mathrm{loc}}(0, T; V')$ such that

$$\int_0^T q(t)\Phi(t)dt = -\int_0^T u(t)\Phi'(t)dt$$

for each $\Phi \in C_0^\infty(0, T)$. This equality has an element of V' at the left-hand side and an element of H at the right-hand side; it can be more explicitly specified by writing

$$\begin{aligned}\Big\langle \int_0^T q(t)\Phi(t)dt, v\Big\rangle &= -\Big\langle \int_0^T u(t)\Phi'(t)dt, v\Big\rangle \\ &= -\int_0^T \langle u(t), v\rangle \Phi'(t)dt \\ &\overset{(\bullet)}{=} -\int_0^T (u(t), v)_H \Phi'(t)dt \qquad \forall\, v \in V ,\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V and V' and $(\cdot, \cdot)_H$ the scalar product in H. Thus

$$\int_0^T \langle q(t), v\rangle \Phi(t)dt = -\int_0^T (u(t), v)_H \Phi'(t)dt \qquad \forall\, v \in V .$$

Therefore, if for $u \in L^1_{\mathrm{loc}}(0, T; H)$ we know that $u' \in L^1_{\mathrm{loc}}(0, T; V')$, we have

$$\int_0^T \langle u'(t), v\rangle \Phi(t)dt = -\int_0^T (u(t), v)_H \Phi'(t)dt \qquad \forall\, v \in V ,$$

which can be also rewritten as

$$\frac{d}{dt}(u(t), v)_H = \langle u'(t), v\rangle \tag{9.1}$$

for almost all $t \in [0, T]$ and all $v \in V$, where $\frac{d}{dt}(u(t), v)_H$ has to be intended as the weak derivative with respect to t of the real valued function $t \mapsto (u(t), v)_H$. From now on the notation u' will always refer to the weak derivative of u with respect to t.

Remark 9.1 A remark on (•). Due to the identification of H' with H, we have that $\omega \in H$ implies $\omega \in V'$ and in particular

$$\langle \omega, v\rangle_{V',V} = \langle \omega, v\rangle_{H',H} = (\omega, v)_H \qquad \forall\, v \in V\,.$$

We are now ready to state the theorem we will often use in the sequel.

Theorem 9.2 *Let H be a separable Hilbert space, V a separable Hilbert space immersed with continuity and dense in H. Let $u \in L^2(0, T; V)$ with $u' \in L^2(0, T; V')$. Then $u \in C^0([0, T]; H)$ and*

$$\|u\|_{C^0([0,T];H)} \leq C_T \left(\|u\|_{L^2(0,T;V)} + \|u'\|_{L^2(0,T;V')}\right)\,.$$

Moreover, if $v \in L^2(0, T; V)$ with $v' \in L^2(0, T; V')$ for each $t, s \in [0, T]$ the integration by parts formula holds

$$\int_s^t \langle u'(\tau), v(\tau)\rangle d\tau = -\int_s^t \langle v'(\tau), u(\tau)\rangle d\tau + (u(t), v(t))_H - (u(s), v(s))_H\,.$$

Also, for almost all $t \in [0, T]$

$$\frac{d}{dt}(u(t), v(t))_H = \langle u'(t), v(t)\rangle + \langle v'(t), u(t)\rangle$$

and

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|_H^2 = \langle u'(t), u(t)\rangle\,.$$

9.2 Abstract Problem

Let us formulate now the abstract problem we want to solve. Suppose we have a separable Hilbert space H, a separable Hilbert space V such that $V \hookrightarrow H$ with continuous and dense immersion. Assume that we are given with $u_0 \in H$ and $F \in$

$L^2(0, T; V')$ and with a family of bilinear forms $a(t; \cdot, \cdot)$, defined in $V \times V$ and valued in $\mathbb{R}$ for almost each $t \in [0, T]$.

We want to find $u \in L^2(0, T; V)$ with $u' \in L^2(0, T; V')$ such that $u(0) = u_0$ (note that from Theorem 9.2 we know that $u \in C^0([0, T]; H)$, thus this equality has a meaning) and

$$\langle u'(t), v\rangle + a(t; u(t), v) = \langle F(t), v\rangle \tag{9.2}$$

for almost all $t \in [0, T]$ and for each $v \in V$. Let us remind that this can be equivalently rewritten as

$$-\int_0^T (u(t), v)_H \Phi'(t)dt = -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T \langle F(t), v\rangle\Phi(t)dt$$

for all $\Phi \in C_0^\infty(0, T)$ and for each $v \in V$.

For showing the existence and uniqueness of such a solution we need some assumptions on the family of bilinear forms $a(t; \cdot, \cdot)$. We suppose that:

(i) $a(t; \cdot, \cdot)$ is uniformly weakly coercive in $V \times V$, namely, there exist a constant $\alpha > 0$ and a constant $\sigma \geq 0$ (both not depending on $t \in [0, T]$) such that

$$a(t; v, v) + \sigma(v, v)_H \geq \alpha\|v\|_V^2 \qquad \forall\, v \in V \text{ and almost all } t \in [0, T]$$

(ii) $a(t; \cdot, \cdot)$ is uniformly bounded in $V \times V$, namely, there exists a constant $\gamma > 0$ (not depending on $t \in [0, T]$) such that

$$|a(t; w, v)| \leq \gamma\|w\|_V\|v\|_V \qquad \forall\, w, v \in V \text{ and almost all } t \in [0, T]$$

(iii) the map $t \mapsto a(t; w, v)$ is measurable in $(0, T)$ for every $w, v \in V$.

The existence and uniqueness theorem reads as follows:

Theorem 9.3 (Existence and Uniqueness) *Let H and V be two separable Hilbert spaces, with $V \hookrightarrow H$ with continuous and dense immersion. Assume $u_0 \in H$ and $F \in L^2(0, T; V')$. Assume that the family of bilinear forms $a(t; \cdot, \cdot)$ is defined in $V \times V$ and valued in $\mathbb{R}$ for almost each $t \in [0, T]$ and satisfies (i), (ii) and (iii). Then there exists a unique solution $u \in L^2(0, T; V)$ of Eq. (9.2), satisfying $u' \in L^2(0, T; V')$ and $u(0) = u_0$. Moreover, for each $\tau \in [0, T]$ the stability estimate*

$$\|u(\tau)\|_H^2 + \alpha\int_0^\tau e^{2\sigma(\tau-t)}\|u(t)\|_V^2 dt \leq e^{2\sigma\tau}\|u_0\|_H^2 + \frac{1}{\alpha}\int_0^\tau e^{2\sigma(\tau-t)}\|F(t)\|_{V'}^2 dt \tag{9.3}$$

holds.

Remark 9.2 In (i) we can always assume that $\sigma = 0$, namely, that $a(t; \cdot, \cdot)$ is uniformly coercive in $V \times V$. In fact, if we set $\hat{u} = e^{-\sigma t} u$, we see that $\hat{u}$ is a solution to

$$\langle \hat{u}'(t), v\rangle + a(t; \hat{u}(t), v) + \sigma(\hat{u}(t), v)_H = \langle e^{-\sigma t} F(t), v\rangle \,,$$

and now the bilinear forms $a(t; \cdot, \cdot) + \sigma(\cdot, \cdot)_H$ are uniformly coercive in $V \times V$.

Proof The proof of the theorem requires several steps. For the proof of uniqueness and existence we assume $\sigma = 0$ in (i) (see Remark 9.2).

First Step Let us start from the uniqueness. It is enough to show that the only solution for $F = 0$ and $u_0 = 0$ is $u = 0$. Let $t \in [0, T]$ be a value for which Eq. (9.2) is satisfied. Take $v = u(t)$. Then

$$\langle u'(t), u(t)\rangle + a(t; u(t), u(t)) = 0 \,.$$

On the other hand we have

$$\frac{d}{dt}\|u(t)\|_H^2 = 2\langle u'(t), u(t)\rangle$$

and

$$a(t; u(t), u(t)) \geq \alpha \|u(t)\|_V^2 \,,$$

thus

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|_H^2 + \alpha\|u(t)\|_V^2 \leq 0 \qquad \text{for almost all } t \in [0, T] \,.$$

As a consequence, integrating in $[0, \tau]$ we find

$$\|u(\tau)\|_H \leq \|u(0)\|_H = \|u_0\|_H = 0 \qquad \text{for all } \tau \in [0, T] \,.$$

Second Step The proof of the existence of a solution is based on an approximation procedure (Galerkin method for a time-dependent problem). Since V is separable, we have a countable orthonormal Hilbertian basis $\{\varphi_m\} \subset V$ (see, e.g., Brezis [4, Théor. V.10, p. 86]). Define $V_N = \text{span}\{\varphi_1, \ldots, \varphi_N\} \subset V$. We want to find an approximate solution u_N in V_N. Since V is dense in H, we can find a sequence $u_{0,N} \in V_N$ such that $u_{0,N}$ converges to u_0 in H. Then we look for an approximate solution u^N of the form

$$u^N(t) = \sum_{j=1}^{N} u_j^N(t)\varphi_j$$

that has to satisfy $u^N(0) = u_{0,N}$ (this means $u_j^N(0) = (u_{0,N}, \varphi_j)_V$) and

$$\langle (u^N)'(t), \varphi_l \rangle + a(t; u^N(t), \varphi_l) = \langle F(t), \varphi_l \rangle$$

for almost all $t \in [0, T]$ and for all $l = 1, \dots, N$. Inserting the expression of u^N, we find

$$\sum_{j=1}^{N} \langle \varphi_j, \varphi_l \rangle (u_j^N)'(t) + \sum_{j=1}^{N} a(t; \varphi_j, \varphi_l) u_j^N(t) = \langle F(t), \varphi_l \rangle \tag{9.4}$$

for each $l = 1, \dots, N$ and almost all $t \in [0, T]$.
Setting $M_{lj} = \langle \varphi_j, \varphi_l \rangle$, $A_{lj}(t) = a(t; \varphi_j, \varphi_l)$, $U^N(t) = (u_1^N(t), \dots, u_N^N(t))$, $U_{0,N} = ((u_{0,N}, \varphi_1)_V, \dots, (u_{0,N}, \varphi_N)_V)$ and $b(t) = (\langle F(t), \varphi_1 \rangle, \dots, \langle F(t), \varphi_N \rangle)$ we have obtained the linear system of ordinary differential equations

$$\begin{cases} M(U^N)'(t) + A(t)U^N(t) = b(t) \\ U^N(0) = U_{0,N} \,. \end{cases} \tag{9.5}$$

The matrix $M_{lj} = \langle \varphi_j, \varphi_l \rangle$ can be rewritten as $(\varphi_j, \varphi_l)_H$ (take into account that $\varphi_j \in V$ and see Remark 9.1); it is clearly symmetric and moreover it is positive definite. In fact, taking $\eta \in \mathbb{R}^N$ one has

$$\sum_{j,l=1}^{N} (\varphi_j, \varphi_l)_H \eta_j \eta_l = \Big(\sum_{j=1}^{N} \eta_j \varphi_j, \sum_{l=1}^{N} \eta_l \varphi_l \Big)_H = \Big\| \sum_{j=1}^{N} \eta_j \varphi_j \Big\|_H^2 \geq 0$$

and the equality gives $\sum_{j=1}^{N} \eta_j \varphi_j = 0$ in H and thus in V, since V is immersed in H. Since φ_j are linearly independent in V, it follows $\eta_j = 0$ for $j = 1, \dots, N$. Thus the matrix $M_{lj} = \langle \varphi_j, \varphi_l \rangle$ is non-singular, therefore there exists a unique solution $(u_1^N(t), \dots, u_N^N(t))$ of the linear system (9.5) and $u_j^N \in C^0([0, T])$ with $(u_j^N)' \in L^2(0, T)$.

Third Step Now we want to pass to the limit in Eq. (9.4) as $N \to \infty$. We need suitable a-priori estimates, in such a way that we can apply some known results of functional analysis. Precisely, we want to find a subsequence u^{N_k} such that u^{N_k} converges weakly to u in $L^2(0, T; V)$. For this purpose, we need to find uniform estimates for u^N in $L^2(0, T; V)$. Multiplying expression (9.4) by $u_l^N(t)$ and adding over l we get

$$((u^N)'(t), u^N(t))_H + a(t; u^N(t), u^N(t)) = \langle F(t), u^N(t) \rangle \,.$$

Since

$$\frac{1}{2}\frac{d}{dt}\|u^N(t)\|_H^2 = ((u^N)'(t), u^N(t))_H \,,$$

integrating on $(0, \tau)$, we have for each $\tau \in [0, T]$

$$\frac{1}{2}\|u^N(\tau)\|_H^2 + \int_0^\tau a(t; u^N(t), u^N(t))dt = \frac{1}{2}\|u_{0,N}\|_H^2 + \int_0^\tau \langle F(t), u^N(t)\rangle dt \,.$$

By coerciveness we have

$$\int_0^\tau a(t; u^N(t), u^N(t))dt \geq \alpha \int_0^\tau \|u^N(t)\|_V^2 dt \,;$$

moreover, from the inequality $ab \leq \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$, valid for any $a \in \mathbb{R}$, $b \in \mathbb{R}$ and $\varepsilon > 0$, we obtain, with $\varepsilon = \alpha$,

$$\begin{aligned}\int_0^\tau \langle F(t), u^N(t)\rangle dt &\leq \int_0^\tau \|F(t)\|_{V'}\|u^N(t)\|_V dt \\ &\leq \frac{\alpha}{2}\int_0^\tau \|u^N(t)\|_V^2 dt + \frac{1}{2\alpha}\int_0^\tau \|F(t)\|_{V'}^2 dt \,,\end{aligned}$$

and consequently

$$\frac{1}{2}\|u^N(\tau)\|_H^2 + \frac{\alpha}{2}\int_0^\tau \|u^N(t)\|_V^2 dt \leq \frac{1}{2}\|u_{0,N}\|_H^2 + \frac{1}{2\alpha}\int_0^\tau \|F(t)\|_{V'}^2 dt \,.$$

Since $u_{0,N}$ converges to u_0 in H, we have obtained a uniform bound for u^N in $L^2(0, T; V)$. Since $L^2(0, T; V)$ is a Hilbert space, it exists a subsequence u^{N_k} (still denoted u^N) that converges weakly to an element $u \in L^2(0, T; V)$ (see Yosida [28, Theorem 1, p. 126, and Theorem of Eberlein–Shmulyan, p. 141]). Take now $\Phi \in C_0^\infty(0, T)$ and $v \in V$. We have a sequence $v_N \in V_N$ such that $v_N \to v$ in V. If we define $\psi^N : [0, T] \mapsto V$ and $\psi : [0, T] \mapsto V$ by setting

$$\psi^N(t) = \Phi(t)v_N \quad , \quad \psi(t) = \Phi(t)v \,,$$

we have at once $\psi^N \to \psi$ in $L^2(0, T; V)$ and $(\psi^N)' \to \psi'$ in $L^2(0, T; V)$. Rewriting Eq. (9.4) as

$$\langle (u^N)'(t), w^N\rangle + a(t; u^N(t), w^N) = \langle F(t), w^N\rangle \qquad \forall\, w^N \in V_N \tag{9.6}$$

and taking $w^N = \psi^N(t)$, by integrating by parts in $(0, T)$ it follows

$$-\int_0^T (u^N(t), (\psi^N)'(t))_H dt + \int_0^T a(t; u^N(t), \psi^N(t))dt = \int_0^T \langle F(t), \psi^N(t)\rangle dt \,.$$

Since u^N converges weakly to $u \in L^2(0, T; V)$, ψ^N converges to ψ in $L^2(0, T; V)$ and $(\psi^N)'$ converges to ψ' in $L^2(0, T; V)$, we can pass to the limit (see Exercise 9.2) and obtain

$$-\int_0^T (u(t), v)_H \Phi'(t)dt + \int_0^T a(t; u(t), v)\Phi(t)dt = \int_0^T \langle F(t), v\rangle \Phi(t)dt \,,$$

hence $u' \in L^2(0, T; V')$ and u satisfies Eq. (9.2), namely,

$$\langle u'(t), v\rangle + a(t; u(t), v) = \langle F(t), v\rangle \qquad \forall\, v \in V$$

for almost all $t \in [0, T]$.

Fourth Step It remains to show that $u(0) = u_0$. Let $\Phi \in C^\infty([0, T])$, with $\Phi(T) = 0$ and $\Phi(0) \neq 0$. First of all, by integration on $(0, T)$ from Eq. (9.2) it follows

$$\int_0^T \langle u'(t), v\rangle \Phi(t)dt = -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T \langle F(t), v\rangle \Phi(t)dt \,.$$

The integration by parts formula in Theorem 9.2 yields

$$\int_0^T \langle u'(t), v\rangle \Phi(t)dt = -\int_0^T (u(t), v)_H \Phi'(t)dt - (u(0), v)_H \Phi(0) \,,$$

thus

$$-\int_0^T (u(t), v)_H \Phi'(t)dt - (u(0), v)_H \Phi(0) = -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T \langle F(t), v\rangle \Phi(t)dt \,.$$

Now define as before $\psi^N(t) = \Phi(t)v_N$, $\psi(t) = \Phi(t)v$, where $v \in V$ and $v_N \in V_N$ with $v_N \to v$ in V. Taking $w^N = \psi^N(t)$ in (9.6), integration by parts gives

$$-\int_0^T (u^N(t), (\psi^N)'(t))_H dt - (\underbrace{u^N(0)}_{=u_{0,N}}, \psi^N(0))_H + \int_0^T a(t; u^N(t), \psi^N(t))dt$$

$$= \int_0^T \langle F(t), \psi^N(t)\rangle dt \,,$$

and passing to the limit as $N \to \infty$ one gets

$$-\int_0^T (u(t), v)_H \Phi'(t)dt - (u_0, v)_H \Phi(0)$$
$$= -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T \langle F(t), v\rangle \Phi(t)dt \,.$$

Hence for each $v \in V$ we have obtained

$$(u_0, v)_H \Phi(0) = (u(0), v)_H \Phi(0) \,.$$

Since we have assumed $\Phi(0) \neq 0$ and V is dense in H, it follows $u(0) = u_0$.

Fifth Step The last step is related to the stability result (9.3). For a while let us assume again that $\sigma = 0$ in (i). Taking $v = u(t)$ in Eq. (9.2) (we have $u(t) \in V$ for almost all $t \in [0, T]$), it follows

$$\langle u'(t), u(t)\rangle + a(t; u(t), u(t)) = \langle F(t), u(t)\rangle \,,$$

thus proceeding as in the third step

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|_H^2 + \alpha\|u(t)\|_V^2 \leq \|F(t)\|_{V'}\|u(t)\|_V \leq \frac{1}{2\alpha}\|F(t)\|_{V'}^2 + \frac{\alpha}{2}\|u(t)\|_V^2 \,.$$

In conclusion, for each $\tau \in [0, T]$ an integration on $(0, \tau)$ gives

$$\|u(\tau)\|_H^2 + \alpha\int_0^\tau \|u(t)\|_V^2 dt \leq \|u_0\|_H^2 + \frac{1}{\alpha}\int_0^\tau \|F(t)\|_{V'}^2 dt \,,$$

and when $\sigma = 0$ the proof is complete. For the case $\sigma > 0$ it is enough to replace $u(t)$ with $e^{-\sigma t}u(t)$ and $F(t)$ with $e^{-\sigma t}F(t)$ and then (9.3) follows easily.

□

Exercise 9.2 Let V be a Hilbert space, and suppose that $v_k \in V$ converges to v in V and that w_k converges weakly to w in V. Then $(v_k, w_k)_V \to (v, w)_V$.

9.2.1 Application to Parabolic PDEs

We are now in a position to present some examples that are covered by this abstract theory. Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary ∂D. For the operator

$$Lv = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j v) + \sum_{i=1}^{n} b_i \mathcal{D}_i v + a_0 v$$

in the elliptic case we have considered four boundary value problems: Dirichlet, Neumann, mixed, Robin. The related variational spaces and bilinear forms are:

Dirichlet $V = H_0^1(D)$, $H = L^2(D)$,

$$a(w, v) = \int_D \sum_{i,j=1}^{n} a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i w v dx + \int_D a_0 w v dx \,.$$

Neumann $V = H^1(D)$, $H = L^2(D)$, $a(w, v)$ as in the Dirichlet case.
Mixed $V = H^1_{\Gamma_D}(D) = \{v \in H^1(D) \mid v_{|\Gamma_D} = 0\}$, $H = L^2(D)$, $a(w, v)$ as in the Dirichlet case.
Robin $V = H^1(D)$, $H = L^2(D)$,

$$\begin{aligned} a(w, v) = &\int_D \sum_{i,j=1}^{n} a_{ij}\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i \mathcal{D}_i w v dx \\ &+ \int_D a_0 w v dx + \int_{\partial D} \kappa w v dS_x \,. \end{aligned}$$

In the present situation, we have also time dependence; therefore the bilinear forms are more generally given by

$$a(t; w, v) = \int_D \sum_{i,j}^{n} a_{ij}(t)\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i(t) \mathcal{D}_i w v dx + \int_D a_0(t) w v dx$$

and similarly for the Robin problem.

We assume that a_{ij}, b_i, a_0 belong to $L^\infty(D \times (0, T))$ and κ belongs to $L^\infty(\partial D \times (0, T))$ (with $\kappa(x, t) \geq 0$ for almost all $(x, t) \in \partial D \times (0, T)$ and $\int_{\partial D} \kappa(t) dS_x \neq 0$ for almost all $t \in [0, T]$), so that conditions (ii) and (iii) in Theorem 9.3 are satisfied. Moreover we also assume that there exists a constant $\alpha_0 > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x, t)\eta_j \eta_i \geq \alpha_0 |\eta|^2 \qquad \forall\, \eta \in \mathbb{R}^n$$

for almost all $(x, t) \in D \times (0, T)$, i.e., on the operator L we assume ellipticity, uniformly with respect to x and t.

Under these assumptions we have already seen in Sect. 5.3 that condition (i) in the existence and uniqueness theorem is satisfied, with

$$\sigma > \max(0, -\mu)\,,$$

where $\mu = \inf_{D\times(0,T)} a_0 - \frac{1}{2\alpha_0}\|b\|^2_{L^\infty(D\times(0,T))}$ and

$$\alpha = \min\left(\frac{\alpha_0}{2}, \sigma + \mu\right)\,.$$

Thus $a(t; w, v)$ is uniformly weakly coercive in $H^1(D)$.

Then we have to check that V and H satisfy the required properties. First of all, it is well-known that $L^2(D)$ is a separable Hilbert space. Moreover, $H^1_0(D)$ and $H^1_{\Gamma_D}(D)$ are closed subspaces of $H^1(D)$, which is a separable Hilbert space (see Remark 4.9); thus they are separable Hilbert spaces (if M is a closed subspace of a separable Hilbert space X and S is a countable set dense in X, the orthogonal projection $P_M S \subset M$ is a countable set dense in M). We also have that

$$C^\infty_0(D) \hookrightarrow H^1_0(D) \hookrightarrow H^1_{\Gamma_D}(D) \hookrightarrow H^1(D) \hookrightarrow L^2(D)$$

and we know that $C^\infty_0(D)$ is dense in $L^2(D)$; therefore for all the boundary value problems we have $V \hookrightarrow H$ with continuous and dense immersion.

On the data, we assume that $u_0 \in L^2(D)$ and we remember that in the four cases the linear and continuous functional F is defined as follows:

Dirichlet $F(t) \in V'$ is given by $v \to \langle F(t), v\rangle = \int_D f(t)v\,dx.$

Neumann $F(t) \in V'$ is given by $v \to \langle F(t), v\rangle = \int_D f(t)v\,dx + \int_{\partial D} g(t)v\,dS_x.$

Mixed $F(t) \in V'$ is given by $v \to \langle F(t), v\rangle = \int_D f(t)v\,dx + \int_{\Gamma_N} g(t)v\,dS_x.$

Robin $F(t) \in V'$ is given by $v \to \langle F(t), v\rangle = \int_D f(t)v\,dx + \int_{\partial D} g(t)v\,dS_x.$

Thus we assume that $f \in L^2(D \times (0, T))$, $g \in L^2(\partial D \times (0, T))$ (for the Neumann and Robin cases) or $g \in L^2(\Gamma_N \times (0, T))$ (for the mixed case), and we conclude that Theorem 9.3 can be applied.

As a final remark, one easily sees that in some case weaker assumptions would be sufficient, for instance $f \in L^2(0, T; (H^1_0(D))')$ for the Dirichlet boundary value problem.

9.2.2 Application to Linear Navier–Stokes Equations for Incompressible Fluids

The abstract theory we have presented can also be used for the analysis of the incompressible linear Navier–Stokes equations. They read as follows:

$$\begin{cases} \frac{\partial u}{\partial t} - \nu \Delta u + \nabla p = f & \text{in } D \times (0, T) \\ \operatorname{div} u = 0 & \text{in } D \times (0, T) \\ u_{|\partial D} = 0 & \text{on } \partial D \times (0, T) \\ u_{|t=0} = u_0 & \text{in } D\,, \end{cases} \qquad (9.7)$$

where u is the velocity of the fluid, p the pressure (per unit density), f the (acceleration of the) external force field, u_0 the initial velocity and $\nu > 0$ the kinematic viscosity. As usual, $D \subset \mathbb{R}^n$ is assumed to be a bounded, connected, open set with a Lipschitz continuous boundary ∂D.

Multiplying for a test function $w = w(x)$ satisfying $\operatorname{div} w = 0$ in D and $w = 0$ on ∂D and integrating by parts we obtain

$$\begin{aligned} &\int_D f \cdot w dx \\ &\quad = \int_D \frac{\partial u}{\partial t} \cdot w dx - \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i \mathcal{D}_i u_j\, w_j dx + \int_D \nabla p \cdot w dx \\ &\quad = \frac{d}{dt} \int_D u \cdot w dx + \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i u_j\, \mathcal{D}_i w_j dx - \nu \int_{\partial D} \sum_{i,j=1}^n n_i \mathcal{D}_i u_j\, w_j dS_x \\ &\qquad\quad - \int_D p \operatorname{div} w dx + \int_{\partial D} p\, w \cdot n dS_x \\ &\quad = \frac{d}{dt} \int_D u \cdot w dx + \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i u_j\, \mathcal{D}_i w_j dx\,. \end{aligned}$$

Thus the term related to the pressure disappears, and we are left with the variational problem

$$\frac{d}{dt} \int_D u \cdot w dx + \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i u_j\, \mathcal{D}_i w_j dx = \int_D f \cdot w dx \qquad \forall\, w \in V\,, \qquad (9.8)$$

set in

$$V = \{w \in (H_0^1(D))^n \mid \operatorname{div} w = 0 \text{ in } D\}$$

for the only unknown u. Concerning the space H we set

$$H = \{w \in (L^2(D))^n \mid \operatorname{div} w = 0 \text{ in } D, w \cdot n = 0 \text{ on } \partial D\},$$

with the $L^2(D)$-norm.

It is known that for $w \in H$ the normal trace $(w \cdot n)_{|\partial D}$ has a meaning on ∂D, and that the map $w \mapsto (w \cdot n)_{|\partial D}$ is bounded from H to a suitable trace space (see Exercise 5.9). Thus V and H are closed subspaces of $(H_0^1(D))^n$ and $(L^2(D))^n$, respectively, and therefore they are separable Hilbert spaces. Clearly we have $V \hookrightarrow H$, with continuous immersion; moreover, it is known that the space $\mathcal{V} = \{w \in (C_0^\infty(D))^n \mid \operatorname{div} w = 0 \text{ in } D\}$ is dense in H (see, e.g., Girault and Raviart [12, Theor. 2.8, p. 30]), thus V is dense in H.

The bilinear form

$$a(v, w) = \nu \int_D \sum_{i,j=1}^n \mathcal{D}_i v_j \, \mathcal{D}_i w_j$$

is clearly bounded in $V \times V$ and also weakly coercive in $V \times V$, for any constant $\sigma > 0$ (and $\alpha = \min(\nu, \sigma)$). If we assume $u_0 \in H$ and consider the functional

$$w \to \langle F, w\rangle = \int_D f \cdot w dx$$

for $f \in L^2(0, T; (L^2(D))^n)$, we easily check that all the assumptions of Theorem 9.3 are satisfied, and we conclude that there exists a unique solution $u \in L^2(0, T; V)$ of Eq. (9.8), satisfying $u' \in L^2(0, T; V')$ and $u(0) = u_0$.

Recovering the pressure needs some additional work. For that, we refer to the results of Chap. 8. The couple of Hilbert spaces V and M that appear in Sect. 8.1.2 will be denoted here by X and M (the notation V in this section is always used for $\{w \in (H_0^1(D))^n \mid \operatorname{div} w = 0 \text{ in } D\}$), and are given by $X = (H_0^1(D))^n$ and $M = L^2_*(D)$. The bilinear form $b(\cdot, \cdot)$ is given by $b(v, r) = \int_D \operatorname{div} v \, r \, dx$, and the operators $B : X \mapsto M'$, $B^T : M \mapsto X'$ are defined by $\langle Bv, r\rangle = b(v, r) = \langle B^T r, v\rangle$ for $v \in X$ and $r \in M$. In particular, it is easily checked that the kernel of the operator B is $N(B) = V$.

We have seen in Sect. 8.1.2 that with these choices the inf–sup condition 8.7 is satisfied. Thus from Proposition 8.2 and Theorem 8.4 we have that $N(B^T) = \{0\}$, $R(B^T)$ is closed in X' and $R(B^T) = N(B)_\sharp = V_\sharp$.

We have just proved that there exists $u \in L^2(0, T; V) \cap C^0([0, T]; H)$ with $u' \in L^2(0, T; V')$ satisfying $u(0) = u_0$ and (9.8), namely,

$$\frac{d}{dt}(u(t), w)_{L^2(D)} + a(u(t), w) = (f(t), w)_{L^2(D)} \qquad \forall w \in V . \tag{9.9}$$

Defining for $t \in [0, T]$

$$U(t) = \int_0^t u(s)\,ds \;\;, \;\; F(t) = \int_0^t f(s)\,ds\,,$$

we have $U \in C^0([0, T]; V)$ and $F \in C^0([0, T]; (L^2(D))^n)$; thus integrating (9.9) from 0 to t we find

$$(u(t), w)_{L^2(D)} - (u_0, w)_{L^2(D)} + a(U(t), w) - (F(t), w)_{L^2(D)} = 0 \qquad \forall\, w \in V\,.$$

For each $t \in [0, T]$ the functional

$$v \mapsto (u(t), v)_{L^2(D)} - (u_0, v)_{L^2(D)} + a(U(t), v) - (F(t), v)_{L^2(D)} \;\;,\;\; v \in X\,,$$

belongs to X' and vanishes on $V = N(B)$, namely, it belongs to $N(B)_\sharp$. Therefore it belongs to $R(B^T)$, so that for each $t \in [0, T]$ there exists a unique $P(t) \in M$ such that

$$(u(t), v)_{L^2(D)} - (u_0, v)_{L^2(D)} + a(U(t), v) - (F(t), v)_{L^2(D)} = b(v, P(t)) \qquad \forall\, v \in X\,,$$

(and moreover it is easily shown that $P \in C^0([0, T]; M)$). Taking the (weak) time derivative we obtain

$$\frac{d}{dt}[(u(t), v)_{L^2(D)} - b(v, P(t))] = -a(u(t), v) + (f(t), v)_{L^2(D)} \tag{9.10}$$

for all $v \in X$ and almost all $t \in [0, T]$.

A stronger form of this equation can be derived having additional information on the smoothness of u and P (see, e.g., Dautray and Lions [7, Chap. XIX, §2.3]). At a formal level, the pressure p appearing in the Navier–Stokes equations is given by $p(t) = \frac{\partial}{\partial t} P(t)$.

Remark 9.3 It is not difficult to check that a similar analysis can be performed assuming $f \in L^2(0, T; ((H_0^1(D))^n)')$ instead of $f \in L^2(0, T; (L^2(D))^n)$. This is not the case if one assumes $f \in L^2(0, T; V')$, as the space V does not contain $(C_0^\infty(D))^n$, therefore the use of weak derivatives is not always justified and the interpretation of the variational problem does not necessarily lead to the partial differential equations at hand.

9.3 Maximum Principle for Parabolic Problems

The maximum principle also holds in the case of parabolic problems. Let us start with some definitions, that are similar to those given for elliptic problems.

Definition 9.2 We say that $u \in L^2(0, T; V)$ with $u' \in L^2(0, T; V')$ is a subsolution for the operator

$$\frac{\partial u}{\partial t} + Lu$$

if the inequality

$$\langle u'(t), v \rangle + a(t; u(t), v) \leq 0 \tag{9.11}$$

holds for almost all $t \in [0, T]$ and for all $v \in H_0^1(D)$ such that $v \geq 0$ a.e. in D.

A similar definition is given for a supersolution: it is enough to say that $-u$ is a subsolution.

Theorem 9.4 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary ∂D. Let L be the elliptic operator*

$$Lw = -\sum_{i,j=1}^{n} \mathcal{D}_i(a_{ij}\mathcal{D}_j w) + \sum_{i=1}^{n} b_i \mathcal{D}_i w + a_0 w \,,$$

with bounded coefficients $a_{ij} = a_{ij}(x, t)$, $b_i = b_i(x, t)$, $a_0 = a_0(x, t)$. Assume that $a_0(x, t) \geq 0$ a.e. in $D \times (0, T)$. Then if u is a subsolution for L we have

$$\sup_{D\times[0,T]} u \leq \sup_{S_T} u^+ = \max\left(\sup_{S_T} u, 0\right) \,,$$

where $S_T = (\partial D \times [0, T]) \cup (D \times \{0\})$. Similarly, if u is a supersolution for L we have

$$\inf_{D\times[0,T]} u \geq \inf_{S_T}(-u^-) = \min\left(\inf_{S_T} u, 0\right) \,.$$

Proof For the sake of simplicity, the proof we present is somehow formal. The lines of a rigorous proof can be found in Dautray and Lions [5, Theorem 1, p. 252] (indeed, under someway different assumptions on the regularity of u and the coefficients; there a good exercise is also to find out and correct some misprints...); a complete presentation is in Ladyženskaja, Solonnikov and Ural'ceva [17, Chapter III, §7].

Let us start from the case of the subsolution. Set $M = \sup_{S_T} u^+$; we can assume M to be finite, otherwise we have nothing to prove, and clearly $M \geq 0$. Choose $v(t) = \max(u(t) - M, 0)$, so that $v(t) \in H_0^1(D)$ and $v(t) \geq 0$ for almost all $t \in [0, T]$. When considering the maximum principle for the elliptic case, we have already noted that $\nabla v(t) = \nabla u(t)$ in $\{u(t) > M\}$, while $v(t) = 0$ and $\nabla v(t) = 0$ in $\{u(t) \leq M\}$. Thus

$$\int_D \sum_{i,j=1}^n a_{ij}(t)\mathcal{D}_j u(t)\mathcal{D}_i v(t) dx = \int_D \sum_{i,j=1}^n a_{ij}(t)\mathcal{D}_j v(t)\mathcal{D}_i v(t) dx \geq \alpha_0 \int_D |\nabla v(t)|^2 dx\,.$$

Moreover, and similarly to what we have just seen

$$\langle u'(t), v(t)\rangle = \langle v'(t), v(t)\rangle = \frac{1}{2}\frac{d}{dt}\int_D v(t)^2 dx\,,$$

as in $\{u(t) > M\}$ we have $v(t) = u(t) - M$, while in $\{u(t) \leq M\}$ it holds $v(t) = 0$ (here the argument is a little bit formal, but let us go on…; for a detailed proof see Ladyženskaja, Solonnikov and Ural'ceva [17, Theorem 7.2, p. 188]).

Finally

$$\begin{aligned}
-\int_D \sum_{i=1}^n b_i(t)\mathcal{D}_i u(t)\, v(t) dx &= -\int_{\{u(t)>M\}} \sum_{i=1}^n b_i(t)\mathcal{D}_i v(t)\, v(t) dx \\
&\leq \|b\|_{L^\infty(D\times(0,T))} \int_D |\nabla v(t)|\, |v(t)| dx \\
&\leq \frac{\alpha_0}{2}\int_D |\nabla v(t)|^2 dx + \frac{\|b\|^2_{L^\infty(D\times(0,T))}}{2\alpha_0}\int_D |v(t)|^2 dx
\end{aligned}$$

and

$$-\int_D a_0(t)u(t)v(t)dx = -\int_{\{u(t)>M\}} \overbrace{a_0(t)}^{\geq 0}\ \overbrace{u(t)}^{\geq M \geq 0}\ \overbrace{(u(t)-M)}^{\geq 0} dx \leq 0\,.$$

From (9.11) we have thus obtained the following inequality

$$\frac{1}{2}\frac{d}{dt}\int_D v(t)^2 dx + \frac{\alpha_0}{2}\int_D |\nabla v(t)|^2 dx \leq \frac{\|b\|^2_{L^\infty(D\times(0,T))}}{2\alpha_0}\int_D |v(t)|^2 dx$$

for almost all $t \in [0, T]$. Integrating in $[0, \tau]$, $\tau \in [0, T]$, it follows

$$\begin{aligned}
&\|v(\tau)\|^2_{L^2(D)} + \alpha_0 \int_0^\tau \|\nabla v(t)\|^2_{L^2(D)} dt \\
&\qquad \leq \|v(0)\|^2_{L^2(D)} + \frac{\|b\|^2_{L^\infty(D\times(0,T))}}{\alpha_0}\int_0^\tau \|v(t)\|^2_{L^2(D)} dt\,.
\end{aligned} \tag{9.12}$$

Since $v(0) = \max(u(0) - M, 0) = 0$, from Gronwall lemma E.2 it follows $v(\tau) = 0$ and therefore $u(\tau) \le M$ for $\tau \in [0, T]$.

For the supersolution, just note that if u is a supersolution, then $-u$ is a subsolution, and $(-u)^+ = u^-$. □

Remark 9.4 If we have $\frac{\partial u}{\partial t} + Lu = f \ge 0$ in $D \times (0, T)$, $u(t)_{|\partial D} \ge 0$, $u_{|t=0} \ge 0$, by the change of variable $\hat{u}(t) = e^{-kt}u(t)$, $k \ge -\inf_{D\times(0,T)} a_0$, we can easily prove that $u(t) \ge 0$ for all $t \in [0, T]$. In fact, with respect to $\hat{u}$ the problem is related to a bilinear form with the coefficient of the zero order term, say $\hat{a}_0$, that satisfies $\hat{a}_0 \ge 0$. Since $\hat{u}(t)_{|\partial D} \ge 0$, $\hat{u}_{|t=0} \ge 0$ and $\hat{f}(t) = e^{-kt} f(t) \ge 0$, it follows $\hat{u}(t) \ge 0$ and consequently $u(t) \ge 0$.

In other words, if you maintain a positive temperature on the walls of a room in which the temperature was positive at the initial time and in which you are injecting heat, then the temperature in the room will remain positive for all the subsequent time. Do you see the power of mathematics?

Remark 9.5 If $a_0 = 0$, one can substitute $\sup_{S_T} u^+$ with $\sup_{S_T} u$ (and $\inf_{S_T}(-u^-)$ with $\inf_{S_T} u$). In fact, the same proof applies choosing $M = \sup_{S_T} u$ (which now is no longer non-negative) and $v = \max(u - M, 0)$. This yields inequality (9.12) and the thesis follows.

9.4 Exercises

Exercise 9.1 Suppose that V and H are two Hilbert spaces, that V is immersed in H with continuity and that V is dense in H. Then H', the dual space of H, is immersed with continuity in V', the dual space of V. Moreover, H' is dense in V'.

Solution Take an element $\Phi \in H'$, which by the Riesz representation theorem 3.1 can be written as $\Phi(h) = (\omega_\Phi, h)_H$ for each $h \in H$, with $\omega_\Phi \in H$. To this functional we can associate the element $\Psi \in V'$ given by $\Psi(v) = (\omega_\Phi, v)_H$ for each $v \in V$. We want to show that the map from $\Phi \in H'$ to $\Psi \in V'$, which is clearly continuous, is one-to-one. Thus suppose that there exists $\Theta \in H'$ given by $(\omega_\Theta, v)_H$ and such that $(\omega_\Theta, v)_H = (\omega_\Phi, v)_H$ for each $v \in V$. Take $h \in H$: since V is dense in H there exists a sequence $v_k \in V$ such that $v_k \to h$ in H. Therefore $(\omega_\Theta, v_k)_H \to (\omega_\Theta, h)_H$ and $(\omega_\Phi, v_k)_H \to (\omega_\Phi, h)_H$, and consequently $(\omega_\Theta, h)_H = (\omega_\Phi, h)_H$ for each $h \in H$, namely, $\Theta = \Phi$ in H'.

Concerning the density result, by the projection theorem (see, e.g., Yosida [28, Theorem 1, p. 82]) it is enough to show that in V' it holds $(\overline{H'})^\perp = \{0\}$ (or, equivalently, $(H')^\perp = \{0\}$). Take $\Psi \in V'$ that satisfies $(\Psi, \Phi)_{V'} = 0$ for each $\Phi \in H'$. By the Riesz representation theorem in V we have $\Psi(v) = (\omega_\Psi, v)_V$ and $\Phi(v) = (\omega_\Phi, v)_V$ for $\omega_\Psi \in V$, $\omega_\Phi \in V$ and for each $v \in V$, and finally $(\Psi, \Phi)_{V'} = (\omega_\Psi, \omega_\Phi)_V$; moreover, by the same theorem in H we know that $\Phi(h) = (\widehat{\omega}_\Phi, h)_H$ for $\widehat{\omega}_\Phi \in H$ and for each $h \in H$, then $\Phi(v) = (\widehat{\omega}_\Phi, v)_H$

for each $v \in V$. Thus we can write

$$0 = (\Psi, \Phi)_{V'} = (\omega_\Psi, \omega_\Phi)_V = \Phi(\omega_\Psi) = (\widehat{\omega}_\Phi, \omega_\Psi)_H \,.$$

This is true for each $\Phi \in H'$, in particular, fixed any $q \in H$, for the functional $\Phi_\star$ given by $h \mapsto (q, h)_H = \Phi_\star(h)$. Therefore $\widehat{\omega}_{\Phi_\star} = q$, and we conclude that $(q, \omega_\Psi)_H = 0$ for each $q \in H$. Thus $\omega_\Psi = 0$ in H and also $\omega_\Psi = 0$ in V, as V is immersed in H. In conclusion $\Psi = 0$, and the result is proved.

[Here it could be interesting to open a parenthesis: making the identification of H with H', we obtain the chain $V \hookrightarrow H \approx H' \hookrightarrow V'$; with a further step, the identification of V with V' seems to imply that all the four spaces V, H, H' and V' are the same. This is clearly too much for the educated reader: thus, what is wrong?

What we have surely seen is that, if $V \hookrightarrow H$ with continuous and dense immersion, then $H' \hookrightarrow V'$ with continuous and dense immersion. The situation is completely symmetric, thus we can now decide to identify, by means of the Riesz representation theorem in H, the dual H' with H, obtaining the chain $V \hookrightarrow H \approx H' \hookrightarrow V'$, or, alternatively, to identify, by means of the Riesz representation theorem in V, the dual V' with V, obtaining the other chain $H' \hookrightarrow V' \approx V \hookrightarrow H$. We can make only one of these identifications: the "glue" can be used only once, and everything depends on the choice of the "pivot" space: either H or V.

The most typical example we have in mind is $V = H^1_0(D)$, $H = L^2(D)$ and $V' = H^{-1}(D) = (H^1_0(D))'$. Everything should be clear when we think at the identification of $L^2(D)$ with its dual, obtaining the chain

$$H^1_0(D) \hookrightarrow L^2(D) \approx (L^2(D))' \hookrightarrow H^{-1}(D) \,.$$

But what happens when we decide to identify $H^1_0(D)$ with its dual $H^{-1}(D)$? Then the chain becomes

$$(L^2(D))' \hookrightarrow H^{-1}(D) \approx H^1_0(D) \hookrightarrow L^2(D) \,,$$

and again it seems that everything collapses on a single space, as we are used to think that $(L^2(D))'$, the dual of $L^2(D)$, is equal to $L^2(D)$.

Keep calm and carry on: the identification of $H^1_0(D)$ with its dual $H^{-1}(D)$ has been done by means of the Riesz representation theorem in $H^1_0(D)$. This signifies that any element $\Psi \in H^{-1}(D)$ has been represented by means of the scalar product in $H^1_0(D)$ in the following way: Ψ is identified to the element $\omega_\Psi \in H^1_0(D)$ that satisfies

$$\int_D \omega_\Psi v + \int_D \nabla\omega_\Psi \cdot \nabla v = \Psi(v) \qquad \forall\, v \in H^1_0(D) \,.$$

On its side, an element $\Phi \in (L^2(D))'$, our old nice dual space, is a functional $q \mapsto \Phi(q) = \int_D \widehat{\omega}_\Phi q$ with $\omega_\Phi \in L^2(D)$ and $q \in L^2(D)$. In particular, it is also an element of $H^{-1}(D)$, and thus can be identified with the solution $\omega_\Phi \in H_0^1(D)$ of

$$\int_D \omega_\Phi v + \int_D \nabla\omega_\Phi \cdot \nabla v = \Phi(v) = \int_D \widehat{\omega}_\Phi v \qquad \forall\, v \in H_0^1(D)\,,$$

namely, the solution ω_Φ of $-\Delta\omega_\Phi + \omega_\Phi = \widehat{\omega}_\Phi$ in D, with homogeneous Dirichlet boundary condition. In conclusion, if we are using the identification $H_0^1(D) \approx H^{-1}(D)$, an element $\Phi \in (L^2(D))'$ is (identified to) an element of the space

$$Q = \{v \in H_0^1(D) \mid \Delta v \in L^2(D)\}$$

(which, if the boundary ∂D is of class C^2, by Theorem 7.12 coincides with $H^2(D) \cap H_0^1(D)$), and we can rewrite the chain above as

$$Q \hookrightarrow H_0^1(D) \hookrightarrow L^2(D)\,.$$

Do you feel this digression too long and impenetrable? In case, just skip it. . .]

Exercise 9.2 Let V be a Hilbert space, and suppose that $v_k \in V$ converges to v in V and that w_k converges weakly to w in V. Then $(v_k, w_k)_V \to (v, w)_V$.

Solution First of all, let us note that a weakly convergent sequence in a Hilbert space is bounded (see Yosida [28, Theorem 1, p. 120])). Then we have

$$\begin{aligned}|(v_k, w_k)_V - (v, w)_V| &= |(v_k - v, w_k)_V + (v, w_k - w)_V| \\ &\le |(v_k - v, w_k)_V| + |(v, w_k - w)_V| \\ &\le \|v_k - v\|_V \|w_k\|_V + |(v, w_k - w)_V|\,.\end{aligned}$$

Being w_k bounded, the first term goes to 0; since for any $v \in V$ the linear functional $\psi \mapsto (v, \psi)_V = F_v(\psi)$ is bounded, from the weak convergence of w_k to w it follows $F_v(w_k - w) \to 0$, and the result is proved.

Exercise 9.3 Let V be a separable Hilbert space. Then V' is a separable Hilbert space, too.

Solution From Remark 2.6 we already know that V' is a Hilbert space. Since V is separable, we have a countable set $\{v_k\}_{k=1}^\infty$ that is dense in V. Consider the countable set of linear and bounded functionals given by $\Psi_k(v) = (v_k, v)_V$ and take now $\Psi \in V'$; by the Riesz representation theorem it can be written as $\Psi(v) = (\omega_\Psi, v)_V$ for a suitable $\omega_\Psi \in V$ and for each $v \in V$. For each $\epsilon > 0$ there exists an element $v_{k_\star}$ such that $\|\omega_\Psi - v_{k_\star}\|_V < \epsilon$. For $\Psi_{k_\star}$ we have

$$\|\Psi - \Psi_{k_\star}\|_{V'} = \sup_{v\in V, v\neq 0} \frac{|\Psi(v) - \Psi_{k_\star}(v)|}{\|v\|_V} = \sup_{v\in V, v\neq 0} \frac{|(\omega_\Psi - v_{k_\star}, v)_V|}{\|v\|_V} < \epsilon$$

by the Cauchy–Schwarz inequality, and thus the proof is complete.

Exercise 9.4 Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary ∂D. Consider the problem

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = 0 & \text{in } D \times (0, +\infty) \\ u_{|\partial D} = 0 & \text{on } \partial D \times (0, +\infty) \\ u_{|t=0} = u_0 & \text{in } D\,, \end{cases}$$

where $u_0 \in L^2(D)$. Show that:

(i) there exists a unique solution $u \in L^2(0, +\infty; H_0^1(D)) \cap C^0([0, +\infty); L^2(D))$ with $u' \in L^2(0, +\infty; L^2(D))$;
(ii) $\lim_{t \to +\infty} \|u(t)\|_{L^2(D)} = 0$.

Solution

(i) Looking at the proof of Theorem 9.3 we easily see that, for a right hand side $f = 0$, it is possible to prove the existence of a solution $u(t)$ for $t \in [0, +\infty)$, and moreover the estimate

$$\|u(\tau)\|^2_{L^2(D)} + \int_0^\tau \|\nabla u(t)\|^2_{L^2(D)} dt \leq \|u_0\|^2_{L^2(D)} \tag{9.13}$$

holds for each $\tau \in [0, +\infty)$.
(ii) Using the Poincaré inequality (6.2) in (9.13) we find

$$\|u(\tau)\|^2_{L^2(D)} + \sigma \int_0^\tau \|u(t)\|^2_{L^2(D)} dt \leq \|u_0\|^2_{L^2(D)} \tag{9.14}$$

for each $\tau \in [0, +\infty)$, where $\sigma = \frac{1}{C_D}$. Now set $w(t) = e^{\sigma t} u(t)$. We obtain at once $w'(t) = e^{\sigma t} u'(t) + \sigma e^{\sigma t} u(t)$, thus

$$\langle w'(t), v \rangle + a(t; w(t), v) - \sigma (w(t), v)_{L^2(D)} = 0 \quad \forall\, v \in H_0^1(D)\,. \tag{9.15}$$

Since

$$a(t; w(t), w(t)) - \sigma (w(t), w(t))_{L^2(D)} = \int_D |\nabla w(t)|^2 dx - \sigma \int_D w(t)^2 dx \geq 0\,,$$

Eq. (9.15) and the relation $w(0) = u_0$ lead to the estimate

$$\frac{1}{2} \frac{d}{dt} \|w(t)\|^2_{L^2(D)} \leq 0$$

for almost all $t \in [0, T]$ and thus

$$\|w(\tau)\|^2_{L^2(D)} \leq \|u_0\|^2_{L^2(D)}$$

for each $\tau \in [0, +\infty)$. In conclusion $\|u(\tau)\|_{L^2(D)} \leq e^{-\sigma\tau}\|u_0\|_{L^2(D)} \to 0$ as $\tau \to +\infty$.

[From the physical point of view this result says that, if no heat is furnished and the boundary temperature is kept to 0, then the internal temperature goes to 0 as time becomes larger and larger: a well-known situation in our real life experience.]

Exercise 9.5 Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary ∂D. For $u_0 \in L^2(D)$, $f \in L^2(D \times (0,T))$ and $g \in L^2(\partial D \times (0,T))$ consider the Neumann problem for the heat equation

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{in } D \times (0,T) \\ \nabla u \cdot n = g & \text{on } \partial D \times (0,T) \\ u_{|t=0} = u_0 & \text{in } D\,, \end{cases}$$

whose solution $u \in L^2(0,T; H^1(D))$ with $u' \in L^2(0,T;(H^1(D))')$ exists and is unique by Theorem 9.3. Under the assumption $\int_D f(x,t)\,dx + \int_{\partial D} g(x,t)\,dS_x = 0$ for almost all $t \in [0,T]$ show that $\int_D u(x,t)\,dx = \int_D u_0(x)\,dx$ for each $t \in [0,T]$.

Solution The solution u satisfies the weak problem

$$\langle u'(t), v\rangle + \int_D \nabla u(t) \cdot \nabla v\,dx = \int_D f(t)\,v\,dx + \int_{\partial D} g(t)\,v\,dS_x$$

for each $v \in H^1(D)$ and for almost all $t \in [0,T]$. Choosing $v = 1$ it follows

$$\langle u'(t), 1\rangle = \int_D f(t)\,dx + \int_{\partial D} g(t)\,dS_x$$

for almost all $t \in [0,T]$. By the integration by parts formula in Theorem 9.2 we find for all $\tau \in [0,T]$

$$\begin{aligned} \int_0^\tau \Big(\int_D f(t)\,dx + \int_{\partial D} g(t)\,dS_x\Big)dt &= \int_0^\tau \langle u'(t), 1\rangle dt \\ &= \int_D u(\tau)\,dx - \int_D u(0)\,dx = \int_D u(\tau)\,dx - \int_D u_0\,dx\,. \end{aligned}$$

Hence we have obtained the balance equation

$$\int_D u(\tau)\,dx = \int_D u_0\,dx + \int_0^\tau \Big(\int_D f(t)\,dx + \int_{\partial D} g(t)\,dS_x\Big)dt\,,$$

and the condition $\int_D f(x,t)\,dx + \int_{\partial D} g(x,t)\,dS_x = 0$ for almost all $t \in [0,T]$ yields $\int_D u(\tau)\,dx = \int_D u_0\,dx$ for all $\tau \in [0,T]$.

Exercise 9.6 Propose a numerical scheme for finding the approximate solution of a parabolic problem which is based on the Galerkin approximation and on the backward Euler method for discretizing $\frac{\partial u}{\partial t}$.

Solution Let V_M be a finite dimensional subspace of V (not necessarily the space generated by the first M elements of an orthonormal basis of V), whose basis is denoted by $\{\phi_1,\ldots,\phi_M\}$. Choose a time-step $\tau = T/K > 0$, define $t_k = k\tau$, $k = 0,1,\ldots,K$, and consider the backward Euler approximation of the first order derivative:

$$\frac{u^{k+1} - u^k}{\tau} \approx u'(t_{k+1}) \quad , \quad k = 0,1,\ldots,K\,.$$

Then the parabolic equation

$$\langle u'(t), v\rangle + a(t; u(t), v) = \langle F(t), v\rangle$$

can be approximated by means of the following numerical scheme: being given $u_M^0 \in V_M$, a suitable approximation of the initial datum u_0, for each $k = 0,1,\ldots,K-1$ find $u_M^{k+1} \in V_M$, solution of the problem

$$\left(\frac{u_M^{k+1} - u_M^k}{\tau}, \phi_i\right)_H + a(t_{k+1}; u_M^{k+1}, \phi_i) = \langle F(t_{k+1}), \phi_i\rangle \quad , \quad i = 1,\ldots,M\,.$$

More explicitly, at each time step t_{k+1}, $k = 0,1,\ldots,K-1$, one has to solve the discretized elliptic problem

$$\frac{1}{\tau}\,(u_M^{k+1}, \phi_i)_H + a(t_{k+1}; u_M^{k+1}, \phi_i) = \frac{1}{\tau}\,(u_M^k, \phi_i)_H + \langle F(t_{k+1}), \phi_i\rangle \;,\; i = 1,\ldots,M\,.$$

This linear system is associated to the matrix $A_{ij}^{k+1} = \frac{1}{\tau}(\phi_j, \phi_i)_H + a(t_{k+1}; \phi_j, \phi_i)$. Note that if $a(t;\cdot,\cdot)$ is uniformly weakly coercive in $V \times V$, then for τ small enough the bilinear form $\frac{1}{\tau}\,(\cdot,\cdot)_H + a(t;\cdot,\cdot)$ is uniformly coercive in $V \times V$, hence the matrix A^{k+1} is uniformly positive definite for $k = 0,1,\ldots,K-1$.

Chapter 10
Hyperbolic PDEs

Hyperbolic equations have the form

$$\frac{\partial^2 u}{\partial t^2} + Lu = f \qquad \text{in } D \times (0, T)\,,$$

where L is an elliptic operator, whose coefficients can depend on t. The "prototype" is the *wave equation*

$$\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = f \qquad \text{in } D \times (0, T)\,,$$

with speed $c > 0$.

As for the parabolic equations, we have to add a boundary condition (one of those we have considered for elliptic problems: Dirichlet, Neumann, mixed, Robin). Since with respect to time we have a second order derivative, we also need to add *two* initial conditions, namely $u_{|t=0}$ and $\frac{\partial u}{\partial t}_{|t=0}$ have to be assigned in D.

In the first section of the chapter we present the abstract variational theory for second order evolution equations in Hilbert spaces; then the application of this theory to hyperbolic equations and to the Maxwell equations is described. The second section is concerned with an important property of the solutions: the finite propagation speed.

10.1 Abstract Problem

We again assume that we are given with a separable Hilbert space H and a separable Hilbert space V, with $V \hookrightarrow H$ with continuous and dense immersion. Assume that $u_0 \in V$, $u_1 \in H$ and $F \in L^2(0, T; H)$. We look for a solution $u \in L^2(0, T; V)$,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7_10

with $u' \in L^2(0, T; H)$ and $u'' \in L^2(0, T; V')$ of the problem

$$\langle u''(t), v \rangle + a(t; u(t), v) = (F(t), v)_H \tag{10.1}$$

for each $v \in V$ and almost all $t \in [0, T]$, with $u(0) = u_0$ and $u'(0) = u_1$. Let us remind that the derivatives have to be intended in the weak sense. Since $u \in L^2(0, T; V) \subset L^2(0, T; H)$ and $u' \in L^2(0, T; H)$, it follows that $u \in C^0([0, T]; H)$, thus the value $u(0)$ has a meaning; similarly, since $u' \in L^2(0, T; H) \subset L^2(0, T; V')$ and $u'' \in L^2(0, T; V')$, it follows that $u' \in C^0([0, T]; V')$, thus $u'(0)$ has a meaning. Note that, similarly to relation (9.1) valid for the parabolic case, if $u' \in L^2(0, T; H)$ and $u'' \in L^2(0, T; V')$ it holds

$$\frac{d}{dt}(u'(t), v)_H = \langle u''(t), v \rangle \tag{10.2}$$

for almost all $t \in [0, T]$ and each $v \in V$, where $\frac{d}{dt}$ has to be intended as the weak time derivative of the real valued function $t \mapsto (u'(t), v)_H$. Therefore (10.1) can be equivalently rewritten as

$$-\int_0^T (u'(t), v)_H \Phi'(t)dt + \int_0^T a(t; u(t), v)\Phi(t)dt = \int_0^T (F(t), v)_H \Phi(t)dt$$

for all $\Phi \in C_0^\infty(0, T)$ and each $v \in V$. Since under the present assumptions (9.1) can be written as

$$\frac{d}{dt}(u(t), v)_H = \langle u'(t), v \rangle = (u'(t), v)_H \,, \tag{10.3}$$

we also have

$$\frac{d^2}{dt^2}(u(t), v)_H = \frac{d}{dt}(u'(t), v)_H = \langle u''(t), v \rangle \,, \tag{10.4}$$

where $\frac{d^2}{dt^2}$ has to be intended as the second order weak time derivative of the real valued function $t \mapsto (u(t), v)_H$.

Let us now clarify the assumptions on the family of bilinear forms $t \mapsto a(t; \cdot, \cdot)$. We assume that

$$a(t; w, v) = \widehat{a}(t; w, v) + a_1(t; w, v) \,,$$

where $a_1(t; w, v)$, the "lower order part", satisfies

(i) $t \mapsto a_1(t; w, v)$ is measurable in $(0, T)$ for all $w, v \in V$
(ii) $|a_1(t; w, v)| \le C_1 \|w\|_V \|v\|_H$ for all $w, v \in V$ and almost all $t \in [0, T]$, with $C_1 > 0$ independent of $t \in [0, T]$,
whereas $\widehat{a}(t; w, v)$, in some sense the "principal part", satisfies

(iii) $t \mapsto \widehat{a}(t; w, v)$ is differentiable for $t \in [0, T]$ and for all $w, v \in V$ (the derivative of thus map will be denoted by $\widehat{a}'(t; w, v)$)
(iv) $|\widehat{a}'(t; w, v)| \leq \widehat{C}_1 \|w\|_V \|v\|_V$ for all $w, v \in V$ and almost all $t \in [0, T]$, with $\widehat{C}_1 > 0$ independent of $t \in [0, T]$
(v) $|\widehat{a}(t; w, v)| \leq \widehat{C}_0 \|w\|_V \|v\|_V$ for all $w, v \in V$ and almost all $t \in [0, T]$, with $\widehat{C}_0 > 0$ independent of $t \in [0, T]$
(vi) $\widehat{a}(t; v, v) + \sigma(v, v)_H \geq \alpha \|v\|_V^2$ for all $v \in V$ and all $t \in [0, T]$, where $\alpha > 0$ and $\sigma \geq 0$ are independent of $t \in [0, T]$
(vii) $\widehat{a}(t; w, v) = \widehat{a}(t; v, w)$ for all $w, v \in V$ and all $t \in [0, T]$ (symmetry of the principal part).

Let us underline from the very beginning that the symmetry of the principal part is a crucial point. The abstract theorem reads as follows.

Theorem 10.1 (Existence and Uniqueness) *Let H and V be two separable Hilbert spaces, with $V \hookrightarrow H$ with continuous and dense immersion. Assume $u_0 \in V$, $u_1 \in H$ and $F \in L^2(0, T; H)$. Assume that the family of bilinear forms $a(t; \cdot, \cdot)$ satisfies the hypothesis (i)–(vii) listed here above. Then there exists a solution $u \in L^2(0, T; V)$ of Eq. (10.1), with $u' \in L^2(0, T; H)$, $u'' \in L^2(0, T; V')$ and $u(0) = u_0$, $u'(0) = u_1$. Uniqueness also holds, under the additional assumption*

(viii) $|a_1(t; w, v)| \leq C_2 \|w\|_H \|v\|_V$ for all $w, v \in V$ and almost all $t \in [0, T]$, with $C_2 > 0$ independent of $t \in [0, T]$.

Remark 10.1 Note that one can obtain a better result, as it is true that $u \in C^0([0, T]; V)$ and $u' \in C^0([0, T]; H)$. For this result see, e.g., Dautray and Lions [6, Chapter XVIII, §5.5].

Proof The proof is obtained by approximation, by proceeding as in the parabolic case.

First Step Since V is separable, we have a countable orthonormal Hilbertian basis $\{\varphi_m\} \subset V$ (see, e.g., Brezis [4, Théor. V.10, p. 86]). Define $V_N = \text{span}\{\varphi_1 \ldots \varphi_N\} \subset V$. Since V is dense in H, we can select a sequence $u_{1,N} \in V_N$ such that $u_{1,N} \to u_1$ in H. Moreover, we also have $u_{0,N} \in V_N$ such that $u_{0,N} \to u_0$ in V. We look for

$$u^N(t) = \sum_{j=1}^{N} u_j^N(t)\varphi_j$$

such that $u^N(0) = u_{0,N}$ (this means $u_j^N(0) = (u_{0,N}, \varphi_j)_V$), $(u^N)'(0) = u_{1,N}$ (this means $(u_j^N)'(0) = (u_{1,N}, \varphi_j)_V$) and moreover

$$\langle (u^N)''(t), \varphi_l \rangle + a(t; u^N(t), \varphi_l) = (F(t), \varphi_l)_H$$

for almost all $t \in [0, T]$ and for all $l = 1, \dots, N$. Inserting the expression of $u^N(t)$, we find

$$\sum_{j=1}^{N} \langle \varphi_j, \varphi_l \rangle (u_j^N)''(t) + \sum_{j=1}^{N} a(t; \varphi_j, \varphi_l) u_j^N(t) = (F(t), \varphi_l)_H \,. \tag{10.5}$$

We have already verified in Theorem 9.3 that the matrix $\langle \varphi_j, \varphi_l \rangle$ is non-singular (it is symmetric and positive definite), thus this is a linear system of second order ordinary differential equations. Setting $q_j(t) = (u_j^N)'(t)$, it can be rewritten as a standard linear system of first order ordinary differential equations, thus we know that there exists a unique solution $(u_1^N(t), \dots, u_N^N(t))$, with $u_j^N \in C^1([0, T])$ and $(u_j^N)'' \in L^2(0, T)$.

Second Step We must now find suitable a-priori estimates for passing to the limit. Multiply Eq. (10.5) by $(u_l^N)'(t)$ and add over l. It holds

$$\begin{aligned}((u^N)''(t), (u^N)'(t))_H &+ \widehat{a}(t; u^N(t), (u^N)'(t)) \\ &= -a_1(t; u^N(t), (u^N)'(t)) + (F(t), (u^N)'(t))_H \,.\end{aligned}$$

We know that

$$((u^N)''(t), (u^N)'(t))_H = \frac{1}{2} \frac{d}{dt} \|(u^N)'(t)\|_H^2 \,.$$

Moreover

$$\widehat{a}(t; u^N(t), (u^N)'(t)) = \frac{1}{2} \frac{d}{dt} \widehat{a}(t; u^N(t), u^N(t)) - \frac{1}{2} \widehat{a}'(t; u^N(t), u^N(t)) \,,$$

due to the symmetry of $\widehat{a}(t; \cdot, \cdot)$. Finally, from assumption (ii),

$$| - a_1(t; u^N(t), (u^N)'(t))| \le C_1 \|u^N(t)\|_V \|(u^N)'(t)\|_H$$

and moreover

$$|(F(t), (u^N)'(t))_H| \le \|F(t)\|_H \|(u^N)'(t)\|_H \,.$$

Summarizing, for almost all $t \in [0, T]$ we have

$$\begin{aligned}&\frac{1}{2} \frac{d}{dt} \|(u^N)'(t)\|_H^2 + \frac{1}{2} \frac{d}{dt} \widehat{a}(t; u^N(t), u^N(t)) \le \\ &\le \frac{1}{2} |\widehat{a}'(t; u^N(t), u^N(t))| + C_1 \|u^N(t)\|_V \|(u^N)'(t)\|_H + \|F(t)\|_H \|(u^N)'(t)\|_H \\ &\le \frac{1}{2} \widehat{C}_1 \|u^N(t)\|_V^2 + C_1 \|u^N(t)\|_V \|(u^N)'(t)\|_H + \|F(t)\|_H \|(u^N)'(t)\|_H \,,\end{aligned}$$

having used assumption (iv). Integrating with respect to t on $[0, \tau]$ we have

$$
\begin{aligned}
&\frac{1}{2}\|(u^N)'(\tau)\|_H^2 + \frac{1}{2}\widehat{a}(\tau; u^N(\tau), u^N(\tau)) \\
&\quad \le \frac{1}{2}\|(u^N)'(0)\|_H^2 + \frac{1}{2}\widehat{a}(0; u^N(0), u^N(0)) + \frac{1}{2}\widehat{C}_1 \int_0^\tau \|u^N(t)\|_V^2 dt \\
&\qquad + C_1 \int_0^\tau \|u^N(t)\|_V \|(u^N)'(t)\|_H dt + \int_0^\tau \|F(t)\|_H \|(u^N)'(t)\|_H dt\,.
\end{aligned}
$$

Using the weak coerciveness of $\widehat{a}(t; \cdot, \cdot)$ we find

$$
\widehat{a}(\tau, u^N(\tau), u^N(\tau)) \ge \alpha \|u^N(\tau)\|_V^2 - \sigma \|u^N(\tau)\|_H^2\,.
$$

From the inequality $ab \le \frac{1}{2}a^2 + \frac{1}{2}b^2$ and using assumption (v) we get

$$
\begin{aligned}
&\alpha \|u^N(\tau)\|_V^2 + \|(u^N)'(\tau)\|_H^2 \\
&\quad \le \sigma \|u^N(\tau)\|_H^2 + \|u_{1,N}\|_H^2 + \widehat{C}_0 \|u_{0,N}\|_V^2 \\
&\qquad + C_* \Big[\int_0^T \|F(t)\|_H^2 dt + \int_0^\tau \Big(\|u^N(t)\|_V^2 + \|(u^N)'(t)\|_H^2 \Big) dt \Big]\,.
\end{aligned}
$$

Since $u_{0,N} \to u_0$ in V and $u_{1,N} \to u_1$ in H, we have $\|u_{0,N}\|_V^2 + \|u_{1,N}\|_H^2 \le$ const. Moreover, we have

$$
u^N(\tau) = \int_0^\tau (u^N)'(t) dt + \underbrace{u^N(0)}_{=u_{0,N}}\,,
$$

thus, noting that $(a+b)^2 \le 2(a^2+b^2)$ and using the Cauchy–Schwarz inequality, we obtain

$$
\begin{aligned}
\|u^N(\tau)\|_H^2 &\le \left(\left\| \int_0^\tau (u^N)'(t) dt \right\|_H + \|u_{0,N}\|_H \right)^2 \\
&\le 2 \left(\left(\int_0^\tau \|(u^N)'(t)\|_H dt \right)^2 + \|u_{0,N}\|_H^2 \right) \\
&\underbrace{\le}_{\text{Cauchy–Schwarz}} 2 \left(\tau \int_0^\tau \|(u^N)'(t)\|_H^2 dt + \|u_{0,N}\|_H^2 \right)\,.
\end{aligned}
$$

Note that this last series of inequalities is not needed if $\sigma = 0$, namely, if the bilinear form $\widehat{a}(t; \cdot, \cdot)$ is coercive and not only weakly coercive.

In conclusion, setting $Q(\tau) = \|u^N(\tau)\|_V^2 + \|(u^N)'(\tau)\|_H^2$, we have found

$$Q(\tau) \leq K_1 + K_2 \int_0^\tau Q(t)dt \qquad \text{for all } \tau \in [0, T] .$$

From Gronwall lemma E.2 we have

$$Q(\tau) \leq K_1 e^{K_2 \tau} \qquad \text{for all } \tau \in [0, T] ,$$

therefore u^N is bounded in $L^2(0, T; V)$ and $(u^N)'$ is bounded in $L^2(0, T; H)$, respectively (more precisely, in $L^\infty(0, T; V)$ and $L^\infty(0, T; H)$). Since $L^2(0, T; V)$ and $L^2(0, T; H)$ are Hilbert spaces, by known results in functional analysis we can select a subsequence (still denoted by u^N) such that $u^N \to u$ weakly in $L^2(0, T; V)$ and $(u^N)' \to w$ weakly in $L^2(0, T; H)$ (see Yosida [28, Theorem 1, p. 126, and Theorem of Eberlein–Shmulyan, p. 141]). It is an easy task to show that $w = u'$; in fact, for each $h \in H$ and $\eta \in C_0^\infty(0, T)$ by integration by parts we have

$$\int_0^T ((u^N)'(t), h)_H \eta(t)dt = -\int_0^T (u_N(t), h)_H \eta'(t)dt ,$$

and passing to the limit, using the weak convergence of u^N and $(u^N)'$ in $L^2(0, T; H)$, we obtain

$$\int_0^T (w(t), h)_H \eta(t)dt = -\int_0^T (u(t), h)_H \eta'(t)dt ,$$

namely, $u' = w$. Take now $\Phi \in C_0^\infty(0, T)$, $v \in V$ and $v_N \in V_N$ such that $v_N \to v$ in V (remember that $V_N = \text{span}\{\varphi_1, \ldots, \varphi_N\}$, where φ_j is an orthonormal Hilbertian basis of V). Set

$$\psi^N(t) = \Phi(t)v_N \quad , \quad \psi(t) = \Phi(t)v .$$

It is clear that $\psi^N \to \psi$ in $L^2(0, T; V)$ and $(\psi^N)' = \Phi' v_N$ converges to $\psi' = \Phi' v$ in $L^2(0, T; V)$.
Equation (10.5) can be rewritten as

$$\langle (u^N)''(t), w^N \rangle + a(t; u^N(t), w^N) = (F(t), w^N)_H \tag{10.6}$$

for each $w^N \in V_N$ and almost all $t \in [0, T]$; choosing $w^N = \psi^N(t)$ and integrating by parts in $(0, T)$, it follows

$$-\int_0^T \Big((u^N)'(t), (\psi^N)'(t)\Big)_H dt + \int_0^T a(t; u^N(t), \psi^N(t))dt$$
$$= \int_0^T (F(t), \psi^N(t))_H dt\,.$$

Passing to the limit we find

$$-\int_0^T (u'(t), v)_H \Phi'(t)dt + \int_0^T a(t; u(t), v)\Phi(t)dt = \int_0^T (F(t), v)_H \Phi(t)dt\,,$$

thus $u''(t) \in L^2(0, T; V')$ and Eq. (10.1) is satisfied.

Third Step The proof of the existence of a solution is completed if we show that $u(0) = u_0$ and $u'(0) = u_1$. Take $\Phi \in C^\infty([0, T])$ with $\Phi(T) = 0$ and $\Phi'(T) = 0$, and define as before $\psi^N(t) = \Phi(t)v_N$, $\psi(t) = \Phi(t)v$, with $v_N \in V_N$ and $v_N \to v$ in V. Integrating Eq. (10.1) on $(0, T)$ we find

$$\int_0^T \langle u''(t), v\rangle \Phi(t)dt = -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T (F(t), v)_H \Phi(t)dt\,.$$

On the other hand, integrating by parts twice on $(0, T)$ we obtain

$$\int_0^T \langle u''(t), v\rangle \Phi(t)dt$$
$$= \int_0^T (u(t), v)_H \Phi''(t)dt - \langle u'(0), v\rangle \Phi(0) + (u(0), v)_H \Phi'(0)$$

(for a similar computation see (10.8), where some additional explanations on the functional framework are also added). Thus

$$\int_0^T (u(t), v)_H \Phi''(t)dt - \langle u'(0), v\rangle \Phi(0) + (u(0), v)_H \Phi'(0)$$
$$= -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T (F(t), v)_H \Phi(t)dt\,.$$

Inserting $w^N = \psi^N(t)$ in Eq. (10.6), it follows, by integration by parts on $(0, T)$,

$$\int_0^T (u^N(t), (\psi^N)''(t))_H dt - (\underbrace{(u^N)'(0)}_{=u_{1,N}}, \psi^N(0))_H + (\underbrace{(u^N)(0)}_{=u_{0,N}}, (\psi^N)'(0))_H$$

$$= -\int_0^T a(t; u^N(t), \psi^N(t))dt + \int_0^T (F(t), \psi^N(t))_H dt$$

Then passing to the limit as $N \to +\infty$, we obtain

$$\int_0^T (u(t), v)_H \Phi''(t)dt - (u_1, v)_H \Phi(0) + (u_0, v)_H \Phi'(0)$$
$$= -\int_0^T a(t; u(t), v)\Phi(t)dt + \int_0^T (F(t), v)_H \Phi(t)dt\,,$$

and in conclusion

$$-\langle u'(0), v\rangle \Phi(0) + (u(0), v)_H \Phi'(0) = -(u_1, v)_H \Phi(0) + (u_0, v)_H \Phi'(0)\,.$$

Due to the arbitrariness of $\Phi(0)$ and $\Phi'(0)$ and v we conclude $u'(0) = u_1$ and $u(0) = u_0$.

Fourth Step Let us come to the proof of the uniqueness of the solution. It is better to divide the proof in two parts, and consider later the general case. In this step we thus make two additional assumptions: firstly that $a_1(t; \cdot, \cdot) = 0$ and secondly that $\widehat{a}(\cdot, \cdot)$ does not depend on $t \in [0, T]$.

Let us assume $F = 0$, $u_0 = 0$, $u_1 = 0$; thus Eq. (10.1) reads

$$\langle u''(t), v\rangle + \widehat{a}(u(t), v) = 0 \qquad (10.7)$$

for all $v \in V$ and for almost all $t \in [0, T]$. Here one would like to follow the same idea employed for the finite dimensional approximation: select a value t among those for which (10.7) is satisfied, and choose $v = u'(t)$. However, this cannot be done since u' does not belong to $L^2(0, T; V)$ but only to $L^2(0, T; H)$. Thus we adopt a classical procedure proposed by Olga A. Ladyženskaya[1] (see also Dautray and Lions [6, p. 572]), and we choose as a test function an antiderivative of u: precisely, for a fixed $s \in [0, T]$ set

$$v(t) = \begin{cases} \int_t^s u(\tau)d\tau & \text{if } 0 \le t \le s \\ 0 & \text{if } s \le t \le T\,. \end{cases}$$

[1] Ladyženskaya [16].

We have $v(t) \in V$ for every $t \in [0, T]$, $v \in L^2(0, T; V)$ and $v'(t) = -u(t)$ for $0 \leq t \leq s$, thus $v' \in L^2(0, s; V)$. Let us choose this $v = v(t)$ in Eq. (10.7); using the density of $C^\infty([0, s]; H)$ in the space

$$W(0, s; H; V') = \{w \in L^2(0, s; H) \mid w' \in L^2(0, s; V')\}$$

and the density of $C^\infty([0, s]; V)$ in the space

$$H^1(0, s; V) = \{w \in L^2(0, s; V) \mid w' \in L^2(0, s; V)\}$$

(the proof of these density results can be done as in Dautray and Lions [6, Lemma 1, p. 473]), due to the fact that $u' \in W(0, s; H; V')$ and $v \in H^1(0, s; V)$ we have

$$\begin{aligned}
\int_0^s \langle u''(t), v(t)\rangle\, dt &= -\int_0^s (u'(t), v'(t))_H\, dt \\
&\quad + \langle u'(s), v(s)\rangle - \langle u'(0), v(0)\rangle \\
&= \int_0^s (u'(t), u(t))_H\, dt \qquad (\text{since } u'(0) = 0 \text{ and } v(s) = 0) \\
&= \int_0^s \tfrac{1}{2}\tfrac{d}{dt}\|u(t)\|_H^2 dt = \tfrac{1}{2}\|u(s)\|_H^2 \qquad (\text{since } u(0) = 0)\,.
\end{aligned} \tag{10.8}$$

Moreover, for $0 \leq t \leq s$ it holds

$$\widehat{a}(u(t), v(t)) = -\widehat{a}(v'(t), v(t)) = -\frac{1}{2}\frac{d}{dt}[\widehat{a}(v(t), v(t))]\,, \tag{10.9}$$

where the last equality holds as $\widehat{a}(\cdot, \cdot)$ is symmetric and not depending on t. Thus integrating (10.7) over $(0, s)$ it follows

$$\begin{aligned}
0 &= \tfrac{1}{2}\|u(s)\|_H^2 - \int_0^s \tfrac{1}{2}\tfrac{d}{dt}\widehat{a}(v(t), v(t))dt \\
&= \tfrac{1}{2}\left(\|u(s)\|_H^2 + \widehat{a}(v(0), v(0))\right) \qquad (\text{since } v(s) = 0) \\
&\geq \tfrac{1}{2}\left(\|u(s)\|_H^2 + \alpha\|v(0)\|_V^2 - \sigma\|v(0)\|_H^2\right) \\
&\qquad\qquad (\text{since } \widehat{a}(\cdot\,,\,\cdot) \text{ is weakly coercive})\,.
\end{aligned} \tag{10.10}$$

We have $v(0) = \int_0^s u(\tau)d\tau$, thus

$$\|v(0)\|_H^2 \leq \left(\int_0^s \|u(\tau)\|_H d\tau\right)^2 \underbrace{\leq}_{\text{Cauchy–Schwarz}} s\int_0^s \|u(\tau)\|_H^2 d\tau\,, \tag{10.11}$$

and then

$$\|u(s)\|_H^2 + \alpha\|v(0)\|_V^2 \le \sigma s \int_0^s \|u(\tau)\|_H^2 d\tau$$
$$\le \sigma T \int_0^s \|u(\tau)\|_H^2 d\tau \quad \forall\, s \in [0,T]\,.$$

From Gronwall lemma E.2 it follows $\|u(s)\|_H = 0$ for $s \in [0,T]$ and uniqueness is proved.

Fifth Step Repeat now the uniqueness result without assuming that $a_1(t;\cdot,\cdot) = 0$ and $\widehat{a}(t;\cdot,\cdot)$ is independent of $t \in [0,T]$. Instead of (10.7) we have the equation

$$\langle u''(t), v(t)\rangle + \widehat{a}(t;u(t),v(t)) = -a_1(t;u(t),v(t))\,, \tag{10.12}$$

and instead of (10.9) we have

$$\begin{aligned}\widehat{a}(t;u(t),v(t)) &= -\widehat{a}(t;v'(t),v(t)) \\ &= -\frac{1}{2}\frac{d}{dt}[\widehat{a}(t;v(t),v(t))] + \frac{1}{2}\widehat{a}'(t;v(t),v(t))\,,\end{aligned} \tag{10.13}$$

thus

$$\begin{aligned}&\langle u''(t), v(t)\rangle - \frac{1}{2}\frac{d}{dt}[\widehat{a}(t;v(t),v(t))] \\ &= -a_1(t;u(t),v(t)) - \tfrac{1}{2}\widehat{a}'(t;v(t),v(t))\,.\end{aligned} \tag{10.14}$$

Therefore integrating (10.14) over $(0,s)$ and proceeding as in (10.8) and (10.10) (where one has to replace $\widehat{a}(\cdot,\cdot)$ by $\widehat{a}(0;\cdot,\cdot)$) it follows

$$-\int_0^s a_1(t;u(t),v(t))dt - \frac{1}{2}\int_0^s \widehat{a}'(t;v(t),v(t))dt$$
$$\ge \frac{1}{2}\left(\|u(s)\|_H^2 + \alpha\|v(0)\|_V^2 - \sigma\|v(0)\|_H^2\right)\,.$$

Using the boundedness of $\widehat{a}'(t;\cdot,\cdot)$ in $V \times V$ and of $a_1(t;\cdot,\cdot)$ in $H \times V$ (see assumptions (iv) and (viii)), we obtain

$$\|u(s)\|_H^2 + \alpha\|v(0)\|_V^2$$
$$\le \sigma\|v(0)\|_H^2 + 2C_2\int_0^s \|u(t)\|_H\|v(t)\|_V dt + \widehat{C}_1\int_0^s \|v(t)\|_V^2 dt$$
$$\underbrace{\le}_{2ab\le a^2+b^2} \sigma\|v(0)\|_H^2 + C_2\int_0^s \|u(t)\|_H^2 dt + (\widehat{C}_1 + C_2)\int_0^s \|v(t)\|_V^2 dt\,.$$

For $0 \le t \le T$ set now $w(t) = \int_0^t u(\tau)d\tau$. It holds $v(0) = w(s)$ and $v(t) = w(s) - w(t)$, for $0 \le t \le s$. Thus, using that $(a+b)^2 \le 2a^2 + 2b^2$, we can rewrite the last equation as

$$\begin{aligned}
&\|u(s)\|_H^2 + \alpha\|w(s)\|_V^2 \\
&\quad \le \sigma\|w(s)\|_H^2 + C^*\Big(\int_0^s \|u(t)\|_H^2 dt + \int_0^s \|w(s) - w(t)\|_V^2 dt\Big) \\
&\quad \le \sigma\|w(s)\|_H^2 + C^*\Big(\int_0^s \|u(t)\|_H^2 dt + 2\int_0^s \|w(t)\|_V^2 dt + 2s\|w(s)\|_V^2\Big),
\end{aligned}$$

where $C^* = \widehat{C}_1 + C_2$. Since $w(s) = v(0)$, we have already seen in (10.11) that

$$\|w(s)\|_H^2 \le s\int_0^s \|u(\tau\|_H^2 d\tau\,,$$

therefore for $0 \le s \le T_1 \le T$

$$\|w(s)\|_H^2 \le T_1\int_0^s \|u(\tau\|_H^2 d\tau\,,$$

and consequently

$$\begin{aligned}
&\|u(s)\|_H^2 + (\alpha - 2T_1C^*)\|w(s)\|_V^2 \\
&\qquad \le (\sigma T_1 + C^*)\int_0^s \|u(t)\|_H^2 dt + 2C^*\int_0^s \|w(t)\|_V^2 dt\,.
\end{aligned}$$

Choosing $T_1 > 0$ so small that $\alpha - 2T_1C^* \ge \frac{\alpha}{2}$, we can apply Gronwall lemma E.2 on the interval $[0, T_1]$ to the function $\eta(s) = \|u(s)\|_H^2 + \frac{\alpha}{2}\|w(s)\|_V^2$, thus obtaining $\eta(s) = 0$ for $s \in [0, T_1]$. Since T_1 only depends on the data of the problem through C^* and α, we can repeat the same argument on $[T_1, 2T_1]$ and so on. □

10.1.1 Application to Hyperbolic PDEs

Let us show some examples of hyperbolic problems that are covered by this abstract theory. Let $D \subset \mathbb{R}^n$ be a bounded, connected open set with a Lipschitz continuous boundary ∂D. The operator L will be as usual

$$Lw = -\sum_{i,j=1}^n \mathcal{D}_i(a_{ij}(t)\mathcal{D}_j w) + \sum_{i=1}^n b_i(t)\mathcal{D}_i w + a_0(t)w\,,$$

that we assume to be elliptic uniformly in space and time, namely, there exists $\alpha_0 > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\eta_j\eta_i \geq \alpha_0|\eta|^2$$

for all $\eta \in \mathbb{R}^n$, almost all $x \in D$ and all $t \in [0, T]$. The associated bilinear form, depending on the boundary conditions we have to consider, is

$$\begin{aligned} a(t; w, v) = &\int_D \sum_{i,j=1}^{n} a_{ij}(t)\mathcal{D}_j w \mathcal{D}_i v dx + \int_D \sum_{i=1}^{n} b_i(t)\mathcal{D}_i w v dx \\ &+ \int_D a_0(t) w v dx \left[+ \int_{\partial D} \kappa(t) w v dS_x \right], \end{aligned}$$

where the integral inside the square brackets is present only in the case of the Robin boundary condition. We define

$$\widehat{a}(t; w, v) = \int_D \sum_{i,j=1}^{n} a_{i,j}(t)\mathcal{D}_j w \mathcal{D}_i v dx \left[+ \int_{\partial D} \kappa(t) w v dS_x \right],$$

which is the bilinear form associated to the principal part.

We assume that a_{ij}, b_i, a_0 belong to $L^\infty(D \times (0, T))$, and that κ belongs to $L^\infty(\partial D \times (0, T))$, with $\kappa(x, t) \geq 0$ for almost all $x \in D$ and all $t \in [0, T]$ and $\int_{\partial D} \kappa(t) dS_x \neq 0$ for all $t \in [0, T]$. We also assume that $a_{ij}(x, t)$ is differentiable with respect to t in $[0, T]$ for almost all $x \in D$, and that $\frac{\partial a_{ij}}{\partial t}$ belongs to $L^\infty(D \times (0, T))$. Similarly we assume that $\kappa(x, t)$ is differentiable with respect to t in $[0, T]$ for almost all $x \in \partial D$, and that $\frac{\partial \kappa}{\partial t}$ belongs to $L^\infty(\partial D \times (0, T))$. Finally, we assume that the coefficient matrix of the principal part of the operator L is *symmetric*, i.e., that

$$a_{ij}(x, t) = a_{ji}(x, t) \quad \text{for almost all } x \in D \text{ and all } t \in [0, T].$$

With these hypotheses it is an easy task to verify that all the assumptions of the abstract Theorem 10.1 are satisfied, choosing H and V as in the parabolic case: in conclusion, the existence of a solution is assured.

Remark 10.2 Let us note that in the hyperbolic case, due to the presence of the second order time derivative, it is not possible to rewrite the given problem as a hyperbolic problem associated to a coercive bilinear form, by using a suitable change of variable (see Remark 9.2 and Exercise 10.5). However, it is possible to choose $\sigma = 0$ in the weak coerciveness assumption provided that the Poincaré inequality is satisfied (or the generalized Poincaré inequality in the case of the Robin problem); in other words, only in the case of the Neumann problem the principal part of the bilinear form is weakly coercive and not coercive.

Concerning uniqueness, we need to check that there exists $C_2 > 0$ such that

$$|a_1(t; w, v)| \leq C_2 \|w\|_H \|v\|_V \tag{10.15}$$

for all $w, v \in V$ and almost all $t \in [0, T]$, where $\|\cdot\|_H = \|\cdot\|_{L^2(D)}$, $\|\cdot\|_V = \|\cdot\|_{H^1(D)}$ and

$$a_1(t; w, v) = \int_D \sum_{i=1}^n b_i(t) \mathcal{D}_i w v dx + \int_D a_0(t) w v dx \, .$$

The second term satisfies (10.15), thus we only have to verify (10.15) for the first term. Let us integrate by parts formally (we will see here below when this is possible):

$$\begin{aligned} \int_D \sum_{i=1}^n b_i(t) \mathcal{D}_i w v dx &= -\int_D w \sum_{i=1}^n \mathcal{D}_i (b_i(t) v) dx + \int_{\partial D} w\, b(t) \cdot n\, v dS_x \\ &= -\int_D w \operatorname{div} b(t)\, v dx - \int_D w\, b(t) \cdot \nabla v dx + \int_{\partial D} w\, b(t) \cdot n\, v dS_x \, . \end{aligned}$$

Therefore we can easily verify that estimate (10.15) holds if for example:

(i) $\operatorname{div} b \in L^\infty(D \times (0, T))$, $V = H^1_0(D)$ (Dirichlet problem)
(ii) $\operatorname{div} b \in L^\infty(D \times (0, T))$, $b \cdot n = 0$ a.e. on $\partial D \times (0, T)$, $V = H^1(D)$ (Neumann or Robin problem)
(iii) $\operatorname{div} b \in L^\infty(D \times (0, T))$, $b \cdot n = 0$ a.e. on $\Gamma_N \times (0, T)$, $V = H^1_{\Gamma_D}(D)$ (mixed problem).

Thus, concerning the regularity of b, we can simply assume $b \in L^\infty(0, T; W^{1,\infty}(D))$ (so that, by the Sobolev immersion theorem 7.15, $b(t)_{|\partial D}$ and $b(t)_{|\Gamma_N}$ have a meaning). Clearly, all these conditions are satisfied if $b_i = 0$ for $i = 1, \ldots, n$.

10.1.2 Application to Maxwell Equations

The Maxwell equations describe the propagation of electromagnetic waves. In terms of the electric induction D, the electric field E, the magnetic induction B and the magnetic field H they read

$$\begin{cases} \frac{\partial B}{\partial t} + \operatorname{curl} E = 0 & \text{[Faraday]} \\ \frac{\partial D}{\partial t} - \operatorname{curl} H = -J_e & \text{[Maxwell–Ampère]} \\ \operatorname{div} B = 0 & \text{[Gauss magnetic]} \\ \operatorname{div} D = \rho & \text{[Gauss electric]} \\ B = \mu H \\ D = \epsilon E \,, \end{cases}$$

where J_e is the applied current density, ρ is the electric charge density, $\mu > 0$ is the magnetic permeability, $\epsilon > 0$ is the electric permittivity, and the operator curl is defined as

$$\operatorname{curl} Q = \nabla \times Q = \det \begin{bmatrix} i & j & k \\ \mathcal{D}_1 & \mathcal{D}_2 & \mathcal{D}_3 \\ Q_1 & Q_2 & Q_3 \end{bmatrix} .$$

Two initial conditions must be added: $E_{|t=0} = E_0$ and $H_{|t=0} = H_0$. Instead, the two Gauss equations can be left apart; the second one can be seen as a definition of the charge density, and gives as a consequence $\frac{\partial \rho}{\partial t} + \operatorname{div} J_e = 0$, the equation of conservation of the total electric charge (just take the divergence of the Maxwell–Ampère equation); the first one is always satisfied if it is satisfied at the initial time (just take the divergence of the Faraday equation).

Taking the curl of the first equation and the time derivative of the second one easily leads to

$$\frac{\partial^2 E}{\partial t^2} + c^2 \operatorname{curl}\operatorname{curl} E = -\frac{1}{\epsilon}\frac{\partial J_e}{\partial t} .$$

where $c^2 = \frac{1}{\mu\epsilon}$. Similarly, differentiating in time the first equation and taking the curl of the second one gives

$$\frac{\partial^2 H}{\partial t^2} + c^2 \operatorname{curl}\operatorname{curl} H = c^2 \operatorname{curl} J_e .$$

The two equations have the same structure, and from now on we will focus on the first one.

When the boundary of the physical domain $D \subset \mathbb{R}^3$ is a perfect conductor, the boundary condition for the electric field is $E \times n = 0$ on ∂D. We have therefore to solve the following problem:

$$\begin{cases} \dfrac{\partial^2 E}{\partial t^2} + c^2 \operatorname{curl}\operatorname{curl} E = -\dfrac{1}{\epsilon}\dfrac{\partial J_e}{\partial t} & \text{in } D \times (0, T) \\ E \times n = 0 & \text{on } \partial D \times (0, T) \\ E_{|t=0} = E_0 & \text{in } D \\ \dfrac{\partial E}{\partial t}_{|t=0} = E_1 & \text{in } D\,, \end{cases} \tag{10.16}$$

where $E_1 = \frac{1}{\epsilon}(\operatorname{curl} H_0 - J_{e|t=0})$.

Let $D \subset \mathbb{R}^3$ be a bounded, connected open set with a Lipschitz continuous boundary ∂D. A variational formulation is easily devised. Multiply the first equation by v, integrate in D and integrate by parts: using Theorem C.8 gives

$$\int_D \frac{\partial^2 E}{\partial t^2} \cdot v\, dx + c^2 \int_D \operatorname{curl} E \cdot \operatorname{curl} v\, dx + c^2 \int_{\partial D} n \times \operatorname{curl} E \cdot v\, dS_x \\ = -\int_D \frac{1}{\epsilon}\frac{\partial J_e}{\partial t} \cdot v\, dx\,.$$

The boundary integral can be rewritten as

$$\int_{\partial D} (v \times n) \cdot \operatorname{curl} E\, dS_x\,,$$

therefore it vanishes if we assume that the test function v satisfies $v \times n = 0$ on ∂D (as it is assumed for the electric field E).

In Exercise 5.8 we have introduced the space

$$H(\operatorname{curl}; D) = \{v \in (L^2(D))^3 \,|\, \operatorname{curl} v \in (L^2(D))^3\}$$

(the curl being intended in the weak sense), endowed with the scalar product

$$(w, v)_{\operatorname{curl}} = \int_D (\operatorname{curl} w \cdot \operatorname{curl} v + w \cdot v) dx\,, \tag{10.17}$$

and its closed subspace

$$H_0(\operatorname{curl}; D) = \{v \in H(\operatorname{curl}; D) \,|\, v \times n = 0 \text{ on } \partial D\}\,.$$

Both $H(\operatorname{curl}; D)$ and $H_0(\operatorname{curl}; D)$ are Hilbert spaces (see Exercise 10.1). We set $V = H_0(\operatorname{curl}; D)$ and $H = (L^2(D)^3$. Clearly, the immersion $V \hookrightarrow H$ is continuous; moreover, $(C_0^\infty(D))^3$ is dense in $H_0(\operatorname{curl}; D)$ (see, e.g., Monk [22, Theor. 3.33, p. 61]). By adapting the proof given in Adams [1, Theor. 3.5, p. 47], it is also easily verified that $H(\text{curl}; D)$ is separable, thus $H_0(\text{curl}; D)$ is separable, too.

Finally, we define the bilinear form

$$a(v, v) = c^2 \int_D \text{curl}\, v \cdot \text{curl}\, v \, dx \,,$$

which clearly is symmetric, bounded and weakly coercive in $V \times V$.

The Maxwell equations have therefore the following variational formulation:

$$\frac{d^2}{dt^2} \int_D E(t) \cdot v \, dx + a(E(t), v) = -\int_D \frac{1}{\epsilon} \frac{\partial J_e}{\partial t}(t) \cdot v \, dx \tag{10.18}$$

for all $v \in V$ and almost all $t \in [0, T]$. Assuming that $E_0 \in H_0(\text{curl}; D)$, $E_1 \in (L^2(D))^3$ and $\partial_t J_e \in L^2(0, T; (L^2(D))^3)$ we can apply Theorem 10.1 and obtain for this problem a unique solution $E \in L^2(0, T; H_0(\text{curl}; D))$, with $\partial_t E \in L^2(0, T; (L^2(D))^3)$, $\partial_{tt} E \in L^2(0, T; (H_0(\text{curl}; D))')$ and $E(0) = E_0$ in D, $\partial_t E(0) = E_1$ in D.

Remark 10.3 When considering as unknown the magnetic field H the formulation of the problem is

$$\begin{cases} \dfrac{\partial^2 H}{\partial t^2} + c^2 \text{curl}\,\text{curl}\, H = c^2 \text{curl}\, J_e & \text{in } D \times (0, T) \\ (\text{curl}\, H - J_e) \times n = 0 & \text{on } \partial D \times (0, T) \\ H_{|t=0} = H_0 & \text{in } D \\ \dfrac{\partial H}{\partial t}_{|t=0} = H_1 & \text{in } D \,, \end{cases} \tag{10.19}$$

with $H_1 = -\frac{1}{\mu} \text{curl}\, E_0$. The corresponding variational formulation is given by

$$\frac{d^2}{dt^2} \int_D H(t) \cdot w \, dx + a(H(t), w) = c^2 \int_D J_e(t) \cdot \text{curl}\, w \, dx \tag{10.20}$$

for all $w \in V$ and almost all $t \in [0, T]$, where $V = H(\text{curl}; D)$. For devising this formulation one has taken into account that the boundary integral

$$c^2 \int_{\partial D} [n \times (\text{curl}\, H - J_e)] \cdot w \, dS_x$$

vanishes due to the boundary condition in (10.19).

Note that, proceeding formally, the boundary condition $(\text{curl}\, H(t) - J_e(t)) \times n = 0$ says that $\frac{\partial E}{\partial t}(t) \times n = 0$ for each $t \in [0, T]$, therefore assuming $E_0 \times n = 0$ gives $E(t) \times n = 0$ for each $t \in [0, T]$.

Remark 10.4 Sufficient assumptions for applying both existence theorems are $E_0 \in H_0(\text{curl}; D)$, $H_0 \in H(\text{curl}; D)$, $J_e \in L^2(0, T; (L^2(D))^3)$ and $\partial_t J_e \in L^2(0, T; (L^2(D))^3)$.

Exercise 10.1 Prove that $H(\text{curl}; D)$ is a Hilbert space with respect to the scalar product (10.17).

10.2 Finite Propagation Speed

The hyperbolic equations have the property of *finite propagation speed*. This is a general property, but we will give a proof of it only for the wave equation, with velocity $c > 0$.

Consider a point (x_0, t_0), with $x_0 \in \mathbb{R}^n$ and $t_0 > 0$, and for $0 \le t < t_0$ define the sets

$$D_t = \{x \in \mathbb{R}^n \mid |x - x_0| < c(t_0 - t)\}$$
$$W = \{(x, t) \in \mathbb{R}^n \times [0, t_0) \mid x \in D_t\}\,.$$

Let us write for simplicity $u_t = \frac{\partial u}{\partial t}$ and $u_{tt} = \frac{\partial^2 u}{\partial t^2}$. The following result holds true:

Theorem 10.2 *Suppose that u is a (smooth enough) solution of $u_{tt} - c^2 \Delta u = 0$ and that $u = 0$, $u_t = 0$ on D_0. Then $u = 0$ in W.*

Proof Define

$$e(t) = \frac{1}{2} \int_{D_t} (u_t^2 + c^2 |\nabla u|^2) dx\,.$$

We want to compute $e'(t)$. We have, by the Reynolds transport theorem D.1,

$$e'(t) = \frac{1}{2} \int_{D_t} (u_t^2 + c^2 |\nabla u|^2)_t dx + \frac{1}{2} \int_{\partial D_t} (u_t^2 + c^2 |\nabla u|^2) V \cdot n dS_x\,,$$

where V is the velocity of ∂D_t and n is the external unit normal on ∂D_t. Since ∂D_t is the zero level-set of

$$Q(x, t) = |x - x_0| - c(t_0 - t)$$

and $D_t = \{x \in \mathbb{R}^n \mid Q(x, t) < 0\}$, we have

$$n = \frac{\nabla Q}{|\nabla Q|} = \frac{x - x_0}{|x - x_0|}\,.$$

For a particle $x = x(t)$ belonging to ∂D_t we have $|x(t) - x_0| = c(t_0 - t)$; thus differentiating with respect to t we have

$$-c = \frac{d}{dt}|x(t) - x_0| = \frac{x(t) - x_0}{|x(t) - x_0|} \cdot x'(t) = n \cdot V\,.$$

Summing up, using the Cauchy–Schwarz inequality and the fact that for any $a, b \in \mathbb{R}$ it holds $2ab \le a^2 + b^2$, we obtain

$$\begin{aligned}
e'(t) &= \frac{1}{2}\int_{D_t} (2u_t\, u_{tt} + 2c^2 \underbrace{\nabla u \cdot \nabla u_t}_{\text{integrate by parts}})dx + \frac{1}{2}\int_{\partial D_t} (u_t^2 + c^2|\nabla u|^2)(-c)dS_x \\
&= \int_{D_t} u_t\, u_{tt} dx - c^2 \int_{D_t} \Delta u\, u_t dx \\
&\quad + c^2 \int_{\partial D_t} \underbrace{\nabla u \cdot n\, u_t}_{\text{Cauchy–Schwarz}} dS_x - \frac{c}{2}\int_{\partial D_t} (u_t^2 + c^2|\nabla u|^2) dS_x \\
&\le \int_{D_t} u_t \underbrace{(u_{tt} - c^2\Delta u)}_{=0} dx + \frac{c}{2}\int_{\partial D_t} \underbrace{2c|\nabla u||u_t|}_{2ab\le a^2+b^2} dS_x \\
&\quad - \frac{c}{2}\int_{\partial D_t} (u_t^2 + c^2|\nabla u|^2) dS_x \\
&\le \frac{c}{2}\int_{\partial D_t} \left(c^2|\nabla u|^2 + u_t^2\right) dS_x - \frac{c}{2}\int_{\partial D_t} (u_t^2 + c^2|\nabla u|^2) dS_x = 0\,,
\end{aligned}$$

so that $e(t) \le e(0) = 0$ for each $t \in [0, t_0]$. Since $e(t) \ge 0$, it follows $e(t) = 0$ for each $t \in [0, t_0]$. In particular this gives $u_t = 0$ in W and, since $u = 0$ on the basis D_0, it follows $u = 0$ in W. □

Remark 10.5 The real-life interpretation of this result looks clear: if you throw a stone in a pond, the generated wave reaches the other side not immediately but after a little time. Do you see how mathematics is powerful?

10.3 Exercises

Exercise 10.1 Prove that $H(\text{curl}; D)$ is a Hilbert space with respect to the scalar product (10.17).

Solution Take a Cauchy sequence q_k in $H(\text{curl}; D)$: in particular q_k and $\text{curl}\, q_k$ are Cauchy sequences in $(L^2(D))^3$, thus we have that $q_k \to q$ and $\text{curl}\, q_k \to w$ in $(L^2(D))^3$. From the definition of weak curl (see Exercise 5.2) we know that $\text{curl}\, q_k$

satisfies

$$\int_D \operatorname{curl} q_k \cdot v dx = \int_D q_k \cdot \operatorname{curl} v dx \quad \forall\, v \in (C_0^\infty(D))^3 \,.$$

Passing to the limit we find

$$\int_D w \cdot v dx = \int_D q \cdot \operatorname{curl} v dx \quad \forall\, v \in (C_0^\infty(D))^3 \,,$$

which means that $w \in (L^2(D))^3$ is the weak curl of q. As a consequence we have proved that the sequence q_k converges to q in $H(\text{curl}; D)$.

Exercise 10.2 Suppose that u is a smooth solution in $D \times (0, T)$ of the homogeneous Dirichlet boundary value problem associated to the wave equation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0 \qquad \text{in } D \times (0, T) \,.$$

Show that $E(t) = \|u'(t)\|^2_{L^2(D)} + c^2 \|\nabla u(t)\|^2_{L^2(D)}$ is constant for each $t \in [0, T]$.

Solution Fix $t \in (0, T)$, and choose $v = u'(t)$ as test function in the weak formulation of the wave equation. We obtain

$$\langle u''(t), u'(t) \rangle + c^2 \int_D \nabla u(t) \cdot \nabla u'(t) dx = 0 \,.$$

This can be rewritten as

$$\frac{1}{2} \frac{d}{dt} \int_D u'(t)^2 dx + \frac{c^2}{2} \frac{d}{dt} \int_D |\nabla u(t)|^2 dx = 0 \,,$$

therefore $\int_D u'(t)^2 dx + c^2 \int_D |\nabla u(t)|^2 dx$ is constant for each $t \in [0, T]$.

[The physical meaning of this equality is that for an event steered by the wave equation the total energy (kinetic plus potential energy) is conserved.]

Exercise 10.3 Devise a variational formulation for the homogeneous Dirichlet boundary value problem associated to the *damped* wave equation

$$\frac{\partial^2 u}{\partial t^2} + \beta \frac{\partial u}{\partial t} - c^2 \Delta u = f \qquad \text{in } D \times (0, T) \,,$$

where $\beta > 0$ is a given parameter.

Solution The result is quite simple: by proceeding as for the wave equation, we look for $u \in L^2(0,T;H_0^1(D))$, with $u' \in L^2(0,T;L^2(D))$ and $u'' \in L^2(0,T;(H_0^1(D))')$, solution of

$$\begin{aligned}\langle u''(t), v\rangle + \beta(u'(t), v)_{L^2(D)} + c^2(\nabla u(t), \nabla v)_{L^2(D)} \\ = (f(t), v)_{L^2(D)} \quad \forall\, v \in H_0^1(D)\,.\end{aligned}$$

Exercise 10.4 Suppose that u is a smooth solution in $D \times (0,+\infty)$ of the homogeneous Dirichlet boundary value problem associated to the damped wave equation described in the previous exercise, with $f = 0$. Show that the total energy $E(t) = \|u'(t)\|^2_{L^2(D)} + c^2\|\nabla u(t)\|^2_{L^2(D)}$ is decreasing.

Solution First of all, let us note that by proceeding as in Theorem 10.1 one could prove the existence and uniqueness of a solution $u \in L^2(0,T;H_0^1(D))$ of the damped wave equation, with $u' \in L^2(0,T;L^2(D))$ and $u'' \in L^2(0,T;(H_0^1(D))')$. However, this would not permit us to use $u'(t)$ as a test function in the weak formulation, as it does not belong to $H_0^1(D)$ but only to $L^2(D)$. Thus let us proceed formally and assume that u is a smooth solution and set $u(0) = u_0$ and $u'(0) = u_1$. Fix $t \in (0,+\infty)$, and choose $v = u'(t)$ as test function in the weak formulation of the damped wave equation. We have

$$\langle u''(t), u'(t)\rangle + \beta \int_D u'(t)^2 dx + c^2 \int_D \nabla u(t) \cdot \nabla u'(t) dx = 0\,.$$

This can be rewritten as

$$\frac{1}{2}\frac{d}{dt}\int_D u'(t)^2 dx + \frac{c^2}{2}\frac{d}{dt}\int_D |\nabla u(t)|^2 dx + \beta \int_D u'(t)^2 dx = 0\,.$$

Therefore we have

$$E'(t) = -2\beta \int_D u'(t)^2 dx \le 0\,.$$

[The physical meaning of this equality is that for an event steered by the damped wave equation the total energy (kinetic plus potential energy) is dissipated as time increases.]

Exercise 10.5 Show that a suitable change of variable transforms the hyperbolic problem

$$\frac{\partial^2 u}{\partial t^2} + Lu = f \qquad \text{in } D \times (0,T)$$

associated to a weakly coercive bilinear form $B_L(\cdot, \cdot)$ into a damped hyperbolic problem

$$\frac{\partial^2 u}{\partial t^2} + \beta \frac{\partial u}{\partial t} + L_\sharp u = \hat{f} \qquad \text{in } D \times (0, T)$$

associated to a coercive bilinear form $B_{L_\sharp}(\cdot, \cdot)$.

Solution Set $w(t) = e^{-\eta t} u(t)$ where $\eta = \sqrt{\sigma} > 0$ and σ is the constant related to weak coerciveness. Then

$$\begin{aligned} w'(t) &= -\eta e^{-\eta t} u(t) + e^{-\eta t} u'(t) = -\eta w(t) + e^{-\eta t} u'(t) \\ w''(t) &= \eta^2 e^{-\eta t} u(t) - 2\eta e^{-\eta t} u'(t) + e^{-\eta t} u''(t) \\ &= \eta^2 w(t) - 2\eta (w'(t) + \eta w(t)) + e^{-\eta t} u''(t) \\ &= -\eta^2 w(t) - 2\eta w'(t) + e^{-\eta t} u''(t) \,. \end{aligned}$$

Thus from $u'' = f - Lu$ it follows

$$w''(t) + 2\sqrt{\sigma} w'(t) + Lw(t) + \sigma w(t) = e^{-\sqrt{\sigma} t} f(t) \,,$$

thus the desired result with $L_\sharp = L + \sigma I$, $\beta = 2\sqrt{\sigma}$ and $\hat{f}(t) = e^{-\sqrt{\sigma} t} f(t)$.

Exercise 10.6 Propose a numerical scheme for finding the approximate solution of a hyperbolic problem which is based on the Galerkin approximation and on a suitable finite difference scheme for discretizing $\frac{\partial^2 u}{\partial t^2}$.

Solution As in Exercise 9.6, let V_M be a finite dimensional subspace of V (not necessarily the space generated by the first M element of an orthonormal basis of V), whose basis is denoted by $\{\phi_1, \dots, \phi_M\}$. Choose a time-step $\tau = T/K > 0$, define $t_k = k\tau$, $k = 0, 1, \dots, K$, and consider the (second order) centered approximation of the second order derivative:

$$\frac{u^{k+1} - 2u^k + u^{k-1}}{\tau^2} \approx u''(t_k) \quad , \quad k = 1, \dots, K-1 \,.$$

Then the hyperbolic equation

$$\langle u''(t), v\rangle + a(t; u(t), v) = \langle F(t), v\rangle$$

can be approximated by means of the following numerical scheme: being given $u^0_M \in V_M$, a suitable approximation of the initial datum u_0, and $u^1_M \in V_M$, a suitable approximation of $u(t_1)$ constructed in terms of u^0_M and of an approximation $u_{1,M}$ of the initial datum u_1 (for instance, $u^1_M = u^0_M + \tau\, u_{1,M}$, or, better, a higher order approximation), for each $k = 1, \ldots, K - 1$ find $u_M^{k+1} \in V_M$, solution of the problem

$$\left(\frac{u_M^{k+1} - 2u_M^k + u_M^{k-1}}{\tau^2}, \phi_i \right)_H + a(t_k; u_M^k, \phi_i) = \langle F(t_k), \phi_i \rangle \ , \quad i = 1, \ldots, M \, .$$

In the literature, this is often called the (second order) "explicit" Newmark method (see, e.g., Raviart and Thomas [23, Sections 8.5 and 8.6]). Here the term "explicit" is used though at each time step t_{k+1}, $k = 1, \ldots, K - 1$, one has indeed to solve the discretized linear problem

$$\begin{aligned}(u_M^{k+1}, \phi_i)_H &= -\tau^2 a(t_k; u_M^k, \phi_i) + (2u_M^k - u_M^{k-1}, \phi_i)_H \\ &\quad + \tau^2 \langle F(t_k), \phi_i \rangle \ , \quad i = 1, \ldots, M \, ;\end{aligned}$$

this linear system is associated to the so-called mass matrix $M_{ij} = (\phi_j, \phi_i)_H$, where the contribution of the bilinear form $a(t; \cdot, \cdot)$ is not present, thus the operator L is not playing any role.

Appendix A
Partition of Unity

A technical result that have been used in the previous chapters is that of partition of unity. Let us explain which is its meaning.

Let K be a compact set in $\mathbb{R}^n$, covered by a finite union of open sets, $K \subset \bigcup_{i=1}^M V_i$. Define

$$V_{i,\varepsilon} = \{x \in V_i \,|\, \mathrm{dist}(x, \partial V_i) > \varepsilon\} \,.$$

The first result that we want to prove is the following one: we can find other open coverings $\bigcup_{i=1}^M V_{i,\varepsilon_0}$, $\bigcup_{i=1}^M V_{i,2\varepsilon_0}$, for a suitable ε_0. Let us prove this assertion.

Proposition A.1 *If a compact set $K \subset \mathbb{R}^n$ is covered by a finite union of open sets, $K \subset \bigcup_{i=1}^M V_i$, then there exists $\varepsilon_0 > 0$ such that $K \subset \bigcup_{i=1}^M V_{i,\varepsilon_0}$.*

Proof We proceed by contradiction, and suppose that the statement is not true. Then for each $\varepsilon > 0$ we can find $x_\varepsilon \in K$, $x_\varepsilon \notin \bigcup_{i=1}^M V_{i,\varepsilon}$. Since K is compact, we can select a subsequence $x_{\varepsilon_k} \to x_0 \in K$, with $\varepsilon_k \to 0$. Then there exists $i_0 \in \{1, \dots, M\}$ such that $x_0 \in V_{i_0}$. On the other hand, since $x_{\varepsilon_k} \notin \bigcup_{i=1}^M V_{i,\varepsilon_k}$, in particular $x_{\varepsilon_k} \notin V_{i_0,\varepsilon_k}$, and consequently we know that

$$\mathrm{dist}(x_{\varepsilon_k}, \partial V_{i_0}) \leq \varepsilon_k \longrightarrow 0 \,.$$

Thus $\mathrm{dist}(x_0, \partial V_{i_0}) = 0$, a contradiction as V_{i_0} is an open set. □

Now we can state the result concerning the partition of unity.

Proposition A.2 *Let K be a compact set in $\mathbb{R}^n$, covered by a finite union of open sets, $K \subset \bigcup_{i=1}^M V_i$. Then there exist functions $\omega_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \dots, M$, with the following properties:*

(i) $\omega_i \in C_0^\infty(V_i)$ *for each* $i = 1, \dots, M$;
(ii) $0 \leq \omega_i(x) \leq 1$ *for each* $i = 1, \dots, M$ *and for each* $x \in \mathbb{R}^n$;
(iii) $\sum_{i=1}^M \omega_i(x) = 1$ *for each* $x \in K$.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

Proof Take the characteristic function χ_i of $V_{i,2\varepsilon_0}$ and for some fixed $\varepsilon < \varepsilon_0$ consider its mollified version $\zeta_i = \eta_\varepsilon * \chi_i$ defined as

$$\zeta_i(x) = \int_{\mathbb{R}^n} \chi_i(y)\eta_\varepsilon(x-y)dy \quad , \quad x \in \mathbb{R}^n$$

(see Theorem 6.1). We know that $\zeta_i \in C^\infty(\mathbb{R}^n)$ and that $\zeta_i(x) \geq 0$ for all $x \in \mathbb{R}^n$, as both χ_i and η_ε are non-negative functions. Since the integral is indeed computed on $V_{i,2\varepsilon_0} \cap B(x,\varepsilon)$, where $B(x,\varepsilon) = \{y \in \mathbb{R}^n \,|\, |y-x| < \varepsilon\}$, we have $\zeta_i(x) = 0$ for $x \notin V_{i,\varepsilon_0}$, as in this case $V_{i,2\varepsilon_0} \cap B(x,\varepsilon) = \emptyset$; therefore $\zeta_i \in C_0^\infty(V_i)$. More precisely, we can see that $\zeta_i(x) > 0$ for $x \in V_{i,2\varepsilon_0-\varepsilon}$, $\zeta_i(x) = 0$ for $x \notin V_{i,2\varepsilon_0-\varepsilon}$, namely $\operatorname{supp} \zeta_i = \overline{V_{i,2\varepsilon_0-\varepsilon}}$. We now define

$$\omega_i(x) = \begin{cases} \frac{\zeta_i(x)}{\sum_{j=1}^M \zeta_j(x)} & \text{if } x \in V_{i,2\varepsilon_0-\varepsilon} \\ 0 & \text{if } x \in \mathbb{R}^n \setminus V_{i,2\varepsilon_0-\varepsilon} \,. \end{cases}$$

Therefore $\omega_i \in C_0^\infty(\mathbb{R}^n)$, $\operatorname{supp}\omega_i = \overline{V_{i,2\varepsilon_0-\varepsilon}} \subset V_{i,\varepsilon_0} \subset\subset V_i$, $\omega_i(x) \geq 0$ for all $x \in \mathbb{R}^n$ and $\omega_i(x) \leq 1$ for all $x \in \mathbb{R}^n$. Finally, for $x \in K \subset \bigcup_{i=1}^M V_{i,2\varepsilon_0} \subset \bigcup_{i=1}^M V_{i,2\varepsilon_0-\varepsilon}$ let us define

$$I_x = \{i = 1, \dots, M \,|\, x \in V_{i,2\varepsilon_0-\varepsilon}\}\,;$$

then we have

$$\sum_{i=1}^M \omega_i(x) = \sum_{s \in I_x} \omega_s(x) = \sum_{s \in I_x} \left(\frac{\zeta_s(x)}{\sum_{s \in I_x} \zeta_s(x)} \right) = 1\,,$$

and the proof is complete. □

An immediate consequence of this result is the construction of a cut-off function:

Corollary A.1 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set. Let Q be an open subset with $Q \subset\subset D$. Then there exists a cut-off function $\zeta \in C_0^\infty(D)$ satisfying $0 \leq \zeta(x) \leq 1$ for $x \in D$ and $\zeta(x) = 1$ for $x \in Q$.*

Proof It is enough to apply Proposition A.2 with $K = \overline{Q}$ and with a covering V_i satisfying $\cup_{i=1}^M V_i \subset\subset D$. For $x \in D$ the cut-off function is then given by $\zeta(x) = \sum_{i=1}^M \omega_i(x)$, and the property $\zeta(x) \leq 1$ in D follows by the definition of ω_i. □

Appendix B
Lipschitz Continuous Domains and Smooth Domains

In this appendix we clarify the meaning we give to the concept of "regularity" of the boundary of a domain.

First of all we have:

Definition B.1 Let $\mathcal{O} \subset \mathbb{R}^n$ be an open set. We say that a function $q : \overline{\mathcal{O}} \mapsto \mathbb{R}^n$ is a Lipschitz function in $\overline{\mathcal{O}}$, and we write $q \in \mathrm{Lip}(\overline{\mathcal{O}})$, if there exists a constant $L > 0$ such that

$$|q(x) - q(y)| \leq L|x - y|$$

for every $x, y \in \overline{\mathcal{O}}$.

To give an example, it is easily verified that, if $\mathcal{O}$ is a *bounded* open set, then a function $q \in C^1(\overline{\mathcal{O}})$ (namely, the restriction to $\overline{\mathcal{O}}$ of a $C^1(\mathbb{R}^n)$-function) is a Lipschitz function in $\overline{\mathcal{O}}$.

Consider now a bounded, connected, open set $D \subset \mathbb{R}^n$. Then the Lipschitz continuous regularity of its boundary ∂D is defined as follows:

Definition B.2 We say that D is a Lipschitz domain, or equivalently a domain with a Lipschitz continuous boundary, if for every point $p \in \partial D$ there exist an open ball B_p centered at p, an open ball $\widehat{B}_0$ centered at 0, a rigid body motion $R_p : B_p \mapsto \widehat{B}_0$ given by $R_p x = A_p x + b_p$, with $R_p p = 0$, A_p an orthogonal $n \times n$-matrix, $b_p \in \mathbb{R}^n$, and a map $\varphi : Q \mapsto \mathbb{R}$, where $Q = \{\xi \in \widehat{B}_0 \,|\, \xi_n = 0\}$, such that

1. $\varphi \in \mathrm{Lip}(\overline{Q})$ and $\varphi(0) = 0$
2. $R_p(B_p \cap \partial D) = \{(\xi', \xi_n) \in \widehat{B}_0 \,|\, \xi_n = \varphi(\xi'),\ \xi' \in Q\}$
3. $R_p(B_p \cap D) = \{(\xi', \xi_n) \in \widehat{B}_0 \,|\, \xi_n > \varphi(\xi'),\ \xi' \in Q\}$.

The meaning of the second condition is that ∂D coincides locally with the graph of a Lipschitz function; the third condition asserts that D is locally situated on one part of its boundary ∂D.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

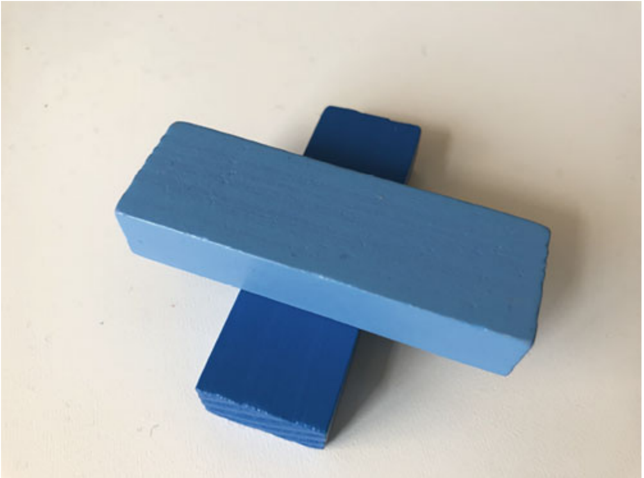


Fig. B.1 A (polyhedral) domain whose boundary is a Lipschitz manifold but it is not locally the graph of a Lipschitz function. The "bad" points are the four vertices of the square that is the interface between the two bricks (Courtesy of Jarno and Beatrice)

In particular, this definition says that a Lipschitz domain is a domain whose boundary is a manifold with a system of local charts that are invertible Lipschitz functions, namely, a Lipschitz manifold.

It can be interesting to note the opposite is not true: if you have a good geometrical intuition you can verify that the boundary of the two-brick set described in Fig. B.1 is an example of a surface that is not locally the graph of a Lipschitz function. On the other hand, it is a Lipschitz manifold (for complete description of this situation, see for instance a recent paper by Licht[1]).

For a Lipschitz domain at almost every point $x \in \partial D$ a tangent (hyper)plane is well defined, together with the unit outward normal vector n.

Definition B.3 We say that D is a domain of class C^k, or equivalently a domain with a C^k-boundary, $k \geq 1$, and we write $\partial D \in C^k$, if the function φ in Definition B.2 belongs to C^k.

[1] Licht [19].

Appendix C
Integration by Parts for Smooth Functions and Vector Fields

This appendix is devoted to various "integration by parts" formulas that have been used several times in the previous chapters.

Let us start from the "fundamental theorem of calculus" (whose proof can be found in any Calculus textbook): the integral of a derivative of a function f can be explicitly expressed by an integral of f over a lower dimensional set.

Theorem C.1 (Fundamental Theorem of Calculus) *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $f : \overline{D} \mapsto \mathbb{R}$ be a function of class $C^1(\overline{D})$. Then*

$$\int_D \mathcal{D}_i f dx = \int_{\partial D} f n_i dS_x \,, \tag{C.1}$$

where n is the unit outward normal vector, defined on ∂D for almost every $x \in \partial D$.

From this theorem we easily obtain many well-known results:

Theorem C.2 (Integration by Parts) *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $f, g : \overline{D} \mapsto \mathbb{R}$ be two functions of class $C^1(\overline{D})$. Then*

$$\int_D \mathcal{D}_i f \, g dx = -\int_D f \, \mathcal{D}_i g dx + \int_{\partial D} f g n_i dS_x \,. \tag{C.2}$$

Proof It is enough to remember that $\mathcal{D}_i(fg) = \mathcal{D}_i f g + f \mathcal{D}_i g$ and to apply Theorem C.1. □

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

Theorem C.3 (Divergence or Gauss Theorem) *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $F : \overline{D} \mapsto \mathbb{R}^n$ be a vector field of class $C^1(\overline{D})$. Then*

$$\int_D divF dx = \int_{\partial D} F \cdot n \, dS_x \, . \tag{C.3}$$

Proof Since div $F = \sum_{i=1}^n \mathcal{D}_i F_i$, one has only to apply Theorem C.2 for $f = F_i$, $g = 1$ and to add over $i = 1, \ldots, n$. □

Theorem C.4 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $F : \overline{D} \mapsto \mathbb{R}^n$ be a vector field of class $C^1(\overline{D})$, $g : \overline{D} \mapsto \mathbb{R}$ be a function of class $C^1(\overline{D})$. Then*

$$\int_D divF \, g dx = -\int_D F \cdot \nabla g dx + \int_{\partial D} F \cdot n \, g \, dS_x \, . \tag{C.4}$$

In particular, taking $F \in C_0^\infty(D)$ and $g \in C_0^\infty(D)$ one verifies that $-\nabla$ is the (formal) transpose operator of div.

Proof It is enough to apply Theorem C.2 to $f = F_i$ and to add over $i = 1, \ldots, n$. □

Theorem C.5 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $f : \overline{D} \mapsto \mathbb{R}$ be a function of class $C^2(\overline{D})$, $g : \overline{D} \mapsto \mathbb{R}$ be a function of class $C^1(\overline{D})$. Then*

$$\int_D (-\Delta f) \, g dx = \int_D \nabla f \cdot \nabla g dx - \int_{\partial D} \nabla f \cdot n \, g \, dS_x \, . \tag{C.5}$$

In particular, taking $g = 1$ it follows

$$\int_D \Delta f dx = \int_{\partial D} \nabla f \cdot n \, dS_x \, . \tag{C.6}$$

Proof Recalling that $-\Delta f = -\text{div}\nabla f$, it is enough to apply Theorem C.4 to $F = -\nabla f$. □

Theorem C.6 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $F : \overline{D} \mapsto \mathbb{R}^n$ be a vector field of class $C^1(\overline{D})$, $G : \overline{D} \mapsto \mathbb{R}^n$ be a vector field of class $C^2(\overline{D})$. Then*

$$\int_D (-\nabla divG) \cdot F dx = \int_D divG \, divF dx - \int_{\partial D} divG \, F \cdot n \, dS_x \, . \tag{C.7}$$

Proof It is enough to apply Theorem C.4 to $g = \text{div}G$. □

Theorem C.7 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $F, G : \overline{D} \mapsto \mathbb{R}^n$ be two vector fields of class $C^1(\overline{D})$. Then*

$$\int_D curl F \cdot G dx = \int_D F \cdot curl G dx + \int_{\partial D} n \times F \cdot G \, dS_x \,. \tag{C.8}$$

In particular, taking $F \in C_0^\infty(D)$ and $G \in C_0^\infty(D)$ one verifies that curl is (formally) equal to its transpose operator.

Proof Recalling that $\text{curl} F$ can be formally computed as the vector product $\nabla \times F$, one has only to apply Theorem C.2 to all the terms of the scalar product $\text{curl} F \cdot G$ and to check that the result follows. □

Theorem C.8 *Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary, and let $M : \overline{D} \mapsto \mathbb{R}^n$ be a vector field of class $C^2(\overline{D})$, $G : \overline{D} \mapsto \mathbb{R}^n$ be a vector field of class $C^1(\overline{D})$. Then*

$$\int_D curl\, curl M \cdot G dx = \int_D curl\, M \cdot curl G dx + \int_{\partial D} n \times curl M \cdot G \, dS_x \,. \tag{C.9}$$

Proof Just take $F = \text{curl} M$ in Theorem C.7. □

Appendix D
Reynolds Transport Theorem

In this appendix we are concerned with a well-known result of differential calculus, which is often useful in continuum mechanics. In the literature we are not aware of a reference presenting its proof in a detailed way (but surely it exists!). Anyway, for the ease of the reader we decided to present the proof here.

We need a preliminary result. Let us denote by $\mathrm{Lip}(\mathbb{R}^n)$ the space of Lipschitz functions on $\mathbb{R}^n$.

Lemma D.1 *Consider* $v = v(t, X) \in L^1(0, +\infty; Lip(\mathbb{R}^n))$*,* $x \in \mathbb{R}^n$*, and let* $\Phi = \Phi(t, x)$ *be the solution of the Cauchy problem*

$$\begin{cases} \dfrac{d}{dt}\Phi(t,x) = v(t, \Phi(t,x)) \ , \quad t > 0 \\ \Phi(0,x) = x \ . \end{cases} \tag{D.1}$$

Defining $j(t,x) = \det \mathrm{Jac}_x \Phi(t,x)$*, it holds*

$$\frac{dj}{dt}(t,x) = [(\mathrm{div}_X v) \circ \Phi](t,x)\, j(t,x) \, . \tag{D.2}$$

Remark D.1 In fluid dynamics one says that v is the velocity of the flow Φ: in other words, the position $\Phi(t,x)$ is determined by integrating the velocity v along the trajectories of the fluid particles. This means that $\Phi(t,x)$ is the position at time t of a particle that at time 0 was at x: then $X = \Phi(t,x)$ is the Lagrangian coordinate, whereas x is the Eulerian coordinate.

Proof Being j a determinant, its derivative is given by

$$\frac{dj}{dt} = \sum_{k=1}^{n} \det M_k \, ,$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

where, for $k = 2, \dots, n-1$, the matrix M_k is given by

$$M_k = \begin{pmatrix} \mathcal{D}_{x_1}\Phi_1 & \dots & \mathcal{D}_{x_n}\Phi_1 \\ \vdots & \vdots & \vdots \\ \mathcal{D}_t\mathcal{D}_{x_1}\Phi_k & \dots & \mathcal{D}_t\mathcal{D}_{x_n}\Phi_k \\ \vdots & \vdots & \vdots \\ \mathcal{D}_{x_1}\Phi_n & \dots & \mathcal{D}_{x_n}\Phi_n \end{pmatrix},$$

with obvious modification for the cases $k = 1$ and $k = n$. For $k, j = 1, \dots, n$ from (D.1) we have

$$\mathcal{D}_t\mathcal{D}_{x_j}\Phi_k = \mathcal{D}_{x_j}\mathcal{D}_t\Phi_k = \mathcal{D}_{x_j}(v_k \circ \Phi),$$

where we have denoted by $g \circ \Phi$ the function $(t, x) \mapsto g(t, \Phi(t, x))$. Moreover, by means of the chain rule we also find, for $k, j = 1, \dots, n$,

$$\mathcal{D}_{x_j}(v_k \circ \Phi) = \sum_{s=1}^{n} \left(\frac{\partial v_k}{\partial X_s} \circ \Phi \right) \mathcal{D}_{x_j}\Phi_s .$$

Take for a while $k = 2, \dots, n-1$. Using the two last results we obtain

$$M_k = \begin{pmatrix} \mathcal{D}_{x_1}\Phi_1 & \dots & \mathcal{D}_{x_n}\Phi_1 \\ \vdots & \vdots & \vdots \\ \displaystyle\sum_{s=1}^{n} \left(\frac{\partial v_k}{\partial X_s} \circ \Phi \right) \mathcal{D}_{x_1}\Phi_s & \dots & \displaystyle\sum_{s=1}^{n} \left(\frac{\partial v_k}{\partial X_s} \circ \Phi \right) \mathcal{D}_{x_n}\Phi_s \\ \vdots & \vdots & \vdots \\ \mathcal{D}_{x_1}\Phi_n & \dots & \mathcal{D}_{x_n}\Phi_n \end{pmatrix} \leftarrow k\text{-th row} .$$

Since the determinant is linear with respect to the rows we find

$$\det M_k = \sum_{s=1}^{n} \left(\frac{\partial v_k}{\partial X_s} \circ \Phi \right) \det \begin{pmatrix} \mathcal{D}_{x_1}\Phi_1 & \dots & \mathcal{D}_{x_n}\Phi_1 \\ \vdots & \vdots & \vdots \\ \mathcal{D}_{x_1}\Phi_s & \dots & \mathcal{D}_{x_n}\Phi_s \\ \vdots & \vdots & \vdots \\ \mathcal{D}_{x_1}\Phi_n & \dots & \mathcal{D}_{x_n}\Phi_n \end{pmatrix} \leftarrow k\text{-th row} .$$

When $s \neq k$ the matrix has two rows that are equal, thus its determinant vanishes; therefore

$$\det M_k = \left(\frac{\partial v_k}{\partial X_k} \circ \Phi\right) \det \begin{pmatrix} \mathcal{D}_{x_1}\Phi_1 & \dots & \mathcal{D}_{x_n}\Phi_1 \\ \vdots & \vdots & \vdots \\ \mathcal{D}_{x_1}\Phi_k & \dots & \mathcal{D}_{x_n}\Phi_k \\ \vdots & \vdots & \vdots \\ \mathcal{D}_{x_1}\Phi_n & \dots & \mathcal{D}_{x_n}\Phi_n \end{pmatrix} = \left(\frac{\partial v_k}{\partial X_k} \circ \Phi\right) j \,.$$

For $k = 1$ and $k = n$ we have the same result, with straightforward modification. Adding over k form 1 to n we find (D.2). □

We are now ready for the main result. Let $D_0 \subset \mathbb{R}^n$ be a bounded, connected, open set with a Lipschitz continuous boundary. For $t > 0$ define

$$D_t = \{X \in \mathbb{R}^n \mid X = \Phi(t, x) \text{ for some } x \in D_0\}$$

and

$$W = \{(t, X) \in (0, +\infty) \times \mathbb{R}^n \mid X \in D_t\} \,.$$

Theorem D.1 (Reynolds Transport Theorem) *Let $f : W \mapsto \mathbb{R}$ be a (smooth enough) scalar function. Then*

$$\frac{d}{dt}\left(\int_{D_t} f \, dX\right) = \int_{D_t} \frac{\partial f}{\partial t} \, dX + \int_{\partial D_t} v \cdot n \, f \, dS_X \,,$$

where v is the velocity of the boundary ∂D_t and n is the unit outward normal vector on ∂D_t.

Proof For any fixed t consider the change of variables $X = \Phi(t, x)$, which yields

$$\int_{D_t} f(t, X) \, dX = \int_{D_0} f(t, \Phi(t, x)) |\det \mathrm{Jac}_x \Phi(t, x)| \, dx \,. \tag{D.3}$$

Since $j(0, x) = \det \mathrm{Jac}_x \Phi(0, x) = \det \mathrm{Jac} \mathrm{Id} = 1$, from (D.2) we find

$$j(t, x) = \exp\left(\int_0^t (\mathrm{div}_X v)(s, \Phi(s, x)) \, ds\right) > 0 \,.$$

Thus in (D.3) we can drop the absolute value of the determinant. Let us now differentiate with respect to t. Since the integral in D_0 is on a fixed set, we can differentiate inside the integral and we find

$$\begin{aligned}\frac{d}{dt}\int_{D_t} f\,dX &= \frac{d}{dt}\int_{D_0} f(t,\Phi(t,x))\ \det\mathrm{Jac}_x\Phi(t,x)\,dx \\ &= \int_{D_0}\frac{d}{dt}[f(t,\Phi(t,x))]\ \det\mathrm{Jac}_x\Phi(t,x)\,dx \\ &\quad + \int_{D_0} f(t,\Phi(t,x))\frac{d}{dt}[\det\mathrm{Jac}_x\Phi(t,x)]\,dx\,.\end{aligned} \tag{D.4}$$

By the chain rule, and taking (D.1) into account, the first factor in the first term of (D.4) can be rewritten as

$$\begin{aligned}\frac{d}{dt}[f(t,\Phi(t,x))] &= \frac{\partial f}{\partial t}(t,\Phi(t,x)) + \sum_{i=1}^{n}\frac{\partial f}{\partial X_i}(t,\Phi(t,x))\frac{d\Phi_i}{dt}(t,x) \\ &= \frac{\partial f}{\partial t}(t,\Phi(t,x)) + \sum_{i=1}^{n}\frac{\partial f}{\partial X_i}(t,\Phi(t,x))v_i(t,\Phi(t,x)) \\ &= \left(\frac{\partial f}{\partial t} + v\cdot\nabla_X f\right)(t,\Phi(t,x))\,.\end{aligned}$$

Using (D.2) in the second term of (D.4) we obtain

$$f(t,\Phi(t,x))\frac{d}{dt}[\det\mathrm{Jac}_x\Phi(t,x)] = (f\,\mathrm{div}_X v)(t,\Phi(t,x))\ \det\mathrm{Jac}_x\Phi(t,x)\,.$$

In conclusion, we have seen that

$$\begin{aligned}&\frac{d}{dt}\int_{D_t} f(t,X)\,dX \\ &\quad= \int_{D_0}\left(\frac{\partial f}{\partial t} + v\cdot\nabla_X f + f\,\mathrm{div}_X\,v\right)(t,\Phi(t,x))\ \det\mathrm{Jac}_x\Phi(t,x)\,dx \\ &\quad= \int_{D_0}\left(\frac{\partial f}{\partial t} + \mathrm{div}_X(fv)\right)(t,\Phi(t,x))\ \det\mathrm{Jac}_x\Phi(t,x)\,dx\,.\end{aligned}$$

Rewriting the integral at the right hand side by means of the change of variable $X = \Phi(t,x)$ we have

$$\frac{d}{dt}\int_{D_t} f(t,X)\,dX = \int_{D_t}\left(\frac{\partial f}{\partial t} + \mathrm{div}_X(fv)\right)(t,X)\,dX\,,$$

hence the thesis by using the divergence theorem C.3. □

Appendix E
Gronwall Lemma

The Gronwall lemma is an useful tool in the analysis of evolution equations. Its statement is the following.

Lemma E.1 (Gronwall Lemma) *Let $f \in L^1(0,T)$ be a non-negative function, g and φ be continuous functions in $[0,T]$. If φ satisfies*

$$\varphi(t) \leq g(t) + \int_0^t f(\tau)\varphi(\tau)d\tau \qquad \forall\, t \in [0,T]\,,$$

then

$$\varphi(t) \leq g(t) + \int_0^t f(s)g(s)\exp\left(\int_s^t f(\tau)d\tau\right)ds \qquad \forall\, t \in [0,T]\,. \tag{E.1}$$

The proof of this lemma will be given below. For the moment let us show some consequences of it.

Corollary E.1 *If g is a non-decreasing function, then*

$$\varphi(t) \leq g(t)\exp\left(\int_0^t f(\tau)d\tau\right) \qquad \forall\, t \in [0,T]\,.$$

Proof If g is non-decreasing, we have $g(s) \leq g(t)$ for $0 \leq s \leq t$, thus from Eq. (E.1)

$$\varphi(t) \leq g(t)\left[1 + \int_0^t f(s)\exp\left(\int_s^t f(\tau)d\tau\right)ds\right].$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

Since

$$\frac{d}{ds}\left(\exp\left(\int_s^t f(\tau)d\tau\right)\right) = -\exp\left(\int_s^t f(\tau)d\tau\right) f(s)\,,$$

we have that

$$\int_0^t f(s)\exp\left(\int_s^t f(\tau)d\tau\right)ds = -\int_0^t \frac{d}{ds}\left(\exp\left(\int_s^t f(\tau)d\tau\right)\right)ds$$
$$= -\left(1-\exp\left(\int_0^t f(\tau)d\tau\right)\right)\,,$$

hence the result. □

Corollary E.2 *If* $g(t) = k_1$ *and* $f(t) = k_2$, *then*

$$\varphi(t) \le k_1 e^{k_2 t} \qquad \forall\, t \in [0, T]\,.$$

Proof Just apply Corollary E.1. □

Proof (of Lemma E.1) For $s \in [0, T]$ set $R(s) = \int_0^s f(\tau)\varphi(\tau)d\tau$. The assumption yields

$$R'(s) = f(s)\varphi(s) \le f(s)[g(s) + R(s)]\,.$$

Then

$$\frac{d}{ds}\left[R(s)\exp\left(-\int_0^s f(\tau)d\tau\right)\right]$$
$$= R'(s)\exp\left(-\int_0^s f(\tau)d\tau\right) - R(s)f(s)\exp\left(-\int_0^s f(\tau)d\tau\right)$$
$$= [R'(s) - R(s)f(s)]\exp\left(-\int_0^s f(\tau)d\tau\right)$$
$$\le f(s)g(s)\exp\left(-\int_0^s f(\tau)d\tau\right)\,.$$

Integrating over $[0, t]$, we find, as $R(0) = 0$,

$$R(t)\exp\left(-\int_0^t f(\tau)d\tau\right) \le \int_0^t f(s)g(s)\exp\left(-\int_0^s f(\tau)d\tau\right)ds\,,$$

thus

$$R(t) \le \int_0^t f(s)g(s)\exp\left(\int_s^t f(\tau)d\tau\right)ds\,,$$

which gives the stated result as a consequence of the assumption $\varphi(t) \le g(t)+R(t)$.

□

Appendix F
Necessary and Sufficient Conditions for the Well-Posedness of the Variational Problem

We present here the well-posedness result for a general variational problem of the form

$$\text{find } u \in V \ : \ B(u, v) = F(v) \quad \forall\, v \in V\,, \tag{F.1}$$

where V is a Hilbert space, $B(\cdot, \cdot) : V \times V \mapsto \mathbb{R}$ is a bounded bilinear form and $F(\cdot) : V \mapsto \mathbb{R}$ is a bounded linear functional.

Theorem F.1 *Problem* (F.1) *is well-posed (namely, it has one and only one solution u for each bounded and linear functional F, and the solution map $F \mapsto u$ is bounded) if and only if the following conditions are satisfied:*

(i) there exists $\alpha > 0$: $\displaystyle\inf_{w\in V, w\neq 0} \sup_{v\in V, v\neq 0} \frac{B(w, v)}{\|w\|_V \|v\|_V} \geq \alpha$

(ii) if $B(w, v) = 0$ *for all* $w \in V$ *then* $v = 0$.

Proof We introduce the linear and bounded functionals $Q : V \mapsto V'$ and $Q^T : V \mapsto V'$ defined as

$$\langle Qw, v\rangle = B(w, v) \quad \forall\, v \in V \ \ , \ \ \langle Q^T v, w\rangle = B(w, v) \quad \forall\, w \in V\,.$$

The well-posedness statement is thus reformulated as: Q is an isomorphism from V onto V'.

($\Rightarrow$) Suppose that Q is an isomorphism from V onto V'. Then $N(Q) = \{0\}$ and $R(Q) = V'$, thus in particular $R(Q)$ is closed. From the closed range theorem (see Yosida [30, Theorem 1, p. 205]) $R(Q) = N(Q^T)_\sharp$, thus $N(Q^T)_\sharp = V'$ and $N(Q^T) = \{0\}$. This means that $Q^T v = 0$ implies $v = 0$, namely, that $\langle Q^T v, w\rangle = B(w, v) = 0$ for each $w \in V$ implies $v = 0$. This is condition (ii). Moreover, since Q is an isomorphism from V onto V', its inverse is bounded, namely, there exists $\alpha > 0$ such that $\|Qw\|_{V'} \geq \alpha\|w\|_V$ for each $w \in V$. This

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

means

$$\sup_{v \in V, v \neq 0} \frac{\langle Qw, v \rangle}{\|v\|_V} = \sup_{v \in V, v \neq 0} \frac{B(w, v)}{\|v\|_V} \geq \alpha \|w\|_V \quad \forall\, w \in V\,,$$

thus condition (i).

(⇐) Let us assume now that (i) and (ii) are satisfied. We can follow the lines of the proof of the Lax–Milgram theorem 2.1. From condition (i) it follows

$$\|Qw\|_{V'} = \sup_{v \in V, v \neq 0} \frac{\langle Qw, v \rangle}{\|v\|_V} \geq \alpha \|w\|_V\,; \tag{F.2}$$

as a consequence we derive that Q is one-to-one, as from $Qw = 0$ it follows at once $w = 0$, and that Q^{-1} is bounded (at the moment, from $R(Q)$ to V). Moreover, we can also prove that $R(Q)$ is closed. In fact, consider a sequence $Qv_k \in R(Q)$ such that $Qv_k \to \omega \in V'$. In particular, Qv_k is a Cauchy sequence in V', and from (F.2) we have that v_k is a Cauchy sequence in V. Therefore we find $v_0 \in V$ such that $v_k \to v_0$ in V, thus $Qv_k \to Qv_0$ in V', which gives $Qv_0 = \omega$.

Since $R(Q)$ is closed, the closed range theorem gives that $R(Q) = N(Q^T)_\sharp$ (see Yosida [30, Theorem 1, p. 205]). Thus for proving that $R(Q) = V'$ it is enough to show that $N(Q^T) = \{0\}$, namely, that $\langle Q^T v, w \rangle = 0$ for all $w \in V$ implies $v = 0$: since $\langle Q^T v, w \rangle = B(w, v)$, this is exactly condition (ii). □

Remark F.1 It is straightforward to verify that the coerciveness of $B(\cdot, \cdot)$ implies both (i) and (ii).

References

1. Adams, R.A.: Sobolev Spaces. Academic, New York (1975)
2. Banach, S.: Théorie des Opérations Linéaires. Instytut Matematyczny Polskiej Akademii Nauk, Warszawa (1932)
3. Bogovskii, M.E.: Solution of the first boundary value problem for an equation of continuity of an incompressible medium [Russian]. Dokl. Akad. Nauk SSSR **248**(5), 1037–1040 (1979) [English translation: Soviet Math. Dokl. **20**, 1094–1098 (1979)]
4. Brezis, H.: Analyse Fonctionnelle. Masson, Paris (1983)
5. Dautray, R., Lions, J.-L.: Mathematical Analysis and Numerical Methods for Science and Technology, vol. 2. Springer, Berlin (1988)
6. Dautray, R., Lions, J.-L.: Mathematical Analysis and Numerical Methods for Science and Technology, vol. 5. Springer, Berlin (1992)
7. Dautray, R., Lions, J.-L.: Mathematical Analysis and Numerical Methods for Science and Technology, vol. 6. Springer, Berlin (1993)
8. Evans, L.C.: Partial Differential Equations. American Mathematical Society, Providence (1998)
9. Friedrichs, K.O.: The identity of weak and strong extensions of differential operators. Trans. Am. Math. Soc. **55**(1), 132–151 (1944)
10. Gazzola, F., Grunau, H.-C., Sweers, G.: Polyharmonic Boundary Value Problems. Springer, Berlin (2010)
11. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Springer, Berlin (1977)
12. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier–Stokes Equations. Springer, Berlin (1986)
13. Hadamard, J.: La Théorie des Équations aux Dérivées Partielles. Éditions Scientifiques, Pekin (1964)
14. Kellogg, O.D.: Foundations of Potential Theory. Ungar, New York (1929)
15. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. SIAM, Philadelphia (2000)
16. Ladyženskaya, O.A.: On integral estimates, convergence, approximate methods, and solution in functionals for elliptic operators [Russian]. Vestnik Leningrad. Univ. **13**(7), 60–69 (1958)
17. Ladyženskaja, O.A., Solonnikov, V.A., Ural'ceva, N.N.: Linear and Quasilinear Equations of Parabolic Type. American Mathematical Society, Providence (1968)
18. Leray, J.: Sur le mouvement d'un liquide visqueux emplissant l'espace. Acta Math. **63**, 193–248 (1934)

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7

19. Licht, M.W.: Smoothed projections over weakly Lipschitz domains. Math. Comput. **88**(315), 179–210 (2019)
20. Maz'ya, V., Shaposhnikova, T.: Jacques Hadamard, a Universal Mathematician. American Mathematical Society, Providence (1999)
21. McLean, W.: Strongly Elliptic Systems and Boundary Integral Equations. Cambridge University Press, Cambridge (2000)
22. Monk, P.: Finite Element Methods for Maxwell's Equations. Oxford University Press, Oxford (2003)
23. Raviart, P.A., Thomas, J.M.: Introduction à l'Analyse Numérique des Équations aux Dérivées Partielles. Masson, Paris (1983)
24. Salsa, S.: Partial Differential Equations in Action, 2nd edn. Springer, Milan (2015)
25. Sobolev, S.L.: Méthode nouvelle à résoudre le problème de Cauchy pour les équations linéaires hyperboliques normales. Mat. Sbornik **1**(43), 39–72 (1936)
26. Sobolev, S.L.: On a boundary value problem for polyharmonic equations. Mat. Sbornik **2**(44), 465–499 (1937) (Russian). [English translation: Amer. Math. Soc. Transl., (2) 33 (1963), 1–40.]
27. Sobolev, Sobolev, S.L.: Sur un théorème d'analyse fonctionnelle. Mat. Sbornik **4**(46), 471–497 (1938) (Russian). [English translation: Amer. Math. Soc. Transl., (2) 34 (1963), 39–68.]
28. Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton University Press, Princeton (1970)
29. Wloka, J.: Partial Differential Equations. Cambridge University Press, Cambridge (1987)
30. Yosida, K.: Functional Analysis, 4th edn. Springer, Berlin (1974)
31. Ziemer, W.P.: Weakly Differentiable Functions. Springer, New York (1989)

Index

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

A. Valli, *A Compact Course on Linear PDEs*, La Matematica per il 3+2 154,
https://doi.org/10.1007/978-3-031-35976-7