**S**pringer Series in
**C**omputational
**M**athematics

**34**

Andrea Toselli
Olof Widlund

# Domain Decomposition Methods – Algorithms and Theory

With 33 Figures

Springer

Andrea Toselli

ETH Zentrum
Seminar for Applied Mathematics
Rämistrasse 101
8092 Zürich
Switzerland
email: toselli@sam.math.ethz.ch

Olof B. Widlund

Courant Institute of Mathematical Sciences
251 Mercer Street
New York, NY 10012
USA
e-mail: widlund@cims.nyu.edu

Library of Congress Control Number: 2004113304

Mathematics Subject Classification (2000): 65F10 65N22 65N30 65N55

ISSN 0179-3632
ISBN 3-540-20696-5 Springer Berlin Heidelberg New York



Cover design: *design&production*, Heidelberg
Typeset by the authors using a Springer LaTeX macro package
Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Printed on acid-free paper 46/3142YL-5 4 3 2 1 0

# Preface

The purpose of this text is to offer a comprehensive and self-contained presentation of some of the most successful and popular domain decomposition methods for partial differential equations. Strong emphasis is put on both algorithmic and mathematical aspects. In addition, we have wished to present a number of methods that have not been treated previously in other monographs and surveys. We believe that this monograph will offer something new and that it will complement those of Smith, Bjørstad, and Gropp [424] and Quarteroni and Valli [392]. Our monograph is also more extensive and broader than the surveys given in Chan and Mathew [132], Farhat and Roux [201], Le Tallec [308], the habilitation thesis by Wohlmuth [469], and the well-known SIAM Review articles by Xu [472] and Xu and Zou [476].

Domain decomposition generally refers to the splitting of a partial differential equation, or an approximation thereof, into coupled problems on smaller subdomains forming a partition of the original domain. This decomposition may enter at the continuous level, where different physical models may be used in different regions, or at the discretization level, where it may be convenient to employ different approximation methods in different regions, or in the solution of the algebraic systems arising from the approximation of the partial differential equation. These three aspects are very often interconnected in practice.

This monograph is entirely devoted to the third aspect of domain decomposition. In practical applications, finite element or other discretizations reduces the problem to the solution of an often huge algebraic system of equations. Direct factorization of such systems might then not be a viable option and the use of basic iterative methods, such as the conjugate gradient algorithm, can result in very slow convergence. The basic idea of domain decomposition is that instead of solving one huge problem on a domain, it may be convenient (or necessary) to solve many smaller problems on single subdomains a certain number of times. Much of the work in domain decomposition relates to the selection of subproblems that ensure that the rate of convergence of the

new iterative method is fast. In other words, domain decomposition methods provide preconditioners that can be accelerated by Krylov space methods.

The development of the field, and the increased interest in domain decomposition methods, is closely related to the growth of high speed computing. We note that in the June 2004 edition of the "Top 500" list, there are no fewer than 242 computer systems sustaining at least 1.0 Teraflop/sec. Scientific computing is therefore changing very fast and many scientists are now developing codes for parallel and distributed systems.

The development of numerical methods for large algebraic systems is central in the development of efficient codes for computational fluid dynamics, elasticity, and other core problems of continuum mechanics. Many other tasks in such codes parallelize relatively easily. The importance of the algebraic system solvers is therefore increasing with the appearance of new computing systems, with a substantial number of fast processors, each with relatively large memory. In addition, robust algebraic solvers for many practical problems and discretizations cannot be constructed by simple algebraic techniques, such as approximate inverses or incomplete factorizations, but the partial differential equation and the discretization must be taken into account. A very desirable feature of domain decomposition algorithms is that they respect the memory hierarchy of modern parallel and distributed computing systems, which is essential for approaching peak floating point performance. The development of improved methods, together with more powerful computer systems, is making it possible to carry out simulations in three dimensions, with quite high resolution, relatively easily. This work is now supported by high quality software systems, such as Argonne's PETSc library, which facilitates code development as well as the access to a variety of parallel and distributed computer systems. In chapters 6 and 9, we will describe numerical experiments with codes developed using this library.

A powerful approach to the analysis and development of domain decomposition is to view the procedure in terms of subspaces, of the original solution space, and with suitable solvers on these subspaces. Typically these subspaces are related to the geometrical objects of the subdomain partition (subdomains, subdomain boundaries, interfaces between subdomains, and vertices, edges, and faces of these interfaces). The abstract Schwarz theory, presented in Chap. 2, relies on these ideas and the convergence of the resulting iterative method is related to the stability of the decomposition into subspaces, certain stability properties of the local solvers, and a measure of the 'orthogonality' of these subspaces. The strong connection between stable decompositions of discrete functions in terms of Sobolev norms and the performance of the corresponding domain decomposition algorithm is not a mere way of giving an elegant mathematical description of a method that already works well in practice, but it is often the way in which new powerful algorithms are actually developed, especially for less standard discretizations such as edge elements for electromagnetic problems.

The book is addressed to mathematicians, computer scientists, and in general to people who are involved in the numerical approximation of partial differential equations, and who want to learn the basic ideas of domain decomposition methods both from a mathematical and an algorithmic point of view. The mathematical tools needed for the type of analysis essentially consist in some basic results on Sobolev spaces, which are reviewed in appendix A. The analysis also employs discrete Sobolev-type inequalities valid for finite element functions and polynomials. These tools are developed in Chap. 4. A basic knowledge of finite element theory and iterative methods is also required and two additional appendices summarize the results that are needed for the full understanding of the analysis of the algorithms developed in the main part of this monograph.

The literature of the field is now quite extensive and it has developed rapidly over the past twenty years. We have been forced to make some important omissions. The most important one is that we do not consider multilevel or multigrid methods, even though many of these algorithms can also be viewed, and then analyzed, using similar techniques as domain decomposition methods; the decomposition into subspaces is now related to a hierarchy of finite element meshes. The inclusion of these methods would have required a large effort and many pages and is likely to have duplicated efforts by real specialists in that field; the authors fully realizes the importance of these algorithms, which provide efficient and robust algorithms for many very large problems.

Other omissions have also been necessary:

- As already mentioned, we only consider domain decomposition as a way of building iterative methods for the solution of algebraic systems of equations.
- While we describe a number of algorithms in such a way as to simplify their implementation, we do not discuss other practical aspects of the development of codes for parallel and distributed computer systems.
- We only consider linear elliptic scalar and vector problems in full detail. Indeed, the methods presented in this monograph can be applied to the solution of linear systems arising from implicit time step discretizations of time-dependent problems or arising from Newton-type iterations for non-linear problems.
- Our presentation and analysis is mainly confined to low-order finite element ($h$ version) and spectral element (a particular $p$ version) approximations. Some domain decomposition preconditioners have also been applied to other types of $p$ and to certain $hp$ approximations and we only briefly comment on some of them in Sect. 7.5. We believe that many important issues remain to be addressed in this field.
- We have not touched the important problems of preconditioning plate and shell problems.

- Our presentation is restricted to conforming approximations. No preconditioner is presented for, e.g., mortar methods or other approximations on nonmatching grids.
- We have also been unable to cover the recent work on domain decomposition methods in time and space which has originated with work by Jacques-Louis Lions and Yvon Maday.

The authors wish to thank, besides the anonymous referees, the many friends that have gone over this monograph or part of it and provided us with important and helpful suggestions, references, and material. They are: Xiao-Chuan Cai, Maksymilian Dryja, Bernhard Hientzsch, Axel Klawonn, Rolf Krause, Frédéric Nataf, Luca Pavarino, Alfio Quarteroni, Marcus Sarkis, Christoph Schwab, Daniel Szyld, Xavier Vasseur, and last, but not least, Barbara Wohlmuth. We would also like to thank Charbel Farhat and Oliver Rheinbach for providing us with several figures.

The authors also wish to thank different funding agencies for their support. In particular, the first author acknowledges the partial support of the Swiss National Science Foundation under Project 20-63397.00. The second author has greatly benefited, over many years, from support from the US Department of Energy and the National Science Foundation. Without this support, for many students and short term visitors, etc., our progress would undoubtedly been much slower. The second author also wishes to thank over a dozen of doctoral students, who has contributed extensively to the development of the field both in graduate school and in their careers.

We end this preface by summarizing the contents of the various chapters in order to facilitate for the reader and to accommodate his/her specific interests.

In Chap. 1, *Introduction*, we present some basic ideas of domain decomposition. In particular, we show how matching conditions for traces and fluxes of the differential problems give rise to conditions on the finite element algebraic system, how simple subdomain iterations can be devised which contain many of the ideas employed by more recent and powerful preconditioners for large scale computations on many subdomains, and how some of the ideas employed in the discussion of the Schwarz alternating method and block Jacobi preconditioners naturally lead up to the abstract Schwarz theory. This is a chapter that requires little in terms of mathematical background. We recommend it to the reader who would like to understand the basic ideas of domain decomposition without entering the specifics of the more complicated, practical algorithms. The last section, Sect. 1.6 contains some less standard and earlier results on overlapping methods and can be bypassed initially.

Chapter 2, *Abstract Theory of Schwarz Methods*, contains the standard abstract theory of additive and multiplicative Schwarz algorithms, together with some additional topics, such as coloring arguments and some hybrid algorithms. The three basic ideas of stable decompositions, strengthened Cauchy inequalities, and stable local solvers contained in three assumptions in Sect.

2.3 are central and therefore recommended in order to prepare for the chapters that follow.

Chapter 3, *Overlapping Methods*, presents overlapping preconditioners in a more general way than is normally done, since we allow for general coarse meshes and problems. In addition, the chapter contains a section on scaling and quotient space arguments, which are routinely employed in the analysis of domain decomposition preconditioners. The sections on restricted algorithms and alternative coarse problems can be bypassed initially.

In Chap. 4, *Substructuring Methods: Introduction*, we present the basic ideas of iterative substructuring methods, which are based on nonoverlapping partitions into subdomains, interior and interface variables, vertex, edge and face variables, Schur complement systems, and discrete harmonic extensions. These notions, at least at a basic level, are necessary in order to understand the iterative substructuring methods developed in the chapters that follow. The last section, Sect. 4.6 contains the Sobolev type inequalities necessary to fully analyze iterative substructuring methods and is necessary for the reader who also wishes to understand the proofs in the chapters that follow.

Chapter 5 is devoted to *Primal Iterative Substructuring Methods* for problems in three dimensions. In Sect. 5.3, we first treat the problem of devising effective local solvers by decoupling degrees of freedom associated with the vertices, edges, and faces of the subdomain partition. In Sect. 5.4, we then consider the problems of devising efficient coarse solvers, which are the key and a quite delicate part of any successful preconditioners for three-dimensional problems.

Chapter 6 is devoted to *Neumann-Neumann and FETI Methods*. We have decided to present these algorithms and their analysis together; recent research has established more and more connections between the two classes of methods. A key ingredient of this analysis is the stability of certain average and interface jump operators (cf. Sect. 6.2.3 and 6.4.3). One of the purposes of this chapter is to present the basics of one-level FETI and the more recent FETI-DP algorithms in a self-contained, sufficiently deep manner. For the reader who is interested in only the basic ideas of these methods, we recommend Sect. 6.3.1 for one-level FETI and Sect. 6.4.1, where the important ideas of FETI-DP can already be appreciated and understood in the more intuitive two-dimensional case.

In Chap. 7, we present generalizations to *Spectral Element Methods*. A basic knowledge of the corresponding algorithms for the $h$ version in the previous chapters is required. The fundamental equivalence between spectral element approximations and some finite element approximations on Gauss-Lobatto meshes is the key ingredient for the development and analysis of both overlapping and nonoverlapping methods; this is the idea underlying the Deville-Mund preconditioners reviewed in Sect. 7.2. Only those parts that are different from the proofs of the $h$ version are treated explicitly in this chapter. We have also added a brief discussion and review of domain decomposition for more general $p$ and $hp$ version finite elements with references to the literature.

In Chap. 8, generalizations to *Linear Elasticity* problems are considered. A basic knowledge of the corresponding algorithms for scalar problems in the previous chapters is required and only those parts that are different from the scalar case are treated explicitly.

In Chap. 9, some selected topics on *Preconditioners for Saddle Point Problems* are presented. They are: some basic ideas about preconditioning saddle-point problems, block-diagonal and block-triangular preconditioners, and some overlapping and iterative substructuring methods. We primarily consider the Stokes system (and briefly the related problem of incompressible elasticity) and flow in porous media. As a general rule, we only review the basic results and refer the reader to the appropriate references for more detailed and thorough presentations.

Chapter 10 is devoted to the field of domain decomposition preconditioners for *Problems in* $H(\mathrm{div}\,;\Omega)$ *and* $H(\mathbf{curl}\,;\Omega)$, which has developed relatively recently. This chapter requires a basic knowledge of the corresponding algorithms for scalar problems. Here, proofs are presented in full detail and this chapter is intended as a self-contained and deep treatment of domain decomposition methods for these problems, the analysis and development of which is in general more technically demanding than for more standard scalar and vector problems. Sections 10.1.1, 10.1.2, and 10.2.1, in particular, contain the technical tools necessary for the analysis and can be bypassed by a reader who is only interested in understanding the algorithms.

Chapter 11 is devoted to *Indefinite and Nonsymmetric Problems*. We first present a generalization of the abstract Schwarz theory to nonsymmetric and/or indefinite problems in detail. We also present some selected topics on domain decomposition preconditioners which are commonly employed in large scale computations but for which very little theory is available. These are algorithms for convection-dominated scalar problems, the Helmholtz equations, eigenvalue and nonlinear problems. This part is only intended as an overview and to provide a collection of relevant references to the literature.

The volume ends with three appendices, references, and an index.

Zürich, New York,
July 2004

*Andrea Toselli*
*Olof Widlund*

# Contents


**1 Introduction** ..... 1
1.1 Basic Ideas of Domain Decomposition ..... 1
1.2 Matrix and Vector Representations ..... 2
1.3 Nonoverlapping Methods ..... 5
1.3.1 An Equation for $u_\Gamma$: the Schur Complement System ... 5
1.3.2 An Equation for the Flux ..... 6
1.3.3 The Dirichlet-Neumann Algorithm ..... 8
1.3.4 The Neumann-Neumann Algorithm ..... 10
1.3.5 A Dirichlet-Dirichlet Algorithm or a FETI Method .... 12
1.3.6 The Case of Many Subdomains ..... 15
1.4 The Schwarz Alternating Method ..... 21
1.4.1 Description of the Method ..... 21
1.4.2 The Schwarz Alternating Method as a Richardson Method ..... 22
1.5 Block Jacobi Preconditioners ..... 24
1.6 Some Results on Schwarz Alternating Methods ..... 27
1.6.1 Analysis for the Case of Two Subdomains ..... 27
1.6.2 The Case of More than Two Subdomains ..... 29

**2 Abstract Theory of Schwarz Methods** ..... 35
2.1 Introduction ..... 35
2.2 Schwarz Methods ..... 35
2.3 Convergence Theory ..... 39
2.4 Historical Remarks ..... 46
2.5 Additional Results ..... 46
2.5.1 Coloring Techniques ..... 46
2.5.2 A Hybrid Method ..... 47
2.5.3 Comparison Results ..... 51
2.6 Remarks on the Implementation ..... 52

**3 Two-Level Overlapping Methods** ..... 55
3.1 Introduction ..... 55
3.2 Local Solvers ..... 56
3.3 A Coarse Problem ..... 59
3.4 Scaling and Quotient Space Arguments ..... 60
3.5 Technical Tools ..... 62
3.6 Convergence Results ..... 67
3.7 Remarks on the Implementation ..... 70
3.8 Numerical Results ..... 73
3.9 Restricted Schwarz Algorithms ..... 75
3.10 Alternative Coarse Problems ..... 75
3.10.1 Convergence Results ..... 76
3.10.2 Smoothed Aggregation Techniques ..... 81
3.10.3 Partition of Unity Coarse Spaces ..... 84

**4 Substructuring Methods: Introduction** ..... 87
4.1 Introduction ..... 87
4.2 Problem Setting and Geometry ..... 88
4.3 Schur Complement Systems ..... 94
4.4 Discrete Harmonic Extensions ..... 96
4.5 Condition Number of the Schur Complement ..... 97
4.6 Technical Tools ..... 99
4.6.1 Interpolation into Coarse Spaces ..... 99
4.6.2 Inequalities for Edges ..... 101
4.6.3 Inequalities for Faces ..... 105
4.6.4 Inequalities for Vertices and Auxiliary Results ..... 111

**5 Primal Iterative Substructuring Methods** ..... 113
5.1 Introduction ..... 113
5.2 Local Design and Analysis ..... 113
5.3 Local Solvers ..... 115
5.4 Coarse Spaces and Condition Number Estimates ..... 117
5.4.1 Vertex Based Methods ..... 118
5.4.2 Wire Basket Based Algorithms ..... 123
5.4.3 Face Based Algorithms ..... 126

**6 Neumann-Neumann and FETI Methods** ..... 131
6.1 Introduction ..... 131
6.2 Balancing Neumann-Neumann Methods ..... 133
6.2.1 Definition of the Algorithm ..... 133
6.2.2 Matrix Form of the Algorithm ..... 137
6.2.3 Condition Number Bounds ..... 139
6.3 One-Level FETI Methods ..... 143
6.3.1 A Review of the One-Level FETI Methods ..... 144
6.3.2 The Case of Nonredundant Lagrange Multipliers ..... 150

6.3.3 The Case of Redundant Lagrange Multipliers ..... 156
6.4 Dual-Primal FETI Methods ..... 160
6.4.1 FETI-DP Methods in Two Dimensions ..... 161
6.4.2 A Family of FETI-DP Algorithms in Three Dimensions 167
6.4.3 Analysis of Three FETI-DP Algorithms ..... 175
6.4.4 Implementation of FETI-DP Methods ..... 185
6.4.5 Computational Results ..... 187

**7 Spectral Element Methods** ..... 193
7.1 Introduction ..... 193
7.2 Deville-Mund Preconditioners ..... 196
7.3 Two-Level Overlapping Schwarz Methods ..... 198
7.4 Iterative Substructuring Methods ..... 200
7.4.1 Technical Tools ..... 202
7.4.2 Algorithms and Condition Number Bounds ..... 206
7.5 Remarks on $p$ and $hp$ Approximations ..... 210
7.5.1 More General $p$ Approximations ..... 210
7.5.2 Extensions to $hp$ Approximations ..... 214

**8 Linear Elasticity** ..... 217
8.1 Introduction ..... 217
8.2 A Two-Level Overlapping Method ..... 219
8.3 Iterative Substructuring Methods ..... 220
8.4 A Wire Basket Based Method ..... 221
8.4.1 An Extension from the Interface ..... 222
8.4.2 An Extension from the Wire Basket ..... 222
8.4.3 A Wire Basket Preconditioner for Linear Elasticity .... 224
8.5 Neumann-Neumann and FETI Methods ..... 225
8.5.1 A Neumann-Neumann Algorithm for Linear Elasticity . 225
8.5.2 One-Level FETI Algorithms for Linear Elasticity ..... 227
8.5.3 FETI-DP Algorithms for Linear Elasticity ..... 227

**9 Preconditioners for Saddle Point Problems** ..... 231
9.1 Introduction ..... 231
9.2 Block Preconditioners ..... 235
9.3 Flows in Porous Media ..... 239
9.3.1 Iterative Substructuring Methods ..... 241
9.3.2 Hybrid-Mixed Formulations and Spectral Equivalencies with Crouzeix-Raviart Approximations .. 246
9.3.3 A Balancing Neumann-Neumann Method ..... 250
9.3.4 Overlapping Methods ..... 255
9.4 The Stokes Problem and Almost Incompressible Elasticity .... 257
9.4.1 Block Preconditioners ..... 258
9.4.2 Iterative Substructuring Methods ..... 261
9.4.3 Computational Results ..... 269

**10 Problems in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$** ..... 271
10.1 Overlapping Methods ..... 274
10.1.1 Problems in $H(\mathbf{curl};\Omega)$ ..... 276
10.1.2 Problems in $H(\mathrm{div};\Omega)$ ..... 283
10.1.3 Final Remarks on Overlapping Methods and Numerical Results ..... 286
10.2 Iterative Substructuring Methods ..... 288
10.2.1 Technical Tools ..... 291
10.2.2 A Face-Based Method ..... 299
10.2.3 A Neumann-Neumann Method ..... 301
10.2.4 Remarks on Two-Dimensional Problems and Numerical Results ..... 305
10.2.5 Iterative Substructuring for Nédélec Approximations in Three Dimensions ..... 308

**11 Indefinite and Nonsymmetric Problems** ..... 311
11.1 Introduction ..... 311
11.2 Algorithms on Overlapping Subregions ..... 314
11.3 An Iterative Substructuring Method ..... 320
11.4 Numerical Results ..... 321
11.4.1 A Nonsymmetric Problem ..... 322
11.4.2 The Helmholtz Equation ..... 324
11.4.3 A Variable-Coefficient, Nonsymmetric Indefinite Problem ..... 324
11.5 Additional Topics ..... 326
11.5.1 Convection-Diffusion Problems ..... 326
11.5.2 The Helmholtz Equation ..... 330
11.5.3 Optimized Interface Conditions ..... 333
11.5.4 Nonlinear and Eigenvalue Problems ..... 334

**A Elliptic Problems and Sobolev Spaces** ..... 337
A.1 Sobolev Spaces ..... 337
A.2 Trace Spaces ..... 341
A.3 Linear Operators ..... 343
A.4 Poincaré and Friedrichs Type Inequalities ..... 343
A.5 Spaces of Vector-Valued Functions ..... 346
A.5.1 The Space $H(\mathrm{div};\Omega)$ ..... 347
A.5.2 The Space $H(\mathbf{curl};\Omega)$ in Two Dimensions ..... 348
A.5.3 The Space $H(\mathbf{curl};\Omega)$ in Three Dimensions ..... 349
A.5.4 The Kernel and Range of the Curl and Divergence Operators ..... 350
A.6 Positive Definite Problems ..... 353
A.6.1 Scalar Problems ..... 355
A.6.2 Linear Elasticity ..... 357
A.6.3 Problems in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$ ..... 360

A.7 Non-Symmetric and Indefinite Problems ..... 362
A.7.1 Generalizations of the Lax-Milgram Lemma ..... 362
A.7.2 Saddle-Point Problems ..... 364
A.8 Regularity Results ..... 369

**B Galerkin Approximations** ..... 371
B.1 Finite Element Approximations ..... 371
B.1.1 Triangulations ..... 371
B.1.2 Finite Element Spaces ..... 372
B.1.3 Symmetric, Positive Definite Problems ..... 374
B.1.4 Non-Symmetric and Indefinite Problems ..... 375
B.2 Spectral Element Approximations ..... 376
B.3 Divergence and Curl Conforming Finite Elements ..... 380
B.3.1 Raviart-Thomas Elements ..... 380
B.3.2 Nédélec Elements in Two Dimensions ..... 382
B.3.3 Nédélec Elements in Three Dimensions ..... 383
B.3.4 The Kernel and Range of the Curl and Divergence Operators ..... 384
B.4 Saddle-Point Problems ..... 386
B.4.1 Finite Element Approximations for the Stokes Problem 387
B.4.2 Spectral Element Approximations for the Stokes Problem ..... 388
B.4.3 Finite Element Approximations for Flows in Porous Media ..... 389
B.5 Inverse Inequalities ..... 389
B.6 Matrix Representation and Condition Number ..... 390

**C Solution of Algebraic Linear Systems** ..... 395
C.1 Eigenvalues and Condition Number ..... 395
C.2 Direct Methods ..... 397
C.2.1 Factorizations ..... 397
C.2.2 Fill-in ..... 398
C.3 Richardson Method ..... 399
C.4 Steepest Descent ..... 402
C.5 Conjugate Gradient Method ..... 403
C.6 Methods for Non-Symmetric and Indefinite Systems ..... 407
C.6.1 The Generalized Minimal Residual Method ..... 407
C.6.2 The Conjugate Residual Method ..... 409

**References** ..... 413

**Index** ..... 447

# 1

# Introduction

## 1.1 Basic Ideas of Domain Decomposition

The basic ideas of domain decomposition are quite natural and simple. Consider the Poisson equation on a region $\Omega$, in two or three dimensions, with zero Dirichlet data given on $\partial\Omega$, the boundary of $\Omega$. Suppose also that $\Omega$ is partitioned into two nonoverlapping subdomains $\Omega_i$ :

$$\overline{\Omega} = \overline{\Omega_1 \cup \Omega_2}, \quad \Omega_1 \cap \Omega_2 = \emptyset, \quad \Gamma = \partial\Omega_1 \cap \partial\Omega_2;$$

see Fig. 1.1. We also assume that

$$\text{measure}(\partial\Omega_1 \cap \partial\Omega) > 0, \quad \text{measure}(\partial\Omega_2 \cap \partial\Omega) > 0,$$

and that the boundaries of the subdomains are Lipschitz continuous, and consider the following problem:

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Under suitable regularity assumptions on $f$ and the boundaries of the subdomains, typically $f$ square-summable and the boundaries Lipschitz, problem (1.1) is equivalent to the following coupled problem:

$$\begin{aligned} -\Delta u_1 &= f && \text{in } \Omega_1, \\ u_1 &= 0 && \text{on } \partial\Omega_1 \setminus \Gamma, \\ u_1 &= u_2 && \text{on } \Gamma, \\ \frac{\partial u_1}{\partial n_1} &= -\frac{\partial u_2}{\partial n_2} && \text{on } \Gamma, \\ -\Delta u_2 &= f && \text{in } \Omega_2, \\ u_2 &= 0 && \text{on } \partial\Omega_2 \setminus \Gamma. \end{aligned} \tag{1.2}$$

Here $u_i$ is the restriction of $u$ to $\Omega_i$ and $\mathbf{n}_i$ the outward normal to $\Omega_i$. This equivalence can be proven by considering the corresponding variational problems; see [392, Sect. 1.2]. The conditions on the interface $\Gamma$ are called *transmission conditions* and they are also equivalent to the equality of any two independent linear combinations of the traces of the functions and their normal derivatives. In the following, we will also refer to the normal derivative as the *flux*.

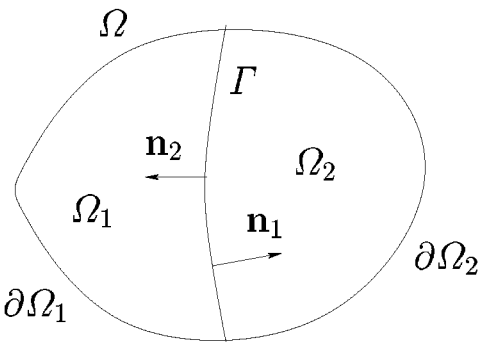


**Fig. 1.1.** Partition into two nonoverlapping subdomains.

*Remark 1.1.* The following one-dimensional example shows that some regularity beyond $f \in H^{-1}(\Omega)$ is required. Let $u$ be the weak solution of

$$\begin{aligned} -\frac{d^2u}{dx^2} &= -2\,\delta \qquad \text{in } (-1,1), \\ u(-1) = u(1) &= 0, \end{aligned} \tag{1.3}$$

where $\delta(x)$ is the delta function. The unique weak solution $u \in H_0^1(-1,1)$ is

$$u(x) = \begin{cases} -1-x & x \le 0, \\ -1+x & x \ge 0, \end{cases}$$

and its derivative has a jump at $x = 0$.

We note that this particular problem is quite relevant to domain decomposition theory. In many algorithms, we will first eliminate all nonzero components of the right hand side, of a finite element approximation, except those on the interface, in this case $x = 0$. We are then left with an equation for the remaining finite element error, which is a direct analog of equation (1.3).

## 1.2 Matrix and Vector Representations

In this section, we consider matrix and vector representations of certain operators and linear functionals; we refer to appendix B for additional details.

Starting with any domain decomposition algorithm written in terms of functions and operators, we will be able to rewrite it in matrix form as a preconditioned iterative method for a certain linear system.

We now consider a triangulation of the domain $\Omega$ and a finite element approximation of problem (1.1). We always assume that subdomains consist of unions of elements or, equivalently, that subdomain boundaries do not cut through any elements. Such an approximation gives rise to a linear system

$$Au = f \tag{1.4}$$

with a symmetric, positive definite matrix which, for a mesh size of $h$, typically has a condition number on the order of $1/h^2$. Here,

$$A = \begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ A_{\Gamma I}^{(1)} & A_{\Gamma I}^{(2)} & A_{\Gamma\Gamma} \end{pmatrix}, \quad u = \begin{pmatrix} u_I^{(1)} \\ u_I^{(2)} \\ u_\Gamma \end{pmatrix}, \quad f = \begin{pmatrix} f_I^{(1)} \\ f_I^{(2)} \\ f_\Gamma \end{pmatrix}, \tag{1.5}$$

where we have partitioned the degrees of freedom into those internal to $\Omega_1$, and to $\Omega_2$, and those of the interior of $\Gamma$.

The stiffness matrix $A$ and the load vector $f$ can be obtained by subassembling the corresponding components contributed by the two subdomains. Indeed, if

$$f^{(i)} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} \end{pmatrix}, \quad A^{(i)} = \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix}, \quad i = 1, 2, \tag{1.6}$$

are the right hand sides and the local stiffness matrices for Poisson problems with a Dirichlet condition on $\partial\Omega_i \setminus \Gamma$ and a Neumann condition on $\Gamma$, we have

$$A_{\Gamma\Gamma} = A_{\Gamma\Gamma}^{(1)} + A_{\Gamma\Gamma}^{(2)}, \quad f_\Gamma = f_\Gamma^{(1)} + f_\Gamma^{(2)}.$$

In view of the transmission conditions in (1.2), we will look for an *approximation* of the normal derivatives on $\Gamma$. Given the local exact solution $u_i$, its normal derivative can be defined as a linear functional by using Green's formula. Thus, if $\phi_j$ is a nodal basis function for a node on $\Gamma$, we have, using (1.2),

$$\int_\Gamma \frac{\partial u_i}{\partial n_i} \phi_j \, ds = \int_{\Omega_i} (\Delta u_i \phi_j + \nabla u_i \cdot \nabla \phi_j) \, dx = \int_{\Omega_i} (-f\phi_j + \nabla u_i \cdot \nabla \phi_j) \, dx.$$

An approximation, $\lambda^{(i)}$, of the functional representing the normal derivative can be found by replacing the exact solution $u_i$ in the right hand side with its finite element approximation. Letting $j$ run over the nodes on $\Gamma$ and using the definition of the local stiffness matrix, we introduce the expression

$$\lambda^{(i)} = A_{\Gamma I}^{(i)} u_I^{(i)} + A_{\Gamma\Gamma}^{(i)} u_\Gamma^{(i)} - f_\Gamma^{(i)}. \tag{1.7}$$

This approximation, not surprisingly, coincides with the residual corresponding to the nodes on $\Gamma$ of a Poisson problem with a Neumann condition on $\Gamma$; see (1.6). We also note that $\lambda^{(i)}$ is not obtained from the normal derivative of the finite element solution but as an approximation of the linear functional of the exact solution.

Using these definitions, we can find an approximation of problem (1.2):

$$\begin{aligned} &A_{II}^{(1)} u_I^{(1)} + A_{I\Gamma}^{(1)} u_\Gamma^{(1)} = f_I^{(1)}, \\ &u_\Gamma^{(1)} = u_\Gamma^{(2)} = u_\Gamma, \\ &(A_{\Gamma I}^{(1)} u_I^{(1)} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^{(1)} - f_\Gamma^{(1)}) = -(A_{\Gamma I}^{(2)} u_I^{(2)} + A_{\Gamma\Gamma}^{(2)} u_\Gamma^{(2)} - f_\Gamma^{(2)}) = \lambda_\Gamma, \\ &A_{II}^{(2)} u_I^{(2)} + A_{I\Gamma}^{(2)} u_\Gamma^{(2)} = f_I^{(2)}. \end{aligned} \tag{1.8}$$

We note that the first and last equations of (1.8) are discretizations of Poisson problems for the interior functions $u_I^{(i)}$ with Dirichlet data which vanishes on $\partial\Omega_i \setminus \Gamma$ and is equal to a common value $u_\Gamma$ on $\Gamma$. Alternatively, the first and third equations provide a discretization of a Poisson problem in $\Omega_1$ for the local function $u_1$ with Neumann data equal to $\lambda_\Gamma$ and vanishing Dirichlet data on $\partial\Omega_1 \setminus \Gamma$. An analogous Neumann problem in $\Omega_2$ is provided by the third and fourth equations. The solution of these local problems with suitable Dirichlet and Neumann data provide the building blocks of the nonoverlapping methods in section 1.3.

We note, finally, that while the equivalence of (1.1) and (1.2) might not be immediate, the equivalence of (1.4) and (1.8) is trivial. If $u_\Gamma^{(1)} = u_\Gamma^{(2)} = u_\Gamma$, then the third equation of (1.8), which ensures the equality of the approximations of the linear functionals representing the normal derivatives, coincides with the third row of the original linear system (1.4) because of (1.5) and (1.6). More importantly, while the continuous problem (1.2) is valid only if the right hand side $f$ is sufficiently regular, its discrete counterpart (1.8) is always valid, since it can be found directly from the finite element problem.

If we, in particular, consider the one-dimensional problem of Remark 1.1 with $\Omega_1 = (-1, 0)$ and $\Omega_2 = (0, 1)$ and a triangulation with a node at $x = 0$, we see that the right hand side $f_\Gamma$ cannot be built by summing two local components. However, problem (1.8) is still equivalent to the original linear system (1.4) since in (1.8) the local components $f_\Gamma^{(i)}$ are not needed but only their sum $f_\Gamma$. A full discussion of the continuous problem in two or more dimensions is relatively complicated. We note that the methods discussed in section 1.3 will have nonzero residuals only at nodal points on the interface after the completion of the first iteration step and this is the analog of a solution of the continuous problem with a nonzero jump in the normal derivative across the interface. However, as we have seen, no genuine technical problems remain once we turn to the finite element case.

## 1.3 Nonoverlapping Methods

We will first consider some simple *iterative substructuring* methods that rely on a partition into nonoverlapping subdomains. We refer to Chap. 4, 5, and 6 for a systematic presentation and further details and generalizations. As we will show, these methods are indeed preconditioned iterative methods for the boundary value $u_\Gamma$ or for the normal derivative $\lambda_\Gamma$. We start with these methods since they are derived directly from the coupled problems (1.2) and (1.8). We note that domain decomposition methods based on overlapping partitions were the first to be devised (namely, the alternating Schwarz method on overlapping subdomains); they will be introduced in Sect. 1.4.

### 1.3.1 An Equation for $u_\Gamma$: the Schur Complement System

Let us consider the linear system (1.4) with $A$, $u$, and $f$ defined in (1.5). In a first step of many iterative domain decomposition methods, the unknowns in the interior of the subdomains ($u_I^{(i)}$) are eliminated. This corresponds to a block factorization of the matrix of (1.5):

$$A = LR = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ A_{\Gamma I}^{(1)} {A_{II}^{(1)}}^{-1} & A_{\Gamma I}^{(2)} {A_{II}^{(2)}}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ 0 & 0 & S \end{pmatrix} \tag{1.9}$$

and a resulting linear system

$$\begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ 0 & 0 & S \end{pmatrix} u = \begin{pmatrix} f_I^{(1)} \\ f_I^{(2)} \\ g_\Gamma \end{pmatrix}. \tag{1.10}$$

Here $I$ is the identity matrix and $S = A_{\Gamma\Gamma} - A_{\Gamma I}^{(1)} {A_{II}^{(1)}}^{-1} A_{I\Gamma}^{(1)} - A_{\Gamma I}^{(2)} {A_{II}^{(2)}}^{-1} A_{I\Gamma}^{(2)}$ is the Schur complement relative to the unknowns on $\Gamma$. By a direct calculation, we see that $S$ and $g_\Gamma$ can be found by subassembling local contributions. In particular, recalling the form of the local matrices (1.6) and defining the local Schur complements by

$$S^{(i)} := A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} A_{I\Gamma}^{(i)}, \quad i = 1, 2, \tag{1.11}$$

we find the Schur complement system for $u_\Gamma$ to be

$$S u_\Gamma = g_\Gamma, \tag{1.12}$$

with

$$S = S^{(1)} + S^{(2)},$$

$$g_\Gamma = (f_\Gamma^{(1)} - A_{\Gamma I}^{(1)} {A_{II}^{(1)}}^{-1} f_I^{(1)}) + (f_\Gamma^{(2)} - A_{\Gamma I}^{(2)} {A_{II}^{(2)}}^{-1} f_I^{(2)}) =: g_\Gamma^{(1)} + g_\Gamma^{(2)}.$$

We note that once $u_\Gamma$ is found, by solving (1.12), the internal components can be found by using (1.10):

$$u_I^{(i)} = A_{II}^{(i)^{-1}}(f_I^{(i)} - A_{I\Gamma}^{(i)}u_\Gamma); \tag{1.13}$$

these are just solutions of two inhomogeneous Dirichlet problems.

The Schur complement system, which provides an equation for the approximation of the trace of the exact solution on $\Gamma$, has been derived purely algebraically by block Gaussian elimination. It is interesting to note that it can also be obtained by using the transmission conditions of the coupled system. Thus, let us write the internal variables $u_I^{(1)}$ and $u_I^{(2)}$ in terms of $u_\Gamma$ by using the first and last equations of (1.8). Substituting these expressions into the third equation of (1.8), we again arrive at equation (1.12). We note that in this last step, we use the flux condition expressed in the third equation of (1.8). A dual procedure that provides an equation for $\lambda_\Gamma$ is given in subsection 1.3.2.

We can also obtain an equation for the trace of the exact solution on $\Gamma$ working directly with the continuous problem (1.2). The corresponding operator is called a *Steklov-Poincaré* operator. The Schur complement system (1.12) is an approximation of the Steklov-Poincaré equation, *determined* directly by the finite element approximation, particularly, by the approximation of the normal derivatives (1.7). (We refer to Chap. 4 for a systematic presentation of Schur complement systems.)

### 1.3.2 An Equation for the Flux

We now derive an equation for the normal derivative $\lambda_\Gamma$ on $\Gamma$ by employing a procedure analogous to that of the previous subsection. We use the unknown common boundary value $\lambda_\Gamma = \lambda_\Gamma^{(1)} = -\lambda_\Gamma^{(2)}$ (see the third equation of problem (1.8)) and solve local Neumann problems to find $u^{(1)}$ and $u^{(2)}$:

$$\begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix} \begin{pmatrix} u_I^{(i)} \\ u_\Gamma^{(i)} \end{pmatrix} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} + \lambda_\Gamma^{(i)} \end{pmatrix}, \quad i = 1, 2. \tag{1.14}$$

Using a block factorization of the local matrices, we find

$$u_\Gamma^{(i)} = S^{(i)-1}(g_\Gamma^{(i)} + \lambda_\Gamma^{(i)}),$$

with $g_\Gamma^{(i)}$ given as in the previous subsection. Using the second equation of problem (1.8), which makes $u_\Gamma^{(1)}$ and $u_\Gamma^{(2)}$ the same, we find

$$F\lambda_\Gamma = d_\Gamma, \tag{1.15}$$

with

$$F = S^{(1)-1} + S^{(2)-1},$$
$$d_\Gamma = d_\Gamma^{(1)} + d_\Gamma^{(2)} := -S^{(1)-1} g_\Gamma^{(1)} + S^{(2)-1} g_\Gamma^{(2)}. \tag{1.16}$$

We note that this is the same linear system which will appear as (6.29) in section 6.3.1 if we specialize that formula to the case of two subdomains. Once $\lambda_\Gamma$ is known, we can find the local solutions $u^{(1)}$ and $u^{(2)}$ by solving the two Neumann problems in (1.14).

In the rest of this section, we will consider some domain decomposition methods of *iterative substructuring* type. We note that many others can be defined by replacing Dirichlet and/or Neumann conditions on $\Gamma$ by more general ones involving linear combinations of $u$ and the normal derivative; see, e.g, [392, Sect. 1.3] for some of them; our presentation is similar to that of that reference. However, we will only work with the coupled differential and algebraic problems (1.2) and (1.8), respectively. Our purpose is to illustrate that many domain decomposition methods derived from (1.2) are indeed preconditioned Richardson methods for the Schur complement system (1.12) or for the equation (1.15). We note that the methods presented in the following can be derived purely algebraically using suitable splittings of the system (1.8), without any reference to the underlying continuous operator, traces, or normal derivatives. However, employing a functional framework does not only give an interpretation of these iteration procedures, but will also turn out to be crucial as a preparation for the analysis of their convergence rates. For many domain decomposition algorithms this will be carried out using equivalences between certain discrete and Sobolev norms and certain Sobolev type inequalities for finite element functions; the systematic development of this theory will begin in Chap. 4.

All the algorithms to be introduced in the next few subsections involve preconditioners in solving equations (1.12) or (1.15). We note that we could also solve these equations by using a Richardson or conjugate gradient method without preconditioning. The evaluation of $S$ times a vector involves the solution of one Dirichlet problem on each subdomain while that of $F$ times a vector requires the solution of one Neumann problem on each subdomain. The rate of convergence of such algorithms is determined by the condition numbers of $S$ and $F$. However, even if the Schur complement has a smaller condition number than the original stiffness matrix, the number of iterations will increase in proportion to $1/h$ when the mesh size $h$ decreases; see Sect. 4.3 for details. We also note that the conjugate gradient method will be faster than the Richardson method and that it requires no a priori spectral information while the optimal choice of the parameter $\theta$ of the Richardson method involves obtaining estimates of the smallest and largest eigenvalues of the operator; see appendix C.3.

### 1.3.3 The Dirichlet-Neumann Algorithm

Methods of this type were first considered in [61, 73, 211, 343, 57, 344]. Extensions, also including global coarse solvers, were considered in [176, 173, 465].

The basic Dirichlet-Neumann algorithm consists of two fractional steps corresponding to the two subregions $\Omega_i, i = 1, 2$. Given an initial guess $u_\Gamma^0$, we first solve a Dirichlet problem in $\Omega_1$ with Dirichlet data $u_\Gamma^0$ on $\Gamma$, and then a mixed Neumann-Dirichlet problem on $\Omega_2$ with a Neumann condition on $\Gamma$ determined by the solution in $\Omega_1$ obtained in the previous step and with Dirichlet conditions on the rest of $\partial\Omega_2$. The new iterate $u_\Gamma^1$ is chosen as the trace of the solution in $\Omega_2$, or, more generally, as a linear combination of this trace and $u_\Gamma^0$, using a suitably chosen relaxation parameter $\theta$; see appendix C.3. In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$(D) \begin{cases} -\Delta u_1^{n+1/2} = f & \text{in } \Omega_1, \\ u_1^{n+1/2} = 0 & \text{on } \partial\Omega_1 \setminus \Gamma, \\ u_1^{n+1/2} = u_\Gamma^n & \text{on } \Gamma, \end{cases}$$

$$(N) \begin{cases} -\Delta u_2^{n+1} = f & \text{in } \Omega_2, \\ u_2^{n+1} = 0 & \text{on } \partial\Omega_2 \setminus \Gamma, \\ \dfrac{\partial u_2^{n+1}}{\partial n_2} = -\dfrac{\partial u_1^{n+1/2}}{\partial n_1} & \text{on } \Gamma, \end{cases} \tag{1.17}$$

$$u_\Gamma^{n+1} = \theta u_2^{n+1} + (1-\theta) u_\Gamma^n \text{ on } \Gamma,$$

with $\theta \in (0, \theta_{\max})$. Using our approximation for the normal derivatives, i.e. (1.7), we can derive the corresponding iteration for the discrete problem. If we define the vectors of internal degrees of freedom as $v_1 = u_I^{(1)}$ and $w_2 = u_I^{(2)}$, cf. (1.7), we find

$$(D)\ A_{II}^{(1)} v_1^{n+1/2} + A_{I\Gamma}^{(1)} u_\Gamma^n = f_I^{(1)},$$

$$(N) \begin{pmatrix} A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ A_{\Gamma I}^{(2)} & A_{\Gamma\Gamma}^{(2)} \end{pmatrix} \begin{pmatrix} w_2^{n+1} \\ \tilde{u}_\Gamma^{n+1} \end{pmatrix} = \begin{pmatrix} f_I^{(2)} \\ f_\Gamma^{(2)} - \lambda_\Gamma^{n+1/2} \end{pmatrix}, \tag{1.18}$$

$$u_\Gamma^{n+1} = \theta \tilde{u}_\Gamma^{n+1} + (1-\theta) u_\Gamma^n,$$

with

$$\lambda_\Gamma^{n+1/2} = A_{\Gamma I}^{(1)} v_1^{n+1/2} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^n - f_\Gamma^{(1)}.$$

It is clear that (1.18) arises from a splitting of the original system (1.8) and thus provides a consistent iterative method for its solution, i.e., the limit of any convergent sequence will satisfy the correct set of equations.

We next eliminate $v_1^{n+1/2}$ from (1.18) and find

$$\lambda_\Gamma^{n+1/2} = -(g_\Gamma^{(1)} - S^{(1)} u_\Gamma^n).$$

Using then a block factorization of the local matrix $A^{(2)}$ and eliminating $w_2^{n+1}$ from (1.18) yields the following equation:

$$S^{(2)}(u_\Gamma^{n+1} - u_\Gamma^n) = \theta(g_\Gamma - S u_\Gamma^n),$$

which shows that the Dirichlet-Neumann algorithm is a preconditioned Richardson iteration for the Schur complement system (1.12), with the preconditioner ${S^{(2)}}^{-1}$. The preconditioned operator is

$${S^{(2)}}^{-1} S = I + {S^{(2)}}^{-1} S^{(1)},$$

the application of which to a vector involves the solution of a Dirichlet problem, (a multiplication by $S^{(1)}$), and a multiplication by ${S^{(2)}}^{-1}$ which corresponds to solving a problem with Neumann conditions on $\Gamma$ and Dirichlet conditions of the rest of $\partial\Omega_2$; see section 4.3 for further details.

We note that the spectral equivalence between $S^{(1)}$ and $S^{(2)}$, and thus a uniform bound for the condition number of ${S^{(2)}}^{-1} S$, is ensured by the existence of stable, discrete harmonic, finite element extensions $\mathcal{H}_i u_\Gamma$ from the interface $\Gamma$ into the subdomains $\Omega_i$; these matters are discussed systematically in section 4.6, in particular in Lemma 4.10. Here, and in what follows, the condition number of the preconditioned operator is the ratio of the largest and smallest eigenvalues of the symmetric generalized eigenvalue problem defined by the operator and its preconditioner, in the case at hand by $S$ and $S^{(2)}$; see appendix C.5. We employ the following definition:

**Definition 1.2 (Optimality).** *An iterative method for the solution of a linear system is said to be optimal, if its rate of convergence to the exact solution is independent of the size of the system.*

We note that for the algebraic systems considered in this monograph optimality is ensured if the rate of convergence is independent of the size of the finite element spaces employed, and therefore of the meshsize $h$ for $h$ approximations or of the polynomial degree for spectral element approximations.

If we denote the appropriate trace seminorm by $|\cdot|_{1/2,\Gamma}$, we will show in Lemma 4.10 that

$$|\mathcal{H}_i u_\Gamma|^2_{1,\Omega_i} = (\mathcal{H}_i u_\Gamma)^T A^{(i)} (\mathcal{H}_i u_\Gamma) \leq C_i \, |u_\Gamma|^2_{1/2,\Gamma},$$

and thus

$$u_\Gamma^T S^{(2)} u_\Gamma = |\mathcal{H}_2 u_\Gamma|^2_{1,\Omega_2} \leq C_2 \, |u_\Gamma|^2_{1/2,\Gamma} \leq C_2 \tilde{C}_1 \, |\mathcal{H}_1 u_\Gamma|^2_{1,\Omega_1} = C_2 \tilde{C}_1 \, u_\Gamma^T S^{(1)} u_\Gamma.$$

Here we have used the relation between the Schur complement and the energy of the discrete harmonic extension given in Lemma 4.9 and the trace estimate for $\Omega_1$ given in Lemma 4.10 with a constant $\tilde{C}_1$. Using similar estimates for $u_\Gamma^T S^{(1)} u_\Gamma$, we find

$$\kappa({S^{(2)}}^{-1}S^{(1)}) \leq \frac{C_1\tilde{C}_2}{C_2\tilde{C}_1}.$$

We note that this bound only depends on having shape regular, quasi-uniform local meshes, and, in particular, on the shape and the relative size of the subdomains. In the special case where $\Omega_1$ and $\Omega_2$ have the same shape and size, are symmetric with respect to the interface $\Gamma$, and have the same triangulations, then $\kappa({S^{(2)}}^{-1}S^{(1)}) = 1$. Generally, the Dirichlet-Neumann method for two subdomains is optimal since stable extensions can be found: the condition number satisfies

$$\kappa({S^{(2)}}^{-1}S) \leq C,$$

with $C$ a constant independent of the dimension of the finite element problem.

### 1.3.4 The Neumann-Neumann Algorithm

Methods of this type were first considered in [65, 309, 154]; see [163, 226, 227, 9] for some earlier closely related studies.

The basic Neumann-Neumann algorithm can be described as follows: we start from an initial guess $u_\Gamma^0$. A step of the Neumann-Neumann algorithm consists in first solving Dirichlet problems on each $\Omega_i$ with data $u_\Gamma^0$ on $\Gamma$, and then a problem on each subdomain, with Neumann data, on $\Gamma$, chosen as the difference of the normal derivatives of the solutions of the two Dirichlet problems. The values on $\Gamma$ of the solutions of these Neumann problems are then employed to correct the initial $u_\Gamma^0$ and find the new iterate $u_\Gamma^1$. In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$(D_i)\left\{\begin{array}{ll} -\Delta u_i^{n+1/2} = f & \text{in } \Omega_i, \\ u_i^{n+1/2} = 0 & \text{on } \partial\Omega_i \setminus \Gamma, \\ u_i^{n+1/2} = u_\Gamma^n & \text{on } \Gamma, \end{array}\right\}, \quad i = 1, 2,$$

$$(N_i)\left\{\begin{array}{ll} -\Delta\psi_i^{n+1} = 0 & \text{in } \Omega_i, \\ \psi_i^{n+1} = 0 & \text{on } \partial\Omega_i \setminus \Gamma, \\ \dfrac{\partial\psi_i}{\partial n_i}^{n+1} = \dfrac{\partial u_1}{\partial n_1}^{n+1/2} + \dfrac{\partial u_2}{\partial n_2}^{n+1/2} & \text{on } \Gamma, \end{array}\right\}, \quad i = 1, 2, \tag{1.19}$$

$$u_\Gamma^{n+1} = u_\Gamma^n - \theta(\psi_1^{n+1} + \psi_2^{n+1}) \text{ on } \Gamma,$$

with a suitable $\theta \in (0, \theta_{\max})$. Using our approximation for the normal derivatives, we can derive an iteration for the discrete problem. If we define the vectors of internal degrees of freedom as $v_i = u_I^{(i)}$ and $w_i = \psi_I^{(i)}$, we find

$$(D_i)\ A_{II}^{(i)} v_i^{n+1/2} + A_{I\Gamma}^{(i)} u_\Gamma^n = f_I^{(i)}, \quad i = 1, 2,$$

$$(N_i)\ \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix} \begin{pmatrix} w_i^{n+1} \\ \eta_i^{n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ r_\Gamma \end{pmatrix}, \quad i = 1, 2, \tag{1.20}$$

$$u_\Gamma^{n+1} = u_\Gamma^n - \theta(\eta_1^{n+1} + \eta_2^{n+1}),$$

where the residual $r_\Gamma$ is defined as

$$r_\Gamma = (A_{\Gamma I}^{(1)} v_1^{n+1/2} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^n - f_\Gamma^{(1)}) + (A_{\Gamma I}^{(2)} v_2^{n+1/2} + A_{\Gamma\Gamma}^{(2)} u_\Gamma^n - f_\Gamma^{(2)});$$

see the third equation in (1.8).

We next eliminate $v_i^{n+1/2}$ and $w_i^{n+1}$ from (1.20). Problems $(D_i)$ give

$$r_\Gamma = -(g_\Gamma - S u_\Gamma^n), \tag{1.21}$$

which shows that the difference $r_\Gamma$ of the local fluxes is equal to minus the residual of the Schur complement system. Using a block factorization of the local matrices $A^{(i)}$, problems $(N_i)$, then give

$$\eta_i^{n+1} = S^{(i)-1} r_\Gamma = -S^{(i)-1}(g_\Gamma - S u_\Gamma^n),$$

Therefore, we find

$$u_\Gamma^{n+1} - u_\Gamma^n = \theta({S^{(1)}}^{-1} + {S^{(2)}}^{-1})(g_\Gamma - S u_\Gamma^n),$$

which shows that the Neumann-Neumann algorithm is also a preconditioned Richardson iteration for the Schur complement system, with the preconditioner ${S^{(1)}}^{-1} + {S^{(2)}}^{-1}$. The preconditioned matrix is

$$FS = ({S^{(1)}}^{-1} + {S^{(2)}}^{-1})S = ({S^{(1)}}^{-1} + {S^{(2)}}^{-1})(S^{(1)} + S^{(2)}), \tag{1.22}$$

the application of which to a vector involves the solution of two Dirichlet problems and two problems with Neumann data on $\Gamma$. We note that the operator $F$ is the same as that in section 1.3.2.

The optimality of this method follows easily from the result on the Dirichlet-Neumann algorithm by using the spectral mapping theorem; the preconditioned matrix in (1.22) can be written as ${S^{(2)}}^{-1} S^{(1)} + 2I + ({S^{(2)}}^{-1} S^{(1)})^{-1}$ where the eigenvalues of ${S^{(2)}}^{-1} S^{(1)}$ are uniformly bounded from above and below.

We note that, given positive weights $\delta_1^\dagger$ and $\delta_2^\dagger$ with

$$\delta_1^\dagger + \delta_2^\dagger = 1,$$

the residual $r_\Gamma$ on the right hand side of the Neumann problems in (1.20) can be replaced by $\delta_1^\dagger r_\Gamma$ and $\delta_2^\dagger r_\Gamma$ for $\Omega_1$ and $\Omega_2$, respectively. Similarly, when finding the new iterate $u_\Gamma^{n+1}$ the sum of the two corrections $\eta_1^{n+1}$ and $\eta_2^{n+1}$ can be replaced by a weighted average:

$$u_\Gamma^{n+1} = u_\Gamma^n - \theta(\delta_1^\dagger \eta_1^{n+1} + \delta_2^\dagger \eta_2^{n+1}).$$

This gives rise to the preconditioner

$$D^{(1)} {S^{(1)}}^{-1} D^{(1)} + D^{(2)} {S^{(2)}}^{-1} D^{(2)},$$

where $D^{(i)}$ is a diagonal matrix, the diagonal entries of which are equal to $\delta_i^\dagger$. Scaling matrices $D^{(i)}$ are commonly employed in practice in order to improve the convergence of Neumann-Neumann algorithms for subdomains with cross points and, in particular, for problems with large changes in the coefficients; see Sect. 6.2.

Generalizations of the Dirichlet-Neumann and Neumann-Neumann algorithms are possible, by using more general Robin conditions. This was already proposed in some early work by Agoshkov and Lebedev [10], Agoshkov [9], and Lions [321]. We note that the use of more general interface conditions is often required for some nonsymmetric or indefinite problems. We refer to Sect. 11.5 for more details and further generalizations of these ideas.

We close this section by introducing an alternative preconditioner for the Schur complement system (1.12). We note that it can be shown that the Poincaré-Steklov operator that corresponds to $S$ is a bijection from $H_{00}^{1/2}(\Gamma)$ to its dual $H_{00}^{-1/2}(\Gamma)$, and that the application of this operator therefore essentially involves the loss of one derivative; see A.2. It is therefore natural to use the inverse of $J$, the square root of minus an appropriate discrete Laplacian on $\Gamma$, as a preconditioner. If the mesh is uniform and has the appropriate number of points, this preconditioner can be implemented using the Fast Fourier transform; see [61]. This idea appears to originate with Dryja [172] and it works quite well. But it also has limitations not shared by the methods just discussed. Thus, in three dimensions, the mesh of the interface $\Gamma$ must have a quite special tensor structure in order to allow use of a two-dimensional Fast Fourier transform.

### 1.3.5 A Dirichlet-Dirichlet Algorithm or a FETI Method

We now consider a method dual to the Neumann-Neumann algorithm. We start from an initial guess $\lambda_\Gamma^0$ of the flux on $\Gamma$; cf. Sect. 1.3.2. We first solve two Neumann problems on $\Omega_i$ with data $\lambda_\Gamma^0$ on $\Gamma$, and then a problem on each subdomain, with Dirichlet data, on $\Gamma$, chosen as the difference of the trace of the solutions of the two Neumann problems on $\Gamma$. The values on $\Gamma$ of the normal derivatives of the solutions of these Dirichlet problems are then employed to correct the initial $\lambda_\Gamma^0$ and find the new iterate $\lambda_\Gamma^1$. We recall

that $\lambda_\Gamma$ is an approximation of the normal derivative in the direction $\mathbf{n}_1$. In order to simplify the notation, we will drop the subscript $\Gamma$ for the normal derivatives and set

$$\lambda^n = \lambda_\Gamma^n, \quad \lambda_1^n = -\lambda_2^n = \lambda^n.$$

In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$(N_i) \left\{ \begin{array}{ll} -\Delta u_i^{n+1/2} = f & \text{in } \Omega_i, \\ u_i^{n+1/2} = 0 & \text{on } \partial\Omega_i \setminus \Gamma, \\ \dfrac{\partial u_i}{\partial n_i}^{n+1/2} = \lambda_i^n & \text{on } \Gamma, \end{array} \right\}, \quad i = 1, 2,$$

$$(D_i) \left\{ \begin{array}{ll} -\Delta \psi_i^{n+1} = 0 & \text{in } \Omega_i, \\ \psi_i^{n+1} = 0 & \text{on } \partial\Omega_i \setminus \Gamma, \\ \psi_i^{n+1} = u_1^{n+1/2} - u_2^{n+1/2} & \text{on } \Gamma, \end{array} \right\}, \quad i = 1, 2, \tag{1.23}$$

$$\lambda^{n+1} = \lambda^n - \theta\left(\frac{\partial \psi_1}{\partial n_1}^{n+1} + \frac{\partial \psi_2}{\partial n_2}^{n+1}\right) \text{ on } \Gamma,$$

with a suitable $\theta \in (0, \theta_{\max})$. Using our approximation of the normal derivatives, we can derive an iteration for the discrete problem. If we define the vectors of internal degrees of freedom by $v_i = u_I^{(i)}$ and $w_i = \psi_I^{(i)}$, we find

$$\begin{aligned} &(N_i) \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix} \begin{pmatrix} v_i^{n+1/2} \\ \gamma_i^{n+1/2} \end{pmatrix} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} + \lambda_i^n \end{pmatrix}, \quad i = 1, 2, \\ &(D_i)\ A_{II}^{(i)} w_i^{n+1} + A_{I\Gamma}^{(i)} r_\Gamma = 0, \quad i = 1, 2, \\ &\lambda^{n+1} = \lambda^n - \theta(\eta_1^{n+1} + \eta_2^{n+1}), \end{aligned} \tag{1.24}$$

where the residual $r_\Gamma$ is defined by

$$r_\Gamma = \gamma_1^{n+1/2} - \gamma_2^{n+1/2}$$

and the fluxes $\eta_i^{n+1}$ by

$$\eta_i^{n+1} = A_{\Gamma I}^{(i)} w_i^{n+1/2} + A_{\Gamma\Gamma}^{(i)} r_\Gamma;$$

cf. Equation (1.7).

We proceed as in the previous section and eliminate $v_i^{n+1/2}$, $\gamma_i^{n+1/2}$, and $w_i^{n+1}$ from (1.24). Using a block factorization of the local matrices $A^{(i)}$, problems $(N_i)$ give

$$r_\Gamma = -(d_\Gamma - F\lambda^n),$$

which shows that the difference $r_\Gamma$ of the local solutions is equal to the negative of the residual of the system (1.15). Problems $(D_i)$ then provide

$$\eta_i^{n+1} = S^{(i)} r_\Gamma = -S^{(i)}(d_\Gamma - F\lambda^n),$$

Therefore, we find

$$\lambda^{n+1} - \lambda^n = \theta(S^{(1)} + S^{(2)})(d_\Gamma - F\lambda^n),$$

which shows that this Dirichlet-Dirichlet algorithm is also a preconditioned Richardson iteration for the system (1.15), with the preconditioner $S^{(1)} + S^{(2)}$. The preconditioned matrix is

$$SF = S({S^{(1)}}^{-1} + {S^{(2)}}^{-1}) = (S^{(1)} + S^{(2)})({S^{(1)}}^{-1} + {S^{(2)}}^{-1}), \tag{1.25}$$

the application of which to a vector involves the solution of two Neumann problems and two problems with Dirichlet data on $\Gamma$. We have called the method presented here the *Dirichlet-Dirichlet* algorithm in analogy to the dual Neumann-Neumann one of the previous section. We remark however that, in the domain decomposition literature, the term Dirichlet-Dirichlet is often used for the Richardson or conjugate gradient method applied to the unpreconditioned Schur complement system (1.12), since the application of $S$ involves the solution of a Dirichlet problem on each subdomain. We also note that the preconditioned operator $SF$ is the same as that of the FETI algorithm with the original Dirichlet preconditioner; cf. Equation (6.36) in Sect. 6.3.1 for the case of two subdomains. In the following, we will refer to the method of this section as the *preconditioned FETI method.* FETI methods will be presented and analyzed systematically in Chap. 6.

We recall that the preconditioned operator for the Neumann-Neumann algorithm is $FS$. It is a trivial matter to prove that the operators $SF$ and $FS$ have the same eigenvalues and thus the same condition number. The condition numbers of the Neumann-Neumann and Dirichlet-Dirichlet methods, and indeed the entire spectra, are *the same* for the case of two subdomains.

As for Neumann-Neumann algorithms, we can also employ weights. With $\delta_1^\dagger$ and $\delta_2^\dagger$ the weights introduced in Sect. 1.3.4, the residual $r_\Gamma$ in the Dirichlet problems in (1.24) is often replaced by $\delta_2^\dagger r_\Gamma$ and $\delta_1^\dagger r_\Gamma$ for $\Omega_1$ and $\Omega_2$, respectively. Similarly, when finding the new iterate $\lambda^{n+1}$ the sum of the two corrections $\eta_1^{n+1}$ and $\eta_2^{n+1}$ can be replaced by an average:

$$\lambda^{n+1} = \lambda^n - \theta(\delta_2^\dagger \eta_1^{n+1} + \delta_1^\dagger \eta_2^{n+1}).$$

This gives rise to the preconditioner

$$D^{(2)} S^{(1)} D^{(2)} + D^{(1)} S^{(2)} D^{(1)};$$

see in particular Sect. 6.3.2 and 6.3.3, for more details, and in particular formulas (6.37) and (6.51). Suitable scaling matrices are commonly employed in practice in order to improve the convergence of FETI algorithms for partitions into subdomains with cross points or for problems with large changes in the coefficients; see Sect. 6.3.

We note that, by proceeding as in Sect. 1.3.3, we can also design a Neumann-Dirichlet algorithm by starting the iteration with an initial guess of the Neumann data; see, e.g., [63, 301]. We would then solve a Neumann problem on one subdomain and use the trace of this solution as Dirichlet data for the problem on the other subdomain, etc. The corresponding preconditioned operator is

$$S^{(2)}F = S^{(2)}(S^{(2)-1} + S^{(1)-1}) = I + S^{(2)}S^{(1)-1},$$

the condition number of which can be immediately bounded using a bound for the Dirichlet-Neumann algorithm in Sect. 1.3.3.

Finally, we mention that ideas similar to those in this and the previous subsection were already proposed in some early work by Glowinski and Wheeler [226, 227] for mixed approximations of elliptic problems.

### 1.3.6 The Case of Many Subdomains

The methods just introduced can be generalized to the case of more than two subdomains. Generalizations of the Neumann-Neumann and FETI methods have become widely used. Here we only give a brief introduction and refer to Chap. 6 for a much more complete description and analysis.

The Dirichlet-Neumann method can also be generalized, under certain restrictions on the partition into subdomains. This method is less widely used than Neumann-Neumann and FETI methods since its performance can deteriorate if the coefficients of the differential equations differ greatly between the subdomains. However, some important ideas underlying much more recent and popular methods, such as the dual-primal FETI algorithms, can be traced back to early work on this family of methods; see Sect. 6.4. In addition, our presentation will show the importance of the cross points of the partition. A cross point is a point common to the boundaries of three or more subregions. If there are no cross points, we essentially have partitioned the region $\Omega$ into strips; that case is in many ways similar to the two subdomain case. We note that while we assume that the original region $\Omega$ has only one connected component, it is quite natural to consider cases where two subregions $\Omega_i$ have several or many components.

Here we will only consider two-dimensional problems. Before proceeding, we need to define some more general operators. We consider problem (1.1) and suppose that $\Omega$ is partitioned into a family of nonoverlapping subdomains $\{\Omega_i,\ 1 \leq i \leq N\}$ with

$$\overline{\Omega} = \bigcup_i \overline{\Omega_i}; \quad \Omega_i \cap \Omega_j = \emptyset \quad i \neq j.$$

If $\Gamma_i = \partial\Omega_i \setminus \partial\Omega$, the interface $\Gamma$ is defined as

$$\Gamma = \bigcup_i \Gamma_i.$$

We note that $\Gamma$ and the $\Gamma_i$ are open. A coupled problem as in (1.2) can be found with transmission conditions imposed along each edge $\partial\Omega_i \cap \partial\Omega_j$.

The linear system (1.4) can now be written as

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ f_\Gamma \end{pmatrix}, \tag{1.26}$$

where we have partitioned the degrees of freedom into those interior to the subdomains and those on $\Gamma$. The stiffness matrix and the right hand side are obtained by subassembling the corresponding components relative to the subdomains; see (1.6).

The unknowns in the interior of the subdomains $u_I$ can again be eliminated by block Gaussian elimination and the resulting linear system is

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ 0 & S \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ g_\Gamma \end{pmatrix}. \tag{1.27}$$

As before the Schur complement $S$ and the vector $g_\Gamma$ can be found by subassembling local contributions; see Chap. 4 for further details. In order to do so, we first define a family of restriction operators. Given a vector of degrees of freedom $u_\Gamma$ on the interface, we define the restriction $R_i u_\Gamma$ as the vector of degrees of freedom of $u_\Gamma$ on $\Gamma_i$. Here $R_i$ is a rectangular matrix of zeros and ones. If for each subdomain the degrees of freedom are partitioned into those internal to $\Omega_i$ and those on $\Gamma_i$, as in (1.6), we have

$$\begin{aligned} S &= \sum_{i=1}^{N} R_i^T S^{(i)} R_i, \\ g_\Gamma &= \sum_{i=1}^{N} R_i^T (f_\Gamma^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} f_I^{(i)}) \end{aligned} \tag{1.28}$$

where the local Schur complements are defined as in (1.11) and $R_i^T$ is the transpose of $R_i$.

### Neumann-Neumann Methods

By examining the preconditioned Neumann-Neumann operator for two subdomains (1.22), we can easily find a generalization to the case of many subdomains. We define

$$S_{NN}^{-1} S = \sum_{i=1}^{N} R_i^T {S^{(i)}}^{-1} R_i \, S. \tag{1.29}$$

We note that the application of this operator to a vector involves the solution, on each subdomain $\Omega_i$, of a Dirichlet problem and a problem with Neumann boundary conditions on $\partial\Omega_i \cap \Gamma$. We also note that for subdomains that do not touch $\partial\Omega$, $S^{(i)}$ is singular and ${S^{(i)}}^{-1}$ in (1.29) should be understood as a pseudo-inverse or the inverse of a regularized problem.

We note that the Neumann-Neumann algorithm can also be defined at the continuous level using (1.19), with $i = 1, 2, \ldots, N$, and Neumann conditions for the problem $N_i$ given on each edge $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$. The new iterate $u_\Gamma^{n+1}$, at a node of the interface, is then built from corrections from all subdomains with that node on their boundaries.

The condition number of the Neumann-Neumann method satisfies a logarithmic bound. If $h$ is the diameter of the finite elements and $H$ that of a typical subdomain, we have

$$\kappa(S_{NN}^{-1}S) \leq \frac{C}{H^2}\left(1 + \log\frac{H}{h}\right)^q. \tag{1.30}$$

The method is not therefore scalable according to the following definition:

**Definition 1.3 (Scalability).** *A domain decomposition iterative method for the solution of a linear system is said to be scalable, if its rate of convergence does not deteriorate when the number of subdomains grows. This typically means that convergence does not deteriorates when $H$, the typical subdomain size, becomes small.*

The dependence on $H^{-2}$ in (1.30) is typical of domain decomposition preconditioned methods without a coarse solver. Since the Green function for elliptic problems does not in general have a compact support, a residual on the linear system of modestly low frequency will result in an error which cannot be neglected in *any* part of the region. Therefore an iterative method for the solution of the resulting linear system in which information is only exchanged between neighboring subregions must necessarily, for certain initial errors, require a number of steps which is at least equal to the diameter of the dual graph corresponding to the subdomain partition. Here, the dual graph is constructed by introducing a vertex for each subregion and an edge between two subregions that share a part of their boundaries. The diameter of the graph is the maximum distance between pairs of vertices, where the distance is defined as the length of the shortest path between the vertices. In case the diameter of the original domain is one, the diameter of the graph is typically $O(H^{-1})$. Using an argument of contradiction and the upper bound for the error of the conjugate gradient iteration in Lemma C.10, we see that the condition number of a domain decomposition preconditioned operator must grow at least as $H^{-2}$. We note that for the same reason, the condition number of the stiffness matrix resulting from the finite element approximation of an elliptic problem must grow at least as $h^{-2}$; see Theorem B.32.

Going back to the bound in (1.30), if the partition does not have cross points, we have $q = 0$ and we have to interpret $H$ as the width of the strips. Otherwise, $q = 3$. A quadratic growth, $q = 2$, is obtained if suitable scaling matrices are incorporated into the preconditioner; see Sect. 6.2 for further details. We note that the Neumann-Neumann method with many subdomains is not in general optimal according to Definition 1.2, since its condition number

may depend on the meshsize $h$. The condition number bound does not however involve $h$ alone but only the ratio $H/h$, which gives a measure of the number of unknowns associated with one subdomain. In addition, the condition number increases slowly (logarithmically) with $H/h$. In such a case, the method is said to be *quasi-optimal*.

We finally note that the preconditioned operator $S_{NN}^{-1}S$ has the typical form of a Schwarz operator; these operators will be discussed in detail in Chap. 2.2. We have

$$S_{NN}^{-1}S = \sum_{i=1}^{N} P_i, \quad P_i = R_i^T {S^{(i)}}^{-1} R_i \, S.$$

Each $P_i$ is associated with a subdomain (or, in other words, to a subspace) and is a projection-like operator; see Sect. 2.2 for further details.

### Dirichlet-Neumann Methods

We now assume that there is a red-black partition of the subdomains into two sets $\mathcal{R}$ and $\mathcal{B}$, such that the intersection between the boundaries of two subregions in the same class is either empty or a vertex; see Fig. 1.2.

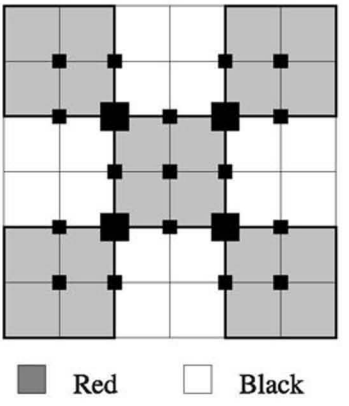


**Fig. 1.2.** Red-black coloring of the subdomains.

We can then decompose the Schur complement as

$$S = \sum_{i \in \mathcal{R}} R_i^T S^{(i)} R_i + \sum_{i \in \mathcal{B}} R_i^T S^{(i)} R_i,$$

and define the preconditioner essentially in terms of Neumann problems on the subdomains in the set $\mathcal{R}$

$$S_{DN}^{-1} = \left( \sum_{i \in \mathcal{R}} R_i^T S^{(i)} R_i \right)^{-1}.$$

In the particular case where there are no cross points, the matrix $S_{DN}$ is block-diagonal, since the intersection between the boundaries of two red regions is empty and there is therefore no coupling between degrees of freedom on two subdomain boundaries in the same class. In particular, for $i, j \in \mathcal{R}$, we then have

$$R_i R_j^T = \begin{cases} 0 & i \neq j, \\ I & i = j, \end{cases}$$

and thus

$$S_{DN}^{-1} = \sum_{i \in \mathcal{R}} R_i^T {S^{(i)}}^{-1} R_i. \tag{1.31}$$

The preconditioned Dirichlet-Neumann operator becomes

$$S_{DN}^{-1} S = I + \left( \sum_{i \in \mathcal{R}} R_i^T {S^{(i)}}^{-1} R_i \right) \left( \sum_{i \in \mathcal{B}} R_i^T S^{(i)} R_i \right), \tag{1.32}$$

and its application involves the solution of a Dirichlet problem on the subdomains in the set $\mathcal{B}$ and a Neumann problem on those in the set $\mathcal{R}$. As for the case of two subdomains, bounds for the condition number of the Dirichlet-Neumann operator rely on finding stable finite element extensions. A uniform bound, with $q = 0$, can be derived for the case of strips while in the general case, discussed below in this section, we have a logarithmic bound (1.30) as for the Neumann-Neumann algorithm. We note however, that in the strip case the algorithm is not scalable and that the bound will contain a factor $1/H^2$ where $H$ is the width of a typical strip of the decomposition; cf. the discussion on the lack of scalability earlier in this subsection.

If there are cross points, then formula (1.31) is no longer valid. Indeed, the inversion of $S_{DN}$ requires the solution of a global problem, which instead of creating a problem turns out to be a blessing since it provides a natural coarse solve, which makes the algorithm scalable, i.e., convergent at a rate independent of the number of subdomains if the subdomains all have good aspect ratios. In particular, we have to solve a Neumann problem on the union of the subdomains of the set $\mathcal{R}$, i.e., on the set glued together at the cross points and given by

$$\bigcup_{i \in \mathcal{R}} \overline{\Omega_i};$$

see the shaded region in Fig. 1.2. This can be done in two stages. All the degrees of freedom of the subdomains of the set $\mathcal{R}$ except for those at cross points (drawn as small squares in Fig. 1.2) are first eliminated in parallel across the subdomains by a step of block Gaussian elimination. All the submatrices involved at this stage are invertible. The resulting Schur complement, which involves only the nodal values at the cross points (drawn as larger squares in Figure 1.2), is sparse since it can be shown that nonzero off-diagonal elements only exist for the pairs of cross points which belong to the boundary of the same subregion. This final step of the elimination typically will involve degrees

of freedom from all parts the region $\Omega$. This mechanism helps explain that bounds can be derived for the condition number of this preconditioned matrix which are independent of the number of subregions. We remark that a similar procedure is employed in FETI-DP methods; see section 6.4. The resulting preconditioned operator is

$$S_{DN}^{-1}S = I + S_{DN}^{-1}\left(\sum_{i\in\mathcal{B}} R_i^T S^{(i)} R_i\right),$$

and its application to a vector requires the solution of Dirichlet problems on the black subdomains, and of the special Neumann problem, just introduced, on the union of the red subdomains, which involves the low-dimensional coarse problem already discussed. The condition number satisfies

$$\kappa(S_{DN}^{-1}S) \leq C\left(1 + \log\frac{H}{h}\right)^2;$$

see [176, 465]. We note that the technical tools needed in deriving such bounds are developed in Chap. 4.

### Methods Involving the Normal Derivative

The expression given in (1.7) of the functional representing the normal derivative $\lambda^{(i)}$ on $\partial\Omega_i$ remains valid for the case of many subdomains and in particular when the partition has cross points. We recall that the coupled system (1.8) is obtained from the original system (1.4) by introducing the additional unknowns $\lambda^{(i)}$. A coupled system for the case of many subdomains can easily be found. This involves the equality of the trace functions $u_\Gamma^{(i)}$ at all the nodes on the interface $\Gamma$. For the normal derivatives $\lambda^{(i)}$, conditions similar to that of the third equation of (1.8) are a consequence of the fact that the stiffness matrix $A$ is obtained by subassembling contributions from neighboring subdomains. For every node $x_h$ on $\Gamma$ the corresponding condition becomes

$$\sum_j \lambda^{(j)}(x_h) = 0,$$

where the sum is taken over all the subdomains that share the node $x_h$. We note that, if $x_h$ is a cross point, the sum is taken over more than two subdomains and it appears to be difficult to give a functional meaning to this condition. In addition, when trying to generalize Equation (1.15) to the case of many subdomains, some of local Neumann problems are not uniquely solvable and therefore the corresponding local Schur complements not invertible, if the subdomain boundaries do not all intersect $\partial\Omega$.

The generalization of what we have called the Dirichlet-Dirichlet algorithm will be carried out systematically in Chap. 6, where we introduce the

important FETI family of methods. These methods provide preconditioned operators which act on suitable approximations of the normal derivative $\lambda_\Gamma$ on the interface $\Gamma$. There, following the development of FETI methods, we will proceed in a purely algebraic way and approximations of the normal derivative will be given by vectors of Lagrange multipliers of suitable constrained systems. As we will show, in Chap. 6, there are more than one way to define an approximation of the normal derivative $\lambda_\Gamma$ when the partition has cross points and the definition will depend on the particular form of the equality constraints chosen for the interface values $u_\Gamma^{(i)}$ across $\Gamma$.

We finally recall that the issue of singular, local Neumann problems was already addressed in [226, 227] for mixed approximations of elliptic problems.

## 1.4 The Schwarz Alternating Method

The earliest domain decomposition method known to the authors is the alternating method of H.A. Schwarz, [416], published in 1870. Schwarz used the algorithm to establish the existence of harmonic functions with prescribed boundary values on regions with nonsmooth boundaries. The regions were constructed recursively by forming unions of pairs of regions starting with regions for which existence could be established by some more elementary means. At the core of Schwarz's work is a proof that this iterative method converges in the maximum norm at a geometric rate.

For more than two subregions, we can in fact define a step of the algorithm by recursion: i) solve on the first subregion; ii) solve on the union of all other subregions, approximately, by recursively invoking a step of the same algorithm.

As pointed out by Pierre-Louis Lions [319], the convergence of this algorithm can be established by two different methods, namely, by a maximum principle and by using Hilbert spaces. The Hilbert space method is the most appropriate here since much of our work relies on the classical calculus of variation and finite elements.

We finally note that between the work of Schwarz, [416], and Lions, [319], there was some quite significant work, in particular by Sobolev [425] and Babuška [29]; see also [319] for additional references.

### 1.4.1 Description of the Method

The classical Schwarz method, for two subregions, can be described as follows: Consider the Poisson problem (1.1) on a bounded Lipschitz region $\Omega$ with zero Dirichlet boundary conditions. There are two fractional steps corresponding to two overlapping subregions, $\Omega_1'$ and $\Omega_2'$ of the original region $\Omega = \Omega_1' \bigcup \Omega_2'$; see Fig. 1.3. Given an initial guess $u^0$, which vanishes on $\partial\Omega$, the iterate $u^{n+1}$ is determined from the previous iterate $u^n$ in two sequential steps in which the approximate solution on the two subregions is updated:

$$\begin{cases} -\Delta u^{n+1/2} = f & \text{in } \Omega_1', \\ u^{n+1/2} = u^n & \text{on } \partial\Omega_1', \\ u^{n+1/2} = u^n & \text{in } \Omega_2 = \Omega_2' \setminus \overline{\Omega_1'}, \end{cases}$$
$$\begin{cases} -\Delta u^{n+1} = f & \text{in } \Omega_2', \\ u^{n+1} = u^{n+1/2} & \text{on } \partial\Omega_2', \\ u^{n+1} = u^{n+1/2} & \text{in } \Omega_1 = \Omega_1' \setminus \overline{\Omega_2'}. \end{cases} \tag{1.33}$$

Thus, the Dirichlet data for these problems are obtained from the original data given on $\partial\Omega \cap \partial\Omega_i'$, and the values from the previous fractional step on the remaining part $\Gamma_i = \partial\Omega_i' \setminus \partial\Omega$ of the subdomain boundaries. We note that this algorithm also can be viewed as a mapping of values on $\Gamma_1$ (or $\Gamma_2$) onto values on the same set; see Sect. 1.6.1.

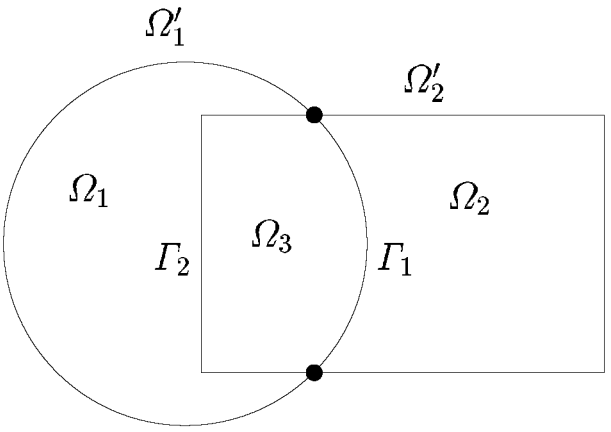


**Fig. 1.3.** Overlapping partition for the Schwarz alternating method with two subdomains

### 1.4.2 The Schwarz Alternating Method as a Richardson Method

In order to show that the algorithm is indeed a Richardson iteration, we rewrite it in variational form. For the original problem, we use the space $H_0^1(\Omega)$ and the bilinear form

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx, \tag{1.34}$$

which defines the $H^1(\Omega)$−seminorm; see appendix A.1. We also introduce a finite element triangulation $\mathcal{T}$, of maximum diameter $h$, such that the local

boundaries $\partial\Omega_i'$ do not cut through any element in $\mathcal{T}$. The mesh $\mathcal{T}$ thus defines two local meshes $\mathcal{T}_i$ on $\Omega_i'$, $i = 1, 2$. We next define the spaces of continuous, piecewise linear finite elements on $\mathcal{T}$, $\mathcal{T}_1$, and $\mathcal{T}_2$, which vanish on $\partial\Omega$, $\partial\Omega_1'$, and $\partial\Omega_2'$, respectively, and denote them by $V$, $V_1$, and $V_2$. We note that there are two natural extension operators

$$R_i^T : V_i \longrightarrow V, \quad i = 1, 2,$$

which take local functions on the $\Omega_i'$ with zero boundary values and extend them by zero to the rest of $\Omega$ to give global functions in $V$. We note that the restriction operators $R_i : V \longrightarrow V_i$ take global functions on $\Omega$ and give local functions in $V_i$ that vanish on $\partial\Omega_i'$ and are equal to the original ones at the nodes *inside* $\Omega_i'$. We also need the two local bilinear forms

$$a_i(u, v) = \int_{\Omega_i'} \nabla u \cdot \nabla v \, dx, \quad u, v \in V_i.$$

We can now write the Schwarz method in terms of two orthogonal projections $P_i$, $i = 1, 2$. For $i = 1, 2$, the projections are

$$P_i = R_i^T \tilde{P}_i,$$

where $\tilde{P}_i : V \longrightarrow V_i$, are defined by

$$a_i(\tilde{P}_i u, v_i) = a(u, R_i^T v_i), \quad v_i \in V_i \, .$$

Indeed, if we consider a finite element discretization of the first problem in (1.33), we see that $u^{n+1/2} - u^n$ vanishes in $\Omega_2 = \Omega_2' \setminus \overline{\Omega_1'}$ and, when restricted to $\Omega_1'$, belongs to $V_1$. With $u$ the exact finite element solution, this function satisfies, for $v_1 \in V_1$,

$$a_1(R_1(u^{n+1/2} - u^n), v_1) = \int_{\Omega_1'} f v_1 \, dx - a_1(u^n, R_1^T v_1) = a(u - u^n, R_1^T v_1),$$

and thus

$$R_1(u^{n+1/2} - u^n) = \tilde{P}_1(u - u^n).$$

Since $(u^{n+1/2} - u^n)$ vanishes in $\Omega \setminus \Omega_1'$, we obtain

$$u^{n+1/2} - u^n = R_1^T R_1(u^{n+1/2} - u^n) = P_1(u - u^n).$$

The error $u^{n+1/2} - u$ is then given by

$$u^{n+1/2} - u = (I - P_1)(u^n - u).$$

Proceeding in a similar way for the finite element discretization of the second problem in (1.33), we find

$$u^{n+1} - u = (I - P_2)(u^{n+1/2} - u) = (I - P_2)(I - P_1)(u^n - u),$$

which shows that the error propagation operator of this *multiplicative* Schwarz method is

$$(I - P_2)(I - P_1) \;=\; I - (P_1 + P_2 - P_2P_1).$$

Therefore, the algorithm can be viewed as a simple iterative method for solving

$$P_{mu}u := (P_1 + P_2 - P_2P_1)u = g.$$

With an appropriate right-hand side $g$, $u \in V$ will be the solution of the original finite element problem. We note that the error propagation operator is not symmetric. However, after the first step, the error will belong to $range(I - P_2)$ and the operator $P_{mu}$ is symmetric on that subspace. The algorithm immediately extends to three and more subspaces; in these more general cases symmetry cannot be recovered in this way.

## 1.5 Block Jacobi Preconditioners

An important variant of the Schwarz methods is the *additive Schwarz method.* We first consider a very simple example, namely a two-block Jacobi method, and try to understand how well it works and how it can be improved. This will give a first example of an additive method.

We will work with the matrix form (1.4) of our finite element problem. The stiffness matrix $A$ is positive definite, symmetric, which are properties inherited by any conforming finite element method, from the bilinear form $a(\cdot,\cdot)$ in (1.34); see appendix B for additional details. Here, and in the following, we exploit the one-to-one correspondence between the finite element space $V$ and the corresponding space of degrees of freedom consisting of the nodal values of a function and use the same notation for finite element spaces and spaces of degrees of freedom, and functions and corresponding vectors of degrees of freedom.

We consider the block-Jacobi/conjugate gradient method: the stiffness matrix $A$ is preconditioned by a matrix $A_J^{-1}$, where $A_J$ is the a direct sum of two diagonal blocks of $A$. Each block corresponds to a set of degrees of freedom, which define a subspace $V_i$. We write the space $V$ as a direct sum of the subspaces $V_i, i = 1, 2$,

$$V = R_1^T V_1 \oplus R_2^T V_2,$$

where $R_i^T$ are the natural extension operators

$$R_i^T : V_i \longrightarrow V, \quad i = 1, 2,$$

which take the sets of local degrees of freedom of $V_i$ and extend them by zero to the remaining degrees of freedom in $\Omega$ to give global vectors of nodal values

in $V$. If the block of $A$ related to $V_i$ is denoted by $A_i$, the preconditioner $A_J^{-1}$ can be written as

$$A_J^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & A_2^{-1} \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}^{-1}. \tag{1.35}$$

The matrix $A_J$ is obtained from $A$ by removing the coupling between the two subspaces $V_1$ and $V_2$. We expect that the weaker the coupling, the better the preconditioner.

We now write $A_J$ in a more compact form. Let the restriction operators

$$R_i : V \longrightarrow V_i, \quad i = 1, 2,$$

be the adjoints of $R_i^T$ with respect to the Euclidean scalar product; $R_i$ takes a vector of global degrees of freedom and extracts the degrees of freedom associated with $V_i$. We immediately see that

$$A_i = R_i A R_i^T, \quad i = 1, 2,$$

and that

$$A_J^{-1} = R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2. \tag{1.36}$$

The choice of the subspaces is a key issue and so is the choice of basis of $V$, in particular in the spectral element case. In order to improve the convergence, we can use subspaces that do not form a direct sum of $V$, e.g., by choosing the subspaces $V_1$ and $V_2$ defined in the previous section, consisting of degrees of freedom associated with the interior of two overlapping subdomains $\Omega_1'$ and $\Omega_2'$. The preconditioner $A_J^{-1}$ can still be written as in (1.36).

In order to establish a connection with the Schwarz alternating method of Sect. 1.4, we will show that the preconditioned additive operator

$$P_{ad} = A_J^{-1} A$$

can also be written using two projections,

$$P_i := R_i^T A_i^{-1} R_i A \ : V \longrightarrow V, \quad i = 1, 2.$$

In Lemma 2.1, we will show that the $P_i$ are the same as those of Sect. 1.4. We immediately see that

$$P_{ad} = P_1 + P_2.$$

Since this operator is the sum of two projections, which are orthogonal in the inner product (1.34), we obtain an upper bound of 2 for the largest eigenvalue. If there are $N$ subspaces, we similarly obtain a bound of $N$. However, this bound can be often be improved by using a coloring technique; see section 2.5.1.

Before concluding this section, we return to the case of two subspaces that form a direct sum decomposition of $V$, and show how in this simple case the

smallest eigenvalue of the block-Jacobi preconditioned system $P_{ad}$ is related to the representation properties of the subspace decomposition. This will be generalized to more general decompositions in section 2.3; see Assumption 2.2 and Lemma 2.5.

Since we now consider a direct sum of two subspaces, there is a unique decomposition $u = R_1^T u_1 + R_2^T u_2$, for every $u \in V$, with

$$u_1 = R_1 u, \quad u_2 = R_2 u.$$

The corresponding block-Jacobi preconditioner is given by (1.35). Let $C_0$ be a positive constant such that, for every $u \in V$,

$$u_1^T A_1 u_1 + u_2^T A_2 u_2 \leq C_0^2 \, u^T A u.$$

We note that $C_0^2$ measures the stability of the splitting and that we use the scalar product induced by the original matrix $A$. Since we work with a direct sum, we can write

$$\begin{aligned} u^T A(P_{ad}^{-1})u &= u^T (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)^{-1} u \\ &= u^T (R_1^T A_1 R_1 + R_2^T A_2 R_2) u \\ &= u_1^T A_1 u_1 + u_2^T A_2 u_2, \end{aligned}$$

which, combined with the stability property of the splitting, gives

$$\sup_{u \in V} \frac{u^T A(P_{ad}^{-1})u}{u^T A u} \leq C_0^2,$$

and thus a lower bound for the smallest eigenvalue of $P_{ad}$. Thus the condition number is bounded from above by $2/C_0^2$.

We conclude this section with some remarks on the two methods presented so far.

*Remark 1.4.*

1. The additive and multiplicative algorithms can be generalized to the case of more than two subspaces in a straightforward way. An additional component, associated with a coarse space $V_0$, can also be added to any of the algorithms. The latter is often necessary for good convergence, in particular, to make the convergence rate independent of the number of subproblems; cf. discussion in subsection 1.3.6 and appendix C.5.
2. Additive and multiplicative algorithms can be determined by the same decomposition into subspaces but they are different. They correspond to different preconditioned operators, $P_{ad}$ and $P_{mu}$, determined by different *polynomial* functions of the projections $\{P_i\}$.
3. More generally, instead of the local solvers $\{A_i^{-1}\}$ in the definition of the preconditioners, we can employ modified operators $\{\tilde{A}_i^{-1}\}$, corresponding, for instance, to *inexact solvers*. We also note that we might consider replacing the sum of the values of $P_1 u$ and $P_2 u$ at the nodes of an overlapping subdomain by a convex combination of the two values.

4. For two subspaces and exact solvers, Bjørstad and Mandel [58] have shown that the eigenvalues of the additive operator $P_1 + P_2$ can be expressed by those of the multiplicative operator $P_1 + P_2 - P_2P_1$, and vice versa.

## 1.6 Some Results on Schwarz Alternating Methods

In this final section, we give two results on the convergence of the Schwarz alternating method. They represent early efforts to prove the convergence of Schwarz overlapping methods before the Schwarz theory of the next chapter was fully developed. While we believe that these results are of interest, this section can also be bypassed since it is not necessary for the understanding of the main part of this monograph.

### 1.6.1 Analysis for the Case of Two Subdomains

In Chap. 2, we will develop a systematic framework to estimate the rate of convergence of the Schwarz alternating method and many other algorithms, and we will also devote Chap. 3 to the description and analysis of a number of algorithms based on overlapping subdomains. Here, we will show, following Bjørstad and Widlund [62], that a, perhaps surprising, connection can be made to the algorithms discussed and analyzed in Sect. 1.3. In this subsection, we will consider the case of two subregions; see Fig. 1.3.

Examining Schwarz's method (1.33), we see that we can view it as a mapping from $\Gamma_1 = \partial\Omega_1' \setminus \partial\Omega$ onto itself; once the correct value of $u_{\Gamma_1}$ has been found, the iteration has converged. This observation is equally valid for the finite element approximation. We note that the values of the iterates on $\Gamma_1$ change only in the second fractional step of the algorithm.

We also see that as soon as the first full step of the algorithm has been completed, the error in (1.33) will be harmonic in the three regions $\Omega_1$, $\Omega_2$, and $\Omega_3 = \Omega_1' \cap \Omega_2'$. Similarly, in the finite element case, we see that the residuals vanish at all nodes inside these three regions after the first step of the algorithm has been completed. We can therefore confine our discussion to the case for which the right hand sides differ from zero only on the interfaces $\Gamma_1$ and $\Gamma_2$. In fact, after the first fractional step the residual will also vanish at all mesh points on $\Gamma_2$, and we see that the error $e^{n+1/2}$ will satisfy the linear system of equations

$$Ae^{n+1/2} = \begin{pmatrix} 0 \\ 0 \\ r_{\Gamma_1}^{n+1/2} \end{pmatrix},$$

where $A$ is the global stiffness matrix on $\Omega$ and we have partitioned the degrees of freedom into subvectors representing nodal values in the interior of $\Omega_1'$, the

interior to $\Omega_2$, and on the interface $\Gamma_1$. Using a block factorization of $A$ (cf. equation (1.12)), we find

$$(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)})e_{\Gamma_1}^{n+1/2} = r_{\Gamma_1}^{n+1/2}. \tag{1.37}$$

Here, $S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}$ is the Schur complement of $A$ with respect to $\Gamma_1$, obtained by subassembling the local Schur complements for $\Omega_1'$ and $\Omega_2$. We can express the same observation by saying that the restriction of the stiffness matrix for the entire region to the space with vanishing residuals on these two subregions can be expressed in terms of the sum of these two Schur complements.

The error on $\Gamma_1$ will only be partially eliminated in the second fractional step, since we are solving a problem on the subregion $\Omega_2'$ rather than on all of $\Omega$. With vanishing residuals in $\Omega_1'$ and $\Omega_2$, we can write the correction in the second fractional step as

$$A_2(u_2^{n+1} - u_2^{n+1/2}) = \begin{pmatrix} 0 \\ 0 \\ r_{\Gamma_1}^{n+1/2} \end{pmatrix},$$

where $A_2$ is the stiffness matrix of $\Omega_2'$ and we have partitioned the degrees of freedom of $\Omega_2'$ into those interior to $\Omega_2$, those interior to $\Omega_3$, and those of the interface $\Gamma_1$. Using a block factorization of $A_2$, we find

$$u_{\Gamma_1}^{n+1} - u_{\Gamma_1}^{n+1/2} = (S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1} r_{\Gamma_1}^{n+1/2}; \tag{1.38}$$

here $S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)}$ is the Schur complement of $A_2$ with respect to $\Gamma_1$, obtained by subassembling the local Schur complements for $\Omega_2$ and $\Omega_3$. We note that $S_{\Gamma_1}^{(2)}$ in (1.38) is the same as in (1.37).

Since $e_{\Gamma_1}^{n+1/2} = e_{\Gamma_1}^{n}$, a simple computation allows us to find the error propagation operator, regarded as a mapping from $\Gamma_1$ to itself:

$$e_{\Gamma_1}^{n+1} = \left(I - (S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1}(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)})\right) e_{\Gamma_1}^{n}. \tag{1.39}$$

It is easy to show that $S_{\Gamma_1}^{(3)} \geq S_{\Gamma_1}^{(1)}$, (an inequality in terms of quadratic forms), since the minimal extension from $\Gamma_1$ to the interior of $\Omega_3$ necessarily has a larger energy than its counterpart for the larger region $\Omega_1'$. Thus, the space of possible extensions is larger in the latter case; see Lemma 4.9. Therefore, the error propagation operator has only positive eigenvalues. The convergence is uniform since we can also show, by arguments quite similar to those for the Neumann-Dirichlet and Neumann-Neumann methods, that the eigenvalues of

$$(S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1}(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}) \tag{1.40}$$

are bounded uniformly away from zero.

### 1.6.2 The Case of More than Two Subdomains

We will now show that the convergence rate of the Schwarz alternating method for three and more subdomains can be estimated in terms of the convergence rates of certain problems on two subdomains; this result was first given in [464]. We note that results of a somewhat different nature for the same problem can be based on the abstract theory of Chap. 2 and the results of Chap. 3. In particular, several methods with coarse spaces will be introduced and analyzed in the latter chapter.

We will consider the case of three and more subregions and a symmetric variant of the algorithm. Exact solvers are employed and the operators $P_i$ are thus projections. For the case of three subdomains, as in Fig. 1.4, the error propagation operator is given by

$$(I - P_1)(I - P_2)(I - P_3)(I - P_2)(I - P_1); \tag{1.41}$$

see Sect. 1.4.2.

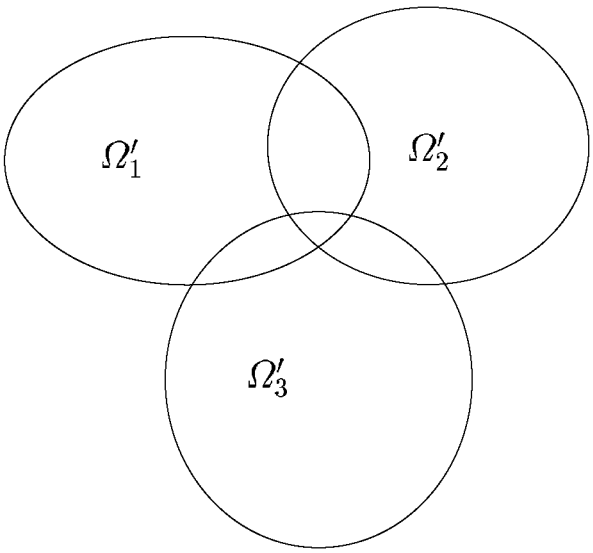


**Fig. 1.4.** Three overlapping subdomains.

This represents an algorithm where we solve Dirichlet problems in $\Omega_1'$, $\Omega_2'$, and $\Omega_3'$, in that order, prior to returning to solve on the second and first subdomains. Since $P_3$ is a projection, we can also write this operator as $E_3^* E_3$, where the transpose $*$ is understood in the sense of the bilinear form $a(\cdot, \cdot)$, and $E_3 = (I - P_3)(I - P_2)(I - P_1)$. This algorithm can clearly be generalized to the case of $k > 3$ subdomains and we will derive a bound for the condition number of the symmetric multiplicative Schwarz operator

$$I - E_k^* E_k \quad \text{where} \quad E_k = (I - P_k) \dots (I - P_1) \tag{1.42}$$

in terms of the condition numbers of problems on pairs of subdomains for which our analysis of subsection 1.6.1 can be used. Since the $P_i$ are orthogonal projections, the maximum eigenvalue of $I - E_k^* E_k$ is 1, and the smallest eigenvalue of this operator equals $1/\kappa$, where $\kappa$ is its condition number. It is now easy to show that $\|E_k\|_a \leq (1 - 1/\kappa)^{1/2}$, with $\|\cdot\|_a$ the norm induced by the scalar product $a(\cdot,\cdot)$. Thus, our bound for the condition number of $I - E_k^* E_k$, given in Theorem 1.6, will immediately translate into an estimate of the rate of convergence of the multiplicative Schwarz method.

Before we formulate and prove the main result of this subsection, we will examine the recursive structure of the algorithm in some detail. We first solve a problem in $\Omega_1'$ resulting in a zero residual in that subregion. For the final, $(2k-1)$th, fractional step, we return to the linear system for the same subregion $\Omega_1'$. The second through $(2k-2)$th fractional steps provide an updated solution in $\Omega \setminus \Omega_1'$. In the final fractional step, we only need the original right hand side and the new values on $\partial\Omega_1'$ obtained in the intermediate steps. All the values in $\Omega_1'$ are overwritten and do not enter the computation in the final fractional step in any other way. Similarly, the third through $(2k-3)$th step can be viewed as solely providing updated values in the set $\Omega \setminus (\Omega_1' \cup \Omega_2')$ including new Dirichlet data on part of $\partial\Omega_2'$. We can therefore view the Schwarz algorithm in terms of two regions, $\Omega_1'$ and $\Omega_2' \cup \ldots \cup \Omega_k'$ where we obtain approximative values in $\Omega \setminus \Omega_1'$ by a recursive call of one step of the symmetric, multiplicative Schwarz algorithm using the subregions $\Omega_2'$ and $\Omega_3' \cup \ldots \cup \Omega_k'$, etc.

We will now give a matrix representation of the process. Given a region $\Psi = \Psi_1' \cup \Psi_2'$, and associating the subvectors $x_1$, $x_2$, and $x_3$ with the degrees of freedom of $\Psi_1'$, $\Psi_2 := \Psi \setminus \overline{\Psi_1'} = \Psi_2' \setminus \overline{\Psi_1'}$, and $\partial\Psi_1' \setminus \partial\Psi$, respectively, we can write the coefficient matrix for the region $\Psi$ as

$$A = \begin{pmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{pmatrix} ; \tag{1.43}$$

cf. (1.5). We can make the substitutions $\Psi_1' = \Omega_i'$, $\Psi_2' = \Omega_{i+1}' \cup \ldots \cup \Omega_k'$, and thus $\Psi_2 = (\Omega_{i+1}' \cup \ldots \cup \Omega_k') \setminus \overline{\Omega_i'}$, for $1 \leq i \leq k-1$, as suggested by our discussion of the recursive formulation of the algorithm.

We now consider the first step of an exact block Cholesky factorization of the matrix (1.43) and obtain

$$A = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ A_{13}^T A_{11}^{-1} & 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} I & 0 & A_{11}^{-1} A_{13} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} . \tag{1.44}$$

As in Sect. 1.2, we write the matrix $A_{33}$ as a sum of $A_{33}^{(1)}$ and $A_{33}^{(2)}$, which represent the contributions from integrals over $\Psi_1'$ and $\Psi_2 = \Psi \setminus \overline{\Psi_1'}$, respectively.

As before, the Schur complement $S^{(1)} = A_{33}^{(1)} - A_{13}^T A_{11}^{-1} A_{13}$ is an intermediate matrix in a block Gaussian elimination of the matrix

$$\begin{pmatrix} A_{11} & A_{13} \\ A_{13}^T & A_{33}^{(1)} \end{pmatrix}, \tag{1.45}$$

which represents the contribution from $\Psi_1'$ to the stiffness matrix of the entire problem. We note that solving a linear system of equations with the Schur complement $S^{(1)}$ as the coefficient matrix gives the same result as solving a system with the matrix (1.45), after extending the right hand side by zero, and then discarding the components of the solution corresponding to the subvector $x_1$. This can easily be seen by considering a block factorization of the matrix (1.45).

As already shown in subsection 1.6.1, the multiplicative Schwarz algorithm for two subregions can be interpreted in terms of the solution of a linear system with a modified factored matrix obtained by replacing the Schur complement $S^{(1)}$, which corresponds to the region $\Psi_1'$, by $S^{(3)}$, which is the Schur complement corresponding to $\Psi_3 = \Psi_1' \cap \Psi_2'$. The matrix, which represents the preconditioner for the symmetric two-region Schwarz algorithm, is then of the form,

$$\hat{A} = \begin{pmatrix} I & 0\,0 \\ 0 & I\,0 \\ A_{13}^T A_{11}^{-1} & 0\,I \end{pmatrix} \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} I\,0 & A_{11}^{-1} A_{13} \\ 0\,I & 0 \\ 0\,0 & I \end{pmatrix}. \tag{1.46}$$

We remark that the second and third factors of the matrix $\hat{A}$ in (1.46) can be modified to obtain a nonsymmetric factorization

$$\begin{pmatrix} I & 0\,0 \\ 0 & I\,0 \\ A_{13}^T A_{11}^{-1} & 0\,I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} A_{11} & 0 & A_{13} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}. \tag{1.47}$$

The three fractional steps, which correspond to the symmetric variant of the Schwarz method for the case of two subregions, are given directly by the three factors of the matrix (1.47).

We have now set the stage for using standard techniques to estimate the condition number of the two subregion Schwarz method in terms of a generalized Rayleigh quotient involving the two matrices (1.44) and (1.46). Since these matrices have the same first and third factors, we might as well consider the generalized Rayleigh quotient

$$\frac{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}. \tag{1.48}$$

As noted in Sect. 1.6.1, it is easy to show, by elementary variational arguments, that $S^{(3)} \geq S^{(1)}$. Therefore the upper bound of the Rayleigh quotient is 1 and this bound is attained for $y_2 = 0$, $y_3 = 0$. The lower right two-by-two principal minors determine the lower bound on the spectrum of the relevant generalized eigenvalue problem. The relevant bound can then be obtained from the reduced generalized Rayleigh quotient

$$\frac{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}} . \tag{1.49}$$

It is important to note that the matrix of the denominator of formula (1.49) is the Schur complement obtained by eliminating the variables of $\Psi_3 = \Psi_1' \cap \Psi_2'$ from the stiffness matrix corresponding to all of $\Psi_2'$. It is equally important to note that the changes in the matrix (1.46) when we go from a symmetric multiplicative Schwarz method on the two subregions $\Omega_1'$ and $\Omega_2' \cup \Omega_3'$ to one on three subregions $\Omega_1'$, $\Omega_2'$, and $\Omega_3'$ are confined to the same lower right two-by-two principal minor since from our discussion of the recursive nature of the algorithm, we have learned that the exact solution on the subregion $\Omega_2' \cup \Omega_3'$ should be replaced by a symmetric multiplicative Schwarz step, with three fractional steps. In both cases, we will solve a problem with zero residuals in $\Psi_1'$ and discard the part of the solution which corresponds to the nodes in that region; the resulting matrices are thus Schur complements obtained by eliminating the variables in the subregion of $\Omega_2' \cup \Omega_3'$ that overlaps $\Omega_1'$. In the case of three subdomains, the relevant generalized Rayleigh quotient is of the form

$$\frac{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{23}^T & \tilde{A}_{33}^{(2)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}} , \tag{1.50}$$

where the elements of the matrix of the denominator is the Schur complement for the three-subdomain case discussed above; we will not need any detailed knowledge on the elements of this matrix.

This generalized Rayleigh quotient can be written as the product of the Rayleigh quotient in formula (1.49) and

$$\frac{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{23}^T & \tilde{A}_{33}^{(2)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}} . \tag{1.51}$$

The minimal and maximal values of the generalized Rayleigh quotient (1.49) can be estimated as in subsection 1.6.1. We also note that the minimal and

maximal values of the generalized Rayleigh quotient (1.50) can be bounded from below by the product of the minimal values of those of (1.49) and (1.51) and that an upper bound can be obtained similarly. Using the following lemma, we can estimate the extremal values of (1.51) by those for the two-subregion symmetric multiplicative Schwarz method on the subregions $\Omega_2'$ and $\Omega_3'$, i.e., again by using the result of subsection 1.6.1.

**Lemma 1.5** *Let two symmetric, positive definite matrices $A$ and $\tilde{A}$ be given with the same block structure,*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{12}^T & \tilde{A}_{22} \end{pmatrix},$$

*and assume that*

$$cA \leq \tilde{A} \leq CA.$$

*Then, their Schur complements, defined by*

$$S = A_{22} - {A_{12}}^T {A_{11}}^{-1} A_{12}$$

*and*

$$\tilde{S} = \tilde{A}_{22} - \tilde{A}_{12}^T \tilde{A}_{11}^{-1} \tilde{A}_{12},$$

*satisfy*

$$cS \leq \tilde{S} \leq CS.$$

*Proof.* We write the matrix $A$ as a sum of two positive semi-definite matrices,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & {A_{12}}^T {A_{11}}^{-1} A_{12} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & S \end{pmatrix}$$

and decompose $\tilde{A}$ in the same way. We note that the first of these matrices is semi-definite: this can be seen easily by carrying out a block Cholesky elimination, which results in a zero Schur complement. Rewriting the assumption of the lemma, using this decomposition, it follows that

$$cx_2{}^T S x_2 \leq c \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & {A_{12}}^T {A_{11}}^{-1} A_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + x_2{}^T S x_2 \right)$$

$$\leq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{12}^T & \tilde{A}_{12}^T \tilde{A}_{11}^{-1} \tilde{A}_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + x_2{}^T \tilde{S} x_2.$$

If we now select

$$x_1 = -\tilde{A}_{11}^{-1} \tilde{A}_{12} x_2,$$

then the first of the quadratic forms of the right hand side vanishes and we obtain one of the inequalities. The other follows in exactly the same way. □

We have now completed all that is required for a proof of the following theorem for the case of three subdomains.

**Theorem 1.6** *The condition number of the symmetric Schwarz operator satisfies,*

$$\kappa(\Omega_1', \Omega_2', \ldots, \Omega_k') \leq \kappa(\Omega_1', \Omega_2' \cup \ldots \cup \Omega_k')\, \kappa(\Omega_2', \Omega_3' \cup \ldots \cup \Omega_k') \ldots \kappa(\Omega_{k-1}', \Omega_k'). \tag{1.52}$$

*Here the expression on the left is the condition number of the operator* $I - E_k^* E_k$. *The first factor on the right hand side is the condition number of the symmetric multiplicative Schwarz operator for the pair of subregions* $\Omega_1'$ *and* $\Omega_2' \cup \Omega_3' \cup \ldots \cup \Omega_k'$, *etc.*

No new ideas are required for the proof of the general case. Thus, we find for the case of four subdomains, by a minor modification of the arguments above and by using the notation of the theorem, that

$$\kappa(\Omega_1', \Omega_2', \Omega_3', \Omega_4') \leq \kappa(\Omega_1', \Omega_2' \cup \Omega_3' \cup \Omega_4')\kappa(\Omega_2', \Omega_3', \Omega_4').$$

We can then use the result for three subdomains to replace the final factor of this estimate by a product of two factors that each only involve two subdomains.

# 2

# Abstract Theory of Schwarz Methods

## 2.1 Introduction

The theory which will be developed in this chapter provides a framework which has proven quite helpful in the design and analysis of a number of new and old iterative methods. It highlights the requirements necessary to turn large and often very ill-conditioned linear systems of algebraic equations into much better conditioned preconditioned systems. We note that success in such an effort can lead to spectral bounds which are independent of the dimension of the problem or which only deteriorate very slowly with decreasing mesh parameters. More precisely, we will look for iterative methods that are scalable (see Definition 1.3) and optimal (see Definition 1.2), whenever possible, or at least quasi optimal (see the discussion after Definition 1.3).

## 2.2 Schwarz Methods

We consider a finite dimensional Hilbert space $V$. Given a symmetric, positive definite bilinear form,

$$a(\cdot,\cdot): V \times V \longrightarrow \mathbb{R},$$

and an element $f \in V'$, we consider the problem of finding $u \in V$, such that

$$a(u,v) = f(v), \quad v \in V. \tag{2.1}$$

Given a basis of $V$, we recall that a function $u \in V$ is uniquely determined by a set of degrees of freedom. Here and in the following, we use the same notation for functional spaces and spaces of degrees of freedom, and functions and corresponding vectors of degrees of freedom. Similarly, we use the same notation for a linear functional $f \in V'$ and the corresponding vector, the load vector, with elements obtained by applying $f$ to the basis functions of $V$.

If $A$ is the stiffness matrix relative to the bilinear form $a(\cdot,\cdot)$ and the given basis, problem (2.1) is equivalent to the linear system

$$Au = f, \tag{2.2}$$

with $A$ symmetric, positive definite; see appendix B.

We next consider a family of spaces $\{V_i,\ i = 0, \ldots, N\}$ and suppose that there exist *interpolation* operators

$$R_i^T :\ V_i \longrightarrow V;$$

instead one often speaks of prolongation or extension operators. We assume that $V$ admits the following decomposition

$$V = R_0^T V_0 + \sum_{i=1}^{N} R_i^T V_i. \tag{2.3}$$

This decomposition will not necessarily be a direct sum of subspaces; in many cases, the representation of an element of $V$ in terms of components of the $V_i$ is not unique. We note that the $V_i$ do not need to be subspaces of $V$, but that, as is customary, we refer to them as 'subspaces' or 'local spaces' in the following. The subspace $V_0$ is usually related to a coarse problem, often built on a coarse mesh, while the remaining spaces are related to a partition into subdomains and are associated with local triangulations; this is the reason they are sometimes referred to as *local spaces*, as opposed to $V_0$ which is a *global space*.

We next introduce local symmetric, positive definite, bilinear forms on the subspaces,

$$\tilde{a}_i(\cdot,\cdot) :\ V_i \times V_i \longrightarrow \mathbb{R}, \quad i = 0, \ldots, N,$$

and the local stiffness matrices associated with them,

$$\tilde{A}_i :\ V_i \longrightarrow V_i.$$

In case we want to use the original bilinear form on the subspaces, we choose

$$\tilde{a}_i(u_i, v_i) = a(R_i^T u_i, R_i^T v_i), \quad u_i, v_i \in V_i \tag{2.4}$$

and find that

$$\tilde{A}_i = R_i A R_i^T. \tag{2.5}$$

In this case, we say that we use *exact local solvers.*

Schwarz operators are defined in terms of projection-like operators

$$P_i = R_i^T \tilde{P}_i :\ V \longrightarrow R_i^T V_i \subset V, \quad i = 0, \ldots, N,$$

where $\tilde{P}_i : V \to V_i$, is defined by

$$\tilde{a}_i(\tilde{P}_i u, v_i) = a(u, R_i^T v_i), \quad v_i \in V_i. \tag{2.6}$$

We note that $\tilde{P}_i$ is well defined since the local bilinear forms are coercive and that, in the case of exact solvers,

$$a(P_i u, R_i^T v_i) = a(u, R_i^T v_i), \quad v_i \in V_i.$$

We have the following lemma.

**Lemma 2.1** *The $P_i$ can be written as*

$$P_i = R_i^T \tilde{A}_i^{-1} R_i A, \quad 0 \leq i \leq N. \tag{2.7}$$

*In addition, the $P_i$ are selfadjoint with respect to the scalar product induced by $a(\cdot,\cdot)$ and positive semi-definite. If moreover the local bilinear form is given by (2.4), then $P_i$ is a projection, i.e.,*

$$P_i^2 = P_i.$$

*Proof.* We first consider the operator $\tilde{P}_i$ defined by (2.6). Writing (2.6) in matrix form, we find

$$v_i^T \tilde{A}_i \tilde{P}_i u = (R_i^T v_i)^T A u,$$

for all $u \in V$ and $v_i \in V_i$. We can then write

$$\tilde{A}_i \tilde{P}_i = R_i A,$$

which gives

$$\tilde{P}_i = \tilde{A}_i^{-1} R_i A$$

and (2.7).

In order to prove that $P_i$ is selfadjoint, let $u, v \in V$. Using (2.7), we easily find

$$a(P_i u, v) = v^T A (R_i^T \tilde{A}_i^{-1} R_i A u) = (R_i^T \tilde{A}_i^{-1} R_i A v)^T A u = a(u, P_i v).$$

The positive semi-definiteness of $P_i$ is a consequence of the coercivity of the local bilinear forms. Indeed, we have

$$a(P_i u, u) = u^T A P_i u = u^T A R_i^T \tilde{A}_i^{-1} R_i A u = v_i^T \tilde{A}_i^{-1} v_i \geq 0,$$

with $v_i = R_i A u$.

For the case of exact solvers, we can easily prove that $P_i$ is a projection by using (2.5). □

Once a set of subspaces and local bilinear forms are given, we can define a number of different Schwarz operators. Each of them is given by a polynomial in the operators $\{P_i\}$ without a zero order term. We note that we can always generate the appropriate right hand side for the resulting operator equation since we can compute $P_i u = R_i^T \tilde{P}_i u$, with $u$ the finite element solution, by noting that $\tilde{a}_i(\tilde{P}_i u, v_i) = a(u, R_i^T v_i) = f(R_i^T v_i)$. Similarly, we can find $P_j P_i u$, once that $P_i u$ has been computed, etc.

The additive operator is defined as previously by

$$P_{ad} = \sum_{i=0}^{N} P_i. \tag{2.8}$$

A multiplicative operator has already been given in Sect. 1.4,

$$P_{mu} = I - E_{mu}, \tag{2.9}$$

where the error propagation operator is defined by

$$E_{mu} = (I - P_N)(I - P_{N-1}) \cdots (I - P_0). \tag{2.10}$$

More general hybrid operators can also be defined. The operator

$$P_{hy1} = I - E_{hy1}, \quad E_{hy1} = (I - P_0)(I - \sum_{i=1}^{N} P_i)(I - P_0),$$

is additive with respect to the local components and multiplicative with respect to the levels. We note that in case an exact solver is employed on $V_0$, Lemma 2.1 ensures that

$$P_{hy1} = P_0 + (I - P_0) \sum_{i=1}^{N} P_i (I - P_0). \tag{2.11}$$

The operator $P_{hy1}$ is due to Mandel [334] and is considered in more detail in subsection 2.5.2; see also Lemma 2.15.

A hybrid method that is nonsymmetric and multiplicative on the local components and additive on the levels is given by

$$P_{hy2} = \alpha P_0 + (I - (I - P_N) \cdots (I - P_1)),$$

with $\alpha$ a relaxation parameter. It exploits the generally more rapid convergence of a multiplicative method, while the special coarse problem can be solved at the same time as the local problems. Thus, one or several processors can work on the coarse problem while the rest of the processors are assigned to the local problems. We note that in a standard multiplicative algorithm, there is a potential bottleneck with many processors idly waiting for the solution of the coarse problem. The operator $P_{hy2}$ is considered and analyzed in [102] for the more general case of a nonsymmetric problem.

We note that the operators $P_{mu}$ and $P_{hy2}$ are not symmetric. Symmetric operators can be obtained by visiting all but one of the subspaces twice. For the multiplicative algorithm, we have, e.g.,

$$E_{sy} = (I - P_0) \cdots (I - P_{N-1})(I - P_N)(I - P_{N-1}) \cdots (I - P_0).$$

We also note that if $P_0$ is a projection then we can save one application of $(I - P_0)$ per step starting with the second step of the iteration.

We end this section by noting that all the Schwarz operators introduced here are preconditioned operators for the original operator $A$ and can be written as the product of a suitable preconditioner and $A$, where the former only involves extensions $\{R_i^T\}$, restrictions $\{R_i\}$, local operators $\{\tilde{A}_i^{-1}\}$, and $A$. This is straightforward for the additive method where

$$P_{ad} = A_{ad}^{-1} A, \quad A_{ad}^{-1} = \sum_{i=0}^{N} R_i^T \tilde{A}_i^{-1} R_i.$$

As it will be shown in Sect. 2.6, it is possible to write simple routines that apply a Schwarz preconditioner itself to a vector even in the multiplicative and hybrid cases.

## 2.3 Convergence Theory

In this section, we consider the additive and the multiplicative Schwarz operators $P_{ad}$ and $P_{mu}$ and the corresponding preconditioned systems for the solution of (2.2).

The additive operator is symmetric and, as we will show in this section, it is also positive definite. We consider then the conjugate gradient algorithm for the solution of

$$P_{ad} u = g_{ad},$$

with $g_{ad} = A_{ad}^{-1} f$, and we will estimate the condition number of $P_{ad}$

$$\kappa(P_{ad}) = \frac{\lambda_{max}(P_{ad})}{\lambda_{min}(P_{ad})},$$

where

$$\lambda_{max}(P_{ad}) = \sup_{u \in V} \frac{a(P_{ad} u, u)}{a(u, u)}, \quad \lambda_{min}(P_{ad}) = \inf_{u \in V} \frac{a(P_{ad} u, u)}{a(u, u)};$$

see appendices B and C. Here, and in the rest of the monograph, we implicitly assume that, given a rational functional $\phi(u)$, $\sup_u \phi(u)$ and $\inf_u \phi(u)$ are taken for $u$ such that the denominator of $\phi(u)$ does not vanish.

The multiplicative operator $P_{mu}$ is not symmetric and we will consider a simple Richardson iteration applied to the corresponding preconditioned system; this is a generalization of the Schwarz alternating method presented in Sect. 1.4. In fact, we will give an upper bound for the norm of the error propagation operator

$$\|E_{mu}\|_a^2 = \sup_{u \in V} \frac{a(E_{mu} u, E_{mu} u)}{a(u, u)},$$

which is strictly less than one; cf. appendix C.

In order to prove bounds for these two Schwarz operators, it is enough to make three assumptions. We note that it is desirable that the parameters be shown to be independent of the dimension of the given problem or to deteriorate only slowly with decreasing mesh parameters.

**Assumption 2.2 (Stable Decomposition)** *There exists a constant $C_0$, such that every $u \in V$ admits a decomposition*

$$u = \sum_{i=0}^{N} R_i^T u_i, \quad \{u_i \in V_i,\ 0 \le i \le N\}$$

*that satisfies*

$$\sum_{i=0}^{N} \tilde{a}_i(u_i, u_i) \le C_0^2 a(u, u).$$

Assumption 2.2 ensures that a *stable splitting* can be found for the family of subspaces and the corresponding bilinear forms. It will allow us to find a strictly positive lower bound for $\lambda_{min}(P_{ad})$ and consequently to ensure that $P_{ad}$ is invertible; see Lemma 2.5 and the discussion about block-Jacobi preconditioners in Sect. 1.5. We remark that in order for Assumption 2.2 to be satisfied it is enough that the subspaces provide a decomposition of $V$. In other words, since $V$ is finite dimensional, the additive Schwarz operator is invertible if the subspaces are able to represent every function in $V$. However, even if we can find a $C_0$ for each space $V$, we typically consider an entire family of finite dimensional spaces and our main concern is to develop uniform bounds for $C_0$ or bounds that only grow slowly with the dimension of the problem. We finally remark that, if the subproblems are all scaled so that the norms of the $P_i$ equal one, a constant $C_0$ close to one is desirable, and consequently orthogonal subspaces would be best in the case of exact solvers.

**Assumption 2.3 (Strengthened Cauchy-Schwarz Inequalities)** *There exist constants $0 \le \epsilon_{ij} \le 1$, $1 \le i, j \le N$, such that*

$$|a(R_i^T u_i, R_j^T u_j)| \le \epsilon_{ij} a(R_i^T u_i, R_i^T u_i)^{1/2}\, a(R_j^T u_j, R_j^T u_j)^{1/2},$$

*for $u_i \in V_i$ and $u_j \in V_j$. We will denote the spectral radius of $\mathcal{E} = \{\epsilon_{ij}\}$ by $\rho(\mathcal{E})$.*

We note that Assumption 2.3 does not involve the space $V_0$. We also note that the theory easily could be modified to accommodate two or any fixed number of specially designated coarse subspaces.

The spectral radius $\rho(\mathcal{E})$ will appear in an upper bound for $\lambda_{max}(P_{ad})$; see Lemma 2.6. The inequalities of the Assumption 2.3 are trivially valid with $\epsilon_{ij} = 1$. However, in this case, we have $\rho(\mathcal{E}) = N$, which would give a poor upper bound. The best bound is obtained for orthogonal spaces $\{R_i^T V_i\}$ in which case $\epsilon_{ij} = 0$, for $i \neq j$ and $\rho(\mathcal{E})$ is equal to one.

**Assumption 2.4 (Local Stability)** *There exists $\omega > 0$, such that*

$$a(R_i^T u_i, R_i^T u_i) \le \omega\, \tilde{a}_i(u_i, u_i), \quad u_i \in range(\tilde{P}_i) \subset V_i, \quad 0 \le i \le N.$$

Assumption 2.4 ensures that the local bilinear forms are coercive and gives a one-sided measure of their approximation properties. We note that Assumption 2.2 ensures a weak type of continuity for the $\tilde{a}_i(\cdot,\cdot)$. The constant $\omega$ provides an upper bound for $\|P_i\|_a$; see Lemma 2.6. If exact, unscaled local solvers are considered, $\omega$ is equal to one. A possible choice for the local bilinear forms is

$$\tilde{a}_i(u_i,u_i) = \frac{1}{\alpha} a(R_i^T u_i, R_i^T u_i), \quad \alpha \in (0,2),$$

where $\alpha$ is a relaxation parameter.

The additive method, accelerated by the conjugate gradient method, is insensitive to a common scale factor, but the use of relaxation parameters is often desirable or even necessary for multiplicative and hybrid methods. We need to assume that $\omega \in (0,2)$ for the case of multiplicative methods. We are not free to scale the local bilinear forms arbitrarily in order to decrease $C_0$; a small value of $\omega$ means that corrections of the error are small. In such a case, $C_0$ of Assumption 2.2 will necessarily be large.

**Lemma 2.5** *Let Assumption 2.2 be satisfied. Then,*

$$a(P_{ad}u,u) \geq C_0^{-2} a(u,u), \quad u \in V, \tag{2.12}$$

*and consequently $P_{ad}$ is invertible. In addition,*

$$a(P_{ad}^{-1}u,u) = \min_{\substack{u_i \in V_i \\ u=\sum R_i^T u_i}} \sum_{i=0}^{N} \tilde{a}_i(u_i,u_i). \tag{2.13}$$

*Proof.* By Assumption 2.2 and the definition of the $P_i$, we have

$$\begin{aligned} a(u,u) &= \sum_{i=0}^{N} a(u, R_i^T u_i) = \sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i u, u_i) \\ &\leq \Big(\sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i u, \tilde{P}_i u)\Big)^{1/2} \Big(\sum_{i=0}^{N} \tilde{a}_i(u_i,u_i)\Big)^{1/2} \\ &\leq \Big(\sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i u, \tilde{P}_i u)\Big)^{1/2} C_0 (a(u,u))^{1/2}, \end{aligned}$$

Here we have used the Cauchy-Schwarz inequality in the $V_i$ and $\ell_2$, and the fact that the local bilinear forms are symmetric and positive definite. Squaring and cancelling a common factor and using the definition of the $\{\tilde{P}_i\}$, we find

$$a(u,u) \leq C_0^2 \sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i u, \tilde{P}_i u) = C_0^2 \sum_{i=0}^{N} a(u, R_i^T \tilde{P}_i u) = C_0^2 a(u, P_{ad} u),$$

which proves (2.12) and thus the invertibility of $P_{ad}$.

In order to prove (2.13), we first find a decomposition of $u \in V$, such that the equality between the energies in (2.13) holds. Since $P_{ad}$ is invertible, we can choose

$$u_i = \tilde{P}_i P_{ad}^{-1} u, \quad 0 \le i \le N. \tag{2.14}$$

We have

$$u = \sum_{i=0}^{N} R_i^T u_i$$

and

$$\begin{aligned}\sum_{i=0}^{N} \tilde{a}_i(u_i, u_i) &= \sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i P_{ad}^{-1} u, \tilde{P}_i P_{ad}^{-1} u) \\ &= \sum_{i=0}^{N} a(P_{ad}^{-1} u, R_i^T \tilde{P}_i P_{ad}^{-1} u) = a(P_{ad}^{-1} u, u).\end{aligned}$$

We next consider any decomposition

$$u = \sum_{i=0}^{N} R_i^T u_i.$$

We can write

$$\begin{aligned}a(P_{ad}^{-1} u, u) &= \sum_{i=0}^{N} a(P_{ad}^{-1} u, R_i^T u_i) = \sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i P_{ad}^{-1} u, u_i) \\ &\le \Big(\sum_{i=0}^{N} \tilde{a}_i(\tilde{P}_i P_{ad}^{-1} u, \tilde{P}_i P_{ad}^{-1} u)\Big)^{1/2} \Big(\sum_{i=0}^{N} \tilde{a}_i(u_i, u_i)\Big)^{1/2} \\ &= \Big(\sum_{i=0}^{N} a(P_{ad}^{-1} u, R_i^T \tilde{P}_i P_{ad}^{-1} u)\Big)^{1/2} \Big(\sum_{i=0}^{N} \tilde{a}_i(u_i, u_i)\Big)^{1/2} \\ &= a(P_{ad}^{-1} u, u)^{1/2} \Big(\sum_{i=0}^{N} \tilde{a}_i(u_i, u_i)\Big)^{1/2}.\end{aligned}$$

We have thus shown that for every decomposition of $u$, we have

$$a(P_{ad}^{-1} u, u) \le \sum_{i=0}^{N} \tilde{a}_i(u_i, u_i).$$

Therefore, the partition given in (2.14) provides the minimum and the proof is complete. □

The following lemma is also crucial in many cases.

**Lemma 2.6** *Let Assumptions 2.4 and 2.3 be satisfied. Then, for* $i = 0, \ldots, N$,

$$\|P_i\|_a \le \omega.$$

*In addition,*

$$a(P_{ad} u, u) \le \omega(\rho(\mathcal{E}) + 1)\, a(u, u).$$

*Proof.* We start with the first bound. Using Assumption 2.4, we find

$$\begin{aligned} a(P_i u, P_i u) &= a(R_i^T \tilde{P}_i u, R_i^T \tilde{P}_i u) \leq \omega \tilde{a}_i(\tilde{P}_i u, \tilde{P}_i u) \\ &= \omega a(u, R_i^T \tilde{P}_i u) = \omega a(u, P_i u), \end{aligned}$$

i.e.,

$$a(P_i u, P_i u) \leq \omega a(u, P_i u), \tag{2.15}$$

and thus

$$a(P_i u, P_i u) \leq \omega^2 a(u, u), \quad u \in V,$$

which gives the first inequality.

In order to prove the second inequality, we first consider the operator

$$\hat{P} = \sum_{i=1}^{N} P_i.$$

From Assumption 2.3, we have

$$a(\hat{P}u, \hat{P}u) = \sum_{1 \leq i,j \leq N} a(P_i u, P_j u) \leq \sum_{1 \leq i,j \leq N} \epsilon_{ij} a(P_i u, P_i u)^{1/2} a(P_j u, P_j u)^{1/2}.$$

Using (2.15) and the fact that the matrix $\mathcal{E}$ is symmetric, and that therefore its $\ell_2-$norm is equal to its spectral radius, we can write

$$\begin{aligned} a(\hat{P}u, \hat{P}u) &\leq \sum_{1 \leq i,j \leq N} \epsilon_{ij} \omega^{1/2} a(u, P_i u)^{1/2} \omega^{1/2} a(u, P_j u)^{1/2} \\ &\leq \omega \rho(\mathcal{E}) \sum_{1 \leq i \leq N} a(u, P_i u) \\ &= \omega \rho(\mathcal{E}) a(u, \hat{P}u) \leq \omega \rho(\mathcal{E}) a(u, u)^{1/2} \, a(\hat{P}u, \hat{P}u)^{1/2}. \end{aligned}$$

We then have

$$a(\hat{P}u, u) \leq a(\hat{P}u, \hat{P}u)^{1/2} \, a(u, u)^{1/2} \leq \omega \rho(\mathcal{E}) a(u, u). \tag{2.16}$$

Using this last estimate and the bound for $\|P_0\|_a$, we finally find

$$a(P_{ad}u, u) = a(P_0 u, u) + a(\hat{P}u, u) \leq \omega a(u, u) + \omega \rho(\mathcal{E}) a(u, u).$$

□

Combining Lemmas 2.5 and 2.6, we find a bound for the condition number of $P_{ad}$.

**Theorem 2.7** *Let Assumptions 2.2, 2.3, and 2.4 be satisfied. Then the condition number of the additive Schwarz operator satisfies*

$$\kappa(P_{ad}) \leq C_0^2 \omega(\rho(\mathcal{E}) + 1).$$

We now prove a bound for the multiplicative operator. We first need a technical lemma.

**Lemma 2.8** *Let Assumptions 2.4 and 2.3 be satisfied. Then, for* $0 \leq i,k \leq N$ *and for* $u,v \in V$,

$$a(P_i u, v) \leq a(P_i u, u)^{1/2}\, a(P_i v, v)^{1/2},$$
$$a(P_i u, P_k v) \leq \omega \epsilon_{ik}\, a(P_i u, u)^{1/2}\, a(P_k v, v)^{1/2}.$$

*Proof.* For the first inequality, we use the fact that $P_i$ is self-adjoint, the definition of $\tilde{P}_i$, and the Cauchy-Schwarz inequality for the bilinear form $\tilde{a}_i(\cdot,\cdot)$ :

$$a(P_i u, v) = a(u, P_i v) = a(u, R_i^T \tilde{P}_i v) = \tilde{a}_i(\tilde{P}_i u, \tilde{P}_i v)$$
$$\leq \tilde{a}_i(\tilde{P}_i u, \tilde{P}_i u)^{1/2}\, \tilde{a}_i(\tilde{P}_i v, \tilde{P}_i v)^{1/2} = a(u, P_i u)^{1/2}\, a(v, P_i v)^{1/2}.$$

The second expression can be bounded by using the Cauchy-Schwarz inequality and the definitions of the $\tilde{P}_j$, $\omega$, and $\mathcal{E}$:

$$a(P_j u, P_k v) \leq \epsilon_{jk} a(P_j u, P_j u)^{1/2}\, a(P_k v, P_k v)^{1/2}$$
$$\leq \omega \epsilon_{jk}\, \tilde{a}_j(\tilde{P}_j u, \tilde{P}_j u)^{1/2}\, \tilde{a}_k(\tilde{P}_k v, \tilde{P}_k v)^{1/2}$$
$$= \omega \epsilon_{jk}\, a(u, P_j u)^{1/2}\, a(v, P_k v)^{1/2}.$$

□

**Theorem 2.9** *Let Assumptions 2.2, 2.3, and 2.4 be satisfied. Let* $\omega$ *be the constant defined in Assumption 2.4 and suppose that* $\omega \in (0,2)$. *Then the error propagation operator of the multiplicative Schwarz method satisfies*

$$\|E_{mu}\|_a^2 = \|I - P_{mu}\|_a^2 \leq 1 - \frac{2-\omega}{(2\hat{\omega}^2 \rho(\mathcal{E})^2 + 1) C_0^2} < 1,$$

*where* $\hat{\omega} = \max(1,\omega)$ *and the constants* $C_0$ *and* $\rho(\mathcal{E})$ *are defined in Assumptions 2.2 and 2.3.*

*Proof.* Our task is to estimate the norm of the error propagation operator $E_N = E_{mu}$ of the multiplicative Schwarz method. We begin by defining

$$E_j = (I - P_j) \cdots (I - P_0), \quad 0 \leq j \leq N, \quad E_{-1} = I,$$

and

$$Q_j = 2P_j - P_j^2 = (2I - P_j) P_j, \qquad 0 \leq j \leq N.$$

We have

$$E_{j-1}^* E_{j-1} - E_j^* E_j = E_{j-1}^* Q_j E_{j-1}, \quad 0 \leq j \leq N.$$

Here and in what follows, the star $*$ denotes the adjoint with respect to the bilinear form $a(\cdot,\cdot)$. Summing over $j$, we find

$$I - E_N^* E_N = \sum_{j=0}^{N} E_{j-1}^* Q_j E_{j-1}. \tag{2.17}$$

We note that, for $\omega = \max \|P_j\|_a < 2$, the operators $Q_j$ are positive semi-definite since

$$Q_j \geq (2 - \omega) P_j$$

and the $P_j$ are positive semi-definite. Therefore,

$$I - E_N^* E_N \geq (2 - \omega) \sum_{j=0}^{N} E_{j-1}^* P_j E_{j-1}. \tag{2.18}$$

A satisfactory upper bound for $\|E_N\|_a$ is obtained by showing that the operator on the right hand side of (2.18) is sufficiently positive definite, i.e., by giving a lower bound for its smallest eigenvalue.

A direct consequence of the definition of the operator $E_j$ is that

$$I = E_{j-1} + \sum_{k=0}^{j-1} P_k E_{k-1} = E_{j-1} + P_0 + \sum_{k=1}^{j-1} P_k E_{k-1}. \tag{2.19}$$

For $j > 0$, we find

$$a(P_j u, u) = a(P_j u, E_{j-1} u) + a(P_j u, P_0 u) + \sum_{k=1}^{j-1} a(P_j u, P_k E_{k-1} u).$$

This expression can be bounded using Lemma 2.8:

$$\begin{aligned} a(P_j u, u) \leq\ & a(P_j u, u)^{1/2} \big( a(P_j E_{j-1} u, E_{j-1} u)^{1/2} \\ & + a(P_j P_0 u, P_0 u)^{1/2} + \omega \sum_{k=1}^{j-1} \epsilon_{jk} a(P_k E_{k-1} u, E_{k-1} u)^{1/2} \big). \end{aligned}$$

Let $c$ denote a vector with the components

$$c_k = a(P_k E_{k-1} u, E_{k-1} u)^{1/2}, \quad k = 1, \cdots, N.$$

We note that $\epsilon_{jj} = 1$ and that we can combine the first and third terms. Canceling a common factor and using elementary arguments, we find that

$$a(P_j u, u) \leq 2\hat{\omega}^2 (\mathcal{E} c)_j^2 + 2a(P_j P_0 u, P_0 u),$$

where $\hat{\omega} = \max(1, \omega)$. Summing for $j = 1, \ldots, N$, using formula (2.16) and Lemma 2.6, and adding the term $a(P_0 u, u)$ to both sides, we obtain

$$a(P_{ad} u, u) \leq 2\hat{\omega}^2 \rho(\mathcal{E})^2 \|c\|_{\ell^2}^2 + (1 + 2\omega^2 \rho(\mathcal{E})) a(P_0 u, u),$$

and finally,

$$a(P_{ad}u,u) \leq (1+2\hat{\omega}^2\rho(\mathcal{E})^2)\sum_{j=0}^{N} a(E_{j-1}^{*}P_jE_{j-1}u,u).$$

The proof can now be completed by using (2.18) and the lower bound of Lemma 2.5. □

## 2.4 Historical Remarks

Much of the development of the abstract Schwarz theory can be traced back to the work of Pierre-Louis Lions [318, 319, 320, 321] on Schwarz alternating methods, where the importance of certain projection operators into local spaces and of stable decompositions was pointed out. There was also pioneering work on additive Schwarz methods by Matsokin and Nepomnyaschikh [348, 360]. In particular, the importance of Assumption 2.2 was recognized by them; cf. also Lions [319].

The abstract theory as presented in this chapter is due in part to Dryja and Widlund [180, 181]. The development of the theory for multiplicative methods is due to Bramble, Pasciak, Wang and Xu [80]. The abstract theory is also outlined in Oswald [368, Sect. 4.1] and Xu [472]; cf. also Griebel [234] and Griebel and Oswald [235, 236]. Early work using these techniques for the analysis of multigrid methods are given by Xu in that paper and also by Bramble, Pasciak, Wang and Xu [79] and by Zhang [479, 481]. More recent work on the general theory is given by Xu and Zikatanov [475, 474].

The hybrid operator $P_{hy1}$ is due to Mandel [334], while $P_{hy2}$ was introduced by Cai [102].

More recently, an algebraic theory of Schwarz method has been developed by Frommer and Szyld [209, 210], Frommer and Schwandt [208], Benzi, Frommer, Nabben, and Szyld [44] and Nabben [356].

## 2.5 Additional Results

### 2.5.1 Coloring Techniques

We often wish to use many subspaces. To avoid a high degree of the polynomials, which define multiplicative and some hybrid Schwarz methods and which can make the algorithm quite sequential, we use a simple graph theory tool called *coloring*. Orthogonal subspaces, such as those corresponding to disjoint subregions, can be grouped together into classes of subspaces each of which can be regarded as one subspace. This is done by coloring an undirected graph, with a vertex for each subspace and an edge for each pair of subspaces

that are not orthogonal, by a minimal or good coloring. Subspaces of the same color can then be merged into one.

More precisely, given a decomposition of $V$, we color the subspaces $\{V_i,\ 1 \le i \le N\}$ in such a way that if two subspaces $V_k$ and $V_j$ have the same color they are orthogonal, i.e.,

$$a(R_k^T u_k, R_j^T u_j) = 0, \quad u_k \in V_k, \quad u_j \in V_j.$$

A coloring thus provides a partition of the set of subspaces $\mathcal{C} = \{\mathcal{C}_j,\ 1 \le j \le N^c\}$ into $N^c$ classes. We note that, if $V_k$ and $V_j$ have the same color, then the corresponding operators satisfy

$$P_k P_j = P_j P_k = 0.$$

Consequently, after a possible reordering of the subspaces, the error propagation operator of the multiplicative method can be written as

$$E_{mu} = \left(I - \sum_{i \in \mathcal{C}_{N^c}} P_i\right) \cdots \left(I - \sum_{i \in \mathcal{C}_1} P_i\right)(I - P_0),$$

reducing the number of sequential steps from $N+1$ to $N^c+1$.

The next lemma relates the spectral radius of $\mathcal{E}$ to the number of colors.

**Lemma 2.10** *Let $\mathcal{E} = \{\epsilon_{ij}\}$ be given as in Assumption 2.3. Suppose that are at most $N^c$ nonzeros in each row of $\mathcal{E}$. Then,*

$$\rho(\mathcal{E}) \le N^c.$$

*Proof.* This result follows by estimating the spectral radius of $\mathcal{E}$ by its $\ell_\infty$-norm and noting that $0 \le \epsilon_{ij} \le 1$. □

We note that upper bounds for one- and two-level methods are often found without using Assumption 2.3 and Lemma 2.10 by a direct coloring argument. In this case a somewhat weaker assumption can be employed on the coloring of the subspaces; see Sect. 3.6.

### 2.5.2 A Hybrid Method

We now consider the hybrid operator $P_{hy1}$ given in (2.11), which plays an important role in certain iterative substructuring methods of Neumann–Neumann type. This will also give us the opportunity of presenting an example of a *projected conjugate gradient* algorithm, which is also employed in the implementation of certain Neumann-Neumann and FETI methods; see Chap. 6. Throughout, we require that an exact solver be used on the coarse space $V_0$. Then, $P_0$ is an orthogonal projection. By defining

$$T_i = (I - P_0)P_i(I - P_0), \quad 1 \le i \le N,$$

we can write

$$P_{hy1} = P_0 + \sum_{i=1}^{N} T_i = A_{hy1}^{-1} A.$$

We will consider the solution of the preconditioned system

$$P_{hy1} u = g, \quad g = A_{hy1}^{-1} f,$$

in detail. By decomposing the exact solution $u$ as

$$u = u_0 + v, \quad u_0 = P_0 u, \quad v \in range(I - P_0), \tag{2.20}$$

we obtain the following operator equation for $v \in range(I - P_0)$

$$\sum_{i=1}^{N} T_i v = g - u_0, \tag{2.21}$$

with

$$\begin{aligned} \sum_{i=1}^{N} T_i &= (I - P_0) \sum_{i=1}^{N} P_i (I - P_0) \\ &= \left[ (I - P_0) \sum_{i=1}^{N} R_i^T \tilde{A}_i^{-1} R_i \, (I - P_0^T) \right] A. \end{aligned} \tag{2.22}$$

We note that $u_0$ can easily be computed by solving a coarse problem. Indeed, according to Lemma 2.1,

$$u_0 = P_0 A^{-1} f = R_0^T A_0^{-1} R_0 f.$$

This element is a suitable initial approximation for the conjugate gradient computation but it can equally well be replaced by $u_0 + w$ in (2.20), with $w$ an arbitrary vector in $range(I - P_0)$. We also note that the application of the preconditioner defined by the preconditioned operator given in (2.22) involves three operations, namely, the applications of $(I - P_0^T)$, $(I - P_0)$, and the local components $\sum_i R_i^T \tilde{A}_i^{-1} R_i$. We will soon show that, because of the choice of $u_0$, the first step can be omitted and consequently the action of $P_0^T$ is not required in practice.

We consider the conjugate gradient method applied to problem (2.21). We have written the algorithm in Fig. 2.1; here $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product. Because of the choice of the initial vector $u^0$, we can easily prove the following lemma.

**Lemma 2.11** *The following properties hold*

$$\begin{aligned} &w^k = q^k, \quad \textit{for } k \geq 0, \\ &u^j - u^0 \in range(I - P_0). \end{aligned}$$

1. Initialize

$$u^0 = R_0^T A_0^{-1} R_0 f + w, \quad w \in range(I - P_0)$$
$$q^0 = f - Au^0$$

2. Iterate $k = 1, 2, \cdots$ until convergence

$$\text{Project: } w^{k-1} = (I - P_0^T) q^{k-1}$$
$$\text{Precondition: } z^{k-1} = \sum_{i=1}^{N} R_i^T \tilde{A}_i^{-1} R_i w^{k-1}$$
$$\text{Project: } y^{k-1} = (I - P_0) z^{k-1}$$
$$\beta^k = \langle y^{k-1}, q^{k-1} \rangle / \langle y^{k-2}, q^{k-2} \rangle \quad [\beta^1 = 0]$$
$$p^k = y^{k-1} + \beta^k p^{k-1} \quad [p^1 = y^0]$$
$$\alpha^k = \langle y^{k-1}, q^{k-1} \rangle / \langle p^k, Ap^k \rangle$$
$$u^k = u^{k-1} + \alpha^k p^k$$
$$q^k = q^{k-1} - \alpha^k Ap^k$$

**Fig. 2.1.** Implementation of the projected preconditioned conjugate gradient algorithm involving the hybrid operator $P_{hy1}$.

We note that, because of Lemma 2.11, the action of $I - P_0^T$ need not be calculated in practice.

We next prove a bound for the condition number of the hybrid operator $P_{hy1}$. In order to do so, we employ Assumptions 2.4 and 2.3, but it is convenient to replace Assumption 2.2.

**Assumption 2.12 (Stable Decomposition)** *There exists a constant $C_0$, such that every $u \in range(I - P_0)$ admits a decomposition*

$$u = \sum_{i=1}^{N} R_i^T u_i, \quad \{u_i \in V_i,\ 1 \le i \le N\}$$

*that satisfies*

$$\sum_{i=1}^{N} \tilde{a}_i(u_i, u_i) \le C_0^2 a(u, u).$$

Assumptions 2.12 and 2.2 are essentially the same. By specializing to elements in $range(I - P_0)$, Assumption 2.2 implies 2.12. We also see that

$$w = P_0 w + \sum_{i=1}^{N} R_i^T u_i$$

provides a good decomposition for an arbitrary $w \in V$, with $u = (I - P_0)w$, in the case when an exact solver is used for the coarse space and $P_0$ is thus an orthogonal projection.

We have the following theorem; see also [334].

**Theorem 2.13** *Let Assumptions 2.12, 2.3, and 2.4 be satisfied. Then the condition number of the hybrid operator satisfies*

$$\max\{1, C_0^2\}^{-1}\, a(u,u) \leq a(P_{hy1}u, u) \leq \max\{1, \omega\rho(\mathcal{E})\}\, a(u,u), \quad u \in V.$$

*The projected additive operator satisfies*

$$C_0^{-2}\, a(u,u) \leq a\left(\sum_{i=1}^{N} T_i u, u\right) \leq \omega\rho(\mathcal{E})\, a(u,u), \quad u \in range(I - P_0).$$

*Proof.* We first give a bound for the largest eigenvalue of $P_{hy1}$. By proceeding as in the proof of Lemma 2.6, we can prove

$$a\left(\sum_{i=1}^{N} P_i u, u\right) \leq \omega\rho(\mathcal{E}) a(u,u), \quad u \in range(I - P_0), \tag{2.23}$$

and thus, since $P_0$ is selfadjoint,

$$a\left((I - P_0)\sum_{i=1}^{N} P_i (I - P_0)u, u\right) \leq \omega\rho(\mathcal{E})\, a\left((I - P_0)u, (I - P_0)u\right), \quad u \in V.$$

Since $P_0$ is an orthogonal projection, we have for $u \in V$

$$\begin{aligned} &a(P_{hy1}u, u) \\ &= a(P_0 u, u) + a\left((I - P_0)\sum_{i=1}^{N} P_i (I - P_0)u, u\right) \\ &\leq a(P_0 u, P_0 u) + \omega\rho(\mathcal{E})\, a\left((I - P_0)u, (I - P_0)u\right) \\ &\leq \max\{1, \omega\rho(\mathcal{E})\}\, a(u,u). \end{aligned} \tag{2.24}$$

We next prove a lower bound. By proceeding as in the proof of Lemma 2.5, we find for $u \in range(I - P_0)$ that

$$a(u,u) \leq C_0^2\, a\left(\sum_{i=1}^{N} P_i u, u\right), \tag{2.25}$$

and thus

$$a\left((I - P_0)u, (I - P_0)u\right) \leq C_0^2\, a\left((I - P_0)\sum_{i=1}^{N} P_i (I - P_0)u, u\right), \quad u \in V.$$

As before, we easily obtain

$$\begin{aligned}
&\max\{1, C_0^2\} a(P_{hy1}u, u) \\
&\geq a(P_0u, P_0u) + C_0^2 a\left((I - P_0)\sum_{i=1}^{N} P_i(I - P_0)u, u\right) \\
&\geq a(P_0u, P_0u) + a\left((I - P_0)u, (I - P_0)u\right) = a(u, u).
\end{aligned} \tag{2.26}$$

The bound for $\sum_i T_i$ is a consequence of (2.23) and (2.25). □

### 2.5.3 Comparison Results

The following trivial result ensures that whenever a subspace with a corresponding (properly scaled) local solver is added to a multiplicative method, convergence cannot deteriorate in the positive definite, symmetric case. This is consistent with the idea that introducing an additional correction of the error can only improve the convergence. We remark that this property may no longer be true for nonsymmetric or indefinite problems.

**Lemma 2.14** *Let $E_N$ be the error propagation operator of a multiplicative Schwarz method. Suppose that an additional subspace $V_{N+1}$ is employed with a local solver satisfying Assumption 2.4 with $\omega \in (0, 2)$. Then the error operator $E_{N+1}$ of the modified method satisfies*

$$\|E_{N+1}\|_a \leq \|E_N\|_a.$$

*Proof.* The proof is trivial and results from the fact that, since from Lemma 2.6,

$$\|P_{N+1}\|_a \leq \omega < 2,$$

and $P_{N+1}$ is positive semi-definite, we have

$$\|I - P_{N+1}\|_a \leq 1.$$

□

The following lemma compares the condition number of the additive and the hybrid operators; see [334].

**Lemma 2.15** *In case an exact solver is employed on the coarse space $V_0$, the condition numbers of the additive operator and the projected additive operator defined in the previous section built from the same subspaces and the same local solvers satisfy*

$$\kappa\left(\sum_{i=1}^{N} \tilde{T}_i\right) \leq \kappa(P_{ad}),$$

*where $\tilde{T}_i$ is the restriction of $T_i$ to range$(I - P_0)$.*

*Proof.* The proof is immediate and follows from the extremal properties of the eigenvalues and the following equality

$$\frac{a\left(\sum_{i=1}^{N} T_i u, u\right)}{a(u,u)} = \frac{a\left(\sum_{i=1}^{N} P_i u, u\right)}{a(u,u)} = \frac{a(P_{ad}u, u)}{a(u,u)}, \quad u \in range(I - P_0) \subset V.$$

□

## 2.6 Remarks on the Implementation

As already pointed out in Sect. 2.2, all the Schwarz operators introduced in this chapter are preconditioned operators for the original operator $A$ and can be written as the product of a suitable preconditioner and $A$. The former only involve extensions $\{R_i^T\}$, restrictions $\{R_i\}$, local operators $\{\tilde{A}_i^{-1}\}$, and $A$.

A program implementing an accelerator, like Richardson, conjugate gradients, or generalized minimal residuals, usually requires a routine that applies a preconditioner to a generic vector; see appendix C. This is straightforward for the additive method, but not so obvious for multiplicative and hybrid methods. We recall that

$$P_{mu} = I - (I - P_N) \cdots (I - P_0) = A_{mu}^{-1} A.$$

The following function applies the multiplicative Schwarz preconditioner $A_{mu}^{-1}$ to a vector $x$. Similar routines can be written for the hybrid preconditioners.

y=$P_{mu} A^{-1}(x)$

$y := R_0^T \tilde{A}_0^{-1} R_0 x$

`for` $i = 1, \ldots, N$

$y := y + R_i^T \tilde{A}_i^{-1} R_i (x - Ay)$

`end`

We note that while the application of the additive preconditioner does not involve any product with the original matrix $A$, the application of *each* local component of the multiplicative preconditioner involves a product with $A$. We can reduce the cost by noting that we often do not need all the components of the residual vector in each step and also that the residual often can be updated only locally where a correction has actually just occurred. This will complicate the algorithm and if the application of the original matrix $A$ to a vector is costly, then fully multiplicative preconditioners might not be competitive with additive or hybrid methods. We note that when a Schur complement system is solved (see Chap. 4, 5, and 6) each application of the

Schur complement to a vector involves the solution of a Dirichlet problem on the subdomains of a nonoverlapping partition, which makes a multiplicative preconditioner potentially expensive. However, a coloring technique as described in subsection 2.5.1 reduces the number of sequential steps and can greatly reduce the number of local Dirichlet problems that need to be solved.

# 3

# Two-Level Overlapping Methods

## 3.1 Introduction

In this chapter, we consider some quite general two-level overlapping methods. They are generalizations of the Schwarz alternating method which was first introduced in Sect. 1.4 for two subdomains. These methods are given in terms of overlapping partitions of $\Omega$ into subdomains $\Omega_i'$ which themselves are unions of finite elements and have diameters of order $H_i$, and by a coarse shape-regular mesh. Any element of this coarse mesh should have a diameter on the order of $H_i$ if it intersects $\Omega_i'$. This coarse mesh can otherwise be completely independent of the fine triangulation; towards the end of Sect. 3.6, we will outline some particular design choices employed in practice. We also briefly introduce the restricted Schwarz methods and, in the last section, we discuss alternative coarse solvers, in particular, those defined by aggregation.

We will consider additive methods only, but note that any proof based on Theorem 2.7 also immediately provides proofs for multiplicative and hybrid methods based on Theorems 2.9 and 2.13.

One level additive overlapping preconditioners were originally introduced by Matsokin and Nepomnyaschikh [348] and Nepomnyaschikh [360]. They were further developed in the more powerful two-level form in [179, 174, 180, 181]; these methods have been used extensively. For related work, see also Bjørstad, Moe, and Skogen [54, 59, 60], Cai [99, 100, 101], Mathew [345, 346], Skogen [419], and Zhang [479, 480]. The basic theory presented, in Sect. 3.2-3.6, was originally given in [182] for the case of nested meshes and in [134] for more general coarse meshes.

For a given bounded polygonal or polyhedral domain $\Omega$, we consider the Poisson problem (1.1) with homogeneous Dirichlet boundary conditions. As in subsection 1.4.2, the variational formulation of the problem involves the bilinear form

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx, \quad u, v \in H_0^1(\Omega).$$

We will consider a shape-regular triangulation $\mathcal{T}$ of $\Omega$ and the space $V$ of continuous, piecewise linear functions on $\mathcal{T}$ that vanish on $\partial\Omega$; see appendix B. We note that we do not assume that $\mathcal{T}$ is quasi uniform. $A$ denotes the corresponding stiffness matrix.

We study additive preconditioned operators of the form

$$P_{ad} = \sum_{i=0}^{N} P_i = \sum_{i=0}^{N} R_i^T A_i^{-1} R_i A, \tag{3.1}$$

see Sect. 2.2, where $P_0$ is associated with a problem on the coarse mesh and the remaining $P_i$ with local problems on the subdomains $\{\Omega_i'\}$. These two components will be introduced in the next two sections.

## 3.2 Local Solvers

We first consider an initial partition of $\Omega$ into nonoverlapping subdomains (substructures) $\{\Omega_i,\ 1 \leq i \leq N\}$ with diameters $H_i$. We will always regard the subregions of any partition as open sets and assume that they are shape regular. The maximum diameter of the substructures is $H$. We now extend each substructure $\Omega_i$ to a larger region $\Omega_i'$, such that $\partial\Omega_i'$ does not cut through any fine elements. This can be done by repeatedly adding a layer of elements. We carry out the same construction for the substructures that meet the boundary except that we do not add any elements outside of $\Omega$. The fine mesh $\mathcal{T}$ gives rise to $N$ local meshes $\{\mathcal{T}_i\}$ on the subdomains $\Omega_i'$.

We make the following assumptions; see also Assumption 3.15.

**Assumption 3.1** *For $i = 1, \ldots, N$, there exists $\delta_i > 0$, such that, if $x$ belongs to $\Omega_i'$, then*

$$\operatorname{dist}(x, \partial\Omega_j' \setminus \partial\Omega) \geq \delta_i,$$

*for a suitable $j = j(x)$, possibly equal to $i$, with $x \in \Omega_j'$. The maximum of the ratios $H_i/\delta_i$ is denoted by*

$$\frac{H}{\delta} = \max_{1 \leq i \leq N} \left\{ \frac{H_i}{\delta_i} \right\}.$$

Here, the distance parameters $\delta_i$ measure the width of the regions $\Omega_i' \setminus \Omega_i$. We note that the special case where there exists a constant $c > 0$, independent of $i$, such that $\delta_i \geq cH_i$, is referred to as *generous overlap*. In this case,

$$\frac{H}{\delta} \leq \frac{1}{c}.$$

To simplify the notations, we will, in what follows, also refer to the diameter of the extended $\Omega_i'$ as $H_i$.

We also need to make some hypotheses on the intersections between the subdomains.

**Assumption 3.2 (Finite Covering)** *The partition $\{\Omega_i'\}$ can be colored using at most $N^c$ colors, in such a way that subregions with the same color are disjoint.*

*Remark 3.3.* We note that Assumptions 3.1 and 3.2 only involve the overlapping partition $\{\Omega_i'\}$ and do not require that the latter be obtained from a nonoverlapping one. In addition, one can prove that if $x \in \Omega$, then it belongs to at most $N^c$ subdomains in $\{\Omega_i'\}$. One can easily find partitions for which the converse is not true; see also Lemma 3.11.

The following lemma associates a family of functions with the overlapping partition. We refer to [93, Lem. 2.4] for a proof in the case of generous overlap.

**Lemma 3.4 (Partition of Unity)** *Let $\{\Omega_i'\}$ be an overlapping partition satisfying Assumptions 3.1 and 3.2. Then, there exists a family of functions in $W^{1,\infty}(\Omega)$, $\{\tilde{\theta}_i,\ 1 \le i \le N\}$, such that*

$$\begin{aligned} &0 \le \tilde{\theta}_i(x) \le 1, \quad x \in \bar{\Omega}, \\ &\operatorname{supp}(\tilde{\theta}_i) \subset \overline{\Omega_i'}, \\ &\sum_{i=1}^{N} \tilde{\theta}_i(x) = 1, \quad x \in \bar{\Omega}, \end{aligned} \tag{3.2}$$

*and*

$$\|\nabla \tilde{\theta}_i\|_\infty \le C/\delta_i, \quad 1 \le i \le N, \tag{3.3}$$

*where $C$ is a constant independent of the $\delta_i$ and the $H_i$.*

*Proof.* For $i = 1, \ldots, N$, we define

$$d_i(x) = \begin{cases} \operatorname{dist}(x, \partial\Omega_i' \setminus \partial\Omega), & x \in \Omega_i', \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$\tilde{\theta}_i(x) = \frac{d_i(x)}{\sum_{k=1}^{N} d_k(x)}.$$

A simple inspection shows that these functions are well defined and continuous in $\overline{\Omega}$. In addition conditions (3.2) hold.

We now consider an arbitrary but fixed $i$ and $x \in \Omega_i'$. Since $\tilde{\theta}_i$ is continuous in $\overline{\Omega}$ and is identically zero in $\Omega \setminus \Omega_i'$, in order to prove (3.3), it is enough to show that $|\nabla \tilde{\theta}_i(x)| \le C/\delta_i$, or, equivalently

$$|\tilde{\theta}_i(x) - \tilde{\theta}_i(y)| \le (C/\delta_i)|x - y|,$$

for $y \in \Omega_i'$, sufficiently close to $x$.

We first show that, with $\delta_k(x,y) := d_k(x) - d_k(y)$, $k = 1, \ldots, N$, and $y$ sufficiently close to $x$

$$|\delta_k(x,y)| \le |x - y|. \tag{3.4}$$

If $x \in \Omega_k'$, we can then choose $y \in \Omega_k'$, and, if $z \in \partial\Omega_k' \setminus \partial\Omega$ is such that

$$\operatorname{dist}(y, \partial\Omega_k' \setminus \partial\Omega) = |y - z|,$$

then, using the triangle inequality, we find

$$d_k(x) \leq |x - z| \leq |x - y| + |y - z| = |x - y| + d_k(y),$$

or $\delta_k(x, y) \leq |x - y|$. Exchanging the role of $x$ and $y$, we obtain (3.4).

If $x \in \Omega_i' \setminus \overline{\Omega_k'}$, for an $i \neq k$, then we can choose $y \in \Omega_i' \setminus \overline{\Omega_k'}$ and (3.4) trivially holds. Finally, in case $x \in \partial\Omega_k'$, we can prove (3.4) by combining the last two arguments. This concludes the proof of (3.4).

We next note that

$$\sum_{k=1}^{N} d_k(x) \geq d_j(x) \geq \delta_i, \tag{3.5}$$

with $j = j(x)$ the index in Assumption 3.1.

We set

$$\delta_{r,i}(x, y) = \sum_{\substack{k=1 \\ k \neq i}}^{N} \delta_k(x, y),$$

$$\tilde{\theta}_{r,i}(x) = 1 - \tilde{\theta}_i(x) = \frac{\displaystyle\sum_{\substack{k=1 \\ k \neq i}}^{N} d_k(x)}{\displaystyle\sum_{k=1}^{N} d_k(x)} \in [0, 1],$$

and note that it follows from Assumption 3.2 and (3.4) that

$$|\delta_{r,i}(x, y)| = \Big|\sum_{\substack{k=1 \\ k \neq i}}^{N} \delta_k(x, y)\Big| \leq C\,|x - y|, \tag{3.6}$$

with a constant $C$ that depends on $N^c$.

Using (3.4), (3.5), and simple algebra, we obtain

$$\begin{aligned} |\tilde{\theta}_i(x) - \tilde{\theta}_i(y)| &= \frac{1}{\displaystyle\sum_{k=1}^{N} d_k(y)} \left|\tilde{\theta}_i(x)\delta_{r,i}(y, x) + \tilde{\theta}_{r,i}(x)\delta_i(x, y)\right| \\ &\leq (C/\delta_i)|x - y|, \end{aligned}$$

for $x \in \Omega_i'$ and $y$ sufficiently close to $x$. □

In the following, we will employ a modified partition of unity, obtained from that of Lemma 3.4 by interpolating the local functions on the fine mesh. We define

$$\theta_i = I^h(\tilde{\theta}_i), \quad 1 \leq i \leq N, \tag{3.7}$$

where $I^h$ is the nodal piecewise linear interpolant on the fine mesh. We note that since

$$\|\nabla I^h(\tilde{\theta}_i)\|_\infty \leq \|\nabla \tilde{\theta}_i\|_\infty,$$

the family $\{\theta_i\}$ also satisfies (3.2) and (3.3).

The local spaces of our Schwarz methods are the finite element spaces of functions that are piecewise linear on the local meshes and vanish on the subdomain boundaries:

$$V_i = \{u \in H_0^1(\Omega_i') | \quad u_{|K} \in \mathbb{P}_1, \ K \in \mathcal{T}_i\}, \quad 1 \leq i \leq N.$$

We note that these subspaces are contained in the original space $V$, in the sense that local functions extended by zero to the whole of $\Omega$ belong to $V$. Let

$$R_i^T : V_i \longrightarrow V, \quad 1 \leq i \leq N,$$

represent these zero extensions.

## 3.3 A Coarse Problem

We introduce a shape-regular coarse mesh $\mathcal{T}_H$ on the domain $\Omega$ and the finite element space $V_0 = V_H$ of continuous, piecewise linear functions on $\mathcal{T}_H$, which vanish on $\partial\Omega$. Let $H_K$ be the diameter of $K \in \mathcal{T}_H$. We stress that the fine mesh $\mathcal{T}$ need not be a refinement of $\mathcal{T}_H$. Bilinear or trilinear finite element spaces could also be considered on quadrilateral or hexagonal coarse meshes.

We define an interpolation operator

$$R_0^T : V_0 \longrightarrow V,$$

obtained by interpolating the coarse functions onto the fine mesh. For $u \in V_0$, we define

$$R_0^T u = I^h u.$$

The coarse mesh does not need to be quasi uniform, but a coarse element should not be large in comparison to the subdomains which it intersects. We formulate this condition as an assumption.

**Assumption 3.5** *There exists a constant $C$ independent of $\mathcal{T}_H$ and the $\Omega_i'$, such that, for $i = 1, \ldots, N$,*

$$H_K \leq C H_i,$$

*for any $K \in \mathcal{T}_H$, such that $K \cap \Omega_i' \neq \emptyset$. Here $H_K$ is the diameter of the coarse element $K$ and $H_i$ the diameter of $\Omega_i'$.*

We note that an even more general situation can be considered, such as when the union of the coarse elements does not coincide with $\Omega$. In this case, we need to require that any part of $\partial\Omega$ where Neumann boundary conditions are imposed be contained in the union of the coarse elements. In addition, we

must assume that no coarse element lies completely outside $\Omega$ and that $\mathcal{T}_H$ covers a significant part of $\Omega$. We do not consider this more general situation here but refer to [134] for a complete analysis. We only point out that this more general situation is quite important in practice, e.g., where the domain $\Omega$ is not a polyhedron and if isoparametric elements are employed.

## 3.4 Scaling and Quotient Space Arguments

Two tools that are routinely employed in the analysis of domain decomposition algorithms and that will be used throughout this monograph are *scaling* and *quotient space* arguments. We present them here, where they are about to be used for the first time in this monograph.

Scaling arguments rely on the fact that Sobolev norms are expressed as integrals over a region. Whenever a bound involving certain Sobolev norms is available, it is therefore possible to find an explicit dependence of the constants on the diameter of the region. We show this with an example. Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded Lipschitz domain of diameter $H < 1$. Suppose that there are two functions in $u, v \in H^1(\mathcal{D})$ that satisfy the bound

$$|v|^2_{H^1(\mathcal{D})} \leq C \|u\|^2_{H^1(\mathcal{D})},$$

where $C$ is independent of $u$ and $v$ but may depend on $\mathcal{D}$. Let $x = H\hat{x}$ be the transformation that maps the domain $\hat{\mathcal{D}}$, with the same shape as $\mathcal{D}$ and with unit diameter, into $\mathcal{D}$. The bound for $u$ and $v$ is certainly valid for $\hat{\mathcal{D}}$, with a constant $\hat{C}$ that only depends on $\hat{\mathcal{D}}$ and thus only on the *shape* of $\mathcal{D}$. We can therefore write

$$|v|^2_{H^1(\mathcal{D})} = \int_{\mathcal{D}} |\nabla v|^2 \, dx = \int_{\hat{\mathcal{D}}} |\hat{\nabla}\hat{v}|^2 H^{-2} \, H^n d\hat{x} \leq \hat{C} H^{n-2} \|\hat{u}\|^2_{H^1(\hat{\mathcal{D}})},$$

where $\hat{v}(\hat{x}) = v(H\hat{x})$ and $\hat{u}(\hat{x}) = u(H\hat{x})$ are the transformed functions. Changing variables a second time, we obtain

$$\|\hat{u}\|^2_{H^1(\hat{\mathcal{D}})} = \int_{\mathcal{D}} |\nabla u|^2 H^2 \, H^{-n} dx + \int_{\mathcal{D}} |u|^2 \, H^{-n} dx,$$

and finally

$$|v|^2_{H^1(\mathcal{D})} \leq \hat{C} \left( \int_{\mathcal{D}} |\nabla u|^2 \, dx + \frac{1}{H^2} \int_{\mathcal{D}} |u|^2 \, dx \right) \leq \frac{\hat{C}}{H^2} \|u\|^2_{H^1(\mathcal{D})}. \tag{3.8}$$

We note the factors $H^{-n}$ and $H^n$, arising from Jacobians, cancel because the variables are changed twice. However, additional factors, involving $H$, remain because of the different scaling of the different terms in the full norm.

Bounds as in (3.8) are commonly taken over single subdomains and therefore involve $H$, the typical diameter of the subdomains. Since the number of

subdomains can be measured by $(1/H)^n$, the presence of $1/H$ factors would then give bounds that depend on the number of subdomains and are not acceptable.

We note that the presence of the $1/H^2$ factor in (3.8) stems from the fact that different terms in Sobolev norms scale differently under dilation; the $L^2$ term in the full $H^1$ norm brings in the offending factor. This leads to the need of bounding the $L^2$-*norm* with the appropriate *seminorm* ($H^1$, in this case). Such bounds are not valid for arbitrary functions since seminorms vanish for certain special nonzero functions. This is reviewed in appendix A.4. However, if $u \in H^1(\mathcal{D})$ vanishes on part of $\partial\mathcal{D}$, a *Friedrichs* inequality holds:

$$\|u\|^2_{L^2(\mathcal{D})} \leq \hat{C}_2 H^2 |u|^2_{H^1(\mathcal{D})}; \tag{3.9}$$

cf. Corollary A.15. We note that this factor $H^2$ also arises from a scaling argument. We also note that the inequality (3.9) can be viewed as an estimate of the eigenvalues of the Laplace operator on the region and with a homogeneous Dirichlet condition on part of the boundary. Combining the bound (3.8) with the last inequality, we find

$$|v|^2_{H^1(\mathcal{D})} \leq \hat{C}(1 + \hat{C}_2)|u|^2_{H^1(\mathcal{D})}, \tag{3.10}$$

where the constants involved are now independent of $H$. An analogous bound can be found if $u$ has a zero mean value over $\mathcal{D}$. In this case, the bound (3.9) still holds and it is referred to as a *Poincaré* inequality; see Corollary A.15. We also note that the inequality (3.9) can be viewed as an estimate of all the eigenvalues of the Laplace operator on the region except for the smallest.

For a subdomain with part of its boundary equipped with a Dirichlet condition, we can use a Friedrichs inequality. For the others, we need an argument involving a suitable coarse space, which is able to represent at least constant functions. We note that typically, decompositions, as in Assumption 2.2, for finite element functions are carried out after removing a coarse component $P_H u$. For the particular case of the overlapping algorithms of this chapter, we typically need to bound

$$v = \theta(u - P_H u) =: Bu,$$

with $\theta$ a finite element function supported on the region $\mathcal{D}$; see the proof of Lemma 3.12. It is clear then that, if $P_H u$ is *able to reproduce constants*, then $Bu$ is invariant if a constant is added to $u$:

$$B(u + c) = Bu, \qquad c \text{ constant.}$$

The bound (3.8) then gives, for any constant $c$,

$$\begin{aligned} |v|^2_{H^1(\mathcal{D})} &= |Bu|^2_{H^1(\mathcal{D})} = |B(u + c)|^2_{H^1(\mathcal{D})} \\ &\leq \hat{C}\left(\int_{\mathcal{D}} |\nabla u|^2 \, dx + \frac{1}{H^2}\int_{\mathcal{D}} |u + c|^2 \, dx\right), \end{aligned}$$

and bound (3.10) can also be recovered in this case by choosing $-c$ as the mean value of $u$ over $\mathcal{D}$. We refer in particular to Theorem A.18 in the appendix. The argument presented is usually referred to as *quotient space* argument.

We note that a similar procedure of removing a coarse component from a finite element function is employed for the additive, iterative substructuring methods in Sect. 5.4. For the Neumann-Neumann and FETI methods in Chap. 6, a quotient space argument is made possible by the fact that hybrid preconditioners are employed; see, e.g., Sect. 2.5.2. There, local solvers are applied to suitable subspaces for which certain averages vanish; see in particular Sect. 6.2, 6.3, and 6.4.2, and the proofs of Lemmas 6.2, 6.17, 6.34, and 6.36.

We finally remark that bounds that are independent of the size of the subregions $\mathcal{D}$ are not necessarily good, since factors depending on the shape of the subdomains may produce large constants. This may happen for certain domain decomposition methods if the subdomains fail to be shape regular. Most, but not all, of the analysis in this monograph is carried out for subregions that are cubes or the union of a few regular tetrahedra and the results remain equally valid if the subregions are images of a reference cube, etc., under sufficiently benign mappings, which effectively means that their aspect ratios remain uniformly bounded.

Similarly, for certain iterative substructuring methods, shape-regular substructures with very irregular boundaries may give rise to large stability constants; cf. Remark 4.5. For overlapping methods, on the other hand, there is no real degradation in the bounds.

## 3.5 Technical Tools

We first define a quasi-interpolant $\tilde{I}^H : H_0^1(\Omega) \to V^H$, which is stable under the assumption that the coarse mesh elements are shape regular, but not necessarily all of comparable diameters. Let $y$ be a node of the coarse mesh $\mathcal{T}_H$, and $\omega_y$ be the union of the elements in $\mathcal{T}_H$ that share $y$. For every vertex $y$, we set

$$(\tilde{I}^H u)(y) = \begin{cases} 0, & y \in \partial\Omega, \\ |\omega_y|^{-1} \displaystyle\int_{\omega_y} u(x)\, dx, & \text{otherwise}, \end{cases}$$

where $|\omega_y|$ denotes the measure of $\omega_y$.

Given an element $K \in \mathcal{T}_H$, we define $\omega_K$ as the union of $K$ and its neighboring elements

$$\overline{\omega}_K = \bigcup_{\substack{K' \in \mathcal{T}_H \\ \overline{K'} \cap \overline{K} \neq \emptyset}} \overline{K'}.$$

We only consider two cases: either $\omega_K$ is in the interior of $\Omega$ or there is at least one full face of a coarse element of $\omega_K$ which is part of $\partial\Omega$. In case $\partial\omega_K$

intersects $\partial\Omega$ in one point or along an edge, we can always add one or few elements in such a way that $\omega_K$ shares a full face with the outside boundary. In a more general case where $\tilde{K}$ is the union of coarse elements, we define $\omega_{\tilde{K}}$ as the union of the $\omega_K$ for which $K \subset \tilde{K}$.

We have the following stability result. We refer to [139], [136, Ex. 3.2.3], [458, Lem. 1.4] for similar operators that do not preserve zero values on the boundary. See also Lemma 3.16 for a different interpolant and a similar proof.

**Lemma 3.6** *Let $\mathcal{T}_H$ be shape regular and let $u \in H_0^1(\Omega)$. Then, there exists a constant $C$ such that*

$$\|u - \tilde{I}^H u\|_{L^2(K)} \leq C H_K |u|_{H^1(\omega_K)}, \tag{3.11}$$

$$|\tilde{I}^H u|_{H^1(K)} \leq C |u|_{H^1(\omega_K)}. \tag{3.12}$$

*Proof.* We only consider the three-dimensional case here. We consider a coarse tetrahedron $K \in \mathcal{T}_H$ and first assume that $\omega_K$ does not touch the boundary $\partial\Omega$. Let $\phi_i$, $i = 1, \ldots, 4$, be the basis functions relative to the vertices of $K$. Since,

$$\|\phi_i\|^2_{L^2(K)} \leq C H_K^3,$$

with a constant $C$ that is independent of $K$ (see Lemma B.5), and

$$|\omega_y|^{-1} \int_{\omega_y} u(x)\, dx| \leq C \|u\|_{L^2(\omega_K)} H_K^{-3/2},$$

we have

$$\|\tilde{I}^H u\|_{L^2(K)} \leq \sum_{i=1}^{4} (C\|u\|_{L^2(\omega_K)} H_K^{-3/2}) \|\phi_i\|_{L^2(K)} \leq C \|u\|_{L^2(\omega_K)}. \tag{3.13}$$

We note that this inequality is equally valid when $\omega_K$ touches the boundary. Let now

$$\hat{u} = u - |\omega_K|^{-1} \int_{\omega_K} u\, dx. \tag{3.14}$$

Since $\tilde{I}^H$ reproduces the constant functions on $K$, we obtain by using a Poincaré inequality (see Corollary A.15):

$$\|u - \tilde{I}^H u\|^2_{L^2(K)} = \|\hat{u} - \tilde{I}^H \hat{u}\|^2_{L^2(K)} \leq C \|\hat{u}\|^2_{L^2(\omega_K)} \leq C H_K^2 |u|^2_{H^1(\omega_K)},$$

since $\hat{u}$ has a zero average on $\omega_K$ and $\omega_K$ has a diameter of order $H_K$.

We next consider the case where $\omega_K$ shares at least one full face with $\partial\Omega$. Since $u$ vanishes on $\partial\Omega \cap \partial\omega_K$, (3.13) and a Friedrichs inequality (see Corollary A.15) yield

$$\|u - \tilde{I}^H u\|^2_{L^2(K)} \leq 2\|\tilde{I}^H u\|^2_{L^2(K)} + 2\|u\|^2_{L^2(K)} \leq C\|u\|^2_{L^2(\omega_K)} \leq C H_K^2 |u|^2_{H^1(\omega_K)}.$$

We next prove the second inequality. Let $K$ be such that $\omega_K$ does not touch the boundary and let $\hat{u}$ be defined as in (3.14). Using an inverse inequality (see Sect. B.5), (3.11), and a Poincaré inequality, we find

$$\begin{aligned}|\tilde{I}^H u|^2_{H^1(K)} &= |\tilde{I}^H \hat{u}|^2_{H^1(K)} \leq C\, H_K^{-2}\, \|\tilde{I}^H \hat{u}\|^2_{L^2(K)} \\ &\leq C H_K^{-2} \left( \|\tilde{I}^H \hat{u} - \hat{u}\|^2_{L^2(K)} + \|\hat{u}\|^2_{L^2(K)} \right) \\ &\leq C |u|^2_{H^1(\omega_K)}.\end{aligned}$$

The case when $\omega_K$ touches $\partial\Omega$ can be treated as before by using a Friedrichs inequality. □

*Remark 3.7.* We note that the interpolant $\tilde{I}^H$ and Lemma 3.6 suffice for our study of methods for scalar elliptic problems. When we extend our results to the case of linear elasticity in chapter 8, we will need a quasi-interpolant which reproduces not just constants but also all linear functions. Following Clément [139], we note that the value of $\tilde{I}^H u$ at a node $y$ is that of the $L^2(\omega_y)-$projection of $u$ onto the space of constant functions on $\omega_y$. We can replace this value by that of the projection onto the space of linear functions. No new ideas are then needed to extend the proof of the lemma to this new case; see Sect. 8.2 (in particular Lemma 8.1) for further details.

The following lemma shows the stability of pointwise interpolation from the coarse to the fine mesh. We refer to [134] for an earlier, different proof of the same result. As for the coarse mesh, for $k \in \mathcal{T}_h$, $\omega_k$ is the union of $k$ and its neighbors.

**Lemma 3.8** *There exists a constant $C$ independent of $h$ and $H$, such that*

$$|u_H - I^h u_H|^2_{H^s(k)} \leq C h_k^{2(1-s)} |u_H|^2_{H^1(\omega_k)},$$

*for $k \in \mathcal{T}_h$, $u_H \in V_H$, and $s = 0, 1$.*

*Proof.* Let $u_H \in V_H$. We will, without limiting the generality of our arguments, confine ourselves to tetrahedral elements in three dimensions; other cases only require minor additional arguments. To prove this stability result, we first note that any element $k \in \mathcal{T}_h$ which falls entirely inside an element of the coarse space, contributes exactly the same to the energy of the two functions and gives a vanishing interpolation error, since $I^h u_H = u_H$ on the element. We can therefore confine our study to an element $k$, of diameter $h_k$, which has vertices $A$, $B$, $C$, and $D$ which do not all belong to the same coarse mesh elements; cf. Fig. 3.1 for a two-dimensional picture. We will estimate $|I^h u_H|^2_{H^1(k)}$ in terms of norms of $u_H$ over four tetrahedra, which will soon be introduced, and which each is a subset of an individual coarse element.

Since $k$ is shape regular, the basis functions relative to $k$ satisfy

$$|\phi_i|^2_{H^1(k)} \leq C h_k, \quad i = A, B, C, D,$$

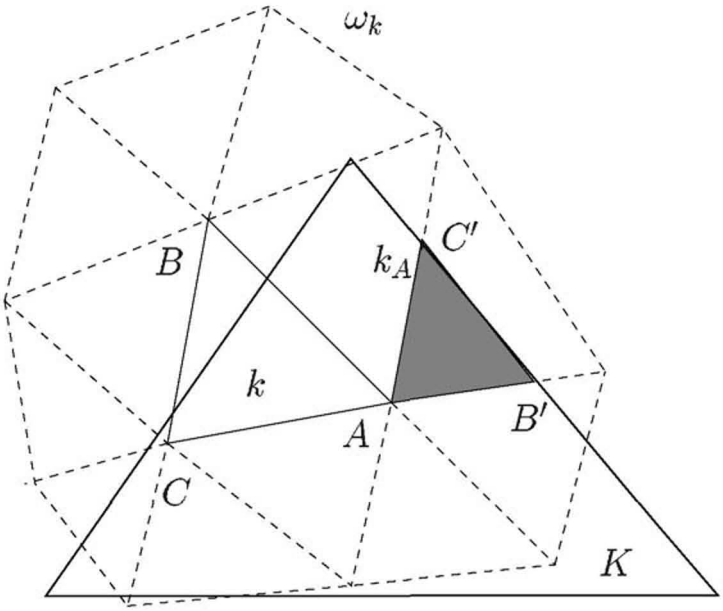


**Fig. 3.1.** Geometry underlying the proof of Lemma 3.8.

(see Lemma B.5) and thus

$$|I^h u_H|^2_{H^1(k)} \leq Ch_k(|u_H(A)|^2 + |u_H(B)|^2 + |u_H(C)|^2 + |u_H(D)|^2).$$

We now bound the four terms on the right hand side separately. Since the elements of the coarse mesh are shape regular, we can find a shape-regular tetrahedron with a diameter comparable to $h_k$, which has a vertex, $A$, $B$, $C$, or $D$, in common with $k$ and which is a subset of one of the coarse mesh elements; see Fig. 3.1 for a two-dimensional example. These tetrahedra, $k_A$, $k_B$, $k_C$, and $k_D$, are not necessarily elements of the fine triangulation but we can always make them subsets of $\omega_k$; see Fig. 3.1.

Let us consider $k_A$ and let $B'$, $C'$, and $D'$ be its other vertices. Since $u_H$ is linear on $k_A$, a scaling argument (see Lemma B.5) gives

$$\begin{aligned} h_k|u_H(A)|^2 &\leq h_k(|u_H(A)|^2 + |u_H(B')|^2 + |u_H(C')|^2 + |u_H(D')|^2) \\ &\leq Ch_k^{-2} \|u_H\|^2_{L^2(k_A)} \leq Ch_k^{-2} \|u_H\|^2_{L^2(\omega_k)}. \end{aligned}$$

Similar bounds can be found for the other three terms and tetrahedra.

We then obtain,

$$|I^h u_H|^2_{H^1(k)} \leq Ch_k^{-2} \|u_H\|^2_{L^2(\omega_k)}$$

and, since $I^h$ reproduces constants,

$$|I^h u_H|^2_{H^1(k)} = |I^h(u_H - \bar{u}_H)|^2_{H^1(k)} \leq C h_k^{-2} \|u_H - \bar{u}_H\|^2_{L^2(\omega_k)},$$

with $\bar{u}_H$ the mean value of $u_H$ on $\omega_k$. The proof of the bound for $s = 1$ is now completed by recalling that the diameter of $\omega_k$ is on the order of $h_k$ and by using a Poincaré inequality.

The proof for $s = 0$ can be carried out in a similar way. Using Lemma B.5, we find

$$\|I^h u_H\|^2_{L^2(k)} \leq C h_k^3 (|u_H(A)|^2 + |u_H(B)|^2 + |u_H(C)|^2 + |u_H(D)|^2),$$

and thus

$$\|I^h u_H\|^2_{L^2(k)} \leq C \|u_H\|^2_{L^2(\omega_k)}.$$

Let $\bar{u}_H$ be defined as before. We can write

$$\|u_H - I^h u_H\|^2_{L^2(k)} = \|(u_H - \bar{u}_H) - I^h(u_H - \bar{u}_H)\|^2_{L^2(k)} \leq C \|u_H - \bar{u}_H\|^2_{L^2(\omega_k)}.$$

The proof is completed by using a Poincaré inequality. □

The following lemma shows the stability of piecewise linear interpolation of piecewise quadratic functions.

**Lemma 3.9** *Let $u_h$ be a continuous, piecewise quadratic function defined on $\mathcal{T}_h$ and let $I^h u_h \in V^h$ be its piecewise linear interpolant on the same mesh. Then, there exists a constant $C$, independent of $h$, such that*

$$|I^h u_h|_{H^1(k)} \leq C |u_h|_{H^1(k)}, \quad k \in \mathcal{T}_h.$$

*Proof.* We have

$$|I^h u_h|^2_{H^1(k)} \leq 2\big(|I^h u_h - u_h|^2_{H^1(k)} + |u_h|^2_{H^1(k)}\big).$$

Consider the contribution of the first term on the right hand. We obtain,

$$|I^h u_h - u_h|^2_{H^1(k)} \leq C h^2 |u_h|^2_{H^2(k)} \leq C |u_h|^2_{H^1(k)},$$

by using a standard error bound (cf. Lemma B.6) and an elementary inverse inequality for quadratic polynomials (cf. Lemma B.27). □

We next develop an estimate of the $L^2$−norm over a strip of width $\delta_i$ along the boundary $\partial\Omega_i'$ of a generic subdomain. Given $\delta_i > 0$, let $\Omega_{i,\delta_i} \subset \Omega_i'$ be the set of points that are within a distance $\delta_i$ of $\partial\Omega_i' \setminus \partial\Omega$ and let $H_i$ be the diameter of $\Omega_i'$. We note that the following result holds for all of $H^1(\Omega_i')$.

**Lemma 3.10** *There exists a constant $C$ such that*

$$\|u\|^2_{L^2(\Omega_{i,\delta_i})} \leq C \delta_i^2 \big((1+H_i/\delta_i)|u|^2_{H^1(\Omega_i')} + 1/(H_i\delta_i)\|u\|^2_{L^2(\Omega_i')}\big), \quad u \in H^1(\Omega_i').$$

*Proof.* Let us cover $\Omega_{i,\delta_i}$ by shape-regular patches with $O(\delta_i)$ diameters. By using a Friedrichs inequality (see Corollary A.15) for each patch and by summing over the patches, we find that

$$\|u\|^2_{L^2(\Omega_{i,\delta_i})} \leq C(\delta_i^2 |u|^2_{H^1(\Omega_{i,\delta_i})} + \delta_i \|u\|^2_{L^2(\partial\Omega_i')}). \tag{3.15}$$

The second term on the right can be estimated by a trace estimate. Indeed, by combining the trace result in Lemma A.6 and the embedding $L^2(\partial\Omega_i') \subset H^{1/2}(\partial\Omega_i')$ with a scaling argument, we find

$$\|u\|^2_{L^2(\partial\Omega_i')} \leq C(H_i |u|^2_{H^1(\Omega_i')} + 1/H_i \|u\|^2_{L^2(\Omega_i')}). \tag{3.16}$$

Combining the last two inequalities concludes the proof. □

## 3.6 Convergence Results

A bound for the largest eigenvalue of the additive operator is given in the next lemma. It is an immediate consequence of Assumptions 2.4, 2.3, and 3.2, and Lemma 2.10.

**Lemma 3.11** *We have*

$$a(P_{ad}u, u) \leq \omega(N^c + 1)\, a(u,u), \quad u \in V^h.$$

We note that an upper bound for the largest eigenvalue can also be found directly using a coloring technique. In this case, a weaker hypothesis on the number of colors can be employed. It is enough to assume that every point $x \in \Omega$ belongs to at most $\hat{N}_c$ overlapping subdomains; see, e.g., [182, Th. 4.1] or [424, p. 165]. We note that it is easy to find partitions with $\hat{N}_c < N^c$, e.g., by considering three subdomains which intersect pairwise but where the three do not have any points in common.

In order to prove a bound for the smallest eigenvalue of the additive operator, we need, according to Lemma 2.5, to find a stable decomposition.

**Lemma 3.12** *If exact solvers are employed on all the subspaces, then there exists a constant $C$, independent of $h$, $H$, and $\delta$, such that*

$$a(P_{ad}^{-1}u, u) \leq C\left(1 + \frac{H}{\delta}\right) a(u,u), \quad u \in V^h.$$

*Proof.* We will find a stable decomposition for every $u \in V$. We first define

$$u_0 := \tilde{I}^H u \in V_0, \tag{3.17}$$

where $R_0^T = \tilde{I}^H$ is the quasi-interpolant of Lemma 3.6. Given the remainder

$$w = u - R_0^T u_0 = u - I^h u_0,$$

we next define the local components by

$$u_i = R_i(I^h(\theta_i w)) \in V_i, \quad 1 \leq i \leq N. \tag{3.18}$$

Here $\{\theta_i\}$ is the piecewise linear partition of unity associated with the overlapping partition and defined after Lemma 3.4. We note that the restriction operator $R_i$ is necessary since, according to our definition, the local spaces consist of functions defined on the subdomains only.

The local functions define a splitting of $u$ since using the definition of $R_0^T$, the linearity of $I^h$, and the assumption that the $\theta_i$ form a partition of unity, we can write

$$R_0^T u_0 + \sum_{i=1}^{N} R_i^T u_i = I^h u_0 + \sum_{i=1}^{N} I^h(\theta_i w) = I^h u_0 + I^h(u - I^h u_0) = u.$$

We now prove the stability of this splitting. A bound for $u_0$ follows from the stability of $\tilde{I}^H$ given in Lemma 3.6:

$$\tilde{a}_0(u_0, u_0) = |u_0|^2_{H^1(\Omega)} \leq C|u|^2_{H^1(\Omega)} = Ca(u,u). \tag{3.19}$$

For $1 \leq i \leq N$, we have

$$\begin{aligned} \tilde{a}_i(u_i, u_i) &= |I^h(\theta_i w)|^2_{H^1(\Omega_i')} \leq C|\theta_i w|^2_{H^1(\Omega_i')} \\ &\leq C\int_{\Omega_i'} |w\nabla\theta_i|^2\,dx + C\int_{\Omega_i'} |\theta_i \nabla w|^2\,dx, \end{aligned} \tag{3.20}$$

where for the first inequality, we have used Lemma 3.9. In order to bound the right hand side of (3.20), we first assume that, for $s = 0, 1$,

$$|w|^2_{H^s(\Omega_i')} = |u - I^h u_0|^2_{H^s(\Omega_i')} \leq C \sum_{\substack{K \in \mathcal{T}_H \\ K \cap \Omega_i' \neq \emptyset}} H_K^{2(1-s)} |u|^2_{H^1(\omega_K)}. \tag{3.21}$$

The second term on the right hand side of (3.20) can be bounded by using Lemma 3.4 and (3.21) for $s = 1$:

$$\int_{\Omega_i'} |\theta_i \nabla w|^2\,dx \leq \int_{\Omega_i'} |\nabla w|^2\,dx \leq C \sum_{\substack{K \in \mathcal{T}_H \\ K \cap \Omega_i' \neq \emptyset}} |u|^2_{H^1(\omega_K)}. \tag{3.22}$$

For the first term on the right hand side of (3.20), we note that $\nabla\theta_i$ differs from zero only in a strip $\Omega_{i,\delta_i}$ of width $\delta_i$ in the vicinity of $\partial\Omega_i'$. Using Lemmas 3.4 and 3.10, we then find that

$$\begin{aligned} \int_{\Omega_i'} |w\nabla\theta_i|^2\,dx &= \int_{\Omega_{i,\delta_i}} |w\nabla\theta_i|^2\,dx \leq \frac{C}{\delta_i^2}\int_{\Omega_{i,\delta_i}} |w|^2\,dx \\ &\leq C\left(\left(1 + \frac{H_i}{\delta_i}\right)|w|^2_{H^1(\Omega_i')} + \frac{1}{H_i\delta_i}\|w\|^2_{L^2(\Omega_i')}\right). \end{aligned} \tag{3.23}$$

The second term in the last sum can be bounded using inequality (3.21) for $s = 0$ and Assumption 3.5:

$$\frac{1}{H_i\delta_i}\|w\|^2_{L^2(\Omega_i')} \leq C \sum_{\substack{K\in\mathcal{T}_H \\ K\cap\Omega_i'\neq\emptyset}} \frac{H_K^2}{H_i\delta_i}|u|^2_{H^1(\omega_K)} \leq C\frac{H_i}{\delta_i} \sum_{\substack{K\in\mathcal{T}_H \\ K\cap\Omega_i'\neq\emptyset}} |u|^2_{H^1(\omega_K)}. \quad (3.24)$$

Combining (3.20), (3.22), (3.23), and (3.24), summing over the subdomains, and using Assumption 3.2, we obtain

$$\sum_{i=1}^{N} \tilde{a}_i(u_i,u_i) \leq C\left(1+\frac{H}{\delta}\right) a(u,u),$$

which, together with (3.19) and Lemma 2.5 gives the desired result.

We still need to provide a proof of inequality (3.21). We write

$$w = u - I^h u_0 = (u-u_0) + (u_0 - I^h u_0).$$

Lemma 3.6 gives

$$|u-u_0|^2_{H^s(\Omega_i')} \leq C \sum_{\substack{K\in\mathcal{T}_H \\ K\cap\Omega_i'\neq\emptyset}} H_K^{2(1-s)}|u|^2_{H^1(\omega_K)},$$

and Lemma 3.8

$$|u_0 - I^h u_0|^2_{H^s(\Omega_i')} \leq C \sum_{\substack{k\in\mathcal{T} \\ k\subset\Omega_i'}} h_k^{2(1-s)}|u_0|^2_{H^1(\omega_k)} \leq C \sum_{\substack{K\in\mathcal{T}_H \\ K\cap\Omega_i'\neq\emptyset}} H_K^{2(1-s)}|u|^2_{H^1(\omega_K)}.$$

These two inequalities, combined with the triangle inequality, give (3.21). □

Combining Lemmas 3.11 and 3.12, we find a bound for the condition number of $P_{ad}$.

**Theorem 3.13** *In case exact solvers are employed on all subspaces, the condition number of the additive Schwarz operator satisfies*

$$\kappa(P_{ad}) \leq C\left(1+\frac{H}{\delta}\right),$$

*where $C$ depends on $N^c$, but is otherwise independent of $h$, $H$, and $\delta$.*

We will now consider some particular cases. A first simplification occurs in the case of a structured fine mesh in which case $\mathcal{T}$ is obtained by refining a coarser triangulation $\mathcal{T}_H$. We can then choose

$$V_0 = V^H \subset V$$

and we note that $I^h u_0 = u_0$. Assumption 3.5 is still required for the proof of the lower bound.

A further simplification is achieved by building the overlapping partition from the coarse mesh. In this case, $\mathcal{T}_H$ coincides with the initial nonoverlapping partition $\{\Omega_i\}$ and Assumption 3.5 is automatically satisfied.

We end this section by noting that the bound given in Theorem 3.13 is sharp. More precisely, we consider the particular case of $\Omega = (0,1)^2$ with $\mathcal{T}$ and $\mathcal{T}_H$ nested, uniform, triangular meshes of diameters $h = 2^{-k}$ and $H = 2^{-K}$, $K > k$, respectively. We also assume that the nonoverlapping partition $\{\Omega_i\}$ coincides with $\mathcal{T}_H$ and that $\delta_i/h = \delta/h$ is bounded from above (small overlap). In this case, there exists a constant $c$ such that

$$\kappa(P_{ad}) \geq c\frac{H}{\delta}.$$

The proof of this result can be found in [86]; we refer in particular to Lemma 3.5 and Remarks 3.7 and 3.8 of that paper.

## 3.7 Remarks on the Implementation

In this section, we make some remarks on some practical aspects of two-level overlapping preconditioners. The preconditioner associated with the additive operator can be immediately deduced from formula (3.1):

$$A_{ad}^{-1} = \sum_{i=1}^{N} R_i^T A_i^{-1} R_i + R_0^T A_0^{-1} R_0. \tag{3.25}$$

In case a multiplicative or a hybrid method is employed, the corresponding preconditioner can be applied using the algorithm in section 2.6. In any case, the restrictions $R_i$ and the operators $A_i$ need to be constructed.

For the local spaces, the $R_i$, $i > 0$, are rectangular matrices of zeros and ones, which simply extract the degrees of freedom that lie inside the subdomains $\Omega_i'$. The restriction matrix $R_0$ for the coarse space is dealt with in more detail in the following. Once the restriction operators are available, the matrices associated with the subspaces can be obtained from the global matrix $A$ through the formula

$$A_i = R_i A R_i^T, \quad i = 0, \ldots, N. \tag{3.26}$$

We note that for the local spaces, this simply means that we extract the blocks of $A$ corresponding to the degrees of freedom associated with the subdomains $\Omega_i'$. Once the matrices $A_i$ are available, exact or inexact factorizations can be found and stored. For overlapping preconditioners, the local and coarse solvers can therefore be obtained from the global matrix $A$. This is not always true for other domain decomposition algorithms; Neumann-Neumann and FETI methods for instance, see Chap. 6, rely on the solution of local problems with Neumann boundary conditions and the corresponding solver cannot be obtained directly from $A$.

We now focus our attention on the restriction $R_0$ onto the coarse space. We first assume that the fine mesh is a refinement of the the coarse one

and therefore $V_0 \subset V$. The restriction operator is defined by the interpolant $R_0^T : V_0 \to V$, from the subspace to the global space, which takes a vector of coarse degrees of freedom and gives the vector of fine degrees of freedom that determines the corresponding coarse function.

More precisely, let $\{\phi_i^H,\ i = 1, \ldots, N_H\}$ and $\{\phi_j^h,\ j = 1, \ldots, N_h\}$ be the basis functions for the fine and coarse spaces, respectively. A coarse vector $u \in V_0 \subset V$ can be represented using either basis:

$$u(x) = \sum_{i=1}^{N_H} u_i^H \, \phi_i^H(x) = \sum_{j=1}^{N_h} u_j^h \, \phi_j^h(x), \tag{3.27}$$

where $u^H$ and $u^h$ are the vectors of coarse and fine degrees of freedom, respectively. The extension matrix $R_0^T$ is such that

$$u^h = R_0^T u^H.$$

Since the coarse basis functions are also contained in the fine space, they can also be represented using the fine basis:

$$\phi_i^H(x) = \sum_{j=1}^{N_h} r_{ij} \, \phi_j^h(x), \quad i = 1, \ldots, N_i^H. \tag{3.28}$$

We note that $r_{ij}$ is the nodal value of the $i$-th coarse basis function at the $j$-th fine node:

$$r_{ij} := \phi_i^H(x_j^h).$$

We now substitute (3.28) into (3.27) and obtain

$$\sum_{i=1}^{N_H} \sum_{j=1}^{N_h} r_{ij} u_i^H \, \phi_j^h(x) = \sum_{j=1}^{N_h} u_j^h \, \phi_j^h(x).$$

Exchanging the sums on the left hand side yields

$$u_j^h = \sum_{i=1}^{N_H} r_{ij} u_i^H.$$

We have therefore shown that the entries of $R_0$ are exactly the coefficients $r_{ij}$ obtained by interpolating the coarse basis functions into the fine mesh. In other words, the $i$-th column of the extension matrix $R_0^T$ consists of the fine nodal values corresponding to the basis function $\phi_i^H$. We note that exactly the same procedure applies when the fine and coarse meshes are not nested.

We end this section by pointing out a somewhat delicate issue for two-level overlapping methods; that of Dirichlet boundary conditions. The presentation in this chapter has been made for the case of homogeneous Dirichlet conditions: the finite element spaces considered are contained in $H_0^1(\Omega)$ and the

corresponding vectors of degrees of freedom consist of nodal values in the interior of $\Omega$. Matrices and operators should in principle be those acting on these degrees of freedom only and can be obtained from the stiffness matrix $A$ for a Neumann problem in $\Omega$ by deleting rows and columns relative to the nodes on the boundary. This is however seldom done in practice since it requires a renumbering of the unknowns. Normally, finite element codes build the global matrix $A$ for the Neumann problem and boundary conditions are taken into account in a second step. For the case of Dirichlet conditions on the entire or part of $\partial\Omega$, the equation (and therefore the corresponding row in $A$ and entry in the right hand side) corresponding to a boundary node $x_k$ is changed into

$$\tilde{a}_{kk}u_k = \tilde{a}_{kk}g_k, \tag{3.29}$$

with $g_k$ the value coming from the Dirichlet condition and $\tilde{a}_{kk}$ a suitable positive scaling parameter (a multiple of the $k$-th diagonal entry of $A$ is sometimes employed). The $k$-th row of the stiffness matrix $A$ is therefore changed and $A$ becomes nonsymmetric (entries in the $k$-th column can be put to zero as well only in the case of homogeneous Dirichlet conditions). If a Krylov-type method employs an initial guess that satisfies the Dirichlet condition, the $k$-th entry of the corresponding residual is zero and the further application of $A$ will give a vector with the same property. Indeed, even though $A$ is not symmetric, it is always applied to residual vectors that are zero in $x_k$ and is therefore symmetric in the corresponding subspace. For these reasons, we need to make sure that the preconditioner as well gives vectors that vanish at the nodes on $\partial\Omega$.

We consider the two-level preconditioner in (3.25) and assume that the local and coarse matrices are obtained through (3.26). The local components do not present any particular problem; the matrices $R_i$ have exactly one entry for each row or column and $A_i$ is simply obtained by extracting the corresponding block from $A$. If we start from a residual vector that vanishes on $\partial\Omega \cap \partial\Omega_i$, the application of each of the three factors in $R_i^T A_i^{-1} R_i$ will give a residual with the same property.

The coarse component of the preconditioner however must be considered more carefully. Indeed, when $A_0$ is obtained from the modified $A$ through (3.26), the row corresponding to a coarse node on $\partial\Omega$ will not be of the type in (3.29) and will not correspond to the correct coarse matrix for a Dirichlet problem. We therefore need to modify those rows accordingly. Alternatively, we can construct $A_0$ directly, by employing the same routine as for $A$ but for the coarse mesh. In either case, when $A_0^{-1}$ is applied to a coarse residual that vanishes on $\partial\Omega$, the same will hold for the resulting vector.

Even with the modifications above, the coarse component of the preconditioner requires an additional change. Since the operator $R_0^T : V_0 \to V$, provides the interpolation from the coarse to the fine mesh, a coarse residual that vanishes on $\partial\Omega$ will result in an interpolated residual in $V$ with the same property. The same does not hold however for the restriction $R_0 : V \to V_0$.

It can easily be seen that the value at a coarse node $X_l$ of a restricted vector $R_0 u^h$ is a linear combination of *all* nodal values of $u^h$ at the fine nodes that lie in the coarse elements that touch $X_l$. A fine residual that vanishes on $\partial\Omega$ will in general not give a coarse one with the same property. If care is not taken, the coarse solve will then give a residual that does not vanish on $\partial\Omega$, leading the Krylov iteration out of the correct subspace. Convergence still occurs to the correct solution but this fact can sometimes cause loss of performance of two-level methods. In practice, a very large value of the diagonal entries in the coarse matrix $A_0$ corresponding to nodes on $\partial\Omega$ can sometimes restore performance, but this is a fix that may depend heavily on the particular problem.

A possible remedy is to put to zero all rows in $R_0$ corresponding to coarse nodes on $\partial\Omega$. The transpose $R_0^T$ of this modified matrix can also be employed for the extension step.

## 3.8 Numerical Results

In this section, we present some numerical results for two-level overlapping preconditioners to show that bounds in Theorem 3.13 are reflected in numerical practice. We will in particular look at how the rate of convergence depends on the relative overlap, the fine mesh, and the number of subdomains. More numerical tests and examples can be found in [424].

The results given here are from Cai, Gropp, and Keyes [108, sect. 4.1], courtesy of the authors, for symmetric, positive definite problems. Others, for nonsymmetric and/or indefinite problems, are reported in Sect. 11.4.

We consider $\Omega = (0,1)^2$ and the Poisson problem

$$-\Delta u = f,$$

with homogeneous Dirichlet conditions. The right hand side $f$ is chosen such that $u = \exp(xy)\sin(\pi x)\sin(\pi y)$ is the solution.

The fine discretization consists of a uniform mesh of diameter $h$ and a five-point central finite difference approximation. Similarly, the coarse problem consists of a finite difference approximation over a coarser uniform mesh of diameter $H$. The overlapping subdomain partition is obtained by extending the coarse elements by adding layers of fine elements; the overlap $\delta$ is thus an integer multiple of $h$. Exact coarse and local solvers are employed.

In order to allow a direct comparison with the nonsymmetric case, the solution of the preconditioned system is accelerated with GMRES without restarts. We note that for the additive preconditioner the conjugate gradient method would be more appropriate. The iterations are stopped after a reduction of the Euclidean norm of the preconditioned residual by a factor $10^{-5}$.

The following algorithms are considered:

1. **MSR**: Richardson's method without relaxation (see appendix C.3) with the nonsymmetric, multiplicative preconditioner (2.9) of Sect. 2.2.
2. **MSM**: GMRES with the same nonsymmetric, multiplicative preconditioner.
3. **ASM**: GMRES with the symmetric, additive preconditioner (3.1).

We note that an upper bound for the error propagation operator for MSR is given in Theorem 2.9.

Iteration counts are shown in Table 3.1 as functions of the fine mesh size and the overlap, for different partitions. Entries that would have required a relative overlap $\delta/H$ greater than 0.5 are left blank. We note that the number of colors $N_c$ is the same for all the cases shown and is equal to 4.

As expected, GMRES performs better than Richardson and the multiplicative preconditioner better than the additive one when the same iterative method is employed; it requires roughly half the number of iterations. In addition, within the same group of columns which correspond to a common value of $H$, the iteration counts are *constant* along the diagonals, corresponding to the same value of $H/\delta$ shown in parenthesis. Comparison with different groups of columns, then reveals that the iteration counts are bounded independently of $h$ and $H$; they appear to grow *linearly* with $H/\delta$ as predicted by the theory. We refer to [108] for additional comments and a discussion of the parallel complexity of the algorithms.

**Table 3.1.** Iteration counts for solving the Poisson equation. The overlapping factors $H/\delta$ for MSR are given in parentheses. The corresponding overlapping factors for MSM and ASM are the same as for MSR and are therefore omitted. The number, as in $(2h)$, which appears next to the name of each method indicates the overlap $\delta$.

| $h^{-1} =$ | 32 | 64 | 128 | 32 | 64 | 128 | 64 | 128 |
|---|---|---|---|---|---|---|---|---|
| Methods | | $H = 1/4$ | | | $H = 1/8$ | | $H = 1/16$ | |
| MSR($h$) | 7 (8) | 11 (16) | 19 (32) | 6 (4) | 7 (8) | 10 (16) | 5 (4) | 6 (8) |
| MSR($2h$) | 6 (4) | 7 (8) | 11 (16) | 5 (2) | 6 (4) | 7 (8) | 4 (2) | 5 (4) |
| MSR($4h$) | 5 (2) | 6 (4) | 7 (8) | | 5 (2) | 6 (4) | | 4 (2) |
| MSR($8h$) | | 5 (2) | 6 (4) | | | 5 (2) | | |
| MSM($h$) | 5 | 6 | 7 | 4 | 4 | 5 | 3 | 3 |
| MSM($2h$) | 5 | 5 | 6 | 4 | 4 | 4 | 3 | 3 |
| MSM($4h$) | 4 | 5 | 5 | | 4 | 4 | | 3 |
| MSM($8h$) | | 4 | 5 | | | 4 | | |
| ASM($h$) | 11 | 13 | 15 | 10 | 10 | 11 | 9 | 8 |
| ASM($2h$) | 11 | 11 | 13 | 10 | 10 | 10 | 8 | 8 |
| ASM($4h$) | 10 | 11 | 11 | | 10 | 10 | | 8 |
| ASM($8h$) | | 10 | 11 | | | 10 | | |

## 3.9 Restricted Schwarz Algorithms

Let us consider the same overlapping subdomain partition as in section 3.2 and the same coarse problems as in Sect. 3.3. A symmetric additive preconditioned operator was defined in (3.1) and we recall that for a subdomain $\Omega_i'$ the restriction $R_i$ is the operator that extracts the degrees of freedom in the interior of $\Omega_i'$, while the extension $R_i^T$ extends a vector of degrees of freedom contained in $\Omega_i'$ by zero. We introduce the following modification:

$$P_{ras} = R_0^T A_0^{-1} R_0 A + \sum_{i=1}^{N} \tilde{R}_i^T A_i^{-1} R_i A,$$

where $A_i$, $A_0$, and $R_0$ are defined as before. Here, the new restriction $\tilde{R}_i$ *only* extracts the degrees of freedom associated with the nonoverlapping subdomain $\Omega_i$ and makes those in the extended subregion $\Omega_i' \setminus \Omega_i$ vanish. In practice, this means that when a local component of the preconditioner is applied to a residual vector, after the solution of a local problem, only the entries of the residual that *only* belong to the subdomain $\Omega_i'$ (and are not shared by others) are updated. This reduces the communication between different subdomains (and between different processors).

This *restricted additive Schwarz* preconditioner is nonsymmetric even for a symmetric $A$ but, surprisingly, for certain symmetric matrices GMRES with a restricted Schwarz preconditioner performs better than the conjugate gradient method with the original symmetric preconditioner, in terms of iteration counts, CPU time, and communication time. We note that in a parallel implementation, half of the communication cost can be saved. A multiplicative version is also available, together with the following variant

$$P_{ash} = R_0^T A_0^{-1} R_0 A + \sum_{i=1}^{N} R_i^T A_i^{-1} \tilde{R}_i A,$$

which is referred to as *additive Schwarz with harmonic overlap*. Variants employing restrictions or extensions with weights are also possible and applications to indefinite and nonsymmetric problems have also been considered.

To our knowledge, a comprehensive theory of these algorithms is still missing. We note however that the restricted additive Schwarz preconditioner is the default parallel preconditioner for nonsymmetric systems in the PETSc library (see [33]) and has been used for the solution of very large problems; see, e.g., [240, 278, 106]. We refer to [115, 409] for a more thorough introduction, additional details, some analysis, and numerical results. See also [356] for an algebraic theory.

## 3.10 Alternative Coarse Problems

So far, we have only considered coarse problems given by a standard finite element approximation on a coarse mesh. Structured meshes are usually con-

structed by refining a coarse triangulation, which can conveniently be employed for the coarse solver. For unstructured meshes, we have shown that a coarse mesh, independent of the fine mesh, can be employed, but only if an interpolation operator from the coarse to the fine finite element space can be found and efficiently implemented; see Sect. 3.3 and 3.7. However, this is not always a simple task, especially in three dimensions. Alternative approaches are realized by (smoothed) aggregation techniques and partition of unity coarse spaces, which provide efficient coarsening procedures without the need of introducing a coarse triangulation. Such procedures, when employed to construct multilevel preconditioners, belong to a class of coarsening techniques called *algebraic multigrid methods*; see, e.g., [454] and the references therein. We also note that it is possible to use the coarse, global spaces which will be developed in Chap. 5 and 6 but that we are not going to pursue that idea here.

The key idea of such techniques is to consider a coarse space that is the span of a set of partition of unity functions associated with an overlapping partition of subdomains, typically the same that we employ for the local components of the preconditioner. In smoothed aggregation methods the basis of the coarse space is found by applying a suitable operator, called a *smoother*, to some initial set of simple functions; in the more properly called *partition of unity coarse spaces* they are constructed by directly assigning explicit nodal values inside the subdomains.

In the following, we assume that the fine triangulation $\mathcal{T}_h$ is quasi uniform. In addition, we only consider quasi-uniform partitions into subdomains. We will only present the most relevant proofs here and refer to the literature for more complete results.

### 3.10.1 Convergence Results

Here we will always assume that we employ exact solvers for the local and coarse problems but approximate solvers could be used as well without any major changes of the theory.

As before, our preconditioner is uniquely defined by two components: an overlapping partition of $\Omega$ into subdomains

$$\mathcal{F} = \{\Omega_i' \subset \Omega \mid \quad 1 \leq i \leq N\},$$

and a set of coarse basis functions $\{\Phi_i \mid 1 \leq i \leq N\}$, that determine a coarse space.

The coarse basis functions are finite element functions and are associated with the subdomains. We first define the index set $\mathcal{I} = \{i = 1, \ldots, N | \ \partial\Omega_i' \cap \partial\Omega = \emptyset\}$; in the construction of the coarse space the subdomains that touch $\partial\Omega$ do not need to contribute any basis functions:

$$V_0 = \text{span}\,\{\Phi_i |\ i \in \mathcal{I}\}.$$

For simplicity, we assume that, if the boundary of a subdomain intersects $\partial\Omega$, this intersection has a nonvanishing $(n-1)$-dimensional measure for problems in $\mathbb{R}^n$. In the more general case, subdomains that intersect $\partial\Omega$ in one or few isolated points, or along a curve, should also contribute to the coarse space.

In practice, a larger coarse space can be employed by adding basis functions related to the other subdomains. The largest eigenvalue of the preconditioned operator will remain the same, and according to Lemmas 2.6 and 2.5, the smallest one can only increase. These extra basis functions are also needed in certain proofs.

We next introduce two sets of sufficient conditions on $\mathcal{F}$ and the coarse functions, given in terms of $H$ and $\delta$, $H > \delta > 0$, which reflect the size of the subdomains and the overlap, respectively. They ensure that the resulting additive preconditioner is optimal and scalable and allow us to derive quantitative bounds which only involve the relative overlap between the subdomains, as in the case of two-level methods with a standard coarse space. We note that the analysis presented here, leading to the quadratic bound in Theorem 3.17, is a modification of that developed in [93] for the case of generous overlap, and in [275] for the case of small overlap. The analysis that follows and leads to Theorem 3.19 is a modification of results from [409].

**Assumption 3.14 (Coarse space I)**

1. $|\Phi_i|^2_{H^1(\Omega)} \leq CH^{(n-1)}/\delta$, $i \in \mathcal{I}$;
2. $\|\Phi_i\|^2_{L^2(\Omega)} \leq CH^n$, $i \in \mathcal{I}$;
3. *There exists* $\Omega_{int} \subset \Omega$, *such that* $\sum_{i\in\mathcal{I}} \Phi_i(x) = 1$ *for* $x \in \Omega_{int}$, *and* $\operatorname{dist}(x, \partial\Omega) \leq CH$ *for* $x \in \Omega \setminus \Omega_{int}$;
4. $\operatorname{supp}\{\Phi_i\} \subset \overline{\Omega_i'}$, $i \in \mathcal{I}$.

We note, that a nonnegative function $\Phi_i$, which equals one in the interior of $\Omega_i'$ and decreases smoothly to zero in a layer of width $\delta$ around $\partial\Omega_i'$, satisfies the given bounds for $|\cdot|_1^2$ and for $\|\cdot\|_0^2$. The additional assumption, that a coarse interpolant reproduces constants everywhere, except in a layer of width $H$ around the boundary, will translate into an error estimate for a suitably defined interpolation operator; see Lemma 3.16.

**Assumption 3.15 (Partition)**

1. $\operatorname{diam}(\Omega_i') \leq CH$;
2. *For every* $x \in \Omega$, *there exists* $\Omega_i' \in \mathcal{F}$, *such that* $x \in \Omega_i'$ *and* $\operatorname{dist}(x, \partial\Omega_i' \setminus \partial\Omega) \geq c\delta$;
3. *The partition* $\mathcal{F}$ *can be colored using at most* $N^c$ *colors, in such a way that subregions with the same color are disjoint;*
4. $|\Omega_i'| \geq CH^n$.

The first and the last assumption together ensure that the subdomains all have diameters of comparable size $H$ and that they are shape regular. According to the second assumption, $\delta$ is a measure of the overlap between the

subdomains; cf. Assumption 3.1. The third assumption is the finite covering assumption of Assumption 3.2.

**Lemma 3.16 (Coarse Interpolant)** *Let Assumption 3.14 hold. Then, the operator* $Q_0 : H_0^1(\Omega) \to V_0$, *defined by*

$$Q_0 u = \sum_{i \in \mathcal{I}} \bar{u}_i \Phi_i, \quad \bar{u}_i = |\Omega_i'|^{-1} \int_{\Omega_i'} u \, dx,$$

*satisfies,*

$$|Q_0 u|^2_{H^1(\Omega)} \leq C \frac{H}{\delta} |u|^2_{H^1(\Omega)},$$
$$\|u - Q_0 u\|^2_{L^2(\Omega)} \leq C H^2 |u|^2_{H^1(\Omega)},$$

*for* $u \in H_0^1(\Omega)$.

*Proof.* In the same way as in the proof of [275, Lem. 2.2] (see also Lemma 3.6), we can prove

$$\begin{aligned} |Q_0 u|^2_{H^1(\Omega_{int})} &\leq C \frac{H}{\delta} |u|^2_{H^1(\Omega)}, \\ \|u - Q_0 u\|^2_{L^2(\Omega_{int})} &\leq C H^2 |u|^2_{H^1(\Omega)}, \end{aligned} \tag{3.30}$$

where $\Omega_{int}$ was defined in Assumption 3.14.

We next define the boundary region $B = \Omega \setminus \Omega_{int}$, collect the indices of the subdomains touching $B$ in

$$\mathcal{B} = \{1 \leq i \leq N | \; B \cap \Omega_i' \neq \emptyset\},$$

and define an extended boundary region by $B' = \bigcup_{i \in \mathcal{B}} \Omega_i'$.

Using Assumption 3.14.1 and the Cauchy-Schwarz inequality, we find

$$|Q_0 u|^2_{H^1(B)} \leq C \sum_{\substack{B \cap \Omega_i' \neq \emptyset \\ i \in \mathcal{I}}} |\bar{u}_i|^2 \frac{H^{d-1}}{\delta} \leq C \frac{1}{\delta H} \|u\|^2_{L^2(B')}.$$

Since $u$ vanishes on $\partial B' \cap \partial \Omega$ and $B'$ has a width on the order of $H$, we can use a Friedrichs inequality (cf. Corollary A.15) and obtain

$$|Q_0 u|^2_{H^1(B)} \leq \frac{H}{\delta} |u|^2_{H^1(B')},$$

which combined with (3.30) proves the first inequality.

In a similar way, we find

$$\begin{aligned} \|u - Q_0 u\|_{L^2(B)} &\leq \|u\|_{L^2(B)} + \|Q_0 u\|_{L^2(B)} \\ &\leq C \|u\|_{L^2(B')} \leq C H |u|_{H^1(B')}. \end{aligned} \tag{3.31}$$

Combining (3.30) and (3.31), we have proven the second inequality. □

We have the following result.

**Theorem 3.17** *Let Assumptions 3.14 and 3.15 hold. Then, in case exact solvers are employed, there exist positive constants c and C, such that*

$$c\left(1+\frac{H}{\delta}\right)^{-2} a(u,u) \leq a(P_{ad}u,u) \leq C\, a(u,u), \quad u \in V,$$

*and thus*

$$\kappa(P_{ad}) \leq C\left(1+\frac{H}{\delta}\right)^{2}.$$

*Proof.* The upper bound is a direct consequence of the finite covering property Assumption 3.15.3; cf. Lemma 3.11. In order to establish the lower bound, we need to find a stable decomposition. For $u \in V$, let

$$\begin{aligned} u_0 &= Q_0 u, \\ u_i &= I^h(\theta_i(u-u_0)), \quad 1 \leq i \leq N, \end{aligned}$$

where $\{\theta_i\}$ is the partition of unity relative to $\mathcal{F}$, defined by Lemma 3.4 and (3.7).

The same arguments as in the proof of Lemma 3.12 give

$$\sum_{i=1}^{N} |u_i|^2_{H^1(\Omega)} \leq C\left(\left(1+\frac{H}{\delta}\right)|u-Q_0u|^2_{H^1(\Omega)} + \frac{1}{H\delta}\|u-Q_0u\|^2_{L^2(\Omega)}\right).$$

The quadratic bound is then found by applying Lemma 3.16. □

The stable decomposition, on which the proof of the previous theorem relies, was originally given in [93] for the case of generous overlap and in [275, Lem. 2.3] for the case of small overlap.

In order to improve the quadratic bound of Theorem 3.17, we replace the partition of unity functions $\{\theta_i\}$ with the coarse basis functions $\{\Phi_i\}$ in the proof of Theorem 3.17. In order to do so, our coarse basis functions need to satisfy additional properties, originally proposed in [409]; cf. Lemma 3.4. They involve bounds on the $L^\infty$-norm of these functions and of their gradients together with the requirement that they form a partition of unity on the whole of $\Omega$.

For the same overlapping partition $\mathcal{F}$, we consider a set of coarse finite element functions as previously introduced and associated with the interior subdomains $I$, and enlarge it by also using functions associated with the subdomains that touch $\partial\Omega$. We introduce

$$\{\Phi_i \mid \quad 1 \leq i \leq N\},$$

such that supp $\{\Phi_i\} \subset \overline{\Omega_i'}$, $1 \leq i \leq N$. It is important to note that some of the $\Phi_i$ may not vanish on $\partial\Omega$ and that they therefore do not belong to the the coarse space $V_0$. We will use them all in our proof and will assume that the family $\{\Phi_i\}_{i=1}^N$ forms a partition of unity on the entire $\overline{\Omega}$. The assumptions on the overlapping partition $\mathcal{F}$ remain the same; see Assumption 3.15. We assume for our coarse basis functions:

**Assumption 3.18 (Coarse space II)**

1. $\|\Phi_i\|_{L^\infty(\Omega)} \leq C$, $1 \leq i \leq N$;
2. $|\Phi_i|_{W^{1,\infty}(\Omega)} \leq C/\delta$, $1 \leq i \leq N$;
3. $\sum_{i=1}^{N} \Phi_i(x) = 1$, *for* $x \in \overline{\Omega}$;
4. *There exists* $\Omega_{int} \subset \Omega$, *such that* $\sum_{i \in \mathcal{I}} \Phi_i(x) = 1$, *for* $x \in \Omega_{int}$, *and* $\operatorname{dist}(x, \partial\Omega) \leq CH$ *for* $x \in \Omega \setminus \tilde{\Omega}_{int}$;
5. $\operatorname{supp}\{\Phi_i\} \subset \overline{\Omega_i'}$, $1 \leq i \leq N$.

Similar assumptions were originally proposed in [409]. Before proceeding, we note that the proof that we present here relies on the same idea as that of [409, Th. 1], but that we treat the boundary coarse functions in a different way, by simply excluding them from the coarse space. We finally note that Assumptions 3.18.1 and 3.18.2 imply 3.14.1 and 3.14.2.

An interpolation operator into the coarse space $V_0$ can be defined in exactly the same way as in Lemma 3.16. We have the following theorem:

**Theorem 3.19** *Let Assumptions 3.18 and 3.15 hold. Then, there exist positive constants c and C, such that,*

$$c\left(1+\frac{H}{\delta}\right)^{-1} a(u,u) \leq a(P_{ad}u,u) \leq C\, a(u,u), \quad u \in V,$$

*and thus*

$$\kappa(P_{ad}) \leq C\left(1+\frac{H}{\delta}\right).$$

*Proof.* The upper bound is a direct consequence of the finite covering assumption in Assumption 3.15.3. In order to find the lower bound, we need to find a stable decomposition. For $u \in V$, let

$$\begin{aligned} u_0 &= Q_0 u, \\ u_i &= I^h(\Phi_i(u-\bar{u}_i)), \quad i \in \mathcal{I}, \\ u_i &= I^h(\Phi_i u), \quad i \notin \mathcal{I}. \end{aligned} \tag{3.32}$$

We have

$$\sum_{i=0}^{N} u_i = \sum_{i\in\mathcal{I}} (\bar{u}_i\Phi_i + I^h(\Phi_i u) - \bar{u}_i\Phi_i) + \sum_{i\notin\mathcal{I}} I^h(\Phi_i u) = \sum_{i=1}^{N} I^h(\Phi_i u) = u,$$

since the $\{\Phi_i\}$ form a partition of unity on the entire $\overline{\Omega}$.

The same arguments, as in the proof of Lemma 3.12, give

$$\begin{aligned} \sum_{i\in\mathcal{I}} |u_i|^2_{H^1(\Omega)} &\leq C \sum_{i\in\mathcal{I}} \left( \left(1+\frac{H}{\delta}\right) |u-\bar{u}_i|^2_{H^1(\Omega_i')} + \frac{1}{H\delta}\|u-\bar{u}_i\|^2_{L^2(\Omega_i')} \right) \\ &\leq C\left(1+\frac{H}{\delta}\right) \sum_{i\in\mathcal{I}} |u|^2_{H^1(\Omega_i')}, \end{aligned} \tag{3.33}$$

where for the last inequality, we can use a Poincaré inequality instead of an error bound for $Q_0$.

Similarly, we find, for $i \notin \mathcal{I}$,

$$\begin{aligned}\sum_{i\notin\mathcal{I}} |u_i|^2_{H^1(\Omega)} &\leq C \sum_{i\notin\mathcal{I}} \left( \left(1+\frac{H}{\delta}\right) |u|^2_{H^1(\Omega_i')} + \frac{1}{H\delta} \|u\|^2_{L^2(\Omega_i')} \right) \\ &\leq C \left(1+\frac{H}{\delta}\right) \sum_{i\in\mathcal{I}} |u|^2_{H^1(\Omega_i')}\,,\end{aligned} \tag{3.34}$$

since $u$ vanishes on $\partial\Omega_i' \cap \partial\Omega$ and we can apply a Friedrichs inequality (cf. Corollary A.15). The linear bound is then found by applying Lemma 3.16 to (3.33) and (3.34). □

We note that the error bound in Lemma 3.16 is not needed for the proof of Theorem 3.19.

*Remark 3.20.* The assumption that the triangulation $\mathcal{T}_h$ is quasi uniform is not used in the proofs of this section. Indeed, Theorems 3.17 and 3.19 are direct consequences of Assumptions 3.14 and 3.15 for any arbitrary shape-regular mesh $\mathcal{T}_h$.

### 3.10.2 Smoothed Aggregation Techniques

The basic ideas of smoothed aggregation are fairly simple and natural: in a first step, the fine mesh points are aggregated to form an initial nonoverlapping partition of the domain $\Omega$, and the characteristic functions associated with this nonoverlapping partition are constructed (aggregation); see below. In a second step, these characteristic functions, with values typically decreasing from one to zero in a layer of width $O(h)$, are smoothed by the application of a suitable smoother. The supports of the smoothed functions define an overlapping partition and corresponding local problems, while their linear span provides a low dimensional coarse problem. This procedure can be applied recursively to obtain additional coarse levels for the construction of a multilevel method. The smoother employs the stencil of the finite element matrix and is typically chosen as a polynomial of degree $q > 0$ of the original stiffness matrix. The overlap is then $\delta \sim qh$. The property that the coarse space represents constant functions, is ensured by exploiting the kernel of the original problem.

An aggregation technique was first introduced in [312] and then used quite extensively for the solution of problems arising in economics; see [339] and the references therein. Smoothed aggregation techniques have been considered in [453, 93] for two levels and in [454, 455, 242, 452] for multiple levels. Extensive studies are reported on certain smoothers and practical procedures proposed for the initial aggregation, i.e., the initial partition into subdomains. Numerical tests have also been performed on a large class of scalar and vector problems. We also mention [306], where smoothed aggregation techniques

are applied to discontinuous Galerkin approximations of advection-diffusion problems.

We will start from an initial partition $\mathcal{F}_0 = \{\Omega_i \mid 1 \leq i \leq N\}$ into nonoverlapping subdomains. We always assume that these nonoverlapping subdomains are shape regular, and that the diameter of each subdomain is on the order of $H$. The method considered will generate both an overlapping partition and coarse basis functions starting from a $\mathcal{F}_0$.

A first choice is to build a coarse space by *aggregation*. We define a set of finite element 'characteristic' functions relative to the initial nonoverlapping partition $\mathcal{F}_0$, $\{\Psi_i\}$ and consider the span of these functions. For every node $x$ of $\mathcal{T}_h$, we set

$$\Psi_i(x) = \begin{cases} 0, & x \in \Omega \setminus \overline{\Omega_i}, \\ \operatorname{card}(\{j \mid x \in \partial\Omega_j\})^{-1}, & x \in \partial\Omega_i, \\ 1, & x \in \Omega_i, \end{cases}$$

where $\operatorname{card}(M)$ denotes the cardinality of a finite set $M$.

We note, that, if the subdomain boundaries do not contain any nodes of the fine mesh $\mathcal{T}_h$, the value of these functions at the nodes is either zero or one, and that they decrease from one to zero across a strip of width $h$. In the general case, they assume values between zero and one, and they decrease from one to zero across a strip of width at most $2h$. Furthermore, the nonvanishing nodal values of $\Psi_i$ cannot be arbitrarily small, since the partition $\mathcal{F}_0$ is shape regular. We also note that these functions form a partition of unity on $\Omega$ except in a strip of width $O(h)$ along $\partial\Omega$.

The set $\{\Psi_i\}$ therefore satisfies Assumption 3.14 with $\delta = h$, and the resulting coarse space can be analyzed within the framework introduced in the previous section. However, the corresponding additive preconditioner would have an unsatisfactory bound of the condition number, which would increase quadratically with $H/h$. Therefore, the coarse functions $\{\Psi_i\}$ need to be smoothed to decrease their energies. In order to do so, we apply a suitable operator, called the smoother,

$$\Phi_i = S\Psi_i, \quad 1 \leq i \leq N.$$

This smoothing process has the effect of increasing the support of the original functions and of creating additional overlap between the supports of the functions. We will then define the overlapping subdomains by

$$\overline{\Omega_i'} = \operatorname{supp}\{\Phi_i\}, \tag{3.35}$$

and we obtain an overlapping partition $\mathcal{F} = \{\Omega_i' \subset \Omega \mid \quad 1 \leq i \leq N\}$.

The smoothing should also exploit the stencil of the operator $A$. If $S = p_q(DA)$, where $p_q$ is a polynomial of degree $q \geq 0$ and $D$ a diagonal matrix, then the support of the initial function $\Psi_i$ is increased by $q$ layers of fine elements, which gives an overlap of order $\delta = qh$. In addition, we need to preserve the property that the modified coarse space $\{\Phi_i\}$ contain constants.

This property is guaranteed by the null space of the original differential operator which consists of constant functions. We note that $A$, the representation of $a(\cdot,\cdot)$ on $V$, is not singular since homogeneous Dirichlet conditions are imposed on $\partial\Omega$, but that when applied to a constant vector, it produces a vector that vanishes everywhere except in a strip, of width $O(h)$, around $\partial\Omega$. If $p_q(0) = 1$, we can then write

$$\sum_{i=1}^{N} \Phi_i = S \sum_{i=1}^{N} \Psi_i = S\mathbf{1} = p_q(DA)\,\mathbf{1} = \mathbf{0}' + p_q(0)\mathbf{1} = \mathbf{1}',$$

where $\mathbf{1}$ is the vector of all ones, $(1,\ldots,1)^T$, while $\mathbf{0}'$ and $\mathbf{1}'$ are vectors of zeros and ones, respectively, except for entries at nodes in a neighborhood, of width $O(qh)$, of $\partial\Omega$. The smoothed coarse functions thus satisfy Assumption 3.14.3 with a $\delta$ on the order of $qh$.

In view of these remarks, we consider the following assumptions on the initial partition $\mathcal{F}_0$ and the smoother $S$.

**Assumption 3.21 (Initial partition and smoother)**

1. *The initial partition $\mathcal{F}_0$ satisfies*

$$cH^n \leq |\Omega_i| \leq CH^n. \tag{3.36}$$

2. *$S$ is equal to $p_q(DA)$, where $p_q$ is a polynomial of degree $q$ and $D$ a diagonal matrix, such that*
   *a) $c\delta \leq qh \leq C\delta \leq C'H$;*
   *b) $p_q(0) = 1$;*
   *c) $\|S\|_2 \leq 1$;*
   *d) $\varrho(S^T AS) \leq Cq^{-2}\varrho(A)$,*
   *where $\|\cdot\|_2$ and $\varrho(\cdot)$ denote the spectral norm and the spectral radius of a matrix, respectively.*

We note that Assumptions 3.21.2.c and 3.21.2.d were originally considered in [453, Lem. 2.8] and [93, Lem. 4.2] for the case when $S$ is a polynomial in $A$. An assumption similar to 3.21.2.a was stated in [93, Ass. 4.1], but in terms of the graph corresponding to the initial partition $\mathcal{F}_0$.

The proof of the following result is given in [305].

**Lemma 3.22** *Let Assumption 3.21 hold. Then, the coarse functions and the overlapping partition obtained by smoothed aggregation using $S$ satisfy Assumptions 3.14 and 3.15. Therefore, there exist positive constants $c$ and $C$, such that the corresponding additive operator satisfies*

$$c\left(1 + \frac{H}{\delta}\right)^{-2} a(u,u) \leq a(P_{ad}u, u) \leq Ca(u,u), \quad u \in V.$$

*Remark 3.23.* If $\|S\|_\infty < C$, then Assumption 3.18.1 holds, but a suitable condition on $S$ that ensures that Assumption 3.18.2 also holds appears to be difficult to find.

Examples of smoothers can be found in [93], [453], [454], [96], and [305]. We note that not all of the Assumptions 3.21.2 can be proved for some of them, but that they have shown comparable performance in some simple test cases; see [305]. We finally note that for linear elasticity problems, the coarse spaces need to be enriched since they need to contain the kernel of the relevant differential operator. We may consider the rigid body modes or even a basis for the space of first-order polynomials and multiply them by the partition of unity functions $\Phi_i$; see section 8.2. Several basis functions are now associated with each subdomain.

### 3.10.3 Partition of Unity Coarse Spaces

Although numerical results on simple test cases in [305] seem to suggest a linear growth, as a function of $H/\delta$, in the condition number when smoothed aggregation coarse spaces are employed, as in Theorem 3.19, for some of the smoothers Assumptions 3.21 do not hold. As was shown in [409], it is also possible to construct coarse basis functions $\{\Phi_i\}$ by directly assigning nodal values inside the subdomains of the overlapping partition $\mathcal{F}$.

As for the smoothed aggregation techniques, we suppose that $\mathcal{F}$ is obtained from an initial nonoverlapping partition $\mathcal{F}_0 = \{\Omega_i\}$ such that

$$cH^n \leq |\Omega_i| \leq CH^n \tag{3.37}$$

We suppose that each subdomain in $\mathcal{F}_0$ consists of elements of $\mathcal{T}_h$. For $i = 1, \ldots, N$, $\Omega_i'$ is obtained from $\Omega_i$ by adding $q$ layers of elements; also let $\Omega_i^0 = \Omega_i$. For $j \geq 1$, the domain $\Omega_i^j$ is obtained from $\Omega_i^{j-1}$ by adding those elements $k \in \mathcal{T}_h$ for which $\partial k \cap \partial\Omega_i^{j-1} \neq \emptyset$. We set $\Omega_i' = \Omega_i^q$ and $\delta = qh$.

We next consider a finite element function $\hat{\Phi}_i$ defined by its nodal values. We set $\hat{\Phi}_i(x) = 1$ if $x$ is a node in $\overline{\Omega_i}$, and $\hat{\Phi}_i(x) = 0$ if $x \in \Omega \setminus \Omega_i'$. For the other nodes, we set $\hat{\Phi}_i(x) = (q-j)/q$ if $x$ belongs to the $j$-th layer, i.e., $x \in \partial\Omega_i^j$. We finally set

$$\Phi_i = I^h \left( \frac{\hat{\Phi}_i}{\sum_{i=1}^N \hat{\Phi}_i} \right).$$

We have the following result; see [409, Th. 1].

**Lemma 3.24** *Let $\{\Phi_i\}$ be the partition of unity functions and let $\mathcal{F}$ be the overlapping partition defined above. Then, Assumptions 3.18 and 3.15 hold. Therefore, there exist positive constants $c$ and $C$, such that the corresponding additive operator satisfies*

$$c\left(1+\frac{H}{\delta}\right)^{-1} a(u,u) \leq a(P_{ad}u,u) \leq Ca(u,u), \quad u \in V.$$

We note that Assumptions 3.18.1 and 3.18.2 can be established in a way similar as in the proof of Lemma 3.4.

# 4

# Substructuring Methods: Introduction

## 4.1 Introduction

In this part of our monograph, which ends with Chap. 6, we will develop and analyze a large number of iterative substructuring methods, which are based on a nonoverlapping partition of the region $\Omega$ on which the original problem is defined. We will consider only scalar, second-order, self adjoint, elliptic problems on $\Omega \subset \mathbb{R}^3$, (and occasionally in $\mathbb{R}^2$) where $\Omega$ is a Lipschitz region of unit diameter. In fact, to make our analysis fully possible, we will make a further assumption, Assumption 4.3. In this chapter and the next, which, in large part, are based on Dryja, Smith, and Widlund [178], we will introduce the Schur complement systems to which the original finite element problems are reduced in a first direct elimination step and also a number of technical tools which are necessary in the analysis of the many iterative substructuring algorithms of this monograph.

Certain iterative substructuring methods have already been introduced in Sect. 1.3. They are based on nonoverlapping decompositions of the region into an often very large set of subdomains, also known as substructures, a term borrowed from structural engineering; see Przemieniecki [389]. We note that many finite element codes are based on similar decompositions, in fact several levels of nested subdivisions are frequently used in such a context. Direct elimination is often used in engineering computations in which the interior variables of elements and larger and larger macro elements, created by merging elements or smaller macro elements, are successively eliminated. We can think of iterative substructuring methods as halting this process at some stage and solving the remaining linear system by a preconditioned Krylov space method. If a parallel computer is used, one or several substructures will be allocated to an individual processor and the elimination of the interior unknowns of the substructures can then be carried out in parallel across the substructures. We note that this type of ordering is also beneficial in that it results in relatively sparse triangular factors; see [220, 221].

## 4.2 Problem Setting and Geometry

Let $\partial\Omega_D \subset \partial\Omega$ be a closed set of positive measure, and let $\partial\Omega_N := \partial\Omega \setminus \partial\Omega_D$ be its complement. We impose homogeneous Dirichlet and general Neumann boundary conditions, respectively, on these two subsets and introduce the Sobolev space $H_0^1(\Omega, \partial\Omega_D) := \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega_D\}$.

For simplicity, we will only consider in full detail the case of a piecewise linear, conforming finite element approximation of the following second order model problem: find $u \in H_0^1(\Omega, \partial\Omega_D)$, such that

$$a(u, v) = f(v) \quad v \in H_0^1(\Omega, \partial\Omega_D), \tag{4.1}$$

where

$$a(u, v) := \int_\Omega \rho(x)\nabla u \cdot \nabla v \, dx, \quad f(v) := \int_\Omega fv \, dx + \int_{\partial\Omega_N} g_N v \, ds. \tag{4.2}$$

Here $g_N$ is the Neumann boundary data defined on $\partial\Omega_N$; it provides, together with the volume load $f$, the contributions to the load vector of the finite element problem. We assume $\rho(x) \geq \rho_{\min} > 0$ for $x \in \Omega$. We note that we could add a zero order term, with a positive coefficient function, to our problem without any real changes in the performance of the algorithms and only minor changes in the theory.

The extension of our results to other conforming finite element methods is routine but we have to expect that the constants in our bounds might depend on the order of the elements. We will consider extensions to spectral finite elements in Chap. 7 and to elliptic systems in Chap. 8.

We decompose $\Omega$ into nonoverlapping subdomains $\Omega_i, i = 1, \ldots, N$, also known as substructures, and each of which is the union of shape-regular elements with the finite element nodes on the boundaries of neighboring subdomains matching across the interface. We will distinguish between subdomains with boundaries which intersect $\partial\Omega_D$ and the floating subdomains with boundaries which do not.

**Definition 4.1.** *A substructure $\Omega_i$ is a* floating subdomain *if the intersection of its boundary with $\partial\Omega_D$ is empty.*

The *interface*

$$\Gamma := \cup_{i \neq j} \partial\Omega_i \cap \partial\Omega_j$$

is the union of

- the interior subdomain *faces*, regarded as open sets, that are shared by two subregions,
- the subdomain *edges*, also regarded as open, that are shared by more than two subregions,
- *vertices*, which are endpoints of edges.

If there is an edge which is a part of $\partial\Omega_N$ and which is common to the boundaries of only two subdomains, we will regard that edge as a part of the face common to this pair of subdomains, while any edge common to at least three subdomains will be included in the set of interface edges. Similarly, we will regard a subdomain vertex on $\partial\Omega_N$ part of an interior edge unless there are several edges that end at the vertex. In the latter case, we will treat the vertex the same as a vertex in the interior of $\Omega$. The subdomain faces which belong to $\partial\Omega$ are not part of the interface $\Gamma$; the nodal values on $\partial\Omega_D$ will always vanish and those on $\partial\Omega_N$ which belong to only one subdomain will effectively belong to the subdomain interior. They will be eliminated together with the interior degrees of freedom when the given linear system is reduced to a Schur complement system involving only the unknowns of the interface $\Gamma$.

We will denote the faces of $\Omega_i$ by $\mathcal{F}^{ij}$, its edges by $\mathcal{E}^{ik}$, its vertices by $\mathcal{V}^{i\ell}$, and its wire basket, i.e., the union of its edges and vertices, by $\mathcal{W}^i$. Occasionally, we will also use faces, edges, and vertices with just one superscript and even drop the superscripts all together.

Let $\mathcal{T}_i$ be the triangulation of subdomain $\Omega_i$, of diameter $h_i$, and let $\mathcal{T} = \mathcal{T}_h$ be the geometrically conforming global mesh on $\Omega$, of diameter $h = \max\{h_i\}$; cf. appendix B.1.1. We denote the standard finite element space of continuous, piecewise linear functions on $\mathcal{T}_i$ by $V^h(\Omega_i)$; we always assume that these functions vanish on $\partial\Omega_i \cap \partial\Omega_D$. We assume that the triangulation of each subdomain is shape regular and quasi uniform. The diameter of $\Omega_i$ is $H_i$, with $H = \max\{H_i\}$. The finite element approximation of the elliptic problem is continuous across $\Gamma$ and we denote the corresponding space by $V^h = V^h(\Omega)$.

**Definition 4.2.** *The sets of nodes on $\Gamma$, $\partial\Omega$, and $\partial\Omega_i$ are denoted by $\Gamma_h$, $\partial\Omega_h$, and $\partial\Omega_{i,h}$, respectively. Analogously, $\mathcal{F}_h^{ij}$, $\mathcal{E}_h^{ik}$, and $\mathcal{W}_h^i$ denote the sets of nodes on a face $\mathcal{F}^{ij}$, an edge $\mathcal{E}^{ik}$, and the wire basket $\mathcal{W}^i$, respectively. Since the faces and edges are open sets, $\mathcal{F}_h^{ij}$ and $\mathcal{W}_h^i$ are disjoint and so are $\mathcal{F}_h^{ij}$ and $\mathcal{E}_h^{ik}$.*

We assume that all possible large jumps of $\rho(x)$ are aligned with the subdomain boundaries and, for simplicity, that on each subregion $\Omega_i$, $\rho(x)$ has a constant value $\rho_i > 0$. Our bilinear form and load vector can then be written in terms of contributions from individual subregions. With

$$a^{(i)}(u,v) = \rho_i \int_{\Omega_i} \nabla u \cdot \nabla v \, dx, \quad i = 1, \ldots, N,$$

we have

$$a(u,v) = \sum_{i=1}^{N} a^{(i)}(u,v), \quad f(v) = \sum_{i=1}^{N} \Big( \int_{\Omega_i} fv \, dx + \int_{\partial\Omega_i \cap \partial\Omega_N} g_N v \, ds \Big). \tag{4.3}$$

We make the following assumption.

**Assumption 4.3** *The partition into subdomains is such that:*

1. *Each subregion $\Omega_i$ is the union of shape-regular coarse tetrahedral elements of a global conforming mesh $\mathcal{T}_H$ and the number of such tetrahedra forming an individual subdomain is uniformly bounded.*
2. *If a face of a subdomain intersects $\partial\Omega_D$, then the measure of this set is comparable to that of $\partial\Omega_i$. Similarly, if an edge of a subdomain intersects $\partial\Omega_D$, the length of this intersection is bounded from below in terms of the diameter of $\Omega_i$.*

*Remark 4.4.* Assumption 4.3.1, which basically ensures that the subregions cannot be very thin, is necessary for our theory of all iterative substructuring methods presented in this monograph. It can easily be shown that the diameters of any pair of neighboring subdomains are comparable. Assumption 4.3.2, on the other hand, appears to be necessary only for balancing Neumann-Neumann methods (see Sect. 6.2) and one-level FETI methods (see Sect. 6.3).

*Remark 4.5.* The theory of iterative substructuring methods does not cover the case of many partitions into substructures that are commonly used in practice such as those obtained using mesh partitioners, where fine elements are aggregated to form subregions; cf. Figure 4.1 for an example. In such cases, the boundaries of the substructures may be very irregular and faces (consisting of points that belong to two substructures) and edges (curves that belong to more than two substructures) are not plane surfaces or straight lines. We note that even in such a case, we can define faces, edges, and vertices of the interface relatively easily. We first associate with each node $x$ of the triangulation the set $\mathcal{N}_x$ of indices of the $\overline{\Omega_i}$ to which it belongs. A node $x$ belongs to a face if $\mathcal{N}_x$ has exactly two elements and $y$ belongs to the same face if $\mathcal{N}_x = \mathcal{N}_y$. Similarly, $x$ is a vertex if no other node has the same index set while $x$ is an edge node if $\mathcal{N}_x$ has at least three elements and there exists at least one additional node $y$ with $\mathcal{N}_y = \mathcal{N}_x$.

We will occasionally use a piecewise linear finite element space $V^H$ on the triangulation consisting of the coarse elements from which the subdomains are composed; there is no need to assume that this coarse triangulation is quasi uniform. We note that many of the estimates in this and the following chapters are given in terms of the ratio $H/h$. Since our analysis will typically be carried out locally for one substructure (and its next neighbors) at a time, we can interpret the factor $H/h$ as $\max_i\{H_i/h_i\}$. The quantity $(H_i/h_i)^n$, $\Omega \subset \mathbb{R}^n$, gives a measure of the number of degrees of freedom associated with $\Omega_i$.

The Sobolev space $H^1(\Omega)$ is closely related to our family of elliptic problems. This space and its norm are defined in appendix A.1 for the case of a region with unit diameter. In the case of a region of diameter $H_i$, such as the substructure $\Omega_i$, we use a norm with different relative weights obtained by a simple dilation argument:

$$\|u\|^2_{H^1(\Omega_i)} = |u|^2_{H^1(\Omega_i)} + \frac{1}{H_i^2}\|u\|^2_{L^2(\Omega_i)}. \tag{4.4}$$

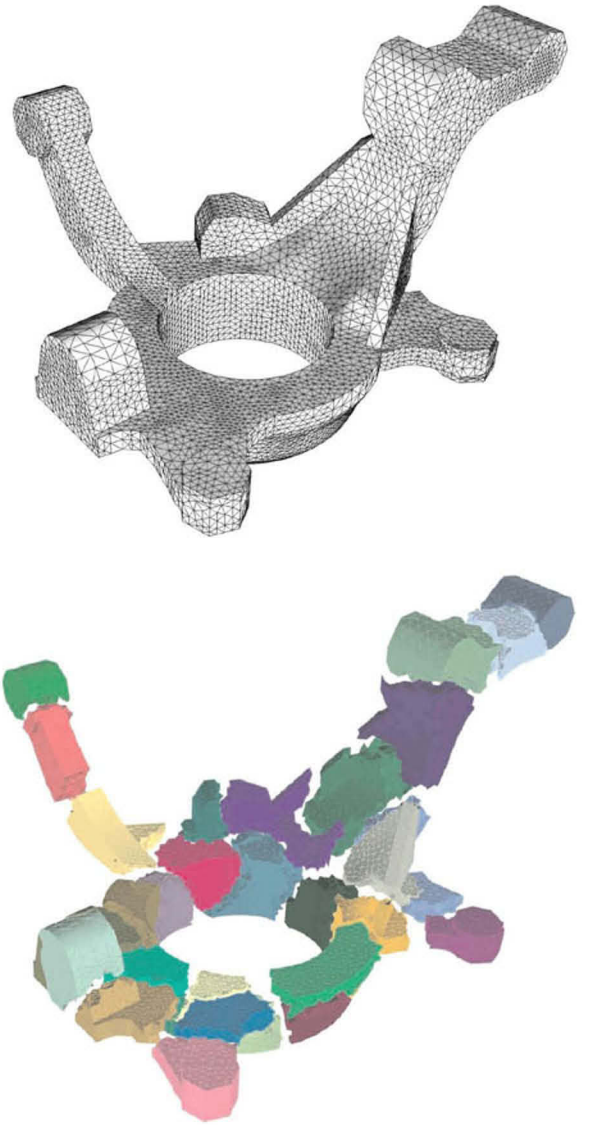

**Fig. 4.1.** Finite element meshing of a mechanical object (*top*) and partition into thirty subdomains (*bottom*). Courtesy of Charbel Farhat.

We will also need to work with traces of functions on the boundaries of the substructures. The trace space $H^{1/2}(\partial\Omega_i)$ is equipped with the scaled norm:

$$\|u\|^2_{H^{1/2}(\partial\Omega_i)} = |u|^2_{H^{1/2}(\partial\Omega_i)} + \frac{1}{H_i}\|u\|^2_{L^2(\partial\Omega_i)}; \tag{4.5}$$

cf. also appendix A.2.

From now on, we will denote the finite element solution by $u$. The discrete problem corresponding to (4.3) is of the form: find $u \in V^h(\Omega)$ such that

$$a(u,v) = f(v) \quad v \in V^h(\Omega). \tag{4.6}$$

We recall that in most discussions of domain decomposition methods, there are technically two spaces: the space of finite element functions $V^h$, and the space of coefficients (degrees of freedom) of the finite element functions; we will make no distinction as far as notations are concerned. Matrix and vector notations for finite element problems were already introduced in Sect. 1.2; see also appendix B.6. Here we will extend our discussion to the case of many subdomains. The contributions to the stiffness matrix $A$ and the right hand side $f$ can be formed one subdomain at a time. The stiffness matrix is then obtained by subassembling these parts. We will order the nodes *interior* to the subdomains first, followed by those on the *interface* $\Gamma$. More precisely, we define as interior degrees of freedom all those that are not on $\Gamma$. Thus, they consist of those that lie in the interior of the subdomains but also of those on faces of $\partial\Omega_N$, in accordance with the precise definition of $\Gamma$ given at the beginning of this section. All the matrices and vectors are expanded by zeros, giving them each the same dimension as the global stiffness matrix and the vector of unknowns; we could also work with the extension and restriction operators of Chap. 1, but from now on, we will often adopt a less cumbersome, shorthand notation. We will otherwise use many of the same notations as in that chapter, e.g., when working with stiffness matrices attributable to subregions, their Schur complements, etc.

We can then write the linear system as

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \sum_{j=1}^{N} \begin{pmatrix} A_{II}^{(j)} & A_{I\Gamma}^{(j)} \\ A_{\Gamma I}^{(j)} & A_{\Gamma\Gamma}^{(j)} \end{pmatrix} \begin{pmatrix} u_I^{(j)} \\ u_\Gamma^{(j)} \end{pmatrix} = \sum_{j=1}^{N} \begin{pmatrix} f_I^{(j)} \\ f_\Gamma^{(j)} \end{pmatrix}. \tag{4.7}$$

Thus, to multiply $A$ by the vector $u$, we first restrict the vectors $u_I$ and $u_\Gamma$ to the substructures, then multiply them by the individual substructure stiffness matrices and, finally, obtain the product $Au$ by padding with zeros and adding the resulting vectors.

The following lemma relates the energy of a finite element function defined on a substructure and the norm of its trace. We note that we use the scaled norms introduced in formulas (4.4) and (4.5).

**Lemma 4.6** *There exists a constant $C_t$, depending only on the shape of $\Omega_i$, such that*

$$\|u\|_{H^{1/2}(\partial\Omega_i)} \leq C_t \|u\|_{H^1(\Omega_i)}, \quad u \in V^h(\Omega_i). \tag{4.8}$$

*In addition, for every $\phi$, which is the trace of a finite element function on $\partial\Omega_i \cap \Gamma$, there exists an extension $u \in V^h(\Omega_i)$, which depends linearly on $\phi$, such that*

$$\|u\|_{H^1(\Omega_i)} \leq C_e \|\phi\|_{H^{1/2}(\partial\Omega_i \cap \Gamma)}. \tag{4.9}$$

*Here $C_e$ depends only on the shape of $\Omega_i$ and the aspect ratio and the quasi-uniformity constant of $\mathcal{T}_i$. The same bounds hold for the $H^1$- and $H^{1/2}$-seminorms provided that the subregion is floating or at least one of its faces intersects the set $\partial\Omega_D$ and the intersection has a measure that is comparable to that of $\partial\Omega_i$.*

*Proof.* The trace estimate (4.8) can be found in Lemma A.6 for regions of unit diameter. The bound involving the scaled norms can be found by a simple rescaling argument, and those for the seminorms by a quotient space argument; we first drop the $L^2$-term on the left hand side of the inequality and note that the remaining seminorm is invariant if a constant is added to $u$. Finally, we use a Poincaré or Friedrichs inequality, see Corollary A.15, to eliminate the $L^2$-term on the right hand side.

As far as (4.9) is concerned, for simplicity, we only give a proof for the case of a substructure that does not touch $\partial\Omega$: $\partial\Omega_i \cap \Gamma = \partial\Omega_i$. We first consider a region $\Omega_i$ of unit diameter and the harmonic extension $v \in H^1(\Omega_i)$ of the boundary data $\phi$, i.e., the solution of the Laplace equation with a Dirichlet boundary condition given by $\phi$. We note that since a step function belongs to $H^s$, $s < 1/2$, the continuous, piecewise linear function $\phi$ belongs to $H^{1+s}(\partial\Omega_i)$ for $s < 1/2$. Therefore, since $\Omega_i$ is a Lipschitz polyhedron, Lemma A.49 ensures that $v \in H^{s+3/2}(\Omega_i)$, for every $s < s_\Omega$, with $s_\Omega > 0$. Thus, by Lemma A.5, the harmonic function $v$ is continuous and the standard Lagrange interpolation operator $I^h$ can be used for both two and three dimensions and for any Lagrangian finite element, resulting in a finite element function $I^h v$; see appendix B.1.2. By using the bound on the interpolation error in Lemma B.6 and the regularity result in Lemma A.49, we now obtain, for $s \in (0, s_\Omega)$,

$$\begin{aligned} |I^h v|_{H^1(\Omega_i)} &\leq |I^h v - v|_{H^1(\Omega_i)} + |v|_{H^1(\Omega_i)} \\ &\leq C h^{s+1/2} |v|_{H^{s+3/2}(\Omega_i)} + C|\phi|_{H^{1/2}(\partial\Omega_i)} \\ &\leq C h^{s+1/2} |\phi|_{H^{s+1}(\partial\Omega_i)} + C|\phi|_{H^{1/2}(\partial\Omega_i)}. \end{aligned}$$

The proof is completed by using an inverse inequality; see Lemma B.27.

Bounds for the scaled norms and seminorms can be found in the same way as before, i.e., by a scaling argument and the use of a Poincaré or Friedrichs inequality. □

*Remark 4.7.* In case $\partial\Omega_i$ intersects $\partial\Omega_D$ but not $\partial\Omega_N$, the $H^{1/2}(\partial\Omega_i)$ and $H^{1/2}(\partial\Omega_i \cap \Gamma)$ norms in (4.8) and (4.9), respectively, can be replaced by the $H^{1/2}_{00}(\partial\Omega_i \cap \Gamma)$ norm. For the case of a substructure that touches $\partial\Omega_N$, a

regularity result for the Laplace problem with boundary conditions of mixed Dirichlet/Neumann type needs to be employed, therefore requiring additional assumptions on $\partial\Omega_i$ (such as convexity, for instance); see [143] for a two-dimensional result.

*Remark 4.8.* The proof of Lemma 4.9 can be extended directly to any Hermite finite element method. In the proof, we just modify the standard interpolation formula by setting all derivative degrees of freedom equal to zero; all that is required from the interpolation formula is that it is stable and that it reproduces constants.

We note that there are other proofs of the main part of Lemma 4.6; cf. [61, 463, 76].

## 4.3 Schur Complement Systems

As in the previous section, we will tacitly assume that vectors defined only on subdomains are extended by zero elsewhere.

In a first step of many iterative substructuring algorithms, the unknowns in the interior of the subdomains are eliminated; cf. Sect. 1.3. In this step, the Schur complements, with respect to the variables associated with the individual substructures and the nodes on $\partial\Omega_i \cap \Gamma$ are introduced. The resulting linear system can be written as

$$\begin{aligned} \begin{pmatrix} A_{II} & A_{I\Gamma} \\ 0 & S \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} \equiv \sum_{j=1}^{N} \begin{pmatrix} A_{II}^{(j)} & A_{I\Gamma}^{(j)} \\ 0 & S_{\Gamma\Gamma}^{(j)} \end{pmatrix} \begin{pmatrix} u_I^{(j)} \\ u_\Gamma^{(j)} \end{pmatrix} = \\ \sum_{j=1}^{N} \begin{pmatrix} f_I^{(j)} \\ f_\Gamma^{(j)} - A_{\Gamma I}^{(j)} {A_{II}^{(j)}}^{-1} f_I^{(j)} \end{pmatrix} \equiv \begin{pmatrix} \tilde{f}_I \\ \tilde{f}_\Gamma \end{pmatrix}, \end{aligned} \tag{4.10}$$

where

$$S^{(j)} = S_{\Gamma\Gamma}^{(j)} = A_{\Gamma\Gamma}^{(j)} - A_{\Gamma I}^{(j)} {A_{II}^{(j)}}^{-1} A_{I\Gamma}^{(j)},$$

and the reduced system is given by

$$S u_\Gamma = \tilde{f}_\Gamma. \tag{4.11}$$

Thus, the matrix $S$ is obtained from the $S^{(j)}$ by subassembly. In practice, the matrix $S$ is often not formed explicitly, since this is a potentially expensive operation. Instead, a sparse representation of the $A_{I\Gamma}^{(j)}$ and the sparse, triangular factors of the $A_{II}^{(j)}$ are stored, and the action of $S$ on a vector is calculated as needed. We note, in particular, that the application of the inverse of $A_{II}^{(j)}$ to a vector corresponds to the solution of a Dirichlet problem on $\Omega_j$ or to a problem with Dirichlet boundary conditions on $\partial\Omega_j \cap \Gamma$ and with homogeneous Neumann data on $\partial\Omega_j \cap \partial\Omega_N$. Similarly, many iterative substructuring

preconditioners are built from inverses of certain Schur complements. These inverses are normally not calculated in practice. Instead, in order to apply the inverse of $S^{(j)}$ to a vector $w_\Gamma^{(j)}$, we can solve a Neumann problem on $\Omega_j$. Indeed, it is easy to show that the vector

$$u_\Gamma^{(j)} = {S^{(j)}}^{-1} w_\Gamma^{(j)}$$

can be found by solving the local system

$$A^{(j)} \begin{pmatrix} u_I^{(j)} \\ u_\Gamma^{(j)} \end{pmatrix} = \begin{pmatrix} 0 \\ w_\Gamma^{(j)} \end{pmatrix}.$$

The matrices $S$ and $S^{(j)}$ can also be viewed as operators defined on interface functions on $\Gamma$ and $\partial\Omega_j \cap \Gamma$, respectively. They can be represented as block matrices with a block for each face, edge, and vertex of $\Gamma$. We will often combine all the edge and vertex blocks of a subregion $\Omega_j$ into a single wire basket block. (We can also merge them all into a single block corresponding to the wire basket of the entire interface.) We then obtain

$$S^{(j)} = \begin{pmatrix} S_{\mathcal{FF}}^{(j)} & S_{\mathcal{FE}}^{(j)} & S_{\mathcal{FV}}^{(j)} \\ {S_{\mathcal{FE}}^{(j)}}^T & S_{\mathcal{EE}}^{(j)} & S_{\mathcal{EV}}^{(j)} \\ {S_{\mathcal{FV}}^{(j)}}^T & {S_{\mathcal{EV}}^{(j)}}^T & S_{\mathcal{VV}}^{(j)} \end{pmatrix} \tag{4.12}$$

and

$$S^{(j)} = \begin{pmatrix} S_{\mathcal{FF}}^{(j)} & S_{\mathcal{FW}}^{(j)} \\ {S_{\mathcal{FW}}^{(j)}}^T & S_{\mathcal{WW}}^{(j)} \end{pmatrix}, \tag{4.13}$$

respectively. Here $S_{\mathcal{FF}}^{(j)}$ is constructed from the blocks that correspond to the individual faces, and to pairs of faces, of $\Omega_j$, etc. We will use both block structures, (4.12) and (4.13), in the description of different algorithms, as appropriate.

All of the algorithms that we consider can be reformulated using inexact interior solvers. We explain briefly how this can be done. After a block Cholesky factorization of $A$, we find that the exact inverse $A^{-1}$ can be written as

$$A^{-1} = \begin{pmatrix} I & -A_{II}^{-1}A_{I\Gamma} \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{II}^{-1} & 0 \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{\Gamma I}A_{II}^{-1} & I \end{pmatrix}.$$

If we have a good preconditioner for $S$, $B_S^{-1}$, and a good preconditioner for $A_{II}$, i.e., an approximate solver $B_{II}^{-1}$ for the interior problems, we can create a preconditioner for $A$ of the form

$$B_A^{-1} = \begin{pmatrix} I & -B_{II}^{-1}A_{I\Gamma} \\ 0 & I \end{pmatrix} \begin{pmatrix} B_{II}^{-1} & 0 \\ 0 & B_S^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{\Gamma I}B_{II}^{-1} & I \end{pmatrix}.$$

We see, by examining the resulting formulas carefully, that an application of $B_A^{-1}$ to a vector need only involve $B_{II}^{-1}$ in two matrix vector multiplies, and

$B_S^{-1}$ in only one; if $f_I$ is the first component of the vector on which $B_A^{-1}$ operates, we can save $B_{II}^{-1} f_I$ and reuse it when multiplying with the second block matrix. It is also possible to use different approximate interior solvers in the three factors of $B_A^{-1}$ and to construct nonsymmetric preconditioners of a similar form.

In the analysis presented here, we will always require exact interior subdomain solvers. Algorithms that use approximate interior solvers have been analyzed, e.g., in Börgers [63], Bramble, Pasciak, and Vassilev [77], and Haase, Langer, and Meyer [250]. Since it is important to first fully understand the case when exact interior solvers are used, we will focus on that case. We can then exclusively work with the space of discrete harmonic functions $\tilde{V}^h$ and the bilinear form $s(\cdot,\cdot)$; cf. Sect. 4.4. Numerical experiments, cf. Börgers [63], Haase, Langer, and Meyer [250], Skogen [419], and Smith [423], indicate that a good rate of convergence can be maintained, e.g., when one multigrid V-cycle is used, instead of an exact solver, to solve the interior problems.

## 4.4 Discrete Harmonic Extensions

The space of *discrete harmonic functions*, is an important subspace which is directly related to the Schur complements and to the values at the nodes on $\Gamma$.

A function $u^{(i)}$ defined on $\Omega_i$ is said to be discrete harmonic on $\Omega_i$ if

$$A_{II}^{(i)} u_I^{(i)} + A_{I\Gamma}^{(i)} u_\Gamma^{(i)} = 0. \tag{4.14}$$

It is easy to see that $u^{(i)} =: \mathcal{H}_i(u_\Gamma^{(i)})$ is completely defined by its values on $\partial\Omega_i \cap \Gamma$ and that it is orthogonal, in the $a_i(\cdot,\cdot)$-inner product, to the space $V^h \cap H_0^1(\Omega_i, \partial\Omega_i \cap \Gamma)$.

The space $\tilde{V}^h \subset V^h$ of global, piecewise discrete harmonic functions consists of functions that are discrete harmonic on each substructure. A function $u$ belongs to $\tilde{V}^h$ if and only if $A_{II}u_I + A_{I\Gamma}u_\Gamma = 0$ and it is completely defined by its values on $\Gamma$. The subspace $\tilde{V}^h$ is orthogonal, in the $a(\cdot,\cdot)$ inner product, to all the interior spaces $V^h \cap H_0^1(\Omega_i, \partial\Omega_i \cap \Gamma)$, $i = 1,\ldots,N$; such a space can also be regarded as the subspace of functions of $H^1(\Omega)$ which vanish on $\Omega \setminus \Omega_i$. We denote the piecewise discrete harmonic extension of $u_\Gamma$ by $\mathcal{H}(u_\Gamma)$.

In the analysis to be given, the inner product defined by the Schur complement $S$ will be important; our preconditioners will be defined with respect to the inner product

$$s(u,v) = u_\Gamma^T S v_\Gamma. \tag{4.15}$$

It follows immediately from the definition of $S$ that $s(\cdot,\cdot)$ is symmetric and coercive; we note that $\lambda_{\min}(S) \geq \lambda_{\min}(A)$.

The following lemma results from elementary variational arguments. It is enough to write the energy ${u^{(i)}}^T A^{(i)} u^{(i)}$ using the $2 \times 2$ block structure in (4.7) and the definition of the discrete harmonic extension in (4.14).

**Lemma 4.9** *Let $u_\Gamma^{(i)}$ be the restriction of a finite element function to $\partial\Omega_i \cap \Gamma$. Then, the discrete harmonic extension $u^{(i)} = \mathcal{H}_i(u_\Gamma^{(i)})$ of $u_\Gamma^{(i)}$ into $\Omega_i$ satisfies*

$$a_i(u^{(i)}, u^{(i)}) = \min_{v^{(i)}|_{\partial\Omega_i\cap\Gamma} = u_\Gamma^{(i)}} a_i(v^{(i)}, v^{(i)})$$

*and*

$$u_\Gamma^{(i)^T} S^{(i)} u_\Gamma^{(i)} = a_i(u^{(i)}, u^{(i)}).$$

*Analogously, if $u_\Gamma$ is the restriction of a finite element function to $\Gamma$, the piecewise discrete harmonic extension $u = \mathcal{H}(u_\Gamma)$ of $u_\Gamma$ into the interior of the subdomains satisfies*

$$a(u,u) = \min_{v|_\Gamma = u_\Gamma} a(v,v)$$

*and*

$$s(u,u) = u_\Gamma^T S u_\Gamma = a(u,u).$$

This lemma ensures that instead of working with functions defined on the interface $\Gamma$, we can consider the corresponding discrete harmonic extensions.

The next lemma ensures that we can equivalently work with norms of local discrete harmonic extensions and traces on the subdomain boundaries. It is a direct consequence of Lemmas 4.6 and 4.9. We note that the result relies on the existence of a stable finite element extension into the subdomains.

**Lemma 4.10** *Let $u$ be discrete harmonic. Then, there exist positive constants $c$ and $C$, independent of $h$ and $H$, such that*

$$c\,\|u_\Gamma\|^2_{H^{1/2}(\partial\Omega_i\cap\Gamma)} \leq \|u\|^2_{H^1(\Omega_i)} \leq C\,\|u_\Gamma\|^2_{H^{1/2}(\partial\Omega_i\cap\Gamma)},$$
$$c\,|u_\Gamma|^2_{H^{1/2}(\partial\Omega_i\cap\Gamma)} \leq |u|^2_{H^1(\Omega_i)} \leq C\,|u_\Gamma|^2_{H^{1/2}(\partial\Omega_i\cap\Gamma)}.$$

*Consequently,*

$$c\,\rho_i |u_\Gamma|^2_{H^{1/2}(\partial\Omega_i\cap\Gamma)} \leq u_\Gamma^{(i)^T} S^{(i)} u_\Gamma^{(i)} \leq C\,\rho_i |u_\Gamma|^2_{H^{1/2}(\partial\Omega_i\cap\Gamma)},$$

*with $u_\Gamma^{(i)}$ the restriction of $u$ to $\partial\Omega_i \cap \Gamma$ and the constants independent of $h$, $H$, and the $\rho_i$.*

## 4.5 Condition Number of the Schur Complement

In this section, we prove an estimate for the condition number of the Schur complement $S$; our discussion follows Brenner [85]. We note that, it follows from the discussion on page 17, that we must expect the condition number to grow at least as $H^{-2}$; if an unpreconditioned conjugate gradient method is applied to the linear system, the sparsity pattern of $S$ makes exchange of

information possible only between neighboring subdomains. However, in this case, the condition number turns out to be even worse.

We assume that $\rho_i = 1$, $i = 1, \ldots, N$. In addition, as in the case of the stiffness matrix $A$, precise estimates in terms of $h$ and $H$ can be found most easily for quasi-uniform triangulations $\mathcal{T}_h$ and subdomain partitions, $\mathcal{T}_H$. For simplicity, we also only consider the case of homogeneous Dirichlet boundary conditions on the whole $\partial\Omega = \partial\Omega_D$. We first introduce an equivalent $L^2$-norm on $\Gamma$:

$$\|u\|_\Gamma^2 = \sum_{i=1}^{N} \|u\|_{L^2(\partial\Omega_i)}^2.$$

**Lemma 4.11** *Let $u$ be the trace of a finite element functions in $V^h$ on $\Gamma$. Assume that $\rho_i = 1$, $i = 1, \ldots, N$, and that the fine mesh $\mathcal{T}_h$ and the coarse partition $\mathcal{T}_H$ are quasi uniform. Then, there exist two positive constants $c$ and $C$, independent of $h$ and $H$, such that*

$$c\,H\|u\|_\Gamma^2 \leq s(u,u) \leq C\,h^{-1}\|u\|_\Gamma^2.$$

*Thus,*

$$\kappa(S) \leq \frac{C}{Hh}.$$

*Proof.* We employ the discrete harmonic extension, still denoted by $u$, into the interior of all the $\Omega_i$. We first consider the lower bound. Using (4.5) and (4.8), we find that

$$H_i\|u\|_{L^2(\partial\Omega_i)}^2 \leq H_i^2\|u\|_{H^{1/2}(\partial\Omega_i)}^2 \leq C_t^2\left(H_i^2|u|_{H^1(\Omega_i)}^2 + \|u\|_{L^2(\Omega_i)}^2\right).$$

Since $\Omega$ has unit diameter, and thus $H_i \leq 1$, summing over the subdomains yields

$$H\|u\|_\Gamma^2 \leq C\,\|u\|_{H^1(\Omega)}^2.$$

The proof of the lower bound is concluded by using a Friedrichs inequality, Lemma A.14, for the discrete harmonic function $u \in H_0^1(\Omega)$ and Lemma 4.9.

In order to prove the upper bound, we use Lemma 4.10 and find

$$s(u,u) = |u|_{H^1(\Omega)}^2 \leq C\sum_{i=1}^{N} |u|_{H^{1/2}(\partial\Omega_i)}^2.$$

The use of an inverse estimate (see Lemma B.27) completes the proof of the upper bound.

In order to prove the bound for the condition number, it is enough to use the estimates

$$c\,h^{n-1}u^Tu \leq \|u\|_\Gamma^2 \leq C\,h^{n-1}u^Tu, \quad \Omega \subset \mathbb{R}^n;$$

see Lemma B.5. □

We note that the estimates given in Lemma 4.11 are sharp; see Brenner [85].

## 4.6 Technical Tools

A number of additional auxiliary results are needed for the analysis of the iterative substructuring and other domain decomposition algorithms. They are mainly based on discrete Sobolev type inequalities for finite element functions. These results will be used extensively in Chap. 5 and 6 as well as in the rest of this monograph; they are crucial in establishing Assumptions 2.2 and 2.4 for a variety of iterative substructuring methods. The relevant norms and seminorms have been introduced in (4.4) and (4.5); these norms contain large multiples of $L^2$-norms.

The analysis and development of efficient iterative substructuring methods in two dimensions mainly rely on three tools. The first is the construction of a stable discrete extension of trace functions from the boundaries of the substructures, which ensures the uniform stability of discrete harmonic extensions; see Lemmas 4.10 and 4.9. The second is the $L^\infty$-bound of Lemma 4.15. Finally, we need a bound for functions associated with single edges, which is the analog of the three-dimensional result given in Lemma 4.24.

We have chosen to present full proofs only for the three-dimensional algorithms. We note that the second bound, just mentioned, ensures the almost uniform stability of the standard coarse interpolant in two dimensions, while the third ensures that a decomposition into functions associated with single edges is almost uniformly stable; see Remarks 4.13 and 5.4. Such estimates are also necessary for the results of Chap. 6, where, in two dimensions, local functions are decomposed into terms associated with substructure edges and vertices; cf. Lemmas 6.3, 6.34, and 6.36.

Similar tools are needed for three-dimensional problems. We need the stability of discrete harmonic extensions and a bound for functions associated with single faces; see Lemma 4.24. Bounds for the $L^\infty$-norm are replaced by a bound involving the $L^2$-norm over single edges or the wire basket. These bounds ensure that a decomposition of local functions into wire basket (edge and vertex) and face components is almost uniformly stable. This is crucial in Sect. 5.4.2 and 5.4.3, and in the proofs of Lemmas 6.3, 6.34, and 6.36 in Chap. 6.

### 4.6.1 Interpolation into Coarse Spaces

The following lemma illustrates the limitations of the interpolation operator $I^H : V^h \to V^H$, in three dimensions. $I^H u$ is the result of piecewise linear interpolation of the finite element function $u$ onto the coarse space $V^H$, built on the coarse triangulation $\mathcal{T}_H$; see appendix B.1.2. In order to prove this lemma, we need an inverse inequality: with $K \in \mathcal{T}_H$,

$$\|u\|^2_{L^\infty(K)} \leq C(1/h)\|u\|^2_{H^1(K)}; \tag{4.16}$$

see Bramble and Xu [81, Lemma 2.3]. Since equality holds for any nodal basis function, we find that this bound is sharp and so are the bounds of the following lemma. The constant $C$ is independent of $H$ and $h$.

**Lemma 4.12** *In three dimensions, for* $u \in V^h(\Omega_i)$,

$$||u - I^H u||^2_{L^2(\Omega_i)} \leq C(H/h)H^2|u|^2_{H^1(\Omega_i)},$$

*and*

$$|I^H u|^2_{H^1(\Omega_i)} \leq C(H/h)|u|^2_{H^1(\Omega_i)}.$$

*Proof.* Let $K \in \mathcal{T}_H$, with $K \subset \Omega_i$, and let $\{\phi_1, \ldots, \phi_4\}$ be the nodal basis functions associated with the nodes $\{P_i\}$ of $K$. Lemma B.5 ensures that

$$|\phi_i|^2_{H^1(K)} \leq CH, \quad \|\phi_i\|^2_{L^2(K)} \leq CH^3$$

and we find,

$$|I^H u|^2_{H^1(K)} \leq C \sum_{i=1}^{4} |u(P_i)|^2 |\phi_i|^2_{H^1(K)}.$$

Using (4.16), we obtain

$$|I^H u|^2_{H^1(K)} \leq C\frac{H}{h}\|u\|^2_{H^1(K)}.$$

Summing over $K \subset \Omega_i$ and using Assumption 4.3 gives

$$|I^H u|^2_{H^1(\Omega_i)} \leq C\frac{H}{h}\|u\|^2_{H^1(\Omega_i)}.$$

The norm on the right hand side can be replaced by the seminorm by using the fact that the left hand side does not change if we add any constant to $u$ and by using a Poincaré or Friedrichs inequality, just as in the proof of Lemma 4.6.

Proceeding similarly, we can prove

$$||u - I^H u||^2_{L^2(\Omega_i)} \leq 2||u||^2_{L^2(\Omega_i)} + 2||I^H u||^2_{L^2(\Omega_i)} \leq C\frac{H}{h}H^2\|u\|^2_{H^1(\Omega_i)}.$$

As before, the norm on the right hand side can be replaced by the seminorm by using the same arguments. □

*Remark 4.13.* The situation is much more favorable in two dimensions. In fact, we have

$$\|u\|^2_{L^\infty(\Omega_i)} \leq C\,(1 + \log(H/h))\,\|u\|^2_{H^1(\Omega_i)},$$

see Lemma 4.15 below, which gives the more benign bounds

$$||u - I^H u||^2_{L^2(\Omega_i)} \leq C\,(1 + \log(H/h))\,H^2|u|^2_{H^1(\Omega_i)},$$
$$|I^H u|^2_{H^1(\Omega_i)} \leq C\,(1 + \log(H/h))\,|u|^2_{H^1(\Omega_i)}.$$

The next lemma concerns an operator for which the bounds are much improved in comparison with Lemma 4.12. We note that the norms are now given in terms of the entire region $\Omega$. In fact, it is not possible to provide bounds as in this lemma for the $H^1$- and $L^2$-norms, weighted by the values $\rho_i$ of the coefficient of the elliptic problem as in (4.3), if we require the constants in the estimates to be independent of the $\rho_i$; see Xu [471]. This fact helps explain why we will resort to quite exotic coarse spaces when designing iterative substructuring methods in Chap. 5 and 6 since our ambition is to obtain algorithms which perform well even for large jumps in the coefficients across the interface $\Gamma$. A proof of Lemma 4.14 and a general discussion are also given in Bramble and Xu [81].

**Lemma 4.14** *Let the coarse mesh $\mathcal{T}_H$ be quasi uniform and let $Q^H u$ be the $L^2$-projection of $u \in H^1(\Omega)$ onto the finite element space $V^H$. Then, in two and three dimensions,*

$$||u - Q^H u||^2_{L^2(\Omega)} \leq CH^2 |u|^2_{H^1(\Omega)},$$

*and*

$$|Q^H u|^2_{H^1(\Omega)} \leq C|u|^2_{H^1(\Omega)}.$$

*Proof.* The first inequality follows directly from Lemma 3.6, since, by definition, $Q^H u$ is the best approximation in $L^2$. To prove the second, we note that, again by Lemma 3.6, it can be reduced to estimating $|\tilde{I}^H u - Q^H u|^2_{H^1(\Omega)}$. We then note that $\tilde{I}^H u - Q^H u = Q^H(\tilde{I}^H u - u)$ and thus

$$|\tilde{I}^H u - Q^H u|^2_{H^1(\Omega)} \leq \frac{C}{H^2} ||Q^H(\tilde{I}^H u - u)||^2_{L^2(\Omega)} \leq C|u|^2_{H^1(\Omega)}.$$

Here we have used an inverse inequality, which holds since the coarse mesh is assumed to be quasi uniform, that $Q^H$ is the $L^2$-projection, and Lemma 3.6. □

We note that the results for the quasi-interpolant $\tilde{I}^H$ given in Sect. 3.5 are stronger since the coarse mesh elements need only be shape regular but not necessarily all of comparable size. We also note that the estimates of Lemma 3.6 cannot be made independent of the $\rho_i$ either but that they are of a more local character.

### 4.6.2 Inequalities for Edges

We now consider inequalities associated with the edges and the wire basket of a substructure.

The following result plays an important role in the theory; see, e.g., [69, 72, 478] where alternative proofs can be found. The bound in this lemma is known to be sharp; cf. Brenner and Sung [91] for an explicit construction. The

sharpness of the bound can also be established by a less direct argument by using lower bounds for iterative substructuring algorithms due to Brenner et al; see [92, 89].

**Lemma 4.15** *Let $\alpha$ be any convex combination of values of $u(x)$, with $x \in \Omega_i$, a bounded subregion in the plane of diameter $H$. Then, for $u \in V^h$,*

$$\|u\|^2_{L^\infty(\Omega_i)} \leq C\,(1+\log(H/h))\|u\|^2_{H^1(\Omega_i)},$$
$$\|u-\alpha\|^2_{L^\infty(\Omega_i)} \leq C\,(1+\log(H/h))|u|^2_{H^1(\Omega_i)}.$$

*The logarithmic factor cannot be removed.*

*Proof.* We will follow a proof given in Thomée [436]. We consider the case of $H = 1$; the general case follows by a scaling argument. We first note that by a Sobolev inequality, we have for any $r \geq 1$

$$\|u\|_{L^r(\Omega_i)} \leq Cr^{1/2}\|u\|_{H^1(\Omega_i)}, \quad u \in H^1(\Omega_i); \tag{4.17}$$

cf. [357, Theorem 1]. We combine this inequality with an inverse inequality

$$\|u\|_{L^\infty(\Omega_i)} \leq Ch^{-2/r}\|u\|_{L^r(\Omega_i)}, \quad u \in V^h; \tag{4.18}$$

cf. [90, Lemma 4.5.3], and obtain

$$\|u\|_{L^\infty(\Omega_i)} \leq Cr^{1/2}\,h^{-2/r}\|u\|_{H^1(\Omega_i)}, \quad r \geq 1.$$

We complete the proof of the first inequality by choosing $r = \log(1/h)+1$:

$$\begin{aligned}\|u\|^2_{L^\infty(\Omega_i)} &\leq C(1+\log(1/h))\ (1/h)^{4/(1+\log(1/h))}\ \|u\|^2_{H^1(\Omega_i)}\\ &\leq C(1+\log(1/h))\ \|u\|^2_{H^1(\Omega_i)}.\end{aligned}$$

The second inequality can then be proven by using the Poincaré inequality in Lemma A.13. □

Results very similar to those of the next lemma can be found in Bramble, Pasciak, and Schatz [76], Bramble and Xu [81], and Dryja [173]. They provide discrete trace inequalities for a single edge or for the wire basket. The proofs are straightforward; it is sufficient to consider the inequalities, given in Lemma 4.15 for two dimensions and to integrate along the third direction.

**Lemma 4.16** *Let $\bar{u}_{\mathcal{E}^k}$ be the average value of $u$ over $\mathcal{E}^k$, an edge of $\Omega_i$. Then,*

$$\|u\|^2_{L^2(\mathcal{E}^k)} \leq C(1+\log(H/h))\|u\|^2_{H^1(\Omega_i)}$$

*and*

$$\|u-\bar{u}_{\mathcal{E}^k}\|^2_{L^2(\mathcal{E}^k)} \leq C(1+\log(H/h))|u|^2_{H^1(\Omega_i)}.$$

*Similar bounds also hold if the edge $\mathcal{E}^k$ is replaced by the wire basket $\mathcal{W}^i$.*

We also give a useful variant of this lemma. We refer to the proof of [173, Lemma 4] for the first inequality. The second can then be proven by using a Poincaré inequality.

**Lemma 4.17** *Let $\bar{u}_{\mathcal{E}^k}$ be the average value of $u$ over $\mathcal{E}^k$, an edge of a face $\mathcal{F}^{ij}$. Then,*

$$||u||^2_{L^2(\mathcal{E}^k)} \leq C(1+\log(H/h))||u||^2_{H^{1/2}(\mathcal{F}^{ij})}$$

*and*

$$||u-\bar{u}_{\mathcal{E}^k}||^2_{L^2(\mathcal{E}^k)} \leq C(1+\log(H/h))|u|^2_{H^{1/2}(\mathcal{F}^{ij})}.$$

*Similar bounds also hold if the edge $\mathcal{E}^k$ and the face $\mathcal{F}^{ij}$ are replaced by the wire basket $\mathcal{W}^i$ and the boundary $\partial\Omega_i$, respectively.*

When we estimate the parameter $C_0^2$, introduced in Assumption 2.2 of the abstract convergence theory, we must demonstrate that all functions in the finite element space can be decomposed into components in the subspaces in such a way that the sum of the resulting energies are uniformly, or almost uniformly, bounded with respect to the parameters $h$, $H$, etc. A main tool for deriving such decompositions is a suitable partition of unity built from a set of functions $\theta_{\mathcal{E}^k}$, associated with edges, $\theta_{\mathcal{F}^j}$, associated with faces, and nodal basis functions, $\theta_{\mathcal{V}^l}$, corresponding to the substructure vertices.

In this subsection, we will consider the functions associated with the subdomain edges.

**Definition 4.18.** *Given an edge $\mathcal{E}^k \subset \Gamma$, let $\vartheta_{\mathcal{E}^k} \in V^h$ be the function that vanishes at all nodes of $\overline{\Omega}_h$ except those of $\mathcal{E}^k_h$ where it takes the value 1. In addition, let $\theta_{\mathcal{E}^k}$ be the discrete harmonic function that coincides with $\vartheta_{\mathcal{E}^k}$ on $\Gamma$.*

We note that, given $u \in V^h$, the function $I^h(\vartheta_{\mathcal{E}^k}u)$ is the extension by zero of the values of $u$ on $\mathcal{E}^k$, while $\mathcal{H}(\theta_{\mathcal{E}^k}u) = \mathcal{H}(\vartheta_{\mathcal{E}^k}u)$ is the piecewise discrete harmonic extension of the same trace. From now on, we will use the notation $\mathcal{H}(\theta_{\mathcal{E}^k}u)$ for the function $\mathcal{H}(I^h(\theta_{\mathcal{E}^k}u))$, etc.

Our next lemma complements Lemma 4.16 and provides bounds for the extensions from a single edge or the wire basket. We note that while $I^h(\theta_{\mathcal{E}^k}u)$ vanishes at the end points of the edge $\mathcal{E}^k$, $u$ in general does not.

**Lemma 4.19** *Let $\mathcal{E}^k$ be an edge of a substructure $\Omega_i$ and let $u \in V^h$. Then,*

$$|\mathcal{H}_i(\theta_{\mathcal{E}^k}u)|^2_{H^1(\Omega_i)} \leq |I^h(\vartheta_{\mathcal{E}^k}u)|^2_{H^1(\Omega_i)} \leq C||I^h(\theta_{\mathcal{E}^k}u)||^2_{L^2(\mathcal{E}^k)},$$
$$|\mathcal{H}_i(\theta_{\mathcal{E}^k}u)|^2_{H^1(\Omega_i)} \leq |I^h(\vartheta_{\mathcal{E}^k}u)|^2_{H^1(\Omega_i)} \leq C||u||^2_{L^2(\mathcal{E}^k)}.$$

*Similar bounds also hold if the edge $\mathcal{E}^k$ is replaced by the wire basket $\mathcal{W}^i$ and/or if the seminorms in $H^1(\Omega_i)$ are replaced by those in $H^{1/2}(\partial\Omega_i)$.*

*Proof.* The first inequalities follow from the minimizing properties of discrete harmonic extensions. For the second inequalities, involving the zero extensions, we note that given a fine element $K \in \mathcal{T}_h$ that shares an edge $(a,b)$ with $\mathcal{E}^k$, Lemma B.5 ensures that

$$|I^h(\vartheta_{\mathcal{E}^k}u)|^2_{H^1(K)} \leq Ch((\vartheta_{\mathcal{E}^k}(a)u(a))^2 + (\vartheta_{\mathcal{E}^k}(b)u(b))^2).$$

Since $\vartheta_{\mathcal{E}^j}$ has values in $[0,1]$, another application of Lemma B.5 ensures that the right hand side can be bounded both by $C\int_a^b (I^h(\theta_{\mathcal{E}^k}u))^2 ds$ and by $C\int_a^b u^2 ds$. The proof of the two inequalities is completed by summing over the elements. Bounds for the wire basket are found in the same way, while those involving the $H^{1/2}$-seminorm on the boundary are a consequence of a trace theorem. □

By combining Lemmas 4.16 and 4.19, we obtain an estimate of the energy of the edge function $\mathcal{H}_i(\theta_{\mathcal{E}^k}u)$:

**Corollary 4.20** *For an edge $\mathcal{E}^k$ of a substructure $\Omega_i$ and $u \in V^h$, we have*

$$|\mathcal{H}_i(\theta_{\mathcal{E}^k}u)|^2_{H^1(\Omega_i)} \leq |I^h(\vartheta_{\mathcal{E}^k}u)|^2_{H^1(\Omega_i)} \leq C(1+\log(H/h))\|u\|^2_{H^1(\Omega_i)}.$$

*A similar bound also holds for the wire basket $\mathcal{W}^i$.*

In the description of our algorithms in Chap. 5, we will use $h_j\|u\|^2_{l^2(\mathcal{E}^k)}$ in addition to $\|u\|^2_{L^2(\mathcal{E}^k)}$. Here $\|u\|_{l^2(\mathcal{E}^k)}$ is the $\ell^2-$norm of the nodal values of $u$ on $\mathcal{E}^k_h$. The first expression is appropriate when defining bilinear forms on a subspace related to the edge $\mathcal{E}^k$. The two are, for all theoretical purposes, interchangeable since the mass matrix related to the second expression is uniformly well conditioned given that the triangulation of each subdomain, by assumption, is quasi uniform; see Lemma B.31.

We finally state and prove a nonstandard Poincaré inequality, which was given in a somewhat different form in [184, Lemma 6]. We note that the logarithmic factors in the estimates are related to the fact that traces over an edge (and therefore averages) are not in general defined for every function in $H^1(\Omega_i)$.

**Lemma 4.21 (Poincaré-Friedrichs inequality)** *Let $\mathcal{E}^k$ be an edge of $\Omega_i$. Then, for any $u \in V^h(\Omega_i)$,*

$$\|u - \bar{u}_{\mathcal{E}^k}\|^2_{L^2(\partial\Omega_i)} \leq CH_i(1+\log(H/h))|u|^2_{H^1(\Omega_i)},$$
$$\|u - \bar{u}_{\mathcal{E}^k}\|^2_{L^2(\Omega_i)} \leq CH_i^2(1+\log(H/h))|u|^2_{H^1(\Omega_i)}.$$

*Proof.* We consider $u - \bar{u}_{\mathcal{E}^k}$. Using the Cauchy-Schwarz inequality, we find

$$|\bar{u}_{\mathcal{E}^k}|^2 \leq CH_i^{-1}\|u\|^2_{L^2(\mathcal{E}^k)};$$

see also Lemma 4.30. The use of Lemma 4.16 then yields

$$\|\bar{u}_{\mathcal{E}^k}\|^2_{L^2(\partial\Omega_i)} \leq CH_i(1+\log(H/h))\|u\|^2_{H^1(\Omega_i)}.$$

A trace theorem gives a similar bound, without a logarithmic factor, for $\|u\|^2_{L^2(\partial\Omega_j)}$. We then note that $u - \bar{u}_{\mathcal{E}^k}$ does not change if we add a constant to $u$ and the proof of the first inequality can be completed by using a

standard Poincaré inequality; see Theorem A.18. The second bound can be proved in a similar way. □

We note that the previous lemma ensures that

$$\|u\|^2_{L^2(\partial\Omega_i)} \leq CH_i(1+\log(H/h))|u|^2_{H^1(\Omega_i)}, \tag{4.19}$$

for a finite element function that vanishes over an edge. Similar bounds can also be found for functions that vanish at a vertex: using Equation (4.16), we find that

$$\|u\|^2_{L^2(\partial\Omega_i)} \leq CH_i(H/h)|u|^2_{H^1(\Omega_i)}.$$

We note that this bound is much less satisfactory; this is the reason why for some iterative substructuring methods of the next two chapters we need to assume that no substructure touches the boundary $\partial\Omega_D$ in just one or few points.

### 4.6.3 Inequalities for Faces

In this section, we will consider functions associated with subdomain faces. In particular, we will partition a function into terms associated with single faces. As for the edges, we will associate a function with each face:

**Definition 4.22.** *Given a face $\mathcal{F}^j \subset \Gamma$, let $\theta_{\mathcal{F}^j} \in V^h$ be the piecewise discrete harmonic function that vanishes at all nodes of $\Gamma_h$ except those of $\mathcal{F}^j_h$ where it takes the value* 1.

As in the previous section, we employ auxiliary functions, $\vartheta_{\mathcal{F}^j}$, that takes the same values as $\theta_{\mathcal{F}^j}$ on $\Gamma$ and for which bounds can be established relatively easily. In the next lemma, we explicitly construct such functions for a shape-regular tetrahedron. Those for a mapped cube can be constructed in the same way; see [126, Lem. 3.3.6]. We note that in the estimate of the first term of (4.25), for a case of a subdomain, which is the union of a finite number of such tetrahedra, we can work with the functions $\vartheta_{\mathcal{F}^j}$ that we will construct in Lemma 4.23 extending them by zero in the additional coarse elements that also are part of the substructure.

**Lemma 4.23** *For any face $\mathcal{F}^j$ of a tetrahedron $\Omega_i$, there exists a finite element function $\vartheta_{\mathcal{F}^j} \in V^h$, that equals* 1 *at the nodal points of $\mathcal{F}^j_h$, vanishes on $\Gamma_h \setminus \mathcal{F}^j_h$, and satisfies*

$$\begin{aligned} &\sum_{\mathcal{F}^j \subset \partial\Omega_i} \vartheta_{\mathcal{F}^j}(x) = 1, \quad x \in (\Omega_{i,h} \cup \partial\Omega_{i,h}) \setminus \mathcal{W}^i_h, \\ &0 \leq \vartheta_{\mathcal{F}^j} \leq 1, \\ &|\nabla\vartheta_{\mathcal{F}^j}(x)| \leq C/r(x). \end{aligned} \tag{4.20}$$

*Here $r = r(x)$ denotes the distance to $\mathcal{W}^i$.*

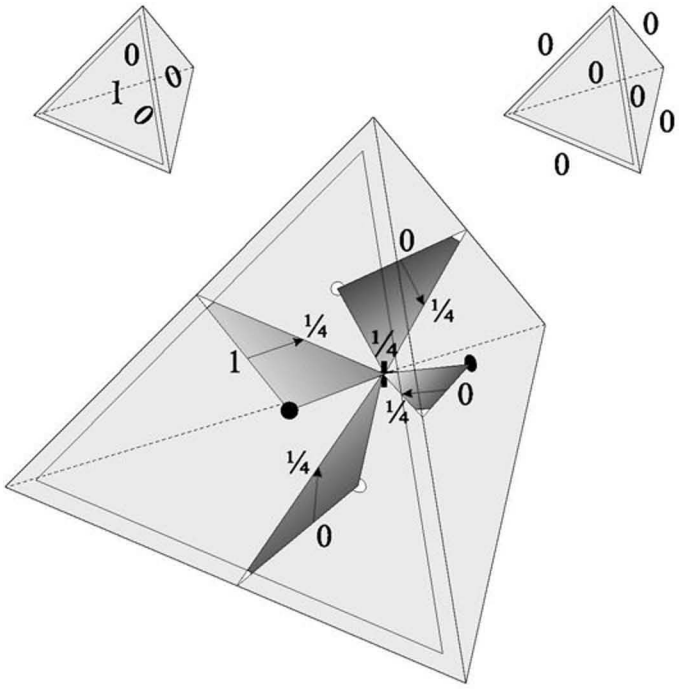


**Fig. 4.2.** Construction of the partition function $\vartheta_{\mathcal{F}^j}$ in a tetrahedron.

*Proof.* We will first construct four functions, that are not finite element functions, and that form a partition of unity, i.e., they satisfy the first equality. The four finite element functions $\vartheta_{\mathcal{F}^j}$ are then obtained by piecewise linear interpolation. It is easy to verify that these new functions also define a partition of unity and that the bound on the gradient is preserved. We will use the same notation for these two sets of functions.

We divide the substructure into four tetrahedra by connecting its centroid $C$, by line segments, to the four vertices of the tetrahedron. Similarly, we divide each triangular face of the substructure into three triangles by extending the bisectors of the three vertices of the triangle until they meet. We denote the resulting points on the faces by $C_k$, see Fig. 4.2. By connecting the $C_k$ with $C$, by line segments, we obtain the wire baskets of twelve tetrahedra.

We construct the function $\vartheta_{\mathcal{F}^j}$, associated with the face $\mathcal{F}^j$, as follows: At $C$ the value is $1/4$. We interpolate linearly between the value $1/4$ and 1 or 0, whichever is appropriate, along the line segments connecting $C$ to the $C_k$. The values elsewhere are constant on the intersection of any plane, through the unique substructure edge that belongs to a specific subtetrahedron, and that same subtetrahedron. This constant value is determined by the value, already known, at the point on the appropriate line segment, which is one of the edges of the same subtetrahedron. We note that $\vartheta_{\mathcal{F}^j}$ is now defined on the

whole closure of $\Omega_i$ except on the edges and vertices. Finally, we modify the function by changing its values in the elements that have at least one vertex on an edge of the substructure. We make the function zero on the wire basket $\mathcal{W}^i$ and continuous, by piecewise linear interpolation, using the previously constructed values at the nodes that are not on, but next to, the edge. We also replace the function elsewhere by its piecewise linear interpolant.

We return for a moment to consider the function prior to the interpolation. The values on any two planes associated with two different substructure edges, which intersect at a point on the appropriate line segment, are the same. The partition functions are therefore continuous across the boundaries of the subtetrahedra. Explicit formulas for the gradient and estimates thereof can, at least in principle, be given. The most important observation is that $|\nabla \vartheta_{\mathcal{F}^j}| \leq C/r$, where $r$ is the distance to the nearest edge of the original tetrahedron.

It is also easy to show that $\{\vartheta_{\mathcal{F}^j}\}$ form a partition of unity on the special line segments, and everywhere else, except in the special elements next to the edges of the original substructure. Finally, by construction, they have values in $[0,1]$. □

The following lemma is an extension to three dimensions of a result of Dryja and Widlund [180]. The present approach makes it possible to prove nontrivial bounds for iterative substructuring algorithms without the use of an extension theorem such as (4.9). We can always work in subspaces of the original finite element spaces, and we never need to use trace and extension theorems. One advantage of this approach is that the constants obtained in the estimates can be calculated explicitly from geometric information. When working with the trace and extension theorems, it is more difficult to determine the exact relationship between the geometry, e.g., the aspect ratios of the substructures, and the constants of the bounds.

**Lemma 4.24** *Let $\vartheta_{\mathcal{F}^j}(x)$ be the functions of Lemma 4.23, where $\mathcal{F}^j$ is a face of the substructure $\Omega_i$, and let $I^h$ denote the interpolation operator associated with the finite element space $V^h$. Then,*

$$\sum_{\mathcal{F}^j} I^h(\vartheta_{\mathcal{F}^j} u)(x) = u(x), \quad x \in \overline{\Omega}_{i,h} \setminus \mathcal{W}_h^i,$$

*and*

$$|I^h(\vartheta_{\mathcal{F}^j} u)|^2_{H^1(\Omega_i)} \leq C(1+\log(H/h))^2 ||u||^2_{H^1(\Omega_i)}.$$

*Proof.* Again we provide a proof only for the case of a tetrahedral substructure. The first formula follows immediately from the first equality of Lemma 4.23. In order to prove the second, we first consider the contributions to the energy from the elements that touch the wire basket $\mathcal{W}^i$. By definition, $\vartheta_{\mathcal{F}^j}$ vanishes on the wire basket. Lemma B.30 ensures that the energy contributed from this small neighborhood of the wire basket can be bounded by $Ch\sum_x |u(x)|^2$, where the sum is taken over all the nodal points that are

within a mesh width of the wire basket. This sum can be bounded by a line integral on paths along the edges by using Lemma B.30 and, as in Lemma 4.16, it can therefore be bounded by $C(1+\log(H/h))||u||^2_{H^1(\Omega_i)}$. We obtain:

$$\sum_{\overline{K}\cap\mathcal{W}^i\neq\emptyset} |I^h(\vartheta_{\mathcal{F}^j}u)|^2_{H^1(K)} \leq C(1+\log(H/h))||u||^2_{H^1(\Omega_i)}. \tag{4.21}$$

We next consider the contributions from the elements that do not touch the wire basket $\mathcal{W}^i$. Let $K$ be such an element. By using elementary considerations, we obtain,

$$|I^h(\vartheta_{\mathcal{F}^j}u)|^2_{H^1(K)} \leq 2|\bar{\vartheta}_{\mathcal{F}^j}u|^2_{H^1(K)} + 2|I^h((\vartheta_{\mathcal{F}^j}-\bar{\vartheta}_{\mathcal{F}^j})u)|^2_{H^1(K)}.$$

Here $0<\bar{\vartheta}_{\mathcal{F}^j}<1$ is the average of $\vartheta_{\mathcal{F}^j}$ over the element $K$.

The bound for the first term is trivial, but that of the second term is more complicated. We first use an inverse inequality and obtain

$$|I^h((\vartheta_{\mathcal{F}^j}-\bar{\vartheta}_{\mathcal{F}^j})u)|^2_{H^1(K)} \leq Ch_i^{-2}||I^h((\vartheta_{\mathcal{F}^j}-\bar{\vartheta}_{\mathcal{F}^j})u)||^2_{L^2(K)}.$$

By using the bound on the gradient of $\vartheta_{\mathcal{F}^j}$ in Lemma 4.23, we can bound $\vartheta_{\mathcal{F}^j}-\bar{\vartheta}_{\mathcal{F}^j}$ by $Ch_i/r$, where $r$ is the distance to the wire basket. Hence, summing over the elements that do not touch the wire basket,

$$\sum_K |I^h((\vartheta_{\mathcal{F}^j}-\bar{\vartheta}_{\mathcal{F}^j})u)|^2_{H^1(K)} \leq C\sum_K r_K^{-2}||u||^2_{L^2(K)},$$

with $r_K$ the distance of the barycenter of $K$ to the wire basket. We partition the elements of $\Omega_i$ into groups, in accordance to the closest edge of $\Omega_i$; the exact rule for the assignment of the elements that are halfway between is of no importance. For each edge of the wire basket, we use a local cylindrical coordinate system with the $z$ axis coinciding with the edge, and the radial direction, $r$, normal to the edge. Using cylindrical coordinates, we estimate the sum by an integral

$$\sum_{K\subset\Omega_i} r_K^{-2}||u||^2_{L^2(K)} \leq C\int_{r=h}^{H}\int_{\phi}\int_{z}|u|^2\frac{r}{r^2}drd\phi dz. \tag{4.22}$$

The integral with respect to $z$ can be bounded using Lemma 4.16. We obtain

$$\sum_{K\subset\Omega_i} r_K^{-2}||u||^2_{L^2(K)} \leq C(1+\log(H_i/h_i))||u||^2_{H^1(\Omega_i)}\int_{r=h}^{H} r^{-1}dr \tag{4.23}$$

and thus

$$\sum_{K\subset\Omega_i} |I^h(\vartheta_{\mathcal{F}^j}u)|^2_{H^1(K)} \leq C(1+\log(H/h))^2||u||^2_{H^1(\Omega_i)}.$$

Combining the last inequality with (4.21) concludes the proof. □

We remark that one of the two logarithmic factors of the previous lemma results from the use of Lemma 4.16 while the other results from the large values of the gradient of $\vartheta_{\mathcal{F}^j}$ when we approach the wire basket; see (4.23). In other words, one factor arises from the fact that we put the values of the function $u$ to zero on the wire basket, while the second comes from the decoupling of the terms associated with single faces.

We also need bounds for the functions $\theta_{\mathcal{F}^k}$.

**Lemma 4.25** *Let $\mathcal{F}^k$ be the face common to two substructures $\Omega_i$ and $\Omega_j$. Then,*

$$|\theta_{\mathcal{F}^k}|^2_{H^1(\Omega_i)} \leq |\vartheta_{\mathcal{F}^k}|^2_{H^1(\Omega_i)} \leq C(1+\log(H/h))H,$$

*and*

$$||\theta_{\mathcal{F}^k}||^2_{L^2(\Omega_i)} \leq CH^3.$$

*The same bounds also hold for the other subdomain $\Omega_j$.*

*Proof.* For the first property, we only need to find a bound for $\vartheta_{\mathcal{F}^k}$; a bound for $\theta_{\mathcal{F}^k}$ then follows from the minimizing property of discrete harmonic extensions given in Lemma 4.9. The proof is similar to that of the second inequality of the previous lemma. We first consider contributions from elements that touch the wire basket. Lemmas B.30 and 4.23 ensure that the energy contributed from this small neighborhood of the wire basket can be bounded by

$$Ch\sum_x |\vartheta_{\mathcal{F}^k}(x)|^2 \leq Ch\sum_x 1 \leq CH,$$

where the sum is taken over all nodal points that are within a mesh width of the wire basket.

To estimate the contribution to the energy from the rest of the substructure, we consider one subtetrahedron $K$ at a time and the bound on the gradient in Lemma 4.23. We obtain:

$$\sum_K |\vartheta_{\mathcal{F}^j}|^2_{H^1(K)} \leq C\sum_K r_K^{-2}\,|K|,$$

with $r_K$ the distance of the barycenter of $K$ to the wire basket and $|K|$ the volume of $K$. The sum is taken over the elements that do not touch the wire basket. The proof can then be concluded as for (4.22) of Lemma 4.24, by replacing the sum with an integral and introducing cylindrical coordinates which use the appropriate substructure edge as the z-axis.

We now turn to the proof of the second inequality. We note that the bound trivially holds for $\vartheta_{\mathcal{F}^k}$ but that a more elaborate argument seems to be required for $\theta_{\mathcal{F}^k}$. To avoid irrelevant scaling factors, we consider the special case of $H=1$, and we also denote the region by $\hat{\Omega}$. In this case, we only have to prove that $||\theta_{\mathcal{F}^k}||_{L^2(\hat{\Omega})}$ is bounded. We introduce an auxiliary function $v \in H^1(\hat{\Omega})$ by solving

$$\begin{aligned} -\Delta v &= \theta_{\mathcal{F}^k}, \quad && x \in \hat{\Omega}, \\ v &= 0, && x \in \partial\hat{\Omega}. \end{aligned}$$

Since $\hat{\Omega}$ is polyhedral, we have

$$|v|_{H^{3/2+s}(\hat{\Omega})} \leq C\|\theta_{\mathcal{F}^k}\|_{L^2(\hat{\Omega})},$$

with $s \in (0, s_{\hat{\Omega}})$; see Lemma A.49. By Green's formula, we find

$$\|\theta_{\mathcal{F}^k}\|^2_{L^2(\hat{\Omega})} = \int_{\hat{\Omega}} \nabla v \cdot \nabla \theta_{\mathcal{F}^k} dx - \int_{\partial\hat{\Omega}} \frac{\partial v}{\partial n} \theta_{\mathcal{F}^k} ds. \tag{4.24}$$

We now find bounds for the two terms on the right hand side of (4.24). Since $\theta_{\mathcal{F}^k}$ is discrete harmonic, we find that

$$\int_{\hat{\Omega}} \nabla v \cdot \nabla \theta_{\mathcal{F}^k} dx = \int_{\hat{\Omega}} \nabla (v - w) \cdot \nabla \theta_{\mathcal{F}^k} dx, \quad w \in V^h(\hat{\Omega}) \cap H^1_0(\hat{\Omega}).$$

The right hand side can be estimated from above using the nodal interpolant $w = I^h v$, which is defined since $v$ is continuous; cf. Lemma A.5. Using the error estimate in Lemma B.6 and the regularity result, we obtain

$$\begin{aligned} |v - w|_{H^1(\hat{\Omega})} |\theta_{\mathcal{F}^k}|_{H^1(\hat{\Omega})} &\leq Ch^{1/2+s} |v|_{H^{3/2+s}(\hat{\Omega})} |\theta_{\mathcal{F}^k}|_{H^1(\hat{\Omega})} \\ &\leq Ch^{1/2+s} \|\theta_{\mathcal{F}^k}\|_{L^2(\hat{\Omega})} |\theta_{\mathcal{F}^k}|_{H^1(\hat{\Omega})}. \end{aligned}$$

By using the bound for $|\theta_{\mathcal{F}^k}|_{H^1(\hat{\Omega})}$, we see that the first term originating from the Green's formula is $o(\|\theta_{\mathcal{F}^k}\|_{L^2(\hat{\Omega})})$.

For the second term, the face integral, we use Cauchy-Schwarz's inequality, a standard trace theorem, and the regularity result. We obtain,

$$\left| \int_{\partial\hat{\Omega}} \frac{\partial v}{\partial n} \theta_{\mathcal{F}^k} ds \right| \leq C\|v\|_{H^{3/2}(\hat{\Omega})} \|\theta_{\mathcal{F}^k}\|_{L^2(\partial\hat{\Omega})} \leq C\|\theta_{\mathcal{F}^k}\|_{L^2(\hat{\Omega})} \|\theta_{\mathcal{F}^k}\|_{L^2(\partial\hat{\Omega})}.$$

The proof can now easily be concluded by observing that $\|\theta_{\mathcal{F}^k}\|_{L^2(\partial\hat{\Omega})} \leq C$. □

As for the edge contributions, we can provide estimates that only involve the values of the functions on the boundary. The following result is a direct consequence of Lemma 4.10 and the Poincaré inequality in Lemma A.17; cf. also Remark A.9.

**Lemma 4.26** *Let $\mathcal{F}^k$ be a face of a substructure $\Omega_i$ and let $u \in V^h$. Then,*

$$|\theta_{\mathcal{F}^k}|^2_{H^{1/2}(\partial\Omega_i)} \leq C(1 + \log(H/h))H$$

*and*

$$\begin{aligned} |I^h(\vartheta_{\mathcal{F}^k} u)|^2_{H^{1/2}(\partial\Omega_i)} &\leq C(1 + \log(H/h))^2 \|u\|^2_{H^{1/2}(\partial\Omega_i)}, \\ |I^h(\vartheta_{\mathcal{F}^k}(u - \bar{u}_{\mathcal{F}^k}))|^2_{H^{1/2}(\partial\Omega_i)} &\leq C(1 + \log(H/h))^2 |u|^2_{H^{1/2}(\partial\Omega_i)}, \\ |I^h(\vartheta_{\mathcal{F}^k}(u - \bar{u}_{\partial\mathcal{F}^k}))|^2_{H^{1/2}(\partial\Omega_i)} &\leq C(1 + \log(H/h))^2 |u|^2_{H^{1/2}(\partial\Omega_i)}, \end{aligned}$$

*where $\bar{u}_{\mathcal{F}^k}$ and $\bar{u}_{\partial\mathcal{F}^k}$ are the averages of $u$ over $\mathcal{F}^k$ and $\partial\mathcal{F}^k$, respectively. The seminorms on the left hand sides can be replaced by the norms of $H^{1/2}_{00}(\mathcal{F}^k)$.*

### 4.6.4 Inequalities for Vertices and Auxiliary Results

We finally consider functions associated with the substructure vertices and provide a few additional results.

**Definition 4.27.** *Given a vertex $\mathcal{V}^k \subset \Gamma$, let $\vartheta_{\mathcal{V}^k} \in V^h$ be the function that vanishes at every node in $\overline{\Omega}_h$ except $\mathcal{V}^k$ where it takes the value* 1. *In addition, let $\theta_{\mathcal{V}^k}$ be the discrete harmonic function that coincides with $\vartheta_{\mathcal{V}^k}$ on $\Gamma$.*

Given $u \in V^h$, the function $u(\mathcal{V}^k)\vartheta_{\mathcal{V}^k}$ is the extension by zero of the value of $u$ at $\mathcal{V}^k$, while $u(\mathcal{V}^k)\theta_{\mathcal{V}^k}$ is the harmonic extension of the resulting boundary values. On a substructure $\Omega_i$, we then have the decomposition

$$u = \sum_{\mathcal{F}\subset\partial\Omega_i} \mathcal{H}_i(\theta_{\mathcal{F}} u) + \sum_{\mathcal{E}\subset\partial\Omega_i} \mathcal{H}_i(\theta_{\mathcal{E}} u) + \sum_{\mathcal{V}\in\partial\Omega_i} u(\mathcal{V})\,\theta_{\mathcal{V}}, \tag{4.25}$$

which is valid for every discrete harmonic function in $V^h(\Omega_i)$.

A bound for the vertex terms can easily be obtained by combining the inverse inequality given in (4.16) with the third bound of Lemma B.5.

**Lemma 4.28** *Let $\mathcal{V}^k$ be a vertex of a substructure $\Omega_i$ and let $u \in V^h$. Then,*

$$|u(\mathcal{V}^k)\,\theta_{\mathcal{V}^k}|^2_{H^{1/2}(\partial\Omega_i)} = |u(\mathcal{V}^k)\,\vartheta_{\mathcal{V}^k}|^2_{H^{1/2}(\partial\Omega_i)} \leq C||u||^2_{H^{1/2}(\partial\Omega_i)}.$$

Corollary 4.20 and Lemmas 4.24 and 4.28 then give:

**Corollary 4.29** *The decomposition (4.25) is stable with a stability constant that grows only as $C(1+\log(H/h))^2$.*

We end this section with some additional auxiliary results. The proof of the following result is an immediate consequence of the definition of the averages and the Cauchy-Schwarz inequality.

**Lemma 4.30** *Let $\bar{u}_{\mathcal{F}^k}$, $\bar{u}_{\partial\Omega_i}$, $\bar{u}_{\partial\mathcal{F}^k}$, $\bar{u}_{\mathcal{E}^l}$, and $\bar{u}_{\mathcal{W}^i}$ be the average of $u$ over $\mathcal{F}^k$, $\partial\Omega_i$, $\partial\mathcal{F}^k$, $\mathcal{E}^l$, and $\mathcal{W}^i$, respectively. Let in addition $H_i$, $H_k$, and $H_l$ be the diameters of $\Omega_i$, $\mathcal{F}^k$, and $\mathcal{E}^l$, respectively. Then,*

$$(\bar{u}_{\mathcal{F}^k})^2 \leq CH_k^{-2}||u||^2_{L^2(\mathcal{F}^k)},$$

$$(\bar{u}_{\partial\Omega_i})^2 \leq CH_i^{-2}||u||^2_{L^2(\partial\Omega_i)},$$

$$(\bar{u}_{\partial\mathcal{F}^k})^2 \leq CH_k^{-1}||u||^2_{L^2(\partial\mathcal{F}^k)},$$

$$(\bar{u}_{\mathcal{E}^l})^2 \leq CH_l^{-1}||u||^2_{L^2(\mathcal{E}^l)},$$

$$(\bar{u}_{\mathcal{W}^i})^2 \leq CH_i^{-1}||u||^2_{L^2(\mathcal{W}^i)}.$$

We will have the occasion of multiplying finite element functions by each other and to estimate their products. It is therefore natural to consider such bounds in a lemma. We note that a partial result has already been given in Lemma 3.9.

**Lemma 4.31** *Let $u$ be a continuous, piecewise quadratic function defined on the finite element triangulation and let $I^h u \in V^h$ be its piecewise linear interpolant on the same mesh. Then, there exists a constant $C$, independent of $h$ and $H$, such that*

$$|I^h u|_{H^1(\Omega_i)} \leq C|u|_{H^1(\Omega_i)} .$$

*The same type of bounds hold for the $L^2$- and $H_{00}^{1/2}$-norms and it can also be extended, with different constants, to the case of piecewise cubic functions, etc.*

*Proof.* The bound involving the $H^1$-seminorms was already proven in Lemma 3.9. For the bound in $L^2$, we note that, for each element $K \in \mathcal{T}_h$, we have

$$\|I^h u\|_{L^2(K)} \leq C\|u\|_{L^2(K)}.$$

This inequality can be proven on the reference element by noting that here $I^h$ is a mapping from and into finite dimensional spaces and that it is therefore bounded. A scaling argument then shows that the constant is independent of the diameter of $K$ and only depends on its aspect ratio. Summing over the elements $K \subset \Omega_i$ then gives the bound in $L^2$. The bounds for the $H_{00}^{1/2}$-norm is obtained by using the K-interpolation method; cf. appendix A.1. Thus,

$$\|u\|^2_{H_{00}^{1/2}} = \|u\|^2_{L^2} + \int_0^\infty t^{-2} K(t,u)^2 dt,$$

where

$$K(t,u) = \inf_{u=u_0+u_1} (\|u_0\|^2_{H_0^1} + t^2\|u_1\|^2_{L^2})^{1/2}.$$

The crucial observation is that

$$K(t, I^h u) \leq CK(t,u),$$

where $C$ can be chosen as the larger of the constants in the $H^1$ and $L^2$ bounds. The proof then follows easily. □

# 5

# Primal Iterative Substructuring Methods

## 5.1 Introduction

In this chapter, which has evolved from Dryja, Smith, and Widlund [178], we will introduce and analyze a number of primal iterative substructuring methods. The mathematical development of this class of methods began in the mid-nineteeneighties with a series of important papers by Bramble, Pasciak, and Schatz; see [72, 74, 75, 76]. For a survey article, which covers similar material, see Xu and Zou [476].

An iterative substructuring method is of primal type if the iteration is carried out in terms of a subset of the original finite element degrees of freedom, namely those on $\Gamma$, the union of the interfaces between the substructures. In Sect. 6.3 and 6.4, we will consider dual methods where the principal unknowns of the iteration are certain Lagrange multipliers which enforces the continuity of the finite element solution across the interface given by the subdomain boundaries.

In this and the following chapter, we will primarily consider problems in three dimensions. In Sect. 5.4, we will consider several families of iterative substructuring methods which differ primarily by the selection of the coarse space. We will also comment on the easier two-dimensional case in several remarks.

## 5.2 Local Design and Analysis

The problems considered in this part of our monograph are of the form (4.3). We will develop our theory for the piecewise constant coefficient case, but all our results are equally valid for the case when the coefficients vary moderately in each subdomain. When all $\rho_j \equiv 1$, we have the special case of the Poisson equation. In order to be successful with problems that have large variations in the coefficients, it is important to be able to carry out a local analysis. This is often done in a Schwarz framework.

Let $V^{(j)}$ be the restriction of the functions in the solution space $V$ to the subdomain $\Omega_j$. We will assume that the space $V$ can be decomposed into a sum of subspaces $\sum R_i^T V_i$. We then decompose $V^{(j)}$ into subspaces $V_i^{(j)}$, with $V_i^{(j)}$ the restriction of $V_i$ to $\Omega_j$, and introduce bilinear forms $a_i^{(j)}(\cdot,\cdot)$ on $V_i^{(j)} \times V_i^{(j)}$, where the bilinear forms

$$\tilde{a}_i(u_i, v_i) = \sum_j \tilde{a}_i^{(j)}(u_i, v_i), \quad u_i, v_i \in V_i,$$

are obtained by subassembly just as the bilinear form $a(u,v)$ can be obtained from

$$a^{(j)}(u,v) = \int_{\Omega_j} \rho_j \nabla u \cdot \nabla v \, dx.$$

The case when the space $V^{(j)}$ is a direct sum of the local subspaces $V_i^{(j)}$ has been considered by Mandel [332, 333]. He showed that if one of the sets of local subspaces contains the null space of $a^{(j)}(\cdot,\cdot)$, for each $\Omega_j$, then bounds on the condition number of the global preconditioned problem can be obtained from bounds for individual subdomains.

We formulate a related result that is also useful when the local subspaces do not form a direct sum decomposition.

**Lemma 5.1** *Assume that there exist constants $C_0^{(j)}$ so that for all $u \in V$ there exists a representation $u = \sum_i R_i^T u_i, u_i \in V_i$, such that*

$$\sum_i \tilde{a}_i^{(j)}(u_i, u_i) \leq C_0^{(j)2} a^{(j)}(u,u),$$

*then the constant $C_0$ of Assumption 2.2 can be given by $C_0 = \max_j C_0^{(j)}$.*

*Proof.* This result follows immediately by summing over the substructures. □

When a local approach to the design of iterative substructuring methods works, then there is no real difference between the Poisson case and that of arbitrary positive values of the $\rho_i$. Once the $\tilde{a}_i^{(j)}(\cdot,\cdot)$ have been constructed for the special case, we can simply scale them by $\rho_j$ and add over $j$.

We note that an algorithm due to Smith [421] (Algorithm 5.10) and the standard vertex based iterative substructuring method (Algorithm 5.2) can be analyzed using substructure by substructure estimates. The overlapping additive Schwarz algorithms of Chap. 3 and [179], [180], [181], [183], [466] and another algorithm due to Smith [422] (Algorithm 5.7), on the other hand, are Schwarz methods for which we have been unable to perform an analysis using only local estimates. To our knowledge, no general results, which are independent of the values of the $\rho_i$, exist for those algorithms.

## 5.3 Local Solvers

Iterative substructuring algorithms with good convergence properties are constructed from two types of components: many local solvers and a coarse grid solver. The coarse model typically has only one or a few degrees of freedom per substructure. In this section, we will primarily describe one basic method of constructing the local solvers; see further Sect. 6.2 for a discussion of the important Neumann-Neumann algorithms which form another important family of primal iterative substructuring methods. We note that, in this and in the following chapter, we consider spaces of discrete harmonic extensions; cf. Sect. 4.4 and, in particular, Lemma 4.9.

Our approach in this section can essentially be viewed as a classical splitting of the Schur complement matrix. For simplicity, we write down the preconditioners only for the additive algorithms; similar, but more complicated, formulas can be given for the multiplicative Schwarz methods, etc.

We first recall that the Schur complement for the entire problem is obtained through subassembly of the matrices given in (4.12). This results in the formula

$$S = \begin{pmatrix} S_{\mathcal{FF}} & S_{\mathcal{FE}} & S_{\mathcal{FV}} \\ S_{\mathcal{FE}}^T & S_{\mathcal{EE}} & S_{\mathcal{EV}} \\ S_{\mathcal{FV}}^T & S_{\mathcal{EV}}^T & S_{\mathcal{VV}} \end{pmatrix}. \tag{5.1}$$

As in the classical theory for iterative methods, cf. Varga [456], a preconditioner for $S$ can be obtained by a splitting, i.e., by dropping certain blocks, or matrix elements. Here we eliminate not only the off-diagonal blocks of (5.1) but also the subblocks representing the coupling between all pairs of faces, edges, and vertices. The resulting preconditioner has the form

$$B^{-1} = \begin{pmatrix} \bar{S}_{\mathcal{FF}}^{-1} & 0 & 0 \\ 0 & \bar{S}_{\mathcal{EE}}^{-1} & 0 \\ 0 & 0 & \bar{S}_{\mathcal{VV}}^{-1} \end{pmatrix}.$$

The matrix $\bar{S}_{\mathcal{FF}}$ is block diagonal with a block for each face, $\bar{S}_{\mathcal{EE}}$ has a block for each edge, and $\bar{S}_{\mathcal{VV}}$ is diagonal. This is a block-Jacobi preconditioner. We note that each block corresponds to a set of adjacent variables on the interface $\Gamma$ and that we decompose the interface functions by cutting from certain mesh points to some of their neighbors, or, equivalently, by partitioning the degrees of freedom of $\Gamma_h$ into disjoint subsets.

We need to introduce some additional notations. Let $S_{\mathcal{F}^i\mathcal{F}^i}$ be the submatrix of $S$ associated with the face $\mathcal{F}^i$, and let $S_{\mathcal{E}^i\mathcal{E}^i}$ be that of the edge $\mathcal{E}^i$. Similarly, $S_{\mathcal{V}^i\mathcal{V}^i}$ is the diagonal element of $S$ associated with the vertex $\mathcal{V}^i$. Let $R_{\mathcal{F}^i}$ be the rectangular restriction matrix which returns only the components of a global vector associated with the face $\mathcal{F}^i$. Similar restriction matrices, $R_{\mathcal{E}^i}$ and $R_{\mathcal{V}^i}$, are introduced for the edges and individual vertices, respectively. We note, e.g., that, $S_{\mathcal{F}^i\mathcal{F}^i} = R_{\mathcal{F}^i} S R_{\mathcal{F}^i}^T$.

The preconditioner $B^{-1}$ can now be rewritten as

$$B^{-1} = \sum_i R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} + \sum_i R_{\mathcal{E}^i}^T S_{\mathcal{E}^i\mathcal{E}^i}^{-1} R_{\mathcal{E}^i} + \sum_i R_{\mathcal{V}^i}^T S_{\mathcal{V}^i\mathcal{V}^i}^{-1} R_{\mathcal{V}^i},$$

and we also find that

$$B^{-1}S = \sum_i R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} S + \sum_i R_{\mathcal{E}^i}^T S_{\mathcal{E}^i\mathcal{E}^i}^{-1} R_{\mathcal{E}^i} S + \sum_i R_{\mathcal{V}^i}^T S_{\mathcal{V}^i\mathcal{V}^i}^{-1} R_{\mathcal{V}^i} S.$$

This preconditioned matrix is the same as that obtained from an additive Schwarz method with the spaces $\tilde{V}_{\mathcal{F}^i}^h$, $\tilde{V}_{\mathcal{E}^i}^h$, $\tilde{V}_{\mathcal{V}^i}^h$. Here $\tilde{V}_{\mathcal{F}^i}^h$ is the space of discrete harmonic functions defined by the degrees of freedom of $\mathcal{F}_h^i$; the operator $R_{\mathcal{F}^i}^T$ represents the extension by zero to all the nodes of $\Gamma_h \setminus \mathcal{F}_h^i$. Similarly $\tilde{V}_{\mathcal{E}^i}^h$ is defined by arbitrary nodal values on $\mathcal{E}_h^i$ with a trivial extension operator. Thus, $R_{\mathcal{E}^i}^T \tilde{V}_{\mathcal{E}^i}^h$ consists of discrete harmonic extensions of functions with degrees of freedom which vanish at the nodes $\Gamma_h \setminus \mathcal{E}_h^i$ and $R_{\mathcal{V}^i}^T \tilde{V}_{\mathcal{V}^i}^h$ consists of discrete harmonic extensions of functions with degrees of freedom which vanish at all nodes in $\Gamma_h$ except at $\mathcal{V}^i$. We note that functions in $R_{\mathcal{F}^i}^T \tilde{V}_{\mathcal{F}^i}^h$ has values different from zero only in the two subdomains which have that face in common. Similarly, the other spaces have elements different from zero only in the subdomains which have that edge or vertex in common.

To decrease the cost and to avoid computing the elements of the Schur complements, we make some further simplifications. We note that the matrices $S_{\mathcal{E}^i\mathcal{E}^i}$ are quite well conditioned; it follows from Lemmas 4.16 and 4.19 that their condition numbers are $O(1 + \log(H/h))$. We therefore replace $S_{\mathcal{E}^i\mathcal{E}^i}^{-1}$, in the preconditioner, by $1/(h_i\rho_{\mathcal{E}^i})I$. Here $I$ is an identity matrix and $\rho_{\mathcal{E}^i} = \sum_{\partial\Omega_j \cap \mathcal{E}^i \neq \emptyset} \rho_j$. In the Schwarz framework, this corresponds to replacing the exact local bilinear forms $s(R_{\mathcal{E}^i}^T u, R_{\mathcal{E}^i}^T u)$ on the spaces $\tilde{V}_{\mathcal{E}^i}$ by $\tilde{a}_{\mathcal{E}^i}(u,u) = h_i\rho_{\mathcal{E}^i}||u||_{\ell^2(\mathcal{E}^i)}^2$. We can also replace the diagonal element $S_{\mathcal{V}^i\mathcal{V}^i}$ by $h_i\rho_{\mathcal{V}^i} = h_i \sum_{\partial\Omega_j \cap \mathcal{V}^i \neq \emptyset} \rho_j$. The modified preconditioner can then be written as

$$B^{-1} = \sum R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} + \sum \frac{1}{h_i\rho_{\mathcal{E}^i}} R_{\mathcal{E}^i}^T R_{\mathcal{E}^i} + \sum \frac{1}{h_i\rho_{\mathcal{V}^i}} R_{\mathcal{V}^i}^T R_{\mathcal{V}^i}.$$

We note that the second and third sums could be combined into one corresponding to the wire basket $\mathcal{W}^{(i)}$ :

$$B^{-1} = \sum R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} + R_{\mathcal{W}}^T (D_{\mathcal{W}}^{\rho})^{-1} R_{\mathcal{W}}. \tag{5.2}$$

Here the elements of the diagonal matrix $D_{\mathcal{W}}^{\rho}$ are constructed from the weights $\rho_{\mathcal{E}^i}$ and $\rho_{\mathcal{V}^i}$ and the local mesh sizes.

We should also provide a relatively inexpensive algorithm for calculating the action of each $S_{\mathcal{F}^k\mathcal{F}^k}^{-1}$. We do so by solving a linear system associated with the two domains $\Omega_i$ and $\Omega_j$ that share the face $\mathcal{F}^k$. Let $A^{(ij)}$ denote the submatrix of $A$ associated with $\Omega_{ij} = \Omega_i \cup \Omega_j \cup \mathcal{F}^k$. Then,

$$S_{\mathcal{F}^k\mathcal{F}^k}^{-1} = \begin{pmatrix} 0 & I \end{pmatrix} \begin{pmatrix} A_{\mathcal{II}}^{(ij)} & A_{\mathcal{IB}}^{(ij)} \\ A_{\mathcal{IB}}^{(ij)^T} & A_{\mathcal{BB}}^{(ij)} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ I \end{pmatrix}; \tag{5.3}$$

cf. Subsect. 1.3.1. Here the subscripts $\mathcal{I}$ and $\mathcal{B}$ represent the nodes of $\Omega_i \cup \Omega_j$ and $\mathcal{F}^k$, respectively. We note that this is a problem very similar to that discussed in Subsect. 1.3.4. Hence, the action of $S^{-1}_{\mathcal{F}^k\mathcal{F}^k}$ can be calculated by solving a homogeneous Dirichlet problem on $\Omega_{ij}$ with a right hand side that differs from zero only on $\mathcal{F}^k_h$. In this construction, we could also replace $\Omega_{ij}$ by any shape-regular region that contains the face $\mathcal{F}^k$ in its interior. We stress that the solution of the local problems never requires the explicit construction of elements of $S$. Instead, in each iteration, independent Dirichlet boundary value problems are solved for regions enclosing the individual faces.

$S_{\mathcal{F}^k\mathcal{F}^k}$ can also be replaced by the $J$ operator, the square root of minus an appropriate discrete Laplacian, introduced in Dryja [172] and at the end of Sect. 1.3.4, in case we have a regular tensor product triangulation of the face. Other preconditioners that are known to be effective for problems on the union of two substructures; cf., e.g., Bjørstad and Widlund [61] and Chap. 1 can also be considered.

In the splittings just considered, we eliminate the coupling between all pairs of faces and all pairs of edges; for many more details, see Sect. 5.4. In our second main approach, we attempt to maintain this coupling. To keep the problems local, we instead eliminate the coupling between neighboring subdomains working with the full Schur complements of the individual substructures; see the discussion of Neumann-Neumann methods in Sect. 6.2.

## 5.4 Coarse Spaces and Condition Number Estimates

In addition to the local solvers discussed in the preceding section, any successful domain decomposition preconditioner must also contain a global space component; see the discussion on page 17. We can either add a coarse solver to a preconditioner based only on local solvers or replace part of the preconditioner. In this section, we will discuss several coarse spaces. The first of them is based on the space $V^H$ of continuous, piecewise linear functions using the substructures as elements. Conceptually this is clearly the simplest, but as will be shown, it can be inadequate in three dimensions, basically because of Lemma 4.12. In the remaining subsections, we discuss wire basket based and face based coarse problems.

Our analysis will be based on the abstract Schwarz theory developed in Chap. 2. We note that for all the algorithms, it is easy to obtain a bound on $\rho(\mathcal{E})$ by using a standard coloring argument; cf. Subsect. 2.5.1. We simply note that the subdomains associated with the local subspaces form an overlapping cover of the domain, and that every point in the domain is covered by a finite, uniformly bounded number of such subregions. The subregions can be grouped into sets, with elements that do not overlap, and the subspaces related to these sets can be merged. The number of subspaces is then reduced to a constant and a uniform upper bound for $\rho(\mathcal{E})$ is obtained.

### 5.4.1 Vertex Based Methods

In this subsection, we will assume, in order to simplify our work, that all the subdomains are tetrahedral; we note that since every $u_H \in V^H$ is piecewise linear, it is harmonic in each substructure, and thus $V^H$ consists of piecewise discrete harmonic functions. To incorporate a global component into the preconditioner, we first represent $S$ in a partially hierarchical basis. The face and edge nodal basis functions are not changed, but those associated with the vertices are replaced by piecewise linear functions on the coarse triangulation. The basis change from the partially hierarchical to the nodal basis is represented by

$$\begin{pmatrix} I & 0 & \hat{R}_{\mathcal{F}}^T \\ 0 & I & \hat{R}_{\mathcal{E}}^T \\ 0 & 0 & I \end{pmatrix},$$

where the operators $\hat{R}_{\mathcal{F}}^T$ and $\hat{R}_{\mathcal{E}}^T$ represent coarse space linear interpolation from the values at the vertices to the faces and edges, respectively.

The Schur complement (5.1) can be rewritten as

$$S = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -\hat{R}_{\mathcal{F}} & -\hat{R}_{\mathcal{E}} & I \end{pmatrix} \begin{pmatrix} S_{\mathcal{F}\mathcal{F}} & S_{\mathcal{F}\mathcal{E}} & \tilde{S}_{\mathcal{F}\mathcal{V}} \\ S_{\mathcal{F}\mathcal{E}}^T & S_{\mathcal{E}\mathcal{E}} & \tilde{S}_{\mathcal{E}\mathcal{V}} \\ \tilde{S}_{\mathcal{F}\mathcal{V}}^T & \tilde{S}_{\mathcal{E}\mathcal{V}}^T & \tilde{S}_{\mathcal{V}\mathcal{V}} \end{pmatrix} \begin{pmatrix} I & 0 & -\hat{R}_{\mathcal{F}}^T \\ 0 & I & -\hat{R}_{\mathcal{E}}^T \\ 0 & 0 & I \end{pmatrix}; \tag{5.4}$$

we note that the inverses of the special block triangular matrices are obtained by switching the sign of the off-diagonal blocks. As in the preceding section, we now drop the off-diagonal blocks as well as the coupling between all pairs of faces and all pairs of edges but we keep the matrix block in the lowest right position intact. We obtain,

$$B = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -\hat{R}_{\mathcal{F}} & -\hat{R}_{\mathcal{E}} & I \end{pmatrix} \begin{pmatrix} \bar{S}_{\mathcal{F}\mathcal{F}} & 0 & 0 \\ 0 & \bar{S}_{\mathcal{E}\mathcal{E}} & 0 \\ 0 & 0 & \tilde{S}_{\mathcal{V}\mathcal{V}} \end{pmatrix} \begin{pmatrix} I & 0 & -\hat{R}_{\mathcal{F}}^T \\ 0 & I & -\hat{R}_{\mathcal{E}}^T \\ 0 & 0 & I \end{pmatrix}.$$

The preconditioner can now be written as

$$B^{-1} = \begin{pmatrix} I & 0 & \hat{R}_{\mathcal{F}}^T \\ 0 & I & \hat{R}_{\mathcal{E}}^T \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} \bar{S}_{\mathcal{F}\mathcal{F}}^{-1} & 0 & 0 \\ 0 & \bar{S}_{\mathcal{E}\mathcal{E}}^{-1} & 0 \\ 0 & 0 & \tilde{S}_{\mathcal{V}\mathcal{V}}^{-1} \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ \hat{R}_{\mathcal{F}} & \hat{R}_{\mathcal{E}} & I \end{pmatrix}$$

and

$$B^{-1}S = \sum R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} S + \sum R_{\mathcal{E}^i}^T S_{\mathcal{E}^i\mathcal{E}^i}^{-1} R_{\mathcal{E}^i} S + R_H^T \tilde{S}_{\mathcal{V}\mathcal{V}}^{-1} R_H S,$$

where $R_H = (\hat{R}_{\mathcal{F}} \;\; \hat{R}_{\mathcal{E}} \;\; I)$. Thus, we obtain an additive Schwarz preconditioner with the same face and edge spaces as before but with a coarse space, $\tilde{V}^H$, in place of the set of individual, local vertex spaces. In the case of piecewise

linear finite elements, the matrix $\tilde{S}_{\mathcal{V}\mathcal{V}}$ is equal to $A^H$, the stiffness matrix obtained by treating the substructures as elements; we note that there are no interior degrees of freedom in this special case. We can therefore replace the two last terms in the preconditioner and obtain

$$B^{-1} = \sum R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} + \sum \frac{1}{h_i \rho_{\mathcal{E}^i}} R_{\mathcal{E}^i}^T R_{\mathcal{E}^i} + R_H^T A^{H^{-1}} R_H.$$

As before, there is no need to form the matrix $S$ explicitly.

**Algorithm 5.2** *Use a Schwarz method with the subspaces $\tilde{V}^H$, $\tilde{V}_{\mathcal{F}^i}^h$, and $\tilde{V}_{\mathcal{E}^j}^h$. For all the $\tilde{V}_{\mathcal{E}^j}^h$ spaces, use the bilinear forms associated with $h_j \rho_{\mathcal{E}^j} ||u||^2_{\ell^2(\mathcal{E}^j)}$.*

**Theorem 5.3** *Algorithm 5.2 satisfies the three assumptions of Sect. 2.3 with*

$$C_0^2 \leq C(H/h)(1 + \log(H/h))^2, \qquad \rho(\mathcal{E}) \leq C, \qquad \omega \leq C.$$

*The constants in the bounds are independent not only of the mesh size and the number of substructures, but also of the values $\rho_i$ of the coefficient of (4.3).*

*Proof.* We first estimate the first stability parameter, $C_0$. We note that we are only going to work with discrete harmonic functions for which $s(u,u) = a(u,u)$. Our subspaces form a direct sum, and there is therefore a unique decomposition for every $u \in \tilde{V}^h(\Omega)$. Let $u_0 = I^H u$. We use Lemma 4.12 and find, by adding over the substructures, that

$$a(R_0^T u_0, R_0^T u_0) \leq C \sum_i \rho_i |I^H u|^2_{H^1(\Omega_i)} \leq C(H/h) a(u,u),$$

where $R_0 = R_H$. We next bound the energy for the parts of the decomposition of the function $u$ that are associated with the faces. This requires the use of Lemmas 4.12 and 4.24. Let $w = u - I^H u$, and let $u_{\mathcal{F}^k} = \mathcal{H}(\vartheta_{\mathcal{F}^k} w)$, where $\vartheta_{\mathcal{F}^k}$ is the function in Lemma 4.23 and $\mathcal{H}(v)$ denotes the discrete harmonic extension of the function $I^h v$ given on the interface $\Gamma$; cf. Sect. 4.4. Then,

$$\begin{aligned} a(R_{\mathcal{F}^k}^T u_{\mathcal{F}^k}, R_{\mathcal{F}^k}^T u_{\mathcal{F}^k}) &= \rho_i |u_{\mathcal{F}^k}|^2_{H^1(\Omega_i)} + \rho_j |u_{\mathcal{F}^k}|^2_{H^1(\Omega_j)} \\ &\leq C(1 + \log(H/h))^2 (\rho_i \|w\|^2_{H^1(\Omega_i)} + \rho_j \|w\|^2_{H^1(\Omega_j)}) \\ &\leq C(H/h)(1 + \log(H/h))^2 (\rho_i |u|^2_{H^1(\Omega_i)} + \rho_j |u|^2_{H^1(\Omega_j)}). \end{aligned}$$

We recall that we use scaled norms on the substructures; cf. Equation (4.4). Therefore,

$$\sum_k a(R_{\mathcal{F}^k}^T u_{\mathcal{F}^k}, R_{\mathcal{F}^k}^T u_{\mathcal{F}^k}) \leq C(H/h)(1 + \log(H/h))^2 a(u,u).$$

Let $u_{\mathcal{E}^i} = R_{\mathcal{E}^i}(u - I^H u)$. Then, by Lemmas 4.12 and 4.16, and the equivalence of the $L^2(\mathcal{E}^i)$- and the scaled $\ell^2(\mathcal{E}^i)$-norm (cf. the discussion after Lemma 4.19), we find

$$\begin{aligned}\sum_i h_i \rho_{\mathcal{E}^i} \|u_{\mathcal{E}^i}\|^2_{\ell^2(\mathcal{E}^i)} &\leq C(1+\log(H/h)) \sum_i \sum_{\partial\Omega_j \cap \mathcal{E}^i \neq \emptyset} \rho_j \|w\|^2_{H^1(\Omega_j)} \\ &\leq C(H/h)(1+\log(H/h)) \sum_j \rho_j |u|^2_{H^1(\Omega_j)} \\ &\leq C(H/h)(1+\log(H/h)) a(u,u).\end{aligned}$$

We have thus found a bound for the stability constant $C_0$. In order to obtain a bound for $\omega$, we only have to consider the edge spaces since exact solvers are employed on all the other subspaces. The constant upper bound follows directly from Lemma 4.19. □

*Remark 5.4.* It is clear that the $H/h$ factor is directly attributable to the potentially large energy of the coarse mesh interpolant. The situation is much more favorable in two dimensions. In that case, we employ the preconditioner

$$B^{-1} = \sum R^T_{\mathcal{E}^i} S^{-1}_{\mathcal{E}^i \mathcal{E}^i} R_{\mathcal{E}^i} + R^T_H A^{H^{-1}} R_H,$$

which consists of a coarse component build on a coarse vertex based space and local components associated with the edges of the partition. The application of each $S^{-1}_{\mathcal{E}^i \mathcal{E}^i}$ requires the solution of a Dirichlet problem on the union of the two substructures that share the edge $\mathcal{E}^i$. A bound for the interpolant $I^H$ is given in Remark 4.13, while a result analogous to Lemma 4.24 holds for each edge; cf. [180]. In this case, we have the following bounds

$$C_0^2 \leq C(1+\log(H/h))^2, \qquad \rho(\mathcal{E}) \leq C, \qquad \omega \leq C.$$

We also note that this is essentially the algorithm developed in an early paper by Bramble, Pasciak, and Schatz [72].

In the proof of Theorem 5.3, we must use $I^H u$ because all functions in the other subspaces vanish at the vertices. In the next algorithm, we add the one-dimensional spaces associated with each vertex and its standard nodal basis function. After doing so, we obtain a much stronger result, but the bounds are no longer independent of the variation of the coefficient of (4.3) across the interface $\Gamma$. The additive Schwarz preconditioner is now given by

$$B^{-1} = \sum R^T_{\mathcal{F}^i} S^{-1}_{\mathcal{F}^i \mathcal{F}^i} R_{\mathcal{F}^i} + \sum \frac{1}{h_i \rho_{\mathcal{E}^i}} R^T_{\mathcal{E}^i} R_{\mathcal{E}^i} + R^T_H A^{H^{-1}} R_H + \sum \frac{1}{h_i \rho_{\mathcal{V}^i}} R^T_{\mathcal{V}^i} R_{\mathcal{V}^i}.$$

We note that, as in (5.2), we can combine the edge and vertex spaces into a single wire basket space, $\tilde{V}^h_{\mathcal{W}}$, with a corresponding restriction operator, $R_{\mathcal{W}}$.

We obtain

$$B^{-1} = \sum R_{\mathcal{F}^i}^T S_{\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} + R_H^T A^{H^{-1}} R_H + R_{\mathcal{W}}^T (D_{\mathcal{W}}^\rho)^{-1} R_{\mathcal{W}}.$$

Here $D_{\mathcal{W}}^\rho$ is the global diagonal matrix constructed from the weights $\rho_{\mathcal{E}^i}$ and $\rho_{\mathcal{V}^i}$, and the local mesh sizes, as in formula (5.2).

**Algorithm 5.5** *Use a Schwarz method with the subspaces $\tilde{V}^H$, $\tilde{V}_{\mathcal{F}^i}^h$, and $\tilde{V}_{\mathcal{W}}^h$. In addition, on the spaces $\tilde{V}_{\mathcal{W}}^h$, use the bilinear form given by $u^T D_{\mathcal{W}}^\rho u$.*

**Theorem 5.6** *Algorithm 5.5 satisfies the three assumptions of Sect. 2.3 with*

$$C_0^2 \leq C(1 + \log(H/h))^2, \qquad \rho(\mathcal{E}) \leq C, \qquad \omega \leq C.$$

*Here we cannot guarantee that the estimate of $C_0^2$ is independent of the jumps in the coefficients of (4.3).*

*Proof.* The proof is almost identical to that given above except that we use $u_0 = \tilde{I}^H u$ and Lemma 3.6 rather than $I^H u$ and Lemma 4.12. □

Conditions on the $\rho_i$ for which the norm of $\tilde{I}^H$, and $C_0^2$, are uniformly bounded are given in Dryja, Sarkis, and Widlund [177].

We can increase the overlap between the subspaces and obtain methods with condition numbers that are uniformly bounded and independent of $H$ and $h$. Such a method was given in Smith [422]. This algorithm, known as the *vertex space algorithm* has much in common with the two-level additive Schwarz methods of Chap. 3 except that we now work with trace spaces defined on the interface $\Gamma$.

To define this algorithm, we first define edge spaces associated with a set $\Gamma^{\mathcal{E}^j}$ that includes all parts of the faces adjacent to the edge $\mathcal{E}^j$ that are within a distance $cH_j$ from the edge, see middle of Fig. 5.1. We also define the vertex

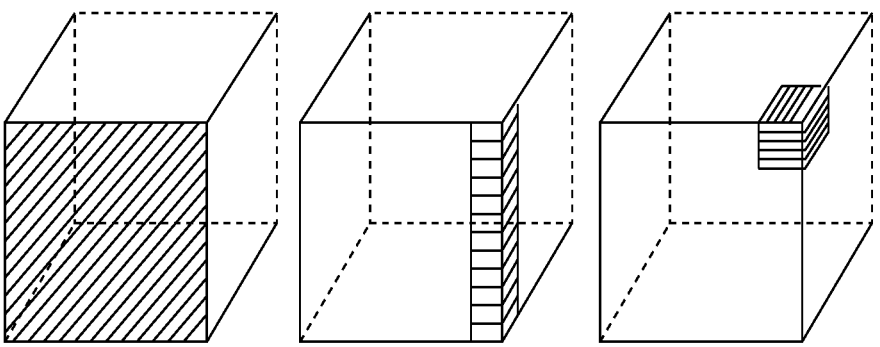

**Fig. 5.1.** Support of face, edge, and vertex spaces.

region $\Gamma^{\mathcal{V}^j}$ as the part of $\Gamma$ that is at a distance less than $cH_j$ from the $j$th vertex of the substructure. The space related to this set, after a trivial extension outside, is

$$R_{\Gamma^{\mathcal{V}^j}} \tilde{V}^h_{\Gamma^{\mathcal{V}^j}} = \{u \in \tilde{V}^h | u(x) = 0, \quad x \in \Gamma_h \setminus \Gamma^{\mathcal{V}^j}\},$$

with a similar definition for $\tilde{V}^h_{\Gamma^{\mathcal{E}^j}}$. For this algorithm, we first consider exact projections; the algorithm is then completely defined by its subspaces.

**Algorithm 5.7** *Use a Schwarz method with the subspaces given by* $\tilde{V}^H$, $\tilde{V}^h_{\mathcal{F}^i}$, $\tilde{V}^h_{\Gamma^{\mathcal{E}^j}}$, *and* $\tilde{V}^h_{\Gamma^{\mathcal{V}^j}}$.

In [422], Smith proved the following result. This proof follows almost immediately from Theorem 3.13; in fact the proof of that theorem can be considerably simplified since we assume that the overlap is generous and the fine mesh is a refinement of the coarse.

**Theorem 5.8** *Algorithm 5.7 satisfies the three assumptions of Sect. 2.3 with*

$$C_0^2 \leq C, \qquad \rho(\mathcal{E}) \leq C, \qquad \omega = 1.$$

*Here we cannot guarantee that the estimate of* $C_0^2$ *is independent of the jumps in the coefficient of (4.3).*

Using the definitions given above, we find that

$$B^{-1} = R_H^T A^{H^{-1}} R_H + \sum_i R^T_{\mathcal{F}^i} S^{-1}_{\mathcal{F}^i\mathcal{F}^i} R_{\mathcal{F}^i} + \sum_j R^T_{\Gamma^{\mathcal{E}^j}} S^{-1}_{\Gamma^{\mathcal{E}^j}\Gamma^{\mathcal{E}^j}} R_{\Gamma^{\mathcal{E}^j}} + \sum_k R^T_{\Gamma^{\mathcal{V}^k}} S^{-1}_{\Gamma^{\mathcal{V}^k}\Gamma^{\mathcal{V}^k}} R_{\Gamma^{\mathcal{V}^k}}.$$

We note that the first term essentially involves solving a system associated with a block of $S$, represented in the partial hierarchical basis, while the other terms involve systems given by blocks of $S$ in the usual nodal basis. In practical implementations, the $S_{\mathcal{F}^i\mathcal{F}^i}$, $S_{\Gamma^{\mathcal{E}^j}\Gamma^{\mathcal{E}^j}}$, and $S_{\Gamma^{\mathcal{V}^k}\Gamma^{\mathcal{V}^k}}$ need not be formed explicitly; we can instead solve problems such as (5.3). Another approach to cutting costs is to use probing to obtain approximations of the blocks of the Schur complement; cf. Chan and Mathew [131] and Chan, Mathew, and Shao [133].

In the analysis given in [422], Smith considered only the case when the overlap is generous, i.e., on the order of $H$. However, numerical experiments suggest that very satisfactory convergence can also be obtained with a quite minimal overlap; cf. Bjørstad et al. [59], [60] and Gropp and Smith [241]. Thus motivated, Dryja and Widlund [182] showed that if the overlap is uniformly on the order of $\delta$, then Algorithm 5.7 satisfies

$$C_0^2 \leq C(1 + \log(H/\delta))^2. \tag{5.5}$$

We note that a case when the overlap is minimal, as for an algorithm based on a direct sum decomposition, corresponds to setting $\delta = h$ in the formula above.

*Remark 5.9.* A similar algorithm with generous overlap can also be developed in two dimensions; see Remark 5.4 for some algorithmic details. In this case, we only have coarse, vertex, and edge components. The logarithmic bound (5.5) can also be found; see [182].

### 5.4.2 Wire Basket Based Algorithms

We will now consider another class of coarse problems based on averages and the wire basket. Methods of this class use a different approach to overcome the difficulties associated with the piecewise linear interpolation over the coarse triangulation, which led to the poor result of Theorem 5.3 or to estimates that are not known to be valid uniformly for all values of the coefficient of (4.3). Instead, we now essentially interpolate using averages of $u$ over the wire basket. These algorithms work quite well for problems with large jumps in the coefficients; cf. Smith [423] which is a report on an early fully parallel implementation of this algorithm. We note that Bramble, Pasciak, and Schatz [76] pioneered the use of similar ideas. Here, we begin by describing a method introduced in Smith [420]; cf. also [421].

For the wire basket based methods, we work with the block matrix (4.13) rather than (4.12). Let $\hat{R}_{\mathcal{W}}^T$ be the operator that maps the values on the wire basket onto the faces by assigning, to each node on a face, the average value of the nodal values on the boundary of the face. This represents an alternative to the change of basis of the space given in equation (5.4). $S$ can now be written as

$$S = \begin{pmatrix} I & 0 \\ -\hat{R}_{\mathcal{W}} & I \end{pmatrix} \begin{pmatrix} S_{\mathcal{FF}} & \bar{S}_{\mathcal{FW}} \\ \bar{S}_{\mathcal{FW}}^T & \bar{S}_{\mathcal{WW}} \end{pmatrix} \begin{pmatrix} I & -\hat{R}_{\mathcal{W}}^T \\ 0 & I \end{pmatrix}.$$

We note the similarity with (5.4), but that we are now using piecewise constant interpolation onto the faces instead of a piecewise linear interpolation onto the faces and edges. Equivalently, the previous transformation is a change of basis, where a standard nodal basis function associated with a degree of freedom on the wire basket is transformed into ones that has constant values on each adjacent face, given by the average of the nodal basis function on the boundary of the face in question. We proceed as in the preceding subsection and drop the coupling between pairs of faces, and the faces and the wire basket. We obtain

$$B^{-1} = \begin{pmatrix} I & \hat{R}_{\mathcal{W}}^T \\ 0 & I \end{pmatrix} \begin{pmatrix} \bar{S}_{\mathcal{FF}}^{-1} & 0 \\ 0 & \bar{S}_{\mathcal{WW}}^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ \hat{R}_{\mathcal{W}} & I \end{pmatrix}$$

and

$$B^{-1}S = R_0^T \bar{S}_{\mathcal{W}\mathcal{W}}^{-1} R_0 S + \sum_i R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} S,$$

where $R_0 = (\hat{R}_{\mathcal{W}} \quad I)$. This is also an additive Schwarz scheme that uses the same face spaces $\tilde{V}_{\mathcal{F}^i}^h$ as the vertex based algorithms. The coarse space, $\tilde{V}_{\mathcal{W}}^h$, can conveniently be defined as the range of an interpolation operator $I_{\mathcal{W}}^h : \tilde{V}^h \to \tilde{V}_{\mathcal{W}}^h$, defined by

$$I_{\mathcal{W}}^h u = \sum_{x_k \in \mathcal{W}_h} u(x_k)\theta_{x_k} + \sum_{\mathcal{F}^k \subset \Gamma} \bar{u}_{\partial\mathcal{F}^k}\theta_{\mathcal{F}^k}. \tag{5.6}$$

Here $\theta_{x_k}$ is the discrete harmonic extension of the standard nodal basis functions $\phi_k$ associated with the node $x_k$; see Definition 4.27. The resulting finite element function is continuous across all substructure interfaces. Therefore, $\tilde{V}_{\mathcal{W}}^h$ is a conforming subspace of $\tilde{V}^h$. We note that we can obtain a much better bound on the norm of this operator than for the standard coarse space interpolant $I^H$ and that this is a key to obtaining better bounds than those in Theorem 5.2.

We note that this coarse space has a relatively large dimension and that the use of an exact solver would be relatively expensive. Instead, we will therefore use the bilinear form given by

$$\tilde{a}_0^{\mathcal{W}}(u,u) = \sum_i (1 + \log(H_i/h_i))h_i\rho_i \min_{\eta_i} \|u - \eta_i\|^2_{\ell^2(\mathcal{W}^i)} \tag{5.7}$$

for this subspace, where the $\eta_i$ are real parameters. The introduction of this bilinear form corresponds to replacing $\bar{S}_{\mathcal{W}\mathcal{W}}$ by a matrix that is built from substructure contributions that each is a simple rank-one perturbation of a multiple of the identity matrix. We remark that the minimum value for the quantity $\eta_i$ is equal to the average of the nodal values of $u$ on the wire basket $\mathcal{W}_i$; see Equation (5.8) below.

In order to solve a linear system involving this bilinear form and a right hand side $r$,

$$\tilde{a}_0^{\mathcal{W}}(u,v) = v^T r, \quad v \in \tilde{V}_{\mathcal{W}}^h,$$

we can use a fast technique suggested by Mandel [331]; cf. also Smith [420], [421]: let $D^{(i)} = (1 + \log(H_i/h_i))h_i\rho_i I$, and rewrite the problem as

$$\min_{u \in \tilde{V}_{\mathcal{W}}^h} \sum_i \min_{\eta_i} \frac{1}{2}(u^{(i)} - \eta_i \mathbf{1}^{(i)})^T D^{(i)}(u^{(i)} - \eta_i \mathbf{1}^{(i)}) - u^T r,$$

where all the components of the vector $\mathbf{1}^{(i)}$ are equal to one. We then take derivatives with respect to $\eta_i$ and $u$ and obtain the linear system

$$\mathbf{1}^{(i)^T} D^{(i)}(u^{(i)} - \mathbf{1}^{(i)}\eta_i) = 0, \tag{5.8}$$

$$Du - \sum_i D^{(i)}\mathbf{1}^{(i)}\eta_i = r. \tag{5.9}$$

Here, $D$ is the diagonal matrix obtained by subassembling the $D^{(i)}$. We then eliminate $u$ and obtain the following system for the averages $\eta_i$ :

$$(\mathbf{1}^{(i)^T} D^{(i)} \mathbf{1}^{(i)})\eta_i - \mathbf{1}^{(i)^T} D^{(i)} D^{-1} \sum_j D^{(j)} \mathbf{1}^{(j)} \eta_j = \mathbf{1}^{(i)^T} D^{(i)} D^{-1} r.$$

We note that this is a sparse linear system with coupling only between next neighboring subdomains. Once the $\eta_i$ are known, $u$ can be found by solving the diagonal system (5.9).

**Algorithm 5.10** *Use a Schwarz method with the subspaces given by $\tilde{V}_{\mathcal{W}}^h$ and $\tilde{V}_{\mathcal{F}^k}^h$ and the bilinear form given by $\tilde{a}_0^{\mathcal{W}}(u,u)$ for the space $\tilde{V}_{\mathcal{W}}^h$.*

**Theorem 5.11** *Algorithm 5.10 satisfies the three assumptions of Sect. 2.3 with*

$$C_0^2 \leq C(1+\log(H/h))^2, \qquad \rho(\mathcal{E}) \leq C, \qquad \omega \leq C.$$

*The constants in the bounds are independent not only of the mesh size and the number of substructures, but also of the values $\rho_i$ of the coefficient of (4.3).*

*Proof.* We begin by estimating $\omega$. For the spaces $\tilde{V}_{\mathcal{F}^k}^h$, $\omega = 1$, trivially, since we use exact projections. Let $u_0 \in \tilde{V}_{\mathcal{W}}^h$ and let $u_0^{(i)}$ be the restriction of $u_0$ to $\bar{\Omega}_i$. Let $w_0^{(i)} = u_0^{(i)} - \bar{u}_0^{(i)}$, where $\bar{u}_0^{(i)}$ is the average of the nodal values of $u_0^{(i)}$ on the wire basket. Since the mesh is quasi uniform, we can replace the $\ell^2(\mathcal{W}^i)$-norm with the $L^2(\mathcal{W}^i)$-norm (cf. the remarks after the proof of Lemma 4.19). We also recall that the minimum values for the quantities $\eta_i$ are equal to the averages of the nodal values of $u$ on the local wire baskets $\mathcal{W}_i$, or, in case the $L^2(\mathcal{W}^i)$-norm is employed, to the averages of $u$ on the $\mathcal{W}_i$.

According to Equation (5.6), we can split $w_0^{(i)}$ into two parts

$$w_0^{(i)} = \sum \bar{w}_{0,\partial\mathcal{F}^k}^h \theta_{\mathcal{F}^k} + w_{\mathcal{W}}^{(i)}.$$

The first has constant values on the faces while the second vanishes there. Then, using Lemmas 4.25, 4.19, and 4.30, we obtain

$$\begin{aligned}
a(R_0^T u_0, R_0^T u_0) &= \sum_i \rho_i |u_0^{(i)}|^2_{H^1(\Omega_i)} \\
&= \sum_i \rho_i |w_0^{(i)}|^2_{H^1(\Omega_i)} \\
&\leq C \sum_i \rho_i \Big(\sum_k (\bar{w}_{0,\partial\mathcal{F}^k})^2 |\theta_{\mathcal{F}^k}|^2_{H^1(\Omega_i)} + |w_{\mathcal{W}}^{(i)}|^2_{H^1(\Omega_i)}\Big) \\
&\leq C \sum_i (1+\log(H/h))\rho_i \|w_0^{(i)}\|^2_{L^2(\mathcal{W}^i)} \\
&\leq C\tilde{a}_0^{\mathcal{W}}(u_0,u_0).
\end{aligned}$$

We now estimate the stability constant $C_0^2$. Let $u$ be an arbitrary discrete harmonic function and let $u_0 := I_{\mathcal{W}}^h u$. If $\bar{u}_{\mathcal{W}^i}$ is the average of $u$ (or, equivalently, of $u_0$) on the wire basket $\mathcal{W}^i$, we find using Lemma 4.16, that

$$\begin{aligned}
\tilde{a}_0^{\mathcal{W}}(u_0, u_0) &\leq C \sum_i (1 + \log(H_i/h_i)) \min_{\hat{\eta}_i} \rho_i ||u_0 - \hat{\eta}_i||^2_{L^2(\mathcal{W}^i)} \\
&= C \sum_i (1 + \log(H_i/h_i)) \rho_i ||u_0 - \bar{u}_{\mathcal{W}^i}||^2_{L^2(\mathcal{W}^i)} \\
&\leq C \sum_i (1 + \log(H_i/h_i))^2 \rho_i |u|^2_{H^1(\Omega_i)} \\
&= C(1 + \log(H/h))^2 a(u, u).
\end{aligned}$$

There remains to estimate the face contributions

$$u_{\mathcal{F}^k} = \mathcal{H}(\vartheta_{\mathcal{F}^k}(u - u_0)) = \mathcal{H}(\vartheta_{\mathcal{F}^k}(u - \bar{u}_{\partial \mathcal{F}^k})).$$

We note that the values of $u_0$ on the wire basket are irrelevant since $\vartheta_{\mathcal{F}^k}$ vanishes on that set. Therefore, by using Lemmas 4.26 and 4.6 We find

$$\begin{aligned}
a(R_{\mathcal{F}^k}^T u_{\mathcal{F}^k}, R_{\mathcal{F}^k}^T u_{\mathcal{F}^k}) &\leq C\rho_i |I^h(\vartheta_{\mathcal{F}^k}(u - \bar{u}_{\partial \mathcal{F}^k}))|^2_{H^{1/2}(\partial \Omega_i)} \\
&\quad + C\rho_j |I^h(\vartheta_{\mathcal{F}^k}(u - \bar{u}_{\partial \mathcal{F}^k}))|^2_{H^{1/2}(\partial \Omega_j)} \\
&\leq C(1 + \log(H/h))^2 \left(\rho_i |u|^2_{H^1(\Omega_i)} + \rho_j |u|^2_{H^1(\Omega_j)}\right).
\end{aligned}$$

We then sum over the faces of the partition to obtain the bound for $C_0^2$. □

We conclude this subsection by mentioning two earlier wire basket based algorithms due to Bramble, Pasciak, and Schatz [76]. Their work has influenced much of the later work in the field. One of their coarse spaces is given in terms of the averages of the nodal values over the entire substructure boundaries $\partial \Omega_i$. The other space is defined by extending the wire basket values as a two-dimensional discrete harmonic function onto the faces, and then as a discrete harmonic function into the interiors of the subdomains. For both methods, Bramble, Pasciak, and Schatz proved the same bounds as in Theorem 5.11; cf. [76].

### 5.4.3 Face Based Algorithms

We have seen that the vertex space method of Algorithm 5.7 has a condition number that is independent of the parameters $h$ and $H$ but that this bound might not be independent of the variations of the coefficients across the interface $\Gamma$. We could explore the possibility of replacing the coarse space $\tilde{V}^H$ by the wire basket space $\tilde{V}_{\mathcal{W}}^h$ and the bilinear form as in Algorithm 5.10. The local spaces could be chosen as in Algorithm 5.7. This leads to an algorithm for which we can prove the same type of bounds as in Theorem 5.11, i.e., that the condition number is bounded by $C(1 + \log(H/h))^2$. We can also show that

a bound with $C(1+\log(H/h))$ holds if we allow the constant to depend on the variation of the $\rho_i$.

However, there are alternative coarse spaces for which it is possible to derive bounds on the condition number that are independent of the values of $\rho_i$ and that are only linear in $(1+\log(H/h))$; these results were first given in [178]. The main ideas behind the first of these algorithms is to expand the coarse space by allowing an additional degree of freedom for each face, rather than specifying the values on the face in terms of values on all or part of the wire basket. We note that some of the coarse spaces that we will explore in the remainder of this section have inspired the recent development of dual-primal FETI methods which is the topic of Sect. 6.4.

The first face based coarse space, $\tilde{V}_M^h$, can be viewed as the range of the following interpolation operator:

$$I_M^h u(x) = \sum_{x_k \in \mathcal{W}_h} u(x_k)\theta_{x_k}(x) + \sum_{\mathcal{F}^k \subset \Gamma} \bar{u}_{\mathcal{F}^k}\theta_{\mathcal{F}^k}(x). \tag{5.10}$$

We note that $I_M^h$ is similar to $I_{\mathcal{W}}^h$, given in (5.6), except that we replace the averages $\bar{u}_{\partial\mathcal{F}^k}$ with $\bar{u}_{\mathcal{F}^k}$. We equip this new space with the bilinear form

$$\tilde{a}_0^M(u,u) = \sum_i \min_{\eta_i} \rho_i \{h_i ||u-\eta_i||^2_{\ell^2(\mathcal{W}^i)} + \sum_{\mathcal{F}^k \subset \partial\Omega_i} H_i(1+\log(H_i/h_i))(\bar{u}_{\mathcal{F}^k} - \eta_i)^2\}.$$

**Algorithm 5.12** *Use a Schwarz method with the subspaces given by $\tilde{V}_M^h$, $\tilde{V}_{\mathcal{F}^i}^h$, $\tilde{V}_{\Gamma\mathcal{E}^j}^h$, and $\tilde{V}_{\Gamma V^j}^h$ and the bilinear form just given by $\tilde{a}_0^M(u,u)$.*

**Theorem 5.13** *Algorithm 5.12 satisfies the three assumptions of Sect. 2.3 with*

$$C_0^2 \leq C(1+\log(H/h)), \qquad \rho(\mathcal{E}) \leq C, \qquad \omega \leq C.$$

*The constants in the bounds are independent not only of the mesh size and the number of substructures, but also of the values $\rho_i$ of the coefficient of (4.3).*

*Proof.* The proof of the first assumption is quite similar to those of the other theorems of this section. We now use $u_0 = I_M^h u$. Instead of Lemma 3.6, we use the following estimates:

$$\|u - I_M^h u\|^2_{L^2(\Omega_i)} \leq C H_i^2 |u|^2_{H^1(\Omega_i)} \tag{5.11}$$

and

$$|I_M^h u|^2_{H^1(\Omega_i)} \leq C(1+\log(H_i/h_i))|u|^2_{H^1(\Omega_i)}. \tag{5.12}$$

Inequality (5.11) follows from Lemmas 4.25, 4.30, a simple trace theorem, which provides a bound of the $L^2(\partial\Omega_i)$-norm in terms of the $H^1(\Omega_i)$-norm, and a Poincaré inequality. The second inequality, (5.12), is established by

using Lemmas 4.19, 4.16, 4.25, 4.30, and, again, the same trace theorem and Poincaré inequality.

The estimate

$$\tilde{a}_0^M(u_0, u_0) \leq C(1 + \log(H/h))a(u, u),$$

with $u_0 := I_M^h u$, follows from Lemmas 4.16, 4.25, 4.30, and the same trace theorem and Poincaré inequality. The estimate is developed one substructure at a time. Instead of the minimizing $\eta_i$, we can use $\bar{u}_{\partial\Omega_i}$, the average over the subdomain boundary. This choice of shift makes it possible to use the Poincaré inequality and also to avoid introducing an additional logarithmic factor.

The bounds for the other components of the decomposition does not increase the value of $C_0$ since we are using spaces which overlap generously. We use the estimates (5.11) and (5.12).

Finally, again working with one subdomain at a time, we use Lemma 4.25 and the fact that $I_M^h$ reproduces constants to establish that

$$a(R_0^T u_0, R_0^T u_0) \leq C\tilde{a}_0^M(u_0, u_0), \quad u_0 \in \tilde{V}_M^h.$$

□

*Remark 5.14.* In Chap. 7, we will show that all the results for the iterative substructuring methods of this chapter and Chap. 6 hold in the spectral element case as well. These results are obtained directly once we have established that almost all the results of Sect. 4.6 are equally valid in the spectral element case. An exception is the $L^2$-bound given in Lemma 4.25 which we have just used in the proof of formula (5.11). However, we can avoid this problem by replacing $\theta_{\mathcal{F}^k}$ by $\vartheta_{\mathcal{F}^k}$ in formula (5.10) for which the $L^2$-bound is trivial and which also results in an interpolant which works equally well.

We can decrease the dimension of the global space just considered. Rather than using the coarse subspace, involving all of the nodal values on the edges, only one degree of freedom per edge, an average value, suffices. The resulting space, denoted by $\tilde{V}_B^h$, is the range of the interpolation operator

$$I_B^h u(x) = \sum_{\mathcal{V}^k \in \Gamma} u(\mathcal{V}^k)\theta_{\mathcal{V}_k}(x) + \sum_{\mathcal{E}^i \subset \mathcal{W}} \bar{u}_{\mathcal{E}^i}\theta_{\mathcal{E}^i}(x) + \sum_{\mathcal{F}^k \subset \Gamma} \bar{u}_{\mathcal{F}^k}\theta_{\mathcal{F}^k}(x). \quad (5.13)$$

Here $\bar{u}_{\mathcal{E}^i}$ is the average of the values of $u$ on $\mathcal{E}_h^i$ and $\theta_{\mathcal{E}^i}$ the discrete harmonic function which equals 1 on that set and vanishes elsewhere on $\Gamma_h$. We define the bilinear form for this space by

$$\tilde{a}_0^B(u, u) = \sum_i \rho_i \min_{\eta_i} \{h_i \sum_{\mathcal{V}^k \in \partial\Omega_i} (u(\mathcal{V}^k) - \eta_i)^2 +$$

$$H_i \sum_{\mathcal{E}^i \subset \partial\Omega_i} (\bar{u}_{\mathcal{E}^i} - \eta_i)^2 + H_i(1 + \log(H_i/h_i)) \sum_{\mathcal{F}^k \subset \partial\Omega_i} (\bar{u}_{\mathcal{F}^k} - \eta_i)^2\}.$$

*Remark 5.15.* Since the number of unknowns of the coarse space is just a small multiple of the number of subregions, we could also consider using an exact solver with this coarse space.

**Algorithm 5.16** *Use a Schwarz method with the subspaces given by* $\tilde{V}_B^h$, $\tilde{V}_{\mathcal{F}^i}^h$, $\tilde{V}_{\Gamma\mathcal{E}^j}^h$, *and* $\tilde{V}_{\Gamma\mathcal{V}^j}^h$ *and the bilinear form associated with* $\tilde{a}_0^B(u,u)$.

**Theorem 5.17** *Algorithm 5.16 satisfies the three assumptions with*

$$C_0^2 \leq C(1+\log(H/h)), \qquad \rho(\mathcal{E}) \leq C, \qquad \omega \leq C.$$

*The constants in the bounds are independent not only of the mesh size and the number of substructures, but also of the values* $\rho_i$ *of the coefficient of (4.3).*

*Proof.* All the estimates needed are close counterparts of those of Theorem 5.13. Thus, we use the interpolant $I_B^h u$ to define the coarse space component $u_0$. Using similar techniques as before, we can establish the estimates

$$\|u - I_B^h u\|_{L^2(\Omega_i)}^2 \leq H_i^2 |u|_{H^1(\Omega_i)}^2$$

and

$$|I_B^h u|_{H^1(\Omega_i)}^2 \leq C(1+\log(H_i/h_i))|u|_{H^1(\Omega_i)}^2.$$

We note that it is easy to show that the operator $I_B^h$ reproduces constants; see (5.13). The proof of the upper bound

$$\tilde{a}_0^B(u_0,u_0) \leq C(1+\log(H/h))a(u,u),$$

also follows by an argument very similar to that in the proof of Theorem 5.13. We can also show that

$$a(R_0^T u_0, R_0^T u_0) \leq C\tilde{a}_0^B(u_0,u_0), \quad u_0 \in \tilde{V}_B^h.$$

□

# 6

# Neumann-Neumann and FETI Methods

## 6.1 Introduction

The FETI and Neumann-Neumann families of algorithms are among the best known and most severely tested domain decomposition methods for elliptic partial differential equations; cf., e.g., [51], [201], and [308]. They are iterative substructuring methods and share many algorithmic components, such as local solvers for both Neumann and Dirichlet problems on the subregions into which the region of the original problem has been partitioned. We note that while the local subspaces of the algorithms of Chap. 5 are related to faces, edges, wire baskets, and vertices of the subdomains, we will encounter subspaces related to the entire boundaries of our subdomains in Sect. 6.2 and 6.3.

We will begin this chapter by a description and analysis of the balancing Neumann-Neumann method previously briefly discussed in Subsect. 1.3.4 and 1.3.6. We establish an auxiliary result, Lemma 6.3, which is also central in the analysis of the FETI methods. We then discuss two families of FETI algorithms namely the one-level FETI and the dual-primal FETI (FETI-DP) methods. The FETI methods are dual iterative substructuring methods where the iterates are discontinuous across the interface; these jumps will only disappear at the convergence of the iteration. The iteration is written in terms of Lagrange multipliers which are dual variables directly associated with the requirement of continuity of the finite element solution across the interface. We note that some of these ideas were already briefly discussed in Subsect. 1.3.5. There the Lagrange multipliers were approximations of the fluxes across the subdomain interface.

As was shown in Subsect. 1.3.4 and 1.3.5, the Neumann-Neumann (FETI, respectively) method provides a preconditioner for the Schur complement system of Sect. 1.3.1 (for the flux equation in Sect. 1.3.2, respectively), by solving Neumann problems using the flux jumps (Dirichlet problems and function jumps, respectively) on each subdomain and then correcting the previous iterate with the corresponding function (flux) values. Similar ideas were already

employed by Glowinski and Wheeler [226, 227] for mixed approximations of elliptic problems.

The term one-level in the FETI context refers to the type of coarse component or solver employed and not to the overall algorithm. It is in contrast to two-level FETI methods, which were developed primarily for biharmonic and shell element problems by Farhat et al [198, 192] and involve satisfying some of the continuity constraints in each step of the iterations. We will not discuss these methods in this monograph but will instead focus on one-level and dual-primal FETI methods. The term dual-primal refers to the idea of enforcing relatively few continuity constraints, across the interface between the subregions, throughout the iteration, as in a primal method, while all other constraints are enforced by using dual variables, i.e., Lagrange multipliers, as in a one-level FETI method. We will see that the FETI-DP methods differ in several important respects from the strictly dual FETI methods, in particular the one-level FETI methods described in Sect. 6.3. In fact, from an algorithmic point of view, the FETI-DP methods are closer to the primal iterative substructuring methods of Chap. 5 than to the one-level FETI methods. While the global part of the preconditioner for a strictly dual FETI method is directly associated with the dual variables, that of a FETI-DP method is not. We note that primal iterative substructuring methods had been studied quite extensively, see, e.g., [177], [178], [184], and Chap. 5, well before a similarly complete, and quite challenging, mathematical theory was developed for the FETI methods; see [87, 88], [340], [435], and [289]. FETI algorithms using inexact subdomain solvers have also been developed and analyzed in [288].

The theory for iterative substructuring methods has been developed for elliptic systems, e.g., in [288] and [383]; see also [291] and Chap. 8 where we will extend the theory of this chapter to the system of linear elasticity. Extensions were also performed to spectral elements in, e.g., [378] and to some *hp* approximations on anisotropic meshes in [444, 445, 446, 447]; see Chap. 7 for a more detailed presentation. Neumann-Neumann type and one-level FETI algorithms have been devised for convection-diffusion problems in, e.g., [2, 5, 440]; see Sect. 11.5 for more details. Neumann-Neumann, one-level FETI, and FETI-DP methods for edge element and Raviart-Thomas approximations have been proposed in, e.g., [438, 443, 441, 393, 447, 448]; see Chap. 10 for more references and details and a presentation of some of them. We note in particular that effective FETI-DP algorithms are now available for three-dimensional edge element approximations as well; see [442]. The Neumann-Neumann and FETI-DP methods have also been extended to mixed finite element methods, in particular, for incompressible Stokes and Navier Stokes equations, see [386, 314, 315, 316], and for almost incompressible elasticity, see [230, 229]; see also Chap. 9. We also note that algorithmically some of the FETI-DP methods that we consider, have certain features in common with very early work on iterative substructuring methods for problems with many substructures; cf. the brief discussion in Subsect. 1.3.6 and the studies on Neumann-Dirichlet algorithms in [176], and two contributions to the first

international symposium on domain decomposition methods, [173] and [465]. We note, in particular, that, as indicated in Chap. 1, the Neumann subsystems of these early algorithms are nonsingular; there are no floating subregions because of a device very similar to that used in the FETI-DP methods. The use of Lagrange multipliers, in a special context, was also suggested in [465].

## 6.2 Balancing Neumann-Neumann Methods

In this section, we will describe and analyze a hybrid Schwarz algorithm known as the balancing Neumann-Neumann method. It is one of a family of Neumann-Neumann methods which were first developed without a coarse space; see early papers by Bourgat, De Roeck, Glowinski, Le Tallec, and Vidrascu [65, 153, 154]. They were later considerably improved by adding a second, coarse level; see Dryja and Widlund [184], Le Tallec [308], and Mandel and Brezina [335]. We note that the work in [184] concerned additive methods and that so does some recent work by Mandel and Dohrmann [336, 164, 165]. In contrast to the work in Chap. 5, we will find that the lower bound on the Schwarz operator (see Sect. 2.3 and 2.5.2) can be derived by algebra alone and that the main effort will go into the proof of an upper bound of the norm of the $P_i$ from which the Schwarz operator of the method is built. We will again confine our study primarily to the three-dimensional model problem given by (4.3). References to other types of problems or approximations have been given in Sect. 6.1.

### 6.2.1 Definition of the Algorithm

In the rest of this chapter, we will denote the standard finite element space of continuous, piecewise linear functions on $\Omega_i$ by $W^h(\Omega_i)$, as is standard practice in the FETI literature. We will reserve the notation $V$ for a space of Lagrange multipliers which is introduced in Sect. 6.3.1. We will always assume that the functions of $W^h(\Omega_i)$ vanish on $\partial\Omega_D$ and that the triangulation of each subdomain is quasi uniform.

We denote the corresponding trace spaces by $W_i := W^h(\partial\Omega_i \cap \Gamma)$, $i = 1, \ldots, N$, and by $W := \prod_{i=1}^{N} W_i$ the associated product space; for a definition of the interface $\Gamma$, see Sect. 4.2. Given $u \in W$, we denote its $i$-th component by $u_i$. We will often consider elements of $W$ which are discontinuous across the interface. Thus, in the FETI methods, the continuity of the finite element solution across the interface will be satisfied only at the convergence of the iteration.

We can naturally define a stiffness matrix $A$ and a Schur complement matrix $S$ on these product spaces as direct sums of the subdomain stiffness and Schur complement matrices, respectively. We recall that for a chosen finite element method and for each subdomain $\Omega_i$, we first assemble the local stiffness matrix $A^{(i)}$ and the local load vector corresponding to single, appropriate

terms in the sums of (4.3). The interior variables of any subdomain are then eliminated by a step of block Gaussian elimination in work that can be parallelized across the subdomains. As in Chap. 4, the resulting matrices are the Schur complements

$$S^{(i)} = A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)}(A_{II}^{(i)})^{-1}A_{I\Gamma}^{(i)}, \quad i = 1, \ldots, N,$$

where $\Gamma$ and $I$ represent the interface and interior, respectively. As before, the $S^{(i)}$ are only needed in terms of matrix-vector products.

Our finite element approximation of the elliptic problem is continuous across $\Gamma$ and we denote the corresponding subspace of $W$ by $\widehat{W}$. We note that while the stiffness matrix $A$ and Schur complement $S$, which correspond to the product space $W$, are singular if we have any floating subdomains, those on $\widehat{W}$ are not. As in Sect. 2.2, we introduce interpolation operators $R_i^T : W_i \longrightarrow \widehat{W}$; the continuous global function $R_i^T w_i \in \widehat{W}$ shares the nodal values with $w_i$ on $\partial\Omega_{i,h} \cap \Gamma_h$ and vanishes on the rest of $\Gamma_h$.

An important role in the description and analysis of the Neumann-Neumann algorithms is played by a family of weighted counting functions $\delta_i \in W_i$, which are associated with the individual $\partial\Omega_i$; cf. [177, 184, 335, 412]. They are defined for $\gamma \in [1/2, \infty)$ by a sum of contributions from $\Omega_i$ and its relevant next neighbors,

$$\delta_i(x) = \frac{\sum_{j \in \mathcal{N}_x} \rho_j^\gamma}{\rho_i^\gamma}, \quad x \in \partial\Omega_{i,h} \cap \Gamma_h. \tag{6.1}$$

Here $\mathcal{N}_x$ is the set of indices $j$ of the subregions such that $x \in \partial\Omega_{j,h}$. We note that this formula can be extended to the case when $\rho_i$ varies over the boundary of the subdomain but that we will only consider the piecewise constant case in our theory. We recall that any node of $\Gamma_h$ belongs either to a face common to two subdomains, to an edge common to at least three subdomains, or is a vertex node common to several substructures; cf. Sect. 4.2. The pseudoinverses $\delta_i^\dagger$ are defined by

$$\delta_i^\dagger(x) = (\delta_i(x))^{-1}, \quad x \in \partial\Omega_{i,h} \cap \Gamma_h. \tag{6.2}$$

They provide a partition of unity:

$$\sum_i R_i^T \delta_i{}^\dagger(x) \equiv 1, \quad x \in \Gamma_h. \tag{6.3}$$

An operator, $E_D : W \longrightarrow \widehat{W}$, will be central to our discussion:

$$E_D u := \sum_{i=1}^N R_i^T I^h(\delta_i^\dagger u_i). \tag{6.4}$$

We note that, at any node $x \in \Gamma_h$, $E_D u(x)$ provides an average of the values $u_i(x)$, with $i \in \mathcal{N}_x$, weighted with the values $\delta_i^\dagger(x)$ and that it defines a

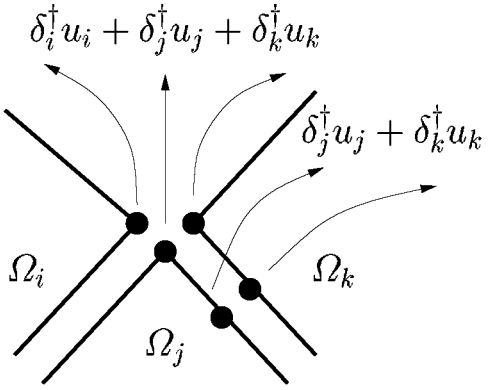


**Fig. 6.1.** Pointwise weighted averages of $u \in W$ provided by the operator $E_D$.

continuous function. Clearly, when restricted to $\widehat{W}$, $E_D$ reduces to the identity mapping because of (6.3); see Fig. 6.1 for an example of three subdomains.

We recall that our iterative substructuring methods determine the value of the finite element solution on the interface $\Gamma$. We also recall that we can regard the traces of an element $w \in W$ as Dirichlet data given on the sets $\partial\Omega_{i,h} \cap \Gamma_h$ and that any discrete harmonic function is completely defined by its values on those sets. Once these values are known, the values inside the substructures can be found by using the first equation of (4.10), i.e., by solving local Dirichlet problems on the substructures or a problem with a Dirichlet condition on $\partial\Omega_{i,h} \cap \Gamma_h$, a homogeneous Dirichlet condition on $\partial\Omega_{i,h} \cap \partial\Omega_{D,h}$, and a Neumann condition on $\partial\Omega_{i,h} \cap \partial\Omega_{N,h}$. Equivalently, we can then say that our iterative substructuring methods determine the piecewise, discrete harmonic part of our solution while the other interior parts, which are orthogonal to the discrete harmonic functions, can be determined immediately from the first equation of (4.10) in the preliminary Gaussian elimination step.

We will now define a Neumann-Neumann algorithm using the framework of Chap. 2. A minimal coarse space $W_0 \subset \widehat{W}$ for this algorithm is defined by the $\delta_i^\dagger$ of the interior, floating substructures, i.e., those with boundaries which do not intersect $\partial\Omega_D$, the part of the boundary of the original region $\Omega$ where we impose zero Dirichlet boundary conditions:

$$W_0 = \text{span}\,\{R_i^T \delta_i^\dagger\,,\ \partial\Omega_i \cap \partial\Omega_D = \emptyset\}. \tag{6.5}$$

We will denote by $R_0^T$ the matrix with columns representing the basis functions of $W_0$ and by $P_0 : W \longrightarrow W_0$, the projection onto the coarse space using an exact solver. We can also include functions related to the other substructures in $W_0$. They are constructed by extending the $\delta_i^\dagger$ by zero on $\partial\Omega_D \cap \partial\Omega_i$. The inclusion of these basis functions often improves the performance of the algorithm. We will make the following assumption:

**Assumption 6.1** *There is no subregion $\Omega_i$ with a boundary that intersects $\partial\Omega_{D,h}$ in only one or a few of its vertices.*

We note that this assumption is not necessary for the theory of the Neumann-Neumann method if extra basis functions for the coarse space are added as just suggested, cf. Dryja and Widlund [184], nor for the FETI-DP theory, cf. Klawonn, Widlund, and Dryja [292], but that we do not know if it can be dispensed with in the proof of the main results for the one-level FETI methods while maintaining good quality bounds on the rate of convergence of the algorithms.

The local spaces $W_i$ of our Schwarz algorithm are simply the $i$-th component of the product space $W$. We recall that an element of $R_i^T W_i \subset \widehat{W}$, associated with the substructure $\Omega_i$, can be regarded as a continuous, piecewise discrete harmonic function defined by its values at the nodes of $\partial\Omega_{i,h} \cap \Gamma_h$ and which vanishes at all points of $\Gamma_h \setminus \partial\Omega_{i,h}$. We also recall that whenever we write $\mathcal{H}_i(\varphi u_i)$, we will assume that we first form $I^h(\varphi u_i)$, i.e., map the product of the two functions $\varphi$ and $u_i$ into the finite element space by interpolation, and then extend the result as a discrete harmonic function.

The bilinear form $\tilde{s}_i(u,v)$ for the subspace $W_i$ is defined by

$$\tilde{s}_i(u,v) := a_{\Omega_i}(\mathcal{H}_i(\delta_i u), \mathcal{H}_i(\delta_i v)) = \rho_i \int_{\Omega_i} \nabla \mathcal{H}_i(\delta_i u) \cdot \nabla \mathcal{H}_i(\delta_i v) dx. \tag{6.6}$$

This scaled Neumann problem defines a projection-like operator $P_i = R_i^T \tilde{P}_i$ given by

$$\tilde{s}_i(\tilde{P}_i u, v_i) \;=\; s(u, R_i^T v_i), \quad v_i \in W_i; \tag{6.7}$$

cf. (2.6). Here $s(\cdot,\cdot)$ is the restriction of the bilinear form $a(\cdot,\cdot)$ to the subspace of piecewise discrete harmonic functions. For any floating subdomain, $\tilde{P}_i u$ is defined only for those $u \in \widehat{W}$ for which $s(u, R_i^T v_i) = 0$ for the $v_i$ for which $\mathcal{H}_i(\delta_i v_i)$ is constant on $\Omega_i$. This condition is satisfied if $s(u, R_i^T \delta_i^\dagger) = 0$; we note that $R_i^T \delta_i^\dagger$ is a basis function for $W_0$. A right hand side of (6.7) satisfying these compatibility conditions is said to be *balanced*. For any floating subdomain, we make the solution $\tilde{P}_i u$ of (6.7) unique by imposing an additional constraint. We could require

$$\int_{\Omega_i} \mathcal{H}_i(\delta_i \tilde{P}_i u) dx = 0, \tag{6.8}$$

which just means that we select the solution orthogonal, with respect to the $L^2$-scalar product, to the null space of the Neumann operator. But it is more convenient to choose the solution that belongs to $range(S^{(i)})$,

$$I^h(\delta_i \tilde{P}_i u) \in range(S^{(i)}), \tag{6.9}$$

i.e., the solution orthogonal, with respect to the $\ell^2$-scalar product, to the null space of the Neumann operator. In any case, $range(\tilde{P}_i)$ has codimension one with respect to the space $W_i$ if $\Omega_i$ is a floating subdomain, and it coincides

with $W_i$ otherwise. We note that the choice of these constraints will not affect the algorithm at all; they are just required for the analysis.

We now use the first hybrid Schwarz operator (2.11) introduced in Sect. 2.2 to define our algorithm. The relevant operator, for which we will establish spectral bounds, is

$$P_{hy1} = P_0 + (I - P_0)(\sum_{i=1}^{N} P_i)(I - P_0). \tag{6.10}$$

We recall that a general framework for hybrid Schwarz operators of this type was introduced in Sect. 2.5.2. In particular, we can employ Theorem 2.13. We note that the right factor $(I - P_0)$ assures us that the right hand side of (6.7) always will be balanced; therefore, the local Neumann problems are all solvable.

The bilinear form of the left hand side of (6.7) is defined by the $H^1(\Omega_i)$-seminorm and a diagonal scaling of the nodal values on $\partial\Omega_{i,h} \cap \Gamma_h$. This scaling has the great advantage that a stable decomposition is immediately available. Since $u = E_D u$, for $u \in \widehat{W}$, formula (6.4) provides a splitting of $u$:

$$u = \sum_{i=1}^{N} R_i^T v_i, \quad v_i := I^h(\delta_i^\dagger u_i) \in W_i.$$

In addition,

$$\sum_{i=1}^{N} \tilde{s}_i(v_i, v_i) = s(u, u), \quad u \in \widehat{W}. \tag{6.11}$$

A consequence of this formula is a lower bound of 1 for the additive Schwarz operator $\sum_{i=1}^{N} P_i$, and thus for $P_{hy1}$, since $P_0$ is an orthogonal projection; cf. Assumption 2.12 and Theorem 2.13.

### 6.2.2 Matrix Form of the Algorithm

We can rewrite our algorithm using matrix notations, denoting the Euclidean scalar product by $\langle \cdot, \cdot \rangle$. Let $D^{(i)}$ be the diagonal matrix with the elements $\delta_i^\dagger(x)$ corresponding to the point $x \in \partial\Omega_{i,h}$. Then,

$$\tilde{s}_i(u, v) = \langle S^{(i)} {D^{(i)}}^{-1} u, {D^{(i)}}^{-1} v \rangle.$$

We also have,

$$P_i = R_i^T \tilde{P}_i = R_i^T D^{(i)} S^{(i)\dagger} D^{(i)} R_i S,$$

where $S^{(i)\dagger}$ is a pseudoinverse of $S^{(i)}$. We recall that the application of $S^{(i)\dagger}$ to a vector can be carried out by solving a Neumann problem on $\Omega_i$ with an appropriate right-hand side; see Sect. 4.3. In case $\Omega_i$ is a floating subdomain, the solution is not unique and different choices for the particular solution

1. Initialize

$$u^0 = R_0^T S_0^{-1} R_0 \tilde{f}_\Gamma + w, \quad w \in range(I - P_0)$$
$$q^0 = \tilde{f}_\Gamma - S u^0$$

2. Iterate $k = 1, 2, \ldots$ until convergence

$$\text{Precondition: } z^{k-1} = \sum_{i=1}^{N} R_i^T D^{(i)} {S^{(i)}}^\dagger D^{(i)} R_i q^{k-1}$$
$$\text{Project: } y^{k-1} = (I - R_0^T S_0^{-1} R_0 S) z^{k-1}$$
$$\beta^k = \langle y^{k-1}, q^{k-1} \rangle / \langle y^{k-2}, q^{k-2} \rangle \quad [\beta^1 = 0]$$
$$p^k = y^{k-1} + \beta^k p^{k-1} \quad [p^1 = y^0]$$
$$\alpha^k = \langle y^{k-1}, q^{k-1} \rangle / \langle p^k, S p^k \rangle$$
$$u^k = u^{k-1} + \alpha^k p^k$$
$$q^k = q^{k-1} - \alpha^k S p^k$$

**Fig. 6.2.** Implementation of the balancing Neumann-Neumann algorithm as a projected preconditioned conjugate gradient method.

provide different pseudoinverses. As in the previous subsection, we will select the solution which belongs to $range(S^{(i)})$ to simplify our analysis. Thus,

$$range(\tilde{P}_i) = range(D^{(i)} S^{(i)}); \tag{6.12}$$

cf. (6.9). In the following, we assume that this choice has been made, but note that similar bounds hold for the first one, and generally for any choice that ensures that a Poincaré inequality holds for vectors in $range({D^{(i)}}^{-1} \tilde{P}_i)$; cf. (6.8). Since for floating subdomains the kernel of $S^{(i)}$ consists of constant vectors, we find that

$$l_i(u_i) := h_i^2 \, (\mathbf{1}_i^T u_i) = 0 \quad \Longleftrightarrow \quad u_i \in range(S^{(i)}), \tag{6.13}$$

where $\mathbf{1}_i \in W_i$ is a vector of ones.

Since an exact solver is employed for the coarse space, we have

$$P_0 = R_0^T S_0^{-1} R_0 S, \quad S_0 := R_0 S R_0^T.$$

Implementation issues have already been discussed in Sect. 2.5.2. We rewrite the projected conjugate gradient algorithm of Sect. 2.5.2, for our particular balancing Neumann-Neumann method for the solution of

$$Su = \tilde{f}_\Gamma, \quad u \in \widehat{W},$$

in Fig. 6.2. We note that because of Lemma 2.11, we are able to eliminate the first projection step. Since we only work with vectors in the continuous space $\widehat{W}$, the expression of the Schur complement $S$ is given by (4.10). The application of $S$ to a vector involves the solution of a Dirichlet problems on each substructure. Each step of the iteration involves one application of $P_0$, and thus one multiplication with $S$ and the solution of a coarse problem involving $S_0$. In addition, we solve a local Neumann problem on each substructure, when applying $S^{(i)\dagger}$, and also a local Dirichlet problem for each subdomain which is required for the second application of $S$ in the calculation of the new residual. Therefore, there is a total of one Neumann and two Dirichlet problems on each substructure and one coarse problem in each step. We note that all the Neumann problems can be solved in parallel and that so can the two groups of Dirichlet problems.

### 6.2.3 Condition Number Bounds

We now provide a condition number estimate; we will use the abstract framework of Sect. 2.5.2 only in part. We first note that $\sum_{i=1}^{N} P_i u$ is obtained by solving a Neumann problem for each subregion $\Omega_i$. If standard unscaled local Neumann solvers are used, we first obtain the vectors $w_i := S^{(i)\dagger} D^{(i)} R_i S u$. These local solutions generally do not match across the interface but they define elements in the $W_i$. They are then averaged using the $\delta_i^\dagger$, i.e., the elements of $D^{(i)}$, as weights and the resulting interface values are then extended as piecewise harmonic functions by solving a Dirichlet problem on each subdomain. Thus, we start with $w \in W$ defined by

$$w_i = {D^{(i)}}^{-1} R_i P_i u = {D^{(i)}}^{-1} \tilde{P}_i u = S^{(i)\dagger} D^{(i)} R_i S u, \quad u \in \widehat{W}. \tag{6.14}$$

We note that, because of our choice for ${S^{(i)}}^\dagger$ (cf. (6.12)), $w \in range(S)$. Using formula (6.4), we find that

$$E_D w = \sum_{i=1}^{N} P_i u. \tag{6.15}$$

We now proceed to find upper and lower bounds on the spectrum of $P_{hyl}$. We have already observed in Chap. 2 that we can view such a hybrid method as an additive Schwarz method on the subspace $range(I - P_0)$; cf. Theorem 2.13. A lower bound of 1 now follows directly from formula (6.11); see Assumption 2.12 and Theorem 2.13.

In order to find an upper bound, we first need certain Poincaré type inequalities.

**Lemma 6.2** *Let $w \in range(S)$. Then, if $\Omega_i$ is a floating subdomain,*

$$\|w_i\|^2_{L^2(\partial\Omega_i)} \leq C H_i \, |w_i|^2_{H^{1/2}(\partial\Omega_i)},$$

*with a constant that is independent of $w_i$, $H_i$, and $h_i$.*

*Proof.* If $\Omega_i$ is a floating subdomain, then (6.13) holds. We note that, if $v$ is constant on $\partial\Omega_i$, then

$$l_i(v) = 0 \quad \Leftrightarrow \quad v = 0. \tag{6.16}$$

Using Lemma B.31 and the scaled norm in (4.5), we find

$$|l_i(w_i)|^2 \leq (h_i^2 \mathbf{1}_i^T \mathbf{1}_i)(h_i^2 w_i^T w_i) \leq C|\partial\Omega_i| \|w_i\|^2_{L^2(\partial\Omega_i)} \leq C H_i^3 \|w_i\|^2_{H^{1/2}(\partial\Omega_i)},$$

with a constant that is independent of $w_i$, $H_i$, and $h_i$. The linear functional $l_i(\cdot)$ can therefore be extended to a functional, still denoted by $l_i(\cdot)$, to all of $H^{1/2}(\partial\Omega_i)$. Property (6.16) still holds. The application of Lemma A.17 and a scaling argument conclude the proof. □

The following lemma is central in the theory for the Neumann-Neumann as well as for the FETI methods. Our estimates are expressed in terms of the norm

$$|w|_S^2 = \langle w, Sw \rangle = \sum_{i=1}^{N} |w_i|^2_{S^{(i)}}, \quad w \in W.$$

We recall that $|w_i|^2_{S^{(i)}}$, $\rho_i |\mathcal{H}_i(w_i)|^2_{H^1(\Omega_i)}$, and $\rho_i |w_i|^2_{H^{1/2}(\partial\Omega_i)}$ provide equivalent seminorms for $w_i \in W_i$, according to Lemmas 4.9 and 4.10.

**Lemma 6.3** *Let $E_D$ be the operator defined in formula (6.4). Then,*

$$|E_D w|_S^2 \leq C(1 + \log(H/h))^2 |w|_S^2, \quad w \in \mathit{range}(S). \tag{6.17}$$

*Proof.* We rewrite the formula for $v := E_D w$ for an arbitrary element $w \in W$ and find that for $i = 1, \ldots, N$,

$$v_i(x) \;:=\; (E_D w(x))_i \;=\; \sum_{j \in \mathcal{N}_x} I^h(\delta_j^\dagger w_j(x)), \quad x \in \partial\Omega_{i,h} \cap \Gamma_h. \tag{6.18}$$

Here $\mathcal{N}_x$ is again the set of indices of the subregions that have $x$ on their boundaries. We note that the coefficients in this expression are constant on each face and edge of $\partial\Omega_i$ and that their values will often differ between different faces, edges, and vertices. It is therefore natural to write $v_i$ as a sum of terms that vanish at all the interface nodes outside individual faces, edges, and vertices, respectively. The norms of the individual terms of this sum are then estimated. We will use the characteristic finite element functions $\theta_{\mathcal{F}}$ of a face $\mathcal{F}$, $\theta_{\mathcal{E}}$ of an edge $\mathcal{E}$, and $\theta_{\mathcal{V}}$ of a vertex $\mathcal{V}$ for this purpose; they were already introduced in Sect. 4.6 (cf. Definitions 4.18, 4.22, and 4.27). They are discrete harmonic functions that vanish at every node of $\partial\Omega_{i,h}$ except at those of $\mathcal{F}_h$, those of $\mathcal{E}_h$, and at $\mathcal{V}$, respectively. As shown in Sect. 4.6, these functions define a partition of unity and we have also already developed a number of technical results pertaining to them in the same section. One of the few things that is really new in this proof is the use of the following simple inequality, which can be proven by an elementary argument:

$$\rho_i \delta_j^{\dagger 2} \leq \min(\rho_i, \rho_j), \tag{6.19}$$

for $\gamma \in [1/2, \infty)$. This inequality and the choice of scaling of the Neumann problems will allow us to develop bounds for the rate of convergence of our algorithms which do not depend on the values of the $\rho_i$.

We only need to estimate the contribution from one subdomain $\Omega_i$ to the norm of $E_D w$. In this proof, we tacitly assume that the number of subdomains to which an edge or a vertex belong is uniformly bounded; this is certainly true if Assumption 4.3 holds.

We have, by using the partition of unity,

$$v_i = \sum_{\mathcal{F} \subset \partial\Omega_i} I^h(\theta_{\mathcal{F}} v_i) + \sum_{\mathcal{E} \subset \partial\Omega_i} I^h(\theta_{\mathcal{E}} v_i) + \sum_{\mathcal{V} \in \partial\Omega_i} v_i(\mathcal{V})\, \theta_{\mathcal{V}}.$$

We first consider the face terms.

**Face Terms.** We note that the sum in (6.18) gives rise to two contribution for a face $\mathcal{F}$ shared by, e.g., $\Omega_i$ and $\Omega_j$, namely,

$$I^h(\theta_{\mathcal{F}} v_i) = I^h(\theta_{\mathcal{F}}(\delta_i^{\dagger} w_i + \delta_j^{\dagger} w_j)).$$

If $\delta_i^{\dagger}(\mathcal{F})$ and $\delta_j^{\dagger}(\mathcal{F})$ are the values of $\delta_i^{\dagger}(x)$ and $\delta_j^{\dagger}(x)$, respectively, at $x \in \mathcal{F}_h$, we can write

$$\begin{aligned}
\rho_i |I^h(\theta_{\mathcal{F}} v_i)|^2_{H^{1/2}(\partial\Omega_i)} &\leq C\rho_i \|I^h(\theta_{\mathcal{F}} v_i)\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&\leq 2C\rho_i \delta_i^{\dagger}(\mathcal{F})^2 \|I^h(\theta_{\mathcal{F}} w_i)\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&+ 2C\rho_i \delta_j^{\dagger}(\mathcal{F})^2 \|I^h(\theta_{\mathcal{F}} w_j)\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&\leq 2C\rho_i \|I^h(\theta_{\mathcal{F}} w_i)\|^2_{H_{00}^{1/2}(\mathcal{F})} + 2C\rho_j \|I^h(\theta_{\mathcal{F}} w_j)\|^2_{H_{00}^{1/2}(\mathcal{F})},
\end{aligned}$$

where we have used (6.19) for the last inequality. Lemma 4.26 and the Poincaré inequality in Lemma 6.2 now provide the bound

$$|I^h(\theta_{\mathcal{F}} v_i)|^2_{S^{(i)}} \leq C(1 + \log(H/h))^2 (|w_i|^2_{S^{(i)}} + |w_j|^2_{S^{(j)}}), \tag{6.20}$$

if both $\Omega_i$ and $\Omega_j$ are floating subdomains. For a subdomain with a boundary that intersects $\partial\Omega_D$, we instead use the inequality

$$\|I^h(\theta_{\mathcal{F}} w_i)\|^2_{H_{00}^{1/2}(\mathcal{F})} \leq 2\|I^h(\theta_{\mathcal{F}}(w_i - \overline{w_i}_{\mathcal{F}})\|^2_{H_{00}^{1/2}(\mathcal{F})} + 2\|\overline{w_i}_{\mathcal{F}} \theta_{\mathcal{F}}\|^2_{H_{00}^{1/2}(\mathcal{F})},$$

with $\overline{w_i}_{\mathcal{F}}$ the average of $w_i$ on $\mathcal{F}$. The first term can be estimated by using Lemma 4.26. The second term can be estimated by

$$C(1 + \log(H/h))\|w_i\|^2_{H^{1/2}(\partial\Omega_i)},$$

by using Lemmas 4.30 and 4.26. We complete the proof by using either a standard Friedrichs inequality or, if the intersection of $\partial\Omega_i$ and $\partial\Omega_D$ is only an edge, Lemma 4.21 and obtain a bound of the same quality as for the floating subdomains. We recall, that under Assumption 6.1, we do not need to consider any other cases.

**Edge Terms.** Similar arguments work for the edge contributions. Let $x$ be any nodal point in $\mathcal{E}_h$. If $\delta_k^\dagger(\mathcal{E})$ is the value of $\delta_k^\dagger(x)$, for $k \in \mathcal{N}_x$, Lemmas 4.19 and 4.9, (6.19), and the triangle inequality yield

$$|I^h(\theta_{\mathcal{E}} v_i)|^2_{S^{(i)}} \leq C\rho_i \| \sum_{k\in\mathcal{N}_x} \delta_k^\dagger(\mathcal{E}) w_k\|^2_{L^2(\mathcal{E})} \leq C \sum_{k\in\mathcal{N}_x} \rho_k \|w_k\|^2_{L^2(\mathcal{E})}.$$

Lemmas 4.17 and 6.2 give, in case all the subdomains are floating,

$$|I^h(\theta_{\mathcal{E}} v_i)|^2_{S^{(i)}} \leq C(1+\log(H/h)) \sum_{k\in\mathcal{N}_x} |w_k|^2_{S^{(k)}}. \tag{6.21}$$

For the other subdomains, we proceed similarly to the estimate of the face terms and write $w_k = (w_k - \overline{w_k}_{\mathcal{F}}) + \overline{w_k}_{\mathcal{F}}$, with $\mathcal{F} \subset \partial\Omega_k$ a face that contains $\mathcal{E}$ on its boundary. We then estimate the two terms separately. For the first, we use Lemma 4.17 and a Poincaré inequality for the face $\mathcal{F}$; cf. Lemma A.17. For the second, we first employ Lemma 4.30 and then either a standard Friedrichs inequality or, if the intersection of $\partial\Omega_i$ and $\partial\Omega_D$ is only an edge, Lemma 4.21. We finally arrive at (6.21).

**Vertex Terms.** We finally consider the vertex components. Lemma 4.28 and (6.19) gives

$$|v_i(\mathcal{V})\,\theta_{\mathcal{V}}|^2_{S^{(i)}} \leq C \sum_{k\in\mathcal{N}_{\mathcal{V}}} \rho_k \|w_k\|^2_{H^{1/2}(\partial\Omega_k)}.$$

The use of a Poincaré inequality gives

$$|v_i(\mathcal{V}^{ij})\,\theta_{\mathcal{V}}|^2_{S^{(i)}} \leq C \sum_{k\in\mathcal{N}_{\mathcal{V}}} |w_k|^2_{S^{(k)}} \tag{6.22}$$

in case all the subdomains are floating. Subdomains next to $\partial\Omega_D$ can be handled with the same techniques as before; if a subdomain only has an edge in common with $\partial\Omega_D$, we have an extra factor $(1+\log(H/h))$ in the estimate.

The proof can now be concluded by using estimates (6.20), (6.21), and (6.22), and summing over the faces, edges, and vertices of $\Omega_i$, and then over the substructures. □

To prove our upper bound for the hybrid Schwarz operator, we could work with Assumption 2.4 and Lemma 2.6 (or, equivalently, Theorem 2.13). However, we will find it more convenient to estimate the norm of the operator $\sum P_i$ directly.

With $w_i$ defined by (6.14), we have already established formula (6.15). We recall that the global function $w \in W$ belongs to $range(S)$. With the $w_i$ as in (6.14), we also find by using the definition of the $\tilde{P}_i$ that

$$\begin{aligned} |w|_S^2 &= \sum_i \langle S^{(i)} w_i, w_i \rangle = \sum_i \langle S^{(i)} {D^{(i)}}^{-1} \tilde{P}_i u, {D^{(i)}}^{-1} \tilde{P}_i u \rangle \\ &= \sum_i \langle Su, R_i^T \tilde{P}_i u \rangle = \langle \sum_i P_i u, u \rangle_S. \end{aligned} \tag{6.23}$$

Therefore, we find, by using (6.15), (6.23), and Lemma 6.3, that

$$\begin{aligned} |\textstyle\sum_i P_i u|_S^2 &= |E_D w|_S^2 \leq C(1 + \log(H/h))^2 |w|_S^2 \\ &= C(1 + \log(H/h))^2 \, \langle \textstyle\sum_i P_i u, u \rangle_S. \end{aligned} \tag{6.24}$$

Therefore

$$\begin{aligned} |\sum_{i=1}^N P_i|_S &\leq C(1 + \log(H/h))^2, \\ \langle \sum_{i=1}^N P_i u, u \rangle_S &\leq C(1 + \log(H/h))^2 |u|_S^2, \end{aligned}$$

and the proof of the following theorem is completed after noting that $P_0$ is an orthogonal projection.

**Theorem 6.4** *The hybrid Schwarz method defined by the operator (6.10) and the spaces and bilinear forms of this section satisfies*

$$s(u,u) \leq s(P_{hy1} u, u) \;\leq\; C(1 + \log(H/h))^2 s(u,u),$$

*where $C$ is independent not only of the mesh size and the number of substructures, but also of the values $\rho_i$ of the coefficient of (4.3).*

*Remark 6.5.* We note that the smallest eigenvalue of $P_{hy1}$ is exactly one, since we have equality in formula (6.11); cf. (2.13) in Lemma 2.5.

## 6.3 One-Level FETI Methods

The finite element tearing and interconnecting (FETI) methods were first introduced by Farhat and Roux [200]; for a detailed introduction, see [201] or [435]. An important advance, which made the rate of convergence of the iteration less sensitive to the number of unknowns of the local problems, was made by Farhat, Mandel, and Roux a few years later in [199] where the *Dirichlet preconditioner* was introduced. The first, quite pioneering theoretical work was carried out by Mandel and Tezaur [340, 435]; see also Brenner [87, 88]. Here we will essentially follow a paper by Klawonn and Widlund [289] in which the family of algorithms and the theory was extended in several different

respects. In this section, we will focus on the older one-level FETI methods; we will introduce and analyze several dual-primal FETI methods in Sect. 6.4. We note that the FETI algorithms have also been extended to some diffusion-reaction problems, for which the local Neumann problems are nonsingular, by Farhat, Chen, and Mandel [191], to some finite element approximations of Maxwell's equations in two dimensions and with discontinuous coefficients by Toselli and Klawonn [443] and Rapetti and Toselli [393], and to some scalar convection-diffusion problems by Toselli [440]. Here we will primarily consider problems in three dimensions, focusing, in particular, on the elliptic problem given in (4.3). More references have been given in Sect. 6.1; see also Sect. 6.4 for FETI-DP algorithms.

When introducing the dual-primal FETI methods, in Sect. 6.4, we will also use additional, intermediate subspaces $\widetilde{W}$ of $W$ for which only a relatively small number of continuity constraints are enforced across the interface in each iteration. One of the benefits of working in $\widetilde{W}$, rather than in $W$, will be that two related Schur complements will be strictly positive definite; see further Sect. 6.4.

### 6.3.1 A Review of the One-Level FETI Methods

In this subsection, we review the FETI method of Farhat and Roux [200, 201], in particular, the variant with a Dirichlet preconditioner introduced in Farhat, Mandel, and Roux [199]. We will also introduce a general family of projections that was first introduced for heterogeneous problems in [201]. Such methods have been tested in large scale numerical experiments; see, e.g., [51]. For a more detailed description and extensions beyond scalar elliptic problems; see [191, 193, 342, 398, 435] and Chap. 8. We also point out that there are other variants of the FETI methods, e.g., those due to Park, Justino, and Felippa [372]. The relation of one of them to the FETI methods developed by Farhat and Roux is discussed in [401], and a convergence analysis of this method can be found in Tezaur's dissertation [435]. Still other variants are due to Brenner [87, 88].

In what follows, we will work almost exclusively with functions in the trace spaces $W_i$ and, whenever convenient, consider such an element as representing a discrete harmonic function in $\Omega_i$; cf. Sect. 4.4. Thus, for $w \in W$, $\mathcal{H}(w)$ denotes the piecewise discrete harmonic extension into all of the $\Omega_i$; we will understand $\mathcal{H}(w)$ as an element in a product space $W$ with the components $\mathcal{H}_i(w_i)$.

The values of the right-hand sides also change when the interior variables are eliminated. We denote the resulting vectors, representing the modified load originating in $\Omega_i$, by $f_i$, and the local vectors of interface nodal values by $u_i$.

We can now reformulate the finite element problem, reduced to the interface $\Gamma$, as a minimization problem with constraints given by the requirement of continuity across $\Gamma$: find $u \in W$ such that

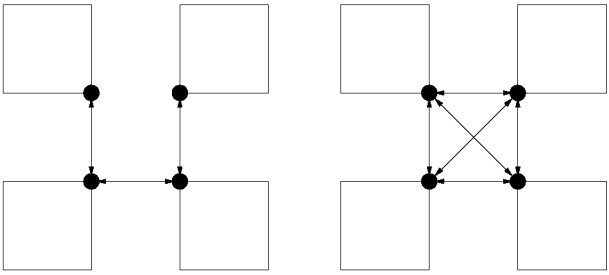

**Fig. 6.3.** *Left*: U-shaped arrangement of Lagrange multipliers for an edge in a nonredundant case. *Right*: Lagrange multipliers for an edge in the fully redundant case.

$$\left.\begin{array}{c} J(u) := \frac{1}{2}\langle Su, u\rangle - \langle f, u\rangle \to \min \\ Bu = 0 \end{array}\right\}, \tag{6.25}$$

where

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{pmatrix}, \text{ and } S = \begin{pmatrix} S^{(1)} & O & \cdots & O \\ O & S^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & S^{(N)} \end{pmatrix}.$$

The matrix

$$B = (B^{(1)}, B^{(2)}, \ldots, B^{(N)}) \tag{6.26}$$

is constructed from $\{0, 1, -1\}$ such that the values of the solution $u$ associated with more than one subdomain coincide when $Bu = 0$. We note that the choice of $B$ is far from unique. While there is little choice of how to write the constraint for a nodal point which belongs to a face, there are many options for a point on an edge or for a vertex. For the face nodes, we only have to choose a sign, but we note that for an edge node, e.g., one common to four subdomains, a minimum set of three constraints can be chosen in many ways to assure continuity at the point in question. In fact, for such an edge node, up to six constraints can be introduced; see Fig. 6.3. The local Schur complements $S^{(i)}$ are positive semidefinite, and they are singular for any subregion with a boundary that does not intersect $\partial\Omega_D$. The problem (6.25) is uniquely solvable since $kernel(S) \cap kernel(B) = \{0\}$, which means that $S$ is invertible on $kernel(B)$.

By introducing a vector of Lagrange multipliers $\lambda$ to enforce the constraints $Bu = 0$, we obtain a saddle point formulation of (6.25): find $(u, \lambda) \in W \times U$ such that

$$\left.\begin{array}{rl} Su + B^T\lambda &= f \\ Bu &= 0 \end{array}\right\} \tag{6.27}$$

We note that the solution $\lambda$ of (6.27) is unique only up to an additive element of $kernel(B^T)$. The space of Lagrange multipliers, $U$, is therefore chosen as $range(B)$. It can then be regarded as the space of jumps of functions in $W$.

We will also use a full-column-rank matrix built from all of the null space elements of $S$; these elements are associated with individual subdomains

$$R = \begin{pmatrix} R^{(1)} & O & \cdots & O \\ O & R^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & R^{(N)} \end{pmatrix}. \tag{6.28}$$

Thus, $range(R) = kernel(S)$. In fact, the subdomains that intersect $\partial\Omega_D$ do not contribute to $kernel(S)$, and therefore those columns of $R$ are void.

*Remark 6.6.* The case of linear elasticity is considerably more complicated. For the interior subregions, there are full six-dimensional null spaces of rigid body motions. There can also be contributions to $R$ from subdomains with boundaries intersecting $\partial\Omega_D$ for which there are not enough essential boundary conditions to fully control the entire space of rigid body motions; see further Chap. 8 and appendix A.6.2.

A solution $u$ of the first equation in (6.27) exists if and only if $f - B^T\lambda \in range(S)$; this constraint will lead to the introduction of a projection $P$. We obtain

$$u = S^\dagger(f - B^T\lambda) - R\alpha \quad \text{if } f - B^T\lambda \perp kernel(S)\,,$$

where $S^\dagger$ is a pseudoinverse of $S$. We will see that $\alpha$ can be determined easily once $\lambda$ has been found. The pseudoinverse is generally not uniquely determined, but it can easily be shown that our algorithms are invariant to the specific choice. Thus, without loss of generality, we can assume in our analysis that $S^\dagger$ is symmetric. We can, e.g., choose the Moore-Penrose generalized inverse; see Golub and Van Loan [231]. We note that another, computationally less expensive alternative has been implemented in Farhat and Roux [200].

Substituting the expression for $u$ into the second equation of (6.27) gives

$$BS^\dagger B^T\lambda = BS^\dagger f - BR\alpha\,, \tag{6.29}$$

and we obtain the system

$$\begin{cases} F\lambda + G\alpha = d \\ G^T\lambda = e \end{cases} \tag{6.30}$$

with $F := BS^\dagger B^T$, $G := BR$, $d := BS^\dagger f$, and $e := R^T f$.

We now introduce a symmetric, positive definite matrix $Q$ and an inner product $\langle\lambda,\mu\rangle_Q := \langle\lambda, Q\mu\rangle$ on $U = range(B)$. As before, $\langle\cdot,\cdot\rangle$ stands for the $\ell^2-$inner product. Let

$$P^T := I - G(G^TQG)^{-1}G^TQ$$

be the projection from $U$ onto the subspace of Lagrange multipliers that are $Q-$orthogonal to $range(G)$; see the definition of $V'$ below. We find that

$$\begin{cases} P^T F\lambda = P^T d \\ G^T\lambda = e \end{cases} \tag{6.31}$$

We note that

$$P = I - QG(G^TQG)^{-1}G^T$$

is a projection from $U$ onto $kernel(G^T)$; this projection is orthogonal in the $Q^{-1}$ inner product, i.e., the inner product defined by $\langle\lambda, Q^{-1}\mu\rangle$. By multiplying (6.29) by $(G^TQG)^{-1}G^TQ$, we find that $\alpha := (G^TQG)^{-1}G^TQ(d - F\lambda)$, which then fully determines the primal variables in terms of $\lambda$. We note that the operators $P$ and $P^T$ represent the only global part of the preconditioner.

There are different successful choices for $Q$. In the case of homogeneous coefficients, it is sufficient to use $Q = I$, while for problems with jumps in the coefficients, we have to make more elaborate choices to make our bounds independent of the variations of the coefficients in Equation (4.3). A full analysis has been developed in [289] for a family of diagonal scaling matrices $Q$ as well as for the case when $Q$ is chosen as the FETI Dirichlet preconditioner; see Sect. 6.3.2 and 6.3.3 and [51, 201]. Here we will only consider the latter choice in detail. The version with a diagonal scaling has been tested experimentally by Rheinbach [396] and it performs very satisfactorily and it also decreases the computational cost. We also note that we can view the introduction of a nontrivial positive definite, symmetric $Q$ in terms of a scaling of the matrix $B$ from the left by the operator $Q^{1/2}$.

We now introduce the spaces

$$\begin{aligned} V &:= \{\lambda \in U \,:\, \langle\lambda, Bz\rangle = 0, \quad z \in kernel(S)\} \\ &= kernel(G^T) = range(P), \end{aligned} \tag{6.32}$$

and

$$V' := \{\mu \in U : \langle\mu, Bz\rangle_Q = 0, \quad z \in kernel(S)\} = range(P^T). \tag{6.33}$$

It can easily be shown that $V'$ is isomorphic to the dual space of $V$. Following Farhat, Chen, and Mandel [191], we call $V$ the space of admissible increments. The one-level FETI method is a preconditioned conjugate gradient method, in the space $V$, applied to

$$P^T F\lambda = P^T d, \quad \lambda \in \lambda_0 + V, \tag{6.34}$$

with an initial approximation $\lambda_0$ chosen such that $G^T\lambda_0 = e$.

The most basic FETI Dirichlet preconditioner, as introduced in Farhat, Mandel, and Roux [199] for $Q = I$, is of the form

$$M^{-1} := BSB^T = \sum_{i=1}^{N} B^{(i)} S^{(i)} {B^{(i)}}^T. \tag{6.35}$$

We note that $S$ can be replaced by a principal minor of $A$, but that such a choice leads to much poorer bounds and it can also lead to an increased number of iterations. To apply $M^{-1}$ to a vector, $N$ independent Dirichlet problems have to be solved, one on each subregion; it is therefore called the Dirichlet preconditioner.

We note that, strictly speaking, the matrix $M^{-1}$ does not have an inverse, but we will show in Lemma 6.11 that $P\widehat{M}^{-1}$ is a one-to-one mapping of $V'$ to $V$; here $\widehat{M}^{-1}$ is our modified preconditioner defined in the next subsection; cf. (6.37). In fact, to keep the search directions of this preconditioned conjugate gradient method in the space $V$, the application of the preconditioner $M^{-1}$ (or $\widehat{M}^{-1}$) is followed by an application of the projection $P$. Hence, the Dirichlet variant of the FETI method is the preconditioned conjugate gradient algorithm applied to the equation

$$PM^{-1}P^TF\lambda = PM^{-1}P^Td, \quad \lambda \in \lambda_0 + V. \tag{6.36}$$

We note that for $\lambda \in V$, $PM^{-1}P^TF\lambda = PM^{-1}P^TP^TFP\lambda$, and we can therefore view the operator on the left-hand side of (6.36) as the product of two symmetric matrices.

We note that several different possibilities of improving the FETI preconditioner $M^{-1}$ have been explored. Some interesting variants are discussed in Rixen and Farhat [399] in a framework of mechanically consistent preconditioners and redundant Lagrange multipliers; see the discussion and analysis in Sect. 6.3.3. A family of improved FETI preconditioners, with nonredundant Lagrange multipliers, will be introduced and analyzed in Subsect. 6.3.2.

We end this subsection with some remarks on implementation. Preconditioned systems of the form (6.36) were introduced in Sect. 2.5.2 and a possible implementation of the conjugate gradient method was also suggested. We have rewritten the projected conjugate gradient algorithm of Sect. 2.5.2 in Fig. 6.4 for a general one-level FETI method. We note that $\lambda_0 = QG(G^TQG)^{-1}e$ satisfies the second condition of (6.31). Indeed, this is a different type of projected algorithm than those that were previously considered. While for the hybrid methods in Sect. 2.5.2 and the Balancing Neumann-Neumann method in Sect. 6.2.2, the coarse and local components of the solution are determined by projecting the original equation into the two subspaces $range(P_0)$ and $range(I - P_0)$ (see the preconditioned operator in equation (6.10)), here the two components are determined by the two independent equations in the FETI system (6.31), which comes from the mixed problem (6.30). We remark that the projection $I - P_0$ employed for the hybrid methods in Sect. 2.5.2 is constructed using the matrix $A$ of the unpreconditioned problem, while the projection $P$ employed here is independent of the original matrix $F$. For this reason, it is a simple matter to check that the first projection step involving $P^T$ is necessary here. Finally, we note that for the calculation of the $\beta_k$ and $\alpha_k$ we can equivalently employ the previous vectors $q^k$ or $r^k$, since $P^T$ is a projection.

1. Initialize

$$\lambda^0 = QG(G^TQG)^{-1}e + \mu, \quad \mu \in range(P)$$
$$r^0 = d - F\lambda^0$$

2. Iterate $k = 1, 2, \ldots$ until convergence

$$\begin{aligned}
\text{Project: } q^{k-1} &= P^T r^{k-1} \\
\text{Precondition: } z^{k-1} &= \widehat{M}^{-1} q^{k-1} \\
\text{Project: } y^{k-1} &= P z^{k-1} \\
\beta^k &= \langle y^{k-1}, q^{k-1} \rangle / \langle y^{k-2}, q^{k-2} \rangle \quad [\beta^1 = 0] \\
p^k &= y^{k-1} + \beta^k p^{k-1} \quad [p^1 = y^0] \\
\alpha^k &= \langle y^{k-1}, q^{k-1} \rangle / \langle p^k, F p^k \rangle \\
\lambda^k &= \lambda^{k-1} + \alpha^k p^k \\
r^k &= r^{k-1} - \alpha^k F p^k
\end{aligned}$$

**Fig. 6.4.** Implementation of the one-level FETI method as a projected preconditioned conjugate gradient method.

As already mentioned, one-level FETI algorithms are determined by the choice of $Q$ and $\widehat{M}^{-1}$. For the choice $Q = \widehat{M}^{-1}$, each step of the corresponding preconditioned conjugate gradient method involves one application of $P^T$ and one of $P$, the solution of local Dirichlet problems on the substructures, necessary for the application of $\widehat{M}^{-1}$, and the solution of local Neumann problems required for the application of $F$ in the calculation of the new residual. Since the applications of $P^T$ and $P$ involve two additional applications of $Q = \widehat{M}^{-1}$ and the solution of two coarse problems, there is a total of one Neumann and three Dirichlet problems on each substructure and two coarse problems in each iteration step. We note that this cost is higher than for a step of the Neumann-Neumann method; see Sect. 6.2. In case a diagonal $Q$ is used, we only need one application of the preconditioner and thus only one Dirichlet problem on each subdomain. We note that the FETI-DP algorithms also allow us to reduce the computational cost of each step.

*Remark 6.7.* It is clear that the algorithms introduced in this section rely on the existence of non-trivial kernels of the local Neumann problems, since they provide a coarse solve which makes the method scalable; see the inverse of $G^TQG$ in the definition of the projection $P$. The characterization of these local kernels is required and may not be a trivial task for some more complicated problems. A modification of one-level FETI algorithms is possible when the local Neumann problems are uniquely solvable, as for the bilinear form

$$a(u,v) + \int_\Omega uv\,dx$$

(see [191]) and for problems arising in electromagnetics (see [443]). Such an algorithm was proposed in [191] and an abstract framework was later given in [443]. This modification leads to a different algorithm and requires a different implementation.

### 6.3.2 The Case of Nonredundant Lagrange Multipliers

In this subsection, we present a one-level FETI preconditioner and prove a spectral bound which is independent of the values of the coefficients $\rho_i$ of (4.3) if the operator $Q$, which enters the definition of $P$, is chosen carefully. In our proofs, we will use a number of arguments developed in [340], but they also differ in several respects. We assume, for the rest of this subsection, that $B$ has full row rank, i.e., the constraints are linearly independent and there are no redundant Lagrange multipliers; cf. Fig. 6.3, left.

The preconditioner is now defined, for any diagonal matrix $D$ with positive elements, as

$$\begin{aligned}\widehat{M}^{-1} &:= (BD^{-1}B^T)^{-1}BD^{-1}SD^{-1}B^T(BD^{-1}B^T)^{-1} \\ &= (BD^{-1}B^T)^{-1}\sum_{i=1}^{N} B^{(i)}{D^{(i)}}^{-1}S^{(i)}{D^{(i)}}^{-1}{B^{(i)}}^T(BD^{-1}B^T)^{-1}\end{aligned} \tag{6.37}$$

It is easy to see that $BD^{-1}B^T$ is a block-diagonal matrix. Its inverse can be computed at essentially no extra cost since the block sizes are $n_x$, where $n_x$ is the number of Lagrange multipliers employed to enforce continuity at the point $x$. To obtain a method which converges at a rate that is independent of the coefficient jumps, we now choose a special family of matrices $D$; a careful choice of the operator $Q$ will also be required. As in Sect. 6.2, a crucial role is played by the functions $\delta_i^\dagger$, which in turn depend of the coefficients $\rho_i$ of the elliptic problem; see (6.2). The diagonal matrices $D^{(i)}$ are the same as in Sect. 6.2. Thus, $D^{(i)}$ has the diagonal entry $\delta_i^\dagger(x)$ corresponding to the point $x \in \partial\Omega_{i,h} \cap \Gamma_h$; see (6.3). Finally, we set

$$D := \begin{pmatrix} D^{(1)} & O & \cdots & O \\ O & D^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & D^{(N)} \end{pmatrix}.$$

We note that this matrix operates on elements in the product space $W$. It can be regarded as a scaling from the right, by $D^{-1/2}$, of the matrix $B$.

An important role will be played by the operator

$$P_D := D^{-1}B^T(BD^{-1}B^T)^{-1}B\,, \tag{6.38}$$

which maps $W$ into itself. This is a projection that is orthogonal in the scaled $\ell_2$-inner product $\langle x, Dy\rangle$, where $x, y \in W$. We note that this operator is invariant if we replace $B$ by $Q^{1/2}B$. In addition, it preserves the jumps, in the sense that

$$BP_{_D} = B. \tag{6.39}$$

The next two lemmas exploit this property.

**Lemma 6.8** *For any $\mu \in U$, there exists a $\tilde{w} \in range(P_{_D})$ such that $\mu = B\tilde{w}$.*

*Proof.* For any $\mu \in U = range(B)$, there exists a $w \in W$ such that $\mu = Bw$. We can then select $\tilde{w} = P_{_D} w \in W$, since, by a simple computation, $B\tilde{w} = Bw = \mu$. □

The next lemma follows directly from (6.39).

**Lemma 6.9** *The projection operator $P_{_D}$ satisfies*

$$w - P_{_D} w \in \widehat{W},$$

*i.e., this function is continuous across $\Gamma$ for all $w \in W$.*

We now show that $P_{_D} u$ can be expressed in terms of the operator $E_{_D}$ defined in (6.4); see also Lemma 6.3. Let $e_x \in \widehat{W}$ be equal to 1 at a point $x \in \Gamma_h$ and vanish at all other points of $\Gamma_h$. It is easy to see that the $D$-weighted average of the components of $u \in W$ at $x$ is equal to $\langle e_x, Du\rangle$ and thus the operator $E_{_D}$ can also be written as

$$E_{_D} u = \sum_{i=1}^{N} R_i^T I^h(\delta_i^\dagger u_i) = \sum_{x \in \Gamma_h} \langle e_x, Du\rangle\, e_x. \tag{6.40}$$

**Lemma 6.10** *Let $u \in W$. Then,*

$$P_{_D} u = u - E_{_D} u\,.$$

*Proof.* By Lemma 6.9, $u - P_{_D} u$ is continuous across the interface and thus

$$E_{_D}(u - P_{_D} u) = u - P_{_D} u. \tag{6.41}$$

Since, for $x \in \Gamma_h$, the vector $e_x$ is continuous across $\Gamma$, $Be_x = 0$. We find, by using the definition of $P_{_D}$ given in (6.38), that the $D$-weighted average of $P_{_D} u$ at $x$ vanishes:

$$\langle e_x, DP_{_D} u\rangle = \langle e_x, B^T(BD^{-1}B^T)^{-1}Bu\rangle = \langle Be_x, (BD^{-1}B^T)^{-1}Bu\rangle = 0,$$

and thus, by using (6.40),

$$E_{_D} P_{_D} u = 0.$$

Combining this equality with (6.41) proves the result. □

The identity of Lemma 6.10 also form a basis for an alternative stopping criterion for one-level FETI and FETI-DP iterations; cf. [400] and [337]. In

addition to the the residual for equation 6.34, we can compute, at a very modest extra cost, a *primal* residual obtained by averaging the displacement vector across the interface and applying the assembled Schur complement.

We note that, since $E_D$ does not depend on $B$, $P_D$ does not depend on the particular choice of $B$ either. For future reference, we write down a formula for $P_D w$ for an arbitrary element $w \in W$. It is an easy consequence of Lemma 6.10. Thus, we find by using formula (6.40), that for $i = 1, \ldots, N$,

$$v_i(x) := (P_D w(x))_i = \sum_{j \in \mathcal{N}_x} \delta_j^\dagger(x)(w_i(x) - w_j(x)), \quad x \in \partial\Omega_{i,h} \cap \Gamma_h, \quad (6.42)$$

where $\mathcal{N}_x$ again is the set of indices of the subregions that have $x$ on their boundaries. We note that the coefficients $\delta_j^\dagger(x)$ in this expression are constant on each face and on each edge of $\partial\Omega_i$ and that they are independent of the particular choice of $B$. In this section, we will use Lemmas 6.3 and 6.10 to estimate the norm of $P_D$. In Sect. 6.4, we will however use the formula (6.42) directly and write the $v_i$ as sums of terms that vanish at all the interface nodes outside individual faces, edges, and vertices, respectively. The norms of the individual terms of this sum are then estimated; see, e.g., [178, 182, 184], the proof of Lemma 6.3 as well as Sect. 6.4.

To prepare for the analysis of the preconditioner, we equip $V'$ with the norm

$$\|\mu\|_{V'}^2 := \left|D^{-1}B^T(BD^{-1}B^T)^{-1}\mu\right|_S^2 = \langle \widehat{M}^{-1}\mu, \mu\rangle, \quad (6.43)$$

where, as before, $|w|_S^2 = \langle Sw, w\rangle$. We have

**Lemma 6.11** *$\|\cdot\|_{V'}$ defines a norm on $V'$.*

*Proof.* Since $\|\cdot\|_{V'}$ is clearly a seminorm, we only need to show that $\|\mu\|_{V'} = 0$ implies $\mu = 0$. Consider any $\mu \in V'$ with $\|\mu\|_{V'} = 0$. By Lemma 6.8, $\mu = B\tilde{w}$ for some $\tilde{w} \in \mathit{range}(P_D)$. Since $P_D\tilde{w} = \tilde{w}$, we obtain

$$0 = \|\mu\|_{V'}^2 = \|B\tilde{w}\|_{V'}^2 = \left|D^{-1}B^T(BD^{-1}B^T)^{-1}B\tilde{w}\right|_S^2 = |P_D\tilde{w}|_S^2 = |\tilde{w}|_S^2.$$

Thus, $\tilde{w} \in \mathit{kernel}(S)$ and by the definition of $V'$, (6.33), we find that $\mu = 0$ since

$$\|\mu\|_Q^2 = \langle \mu, Q\mu\rangle = \langle \mu, QB\tilde{w}\rangle = 0.$$

□

The importance of the operator $P_D$ can be seen by the fact that, for $\mu = Bu$, we have

$$\|\mu\|_{V'}^2 = \langle \widehat{M}^{-1}\mu, \mu\rangle = |P_D u|_S^2.$$

We will use this property extensively in the analysis of all FETI methods in this chapter.

We can now show that

$$P\widehat{M}^{-1} : V' \to V,$$

is symmetric and positive definite. Symmetry is easy to establish and positive definiteness follows immediately from Lemma 6.11 and the fact that, with $\lambda \in V'$, $\langle P\widehat{M}^{-1}\lambda, \lambda\rangle = \langle \widehat{M}^{-1}\lambda, \lambda\rangle = \|\lambda\|^2_{V'}$, since $V' = range(P^T)$.

We equip the space of admissible increments $V$ with a norm

$$\|\lambda\|_V := \sup_{\mu \in V'} \frac{\langle \lambda, \mu\rangle}{\|\mu\|_{V'}}.$$

As we have already observed, $V'$ is isomorphic to the dual space of $V$. By using formula (6.43), we find by a simple computation that

$$\|\lambda\|^2_V = \langle \widehat{M}\lambda, \lambda\rangle, \quad \lambda \in V. \tag{6.44}$$

This formula is legitimate since $V = range(P)$, and we can give a good meaning to

$$P^T\widehat{M} : V \to V',$$

as a symmetric, positive definite operator. We can effectively view $\widehat{M}^{-1}$ and $\widehat{M}$ as symmetric, positive definite operators from $V'$ onto $V$ and $V$ onto $V'$, respectively.

The next result is needed in the proofs of Lemma 6.14 and indirectly in the proof of Theorem 6.15.

**Lemma 6.12** *For any $w \in W$, there exists a unique $z_w \in kernel(S)$ such that $B(w + z_w) \in V'$. Moreover,*

$$\|Bz_w\|_Q \leq \|Bw\|_Q. \tag{6.45}$$

*Proof.* We recall that the property $B(w + z_w) \in V'$ is equivalent to

$$B^TQB(w + z_w) \perp kernel(S).$$

This Galerkin condition ensures that the element $z_w \in kernel(S)$ is a solution of the variational problem:

$$z^T(B^TQB)(w + z_w) = 0, \quad z \in kernel(S).$$

Since $kernel(S) \cap kernel(B) = \{0\}$ and $Q$ is symmetric and positive definite, the operator $(B^TQB)$ is symmetric and positive definite on $kernel(S)$. The problem thus has a unique solution $z_w$ that minimizes $\|B(w + z)\|^2_Q$, over $z \in kernel(S)$. The orthogonality condition also ensures that

$$\|Bz_w\|^2_Q = \langle z_w, (B^TQB)z_w\rangle \leq \langle w, (B^TQB)w\rangle = \|Bw\|^2_Q,$$

which concludes the proof. □

We will now establish an important stability estimate for $P_D$, which is at the core of the proof of our main results. It is closely related to Lemma 6.3 which is the core result in the convergence theory of the Neumann-Neumann

algorithms. For the choice $Q = \widehat{M}^{-1}$, we are then almost ready to prove one of our main results, Theorem 6.15. As noted before equally strong results can be obtained with a careful choice of a diagonal $Q$; the use of such an operator will decrease the cost of the algorithm. We have chosen not to cover the related, quite complicated theory; see [289] for full details.

We note that we do not know how to prove our next result, Lemma 6.13, nor the corresponding result in Sect. 6.3.3, without using Assumption 6.1. We recall that this assumption is not necessary in the Neumann-Neumann theory (cf. Dryja and Widlund [184]) since in the Neumann-Neumann algorithm we are free to increase the coarse space, e.g., by adding basis functions associated with the special boundary subdomains while in the one-level FETI algorithms of this section, we have to work exactly with *kernel*$(S)$ when constructing the projection $P$.

**Lemma 6.13** *For any $w \in$ range$(S)$, we have*

$$|P_D w|_S^2 \leq C(1 + \log(H/h))^2 |w|_S^2 .$$

*Here $C$ is independent of $h$, $H$, $\gamma$, and the values of the $\rho_i$.*

*Proof.* This result follows immediately from Lemmas 6.10 and 6.3. □

We now combine the results of Lemmas 6.12 and 6.13.

**Lemma 6.14** *For any $w \in$ range$(S)$ and with the unique $z_w \in$ kernel$(S)$ given in Lemma 6.12, and for $Q = \widehat{M}^{-1}$, we have*

$$|P_D z_w|_S^2 \leq C(1 + \log(H/h))^2 |w|_S^2 .$$

*Here $C$ is independent of $h$, $H$, $\gamma$, and the values of the $\rho_i$.*

*Proof.* For any $u \in W$ and our choice of $Q$, we have

$$|P_D u|_S^2 = \langle S P_D u, P_D u \rangle = \langle \widehat{M}^{-1} B u, B u \rangle = \|B u\|_Q^2 .$$

According to Lemma 6.12, for any $w \in$ range$(S)$, the unique $z_w \in$ kernel$(S)$ such that $w + z_w \in V'$ satisfies

$$|P_D z_w|_S = \|B z_w\|_Q \leq \|B w\|_Q = |P_D w|_S .$$

The proof is completed by combining this inequality with Lemma 6.13. □

We are now ready to prove a condition number estimate for the preconditioned FETI operator $P\widehat{M}^{-1}P^T F$.

**Theorem 6.15** *The preconditioner $\widehat{M}$, with $Q = \widehat{M}^{-1}$, satisfies*

$$\langle \widehat{M}\lambda, \lambda \rangle \leq \langle F\lambda, \lambda \rangle \leq C(1 + \log(H/h))^2 \langle \widehat{M}\lambda, \lambda \rangle, \quad \lambda \in V. \tag{6.46}$$

*Here $C$ is independent of $h$, $H$, $\gamma$, and the values of the $\rho_i$.*

*Proof.* We will estimate the smallest eigenvalue $\lambda_{\min}(P\widehat{M}^{-1}P^T F)$ from below and the largest eigenvalue $\lambda_{\max}(P\widehat{M}^{-1}P^T F)$ from above.

**Lower Bound:**

We note that this bound is optimal in the sense that it is independent of $h$ and $H$ and possible coefficient jumps. It is derived using purely algebraic arguments.

Following Mandel and Tezaur [340, proof of lemma 3.11], we will use the formula

$$\langle F\lambda, \lambda\rangle = \sup_{w \in range(S)} \frac{\langle \lambda, Bw\rangle^2}{|w|_S^2}, \quad \lambda \in V. \tag{6.47}$$

Our proof of (6.47) is essentially borrowed from Mandel and Tezaur. We first note that $S^{-1/2}B^T\lambda \in range(S)$ has a good meaning since $\lambda \in V$ means that $B^T\lambda \in range(S)$. We find that

$$\begin{aligned}\langle F\lambda, \lambda\rangle = \langle S^\dagger B^T\lambda, B^T\lambda\rangle &= \|S^{-1/2}B^T\lambda\|^2 \\ = \sup_{v \in range(S)} \frac{\langle S^{-1/2}B^T\lambda, v\rangle^2}{\|v\|^2} &= \sup_{w \in range(S)} \frac{\langle B^T\lambda, w\rangle^2}{|w|_S^2}.\end{aligned}$$

Let $\mu \in V'$ be arbitrary. It follows from Lemma 6.8 that there exists a $\tilde{w} \in range(P_D)$ such that $\mu = B\tilde{w}$. We denote by $\tilde{w}_\perp$ the component of $\tilde{w}$ that is orthogonal to $kernel(S)$. Clearly, we have

$$\sup_{w \in range(S)} \frac{\langle \lambda, Bw\rangle^2}{|w|_S^2} \geq \frac{\langle \lambda, B\tilde{w}_\perp\rangle^2}{|\tilde{w}_\perp|_S^2}.$$

We also observe that for all $\tilde{w}$,

$$\langle S\tilde{w}_\perp, \tilde{w}_\perp\rangle = \langle S\tilde{w}, \tilde{w}\rangle, \tag{6.48}$$

and it also follows, from the definition of $V$, that

$$\langle \lambda, B\tilde{w}_\perp\rangle = \langle \lambda, B\tilde{w}\rangle, \quad \lambda \in V. \tag{6.49}$$

Using (6.48) and (6.49), we obtain, since $\tilde{w} = P_D\tilde{w}$,

$$\frac{\langle \lambda, B\tilde{w}_\perp\rangle^2}{|\tilde{w}_\perp|_S^2} = \frac{\langle \lambda, B\tilde{w}\rangle^2}{|\tilde{w}|_S^2} = \frac{\langle \lambda, B\tilde{w}\rangle^2}{|P_D\tilde{w}|_S^2} = \frac{\langle \lambda, \mu\rangle^2}{|D^{-1}B^T(BD^{-1}B^T)^{-1}\mu|_S^2} = \frac{\langle \lambda, \mu\rangle^2}{\|\mu\|_{V'}^2},$$

for $\mu \in V'$. The proof of the left inequality of (6.46) concludes by using the definition of the norm $\|\cdot\|_V$ and formula (6.44).

**Upper Bound:**

We will derive an upper bound for $\langle F\lambda, \lambda\rangle$ that depends only polylogarithmically on $H/h$ and is independent of possible coefficient jumps.

Let $w \in range(S)$ be arbitrary. By Lemma 6.12, there exists a unique $z_w \in kernel(S)$ such that $B(w + z_w) \in V'$. By using Lemmas 6.13 and 6.14, we obtain

$$|P_D(w+z_w)|_S^2 \leq C(1+\log(H/h))^2 |w|_S^2 . \tag{6.50}$$

Combining this formula with (6.47), we obtain, for all $\lambda \in V$,

$$\begin{aligned}
\langle F\lambda, \lambda \rangle &= \sup_{w \in range(S)} \frac{\langle \lambda, Bw \rangle^2}{|w|_S^2} \\
&\leq C\,(1+\log(H/h))^2 \sup_{w \in range(S)} \frac{\langle \lambda, Bw \rangle^2}{|P_D(w+z_w)|_S^2} \\
&= C\,(1+\log(H/h))^2 \sup_{w \in range(S)} \frac{\langle \lambda, B(w+z_w) \rangle^2}{\|B(w+z_w)\|_{V'}^2} \\
&= C\,(1+\log(H/h))^2 \sup_{\substack{\tilde{w} \in W \\ B\tilde{w} \in V'}} \frac{\langle \lambda, B\tilde{w} \rangle^2}{\|B\tilde{w}\|_{V'}^2} \\
&= C\,(1+\log(H/h))^2 \sup_{\mu \in V'} \frac{\langle \lambda, \mu \rangle^2}{\|\mu\|_{V'}^2} \\
&= C\,(1+\log(H/h))^2 \, \|\lambda\|_V^2 .
\end{aligned}$$

The proof of the right inequality of (6.46) concludes by using (6.44). □

It is important to note that the special choice of $Q$ enters the proof of this theorem only via (6.50), which in turn depends on Lemma 6.14. Therefore, if we can prove an equally strong bound for $|P_D z_w|_S^2$ for another choice of $Q$, then we immediately obtain a result as strong as Theorem 6.15. As noted before, an alternative recipe, with a diagonal $Q$, is provided in full detail in Klawonn and Widlund [289, Theorem 4.11]; it gives as strong a result as Theorem 6.15.

### 6.3.3 The Case of Redundant Lagrange Multipliers

In this subsection, we extend our analysis to the case of redundant Lagrange multipliers; cf. Fig. 6.3, right. For a detailed algorithmic description of FETI preconditioners in this case and an analysis based on mechanics, see Rixen and Farhat [398, 399]. We note that, in an implementation, there appears to be real advantages in treating the constraints in a fully symmetric way, i.e., using a maximum number of of Lagrange multipliers. In those papers, $Q_r = I$; to distinguish the redundant from the nonredundant case, we will write $Q_r$ instead of $Q$, etc., in this subsection. We will choose the Dirichlet preconditioner as $Q_r$ and note that the resulting algorithm has proven successful for difficult industrial problems; cf. Bhardwaj et al. [51]. For this choice of $Q_r$, we show, in Theorem 6.21, a condition number estimate which is independent of the jumps in the coefficients. As noted before a carefully designed, diagonal $Q_r$ can also be chosen for which we can prove a condition number estimate that is also independent of the values of the $\rho_i$; see Klawonn and Widlund [289, Theorem 5.7].

Following Rixen and Farhat, we consider the case where a maximum number of redundant Lagrange multipliers are introduced, i.e., when all possible pairs of degrees of freedom of the primal variables $u$, that belong to the same nodal point $x \in \Gamma_h$, are connected by a Lagrange multiplier. Any edge or vertex node, where at least three subregions meet, will then contribute at least one additional Lagrange multiplier in comparison with the nonredundant case. An illustration of an edge common to four subregions is given in Fig. 6.3, right.

We denote the new jump operator, similar to the one given by (6.26) but with additional rows, by

$$B_r = [B_r^{(1)}, B_r^{(2)}, \ldots, B_r^{(N)}],$$

where $B_r^{(i)}$ consists of the columns of $B_r$ attributed to the $i$-th component of the product space $W$. The new vector of Lagrange multipliers is denoted by $\lambda_r$. The space of Lagrange multipliers is chosen as $U_r := range(B_r)$. This guarantees uniqueness of the Lagrange multiplier solution since otherwise the solution of Equation (6.52), given below, would only be unique up to an additive term from $kernel(B_r^T)$. We also introduce diagonal scaling matrices $D_r^{(i)} : U_r \to U_r$, that operate on the Lagrange multiplier space. This is in contrast to the matrix $D$ of the nonredundant case, discussed in Subsect. 6.3.2, which maps the space of primal variables $W$ onto itself. The row of $B_r^{(i)}$ relative to the Lagrange multiplier that enforces continuity between the nodal values of $w_i \in W_i$ and $w_j \in W_j$, at $x \in \partial\Omega_{i,h} \cap \partial\Omega_{j,h}$, is scaled by $\delta_j^\dagger(x)$, cf. (6.3), and this scale factor defines the corresponding element of $D_r^{(i)}$. Finally, we define a scaled jump operator by

$$B_{D_r} := (D_r^{(1)} B_r^{(1)}, \ldots, D_r^{(N)} B_r^{(N)}).$$

The FETI preconditioner is given by

$$\widehat{M}_r^{-1} := B_{D_r} S B_{D_r}^T = \sum_{i=1}^{N} D_r^{(i)} B_r^{(i)} S^{(i)} B_r^{(i)T} D_r^{(i)}. \tag{6.51}$$

This preconditioner, with $\gamma = 1$, and a different scaling, was introduced in Rixen and Farhat [399, Sect. 5]. We also note that in the special case of continuous coefficients, we obtain the multiplicity scaling described in [399, Sect. 3].

The matrix of the reduced linear system can be written as

$$F_r := B_r S^\dagger B_r^T,$$

and we now have to solve the preconditioned system

$$P_r \widehat{M}_r^{-1} P_r^T F_r \lambda_r = P_r \widehat{M}_r^{-1} P_r^T d_r, \tag{6.52}$$

with $P_r := I - Q_r G_r (G_r^T Q_r G_r)^{-1} G_r^T, G_r := B_r R$, and $d_r := B_r S^\dagger f$. Here $Q_r$ will be chosen as $\widetilde{M}_r^{-1}$. We denote the inner product induced by $Q_r$ by $\langle \lambda_r, \mu_r \rangle_{Q_r}$.

The next lemma shows that the redundant and the nonredundant implementations of the Lagrange multiplier methods gives rise to the same operator, which is central in the theory.

**Lemma 6.16** *The operator $B_{D_r}^T B_r$, with its two factors just defined in this subsection, and the operator $P_D$, defined in (6.38), are the same:*

$$B_{D_r}^T B_r = P_D.$$

*Proof.* We first note that *range*$(B_r)$ contains all possible Lagrange multipliers for every $x \in \Gamma_h$. By construction, each nonzero entry of $D_r^{(i)}$ corresponds to a Lagrange multiplier and to a point $x \in \partial\Omega_{i,h} \cap \partial\Omega_{j,h}$, for some other subregion $\Omega_j$, and it is equal to $\delta_j^\dagger(x)$. Applying $(B_r^{(i)})^T$ to the vector given by $D_r^{(i)} B_r w$ yields a vector $v_i \in W_i$, with the components

$$v_i(x) := (B_r^{(i)})^T D_r^{(i)} B_r w(x) = \sum_{j \in \mathcal{N}_x} \delta_j^\dagger(x)(w_i(x) - w_j(x)), \ x \in \partial\Omega_{i,h}. \quad (6.53)$$

This is the same formula as (6.42). □

As a consequence of this result, the operator $B_{D_r}^T B_r$ does not depend on the particular choice of $B_r$ and we can still use Lemma 6.13 in the redundant case. We also note that $B_{D_r}^T B_r$ is not symmetric unless the $\rho_i$ are all the same.

A full analysis of this FETI variant, with redundant Lagrange multipliers, can now be carried out using Lemma 6.16, adapting the arguments of Subsect. 6.3.2 to the current context, step by step. From Lemmas 6.10 and 6.16, we obtain

$$B_r B_{D_r}^T B_r = B_r P_D = B_r.$$

As in subsection 6.3.2, we obtain several results by using such an identity. The proof of the following lemma proceeds exactly as that of Lemma 6.8.

**Lemma 6.17** *For any $\mu_r \in U_r$, there exists a $\tilde{w} \in$ range$(B_{D_r}^T B_r)$, such that $\mu_r = B_r \tilde{w}$.*

As in Subsect. 6.3.2, we define a space of admissible increments,

$$V_r := \{\lambda_r \in U_r : \langle \lambda_r, B_r z \rangle = 0, \quad z \in \textit{kernel}(S)\} = \textit{range}(P_r),$$

and the space

$$V_r' := \{\mu_r \in U_r : \langle \mu_r, B_r z \rangle_{Q_r} = 0, \quad z \in \textit{kernel}(S)\} = \textit{range}(P_r^T).$$

We equip $V_r'$ with the norm

$$\|\mu_r\|_{V_r'} := |B_{D_r}^T \mu_r|_S, \quad \mu_r \in V_r',$$

and $V_r$ with the norm

$$\|\lambda_r\|_{V_r} := \sup_{\mu_r \in V_r'} \frac{\langle \lambda_r, \mu_r \rangle}{\|\mu_r\|_{V_r'}}.$$

The fact that $\|\cdot\|_{V_r'}$ is a norm is established exactly as in the nonredundant case by using Lemmas 6.16 and 6.17.

**Lemma 6.18** *$\|\cdot\|_{V_r'}$ defines a norm on $V_r'$.*

As in the nonredundant case, we easily find that, for $\mu_r = B_r u$,

$$\|\mu_r\|_{V_r'}^2 = \langle \widehat{M}_r^{-1} \mu_r, \mu_r \rangle = |B_{D_r}^T B_r u|_S^2 = |P_D u|_S^2.$$

and that

$$\langle \widehat{M}_r \lambda_r, \lambda_r \rangle = \|\lambda_r\|_{V_r}^2 \quad \lambda_r \in V_r, \tag{6.54}$$

by a simple computation. Again, as in the nonredundant case, it is immediate that $P_r \widehat{M}_r^{-1} : V_r' \to V_r$, and $P_r^T \widehat{M}_r : V_r \to V_r'$, are symmetric, positive definite operators. Thus, we can view $\widehat{M}_r^{-1}$ and $\widehat{M}_r$ as symmetric, positive definite operators from $V_r'$ onto $V_r$ and $V_r$ onto $V_r'$, respectively. We now formulate a result analogous to Lemma 6.12. The proof is the same.

**Lemma 6.19** *For any $w \in W$, there exists a unique $z_w \in kernel(S)$, such that $\tilde{w} := w + z_w$ with $B_r \tilde{w} \in V_r'$. Moreover,*

$$\|B_r z_w\|_{Q_r} \leq \|B_r w\|_{Q_r}.$$

Using Lemmas 6.13, 6.16, and 6.19, we obtain an exact counterpart of Lemma 6.14

**Lemma 6.20** *For any $w \in range(S)$, and the unique $z_w \in kernel(S)$ given in Lemma 6.19, and for $Q_r = \widehat{M}_r^{-1}$, we have*

$$|B_{D_r}^T B_r z_w|_S^2 \leq C\,(1 + \log(H/h))^2 |w|_S^2.$$

*Here $C$ is independent of $h, H, \gamma$, and the values of the $\rho_i$.*

Continuing as in the nonredundant case, after substituting $B_{D_r}^T B_r$ for $P_D$, and the auxiliary results of this subsection for those of the previous subsection, we obtain

**Theorem 6.21** *The preconditioner $\widehat{M}_r$, with $Q = \widehat{M}_r^{-1}$, satisfies*

$$\langle \widehat{M}_r \lambda, \lambda \rangle \leq \langle F_r \lambda, \lambda \rangle \leq C(1 + \log(H/h))^2 \langle \widehat{M}_r \lambda, \lambda \rangle, \quad \lambda \in V.$$

*Here $C$ is independent of $h, H, \gamma$, and the values of the $\rho_i$.*

*Proof.* The proof proceeds, line by line, as the proof of Theorem 6.15. □

## 6.4 Dual-Primal FETI Methods

The dual-primal FETI (FETI-DP) methods were introduced more recently, than the one-level FETI methods, by Farhat, Lesoinne, Le Tallec, Pierson, and Rixen [194]. Their work was soon followed by a significant contribution to the theory for two dimensional second and fourth order problems by Mandel and Tezaur [341], by a paper by Farhat, Lesoinne, and Pierson [195], which specifically addresses an algorithm for three-dimensional problems, and by Pierson's doctoral dissertation [388]. In all FETI-DP algorithms, we enforce a relatively small number of continuity constraints across the interface in each iteration step. In the first algorithms of this kind only the values at the subdomain vertices were classified as primal, i.e., as having unique values. In addition, the algorithms presented in [195] and [388] use constraints on the averages over edges and faces, similarly to those of the algorithms considered in this section. Since the face and edge constraints were originally introduced to enhance the rate of convergence of the iteration, they are referred to as *optional.* We will see that we can replace all or most of the primal vertex constraints by constraints written in terms of averages over edges (and faces). Such primal edge constraints generally provide better rates of convergence of FETI-DP algorithms than when only primal vertex constraints are used.

Our presentation is, in large part, based on Klawonn, Widlund, and Dryja [292, 293] and Klawonn, Rheinbach, and Widlund [286] and, after a discussion, in the next subsection, of the original FETI-DP method in two dimensions, it will focus on algorithms for problems in three dimensions. We will show that good convergence bounds can be maintained even for quite general coefficients such as those of equation (4.3), which can model highly heterogeneous materials. The work in [292] was inspired by that of Mandel and Tezaur and also based on other earlier work, in particular [177], [178], and [289].

We recall that domain decomposition algorithms for elliptic problems cannot be scalable, i.e., have a rate of convergence which is independent of the number of subregions, unless a coarse space component is included. As we have seen in Chap. 5, the underlying coarse spaces for three dimensional problems can be quite complicated; see also [467]. We will construct two FETI-DP methods, namely Algorithms B and C, which are inspired by relatively exotic coarse spaces such as those of Subsect. 5.4.3; see also [177, 178]. Both of these methods have relatively large global, primal subspaces and we note that a main issue when searching for competitive preconditioners is to try to design methods with a coarse component of a small dimension while, at the same time, maintaining a good rate of convergence. We then present and analyze Algorithm D, which can have a very satisfactory ratio between the dimension of the coarse space and the number of substructures as well as a good rate of convergence. We finally discuss Algorithm E, which would first appear to be a useful alternative to Algorithm C, but for which additional conditions on the coefficients are required to obtain as good bounds as for Algorithms B, C, and D.

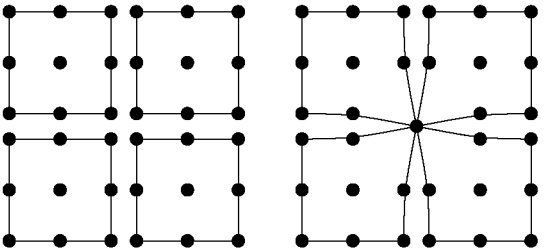

**Fig. 6.5.** One-level FETI and FETI-DP. Degrees of freedom of the spaces $W$, *left*, and $\widetilde{W}$, *right*.

As already mentioned in the beginning of Sect. 6.3, we will use subspaces $\widetilde{W}$ which are intermediary between $\widehat{W}$ and $W$ and for which only a relatively small number of continuity constraints are enforced across the interface. An important benefit of working in $\widetilde{W}$, rather than in $W$, is that the Schur complements that arise in the computation will all be strictly positive definite. The underlying variational problem can be written as in (6.25) except that the space $W$ is replaced by its subspace $\widetilde{W}$.

Since the three dimensional case is complicated, we will begin our discussion with the case of two dimensions for which the primal unknowns are only associated with the subdomain vertices. The generalization to two-dimensional elasticity problems is straightforward while three-dimensional elasticity is treated in Sect. 8.5. Algorithms for edge element approximations have been proposed in [447, 448] for two dimensions and, more recently, in [442] for three dimensions. Recently Li has extended the FETI-DP algorithms to incompressible Stokes and Navier-Stokes equations; see [314, 315, 316].

We finally note that in recent work, Mandel, Dohrmann, and Tezaur [337] have established very interesting and close connections between FETI-DP algorithms and a new family of additive Schwarz methods of balancing type, see [164, 165, 336]. They have established that almost the entire spectrum of a FETI-DP algorithm coincides with that of a related balancing algorithm. Experimental evidence of this phenomenon had previously been presented by Fragakis and Papadrakakis [206]. We note that our emphasis in this section is different; we focus on establishing selection criteria for the primal constraints which will guarantee rapid convergence of the iterations.

### 6.4.1 FETI-DP Methods in Two Dimensions

In early studies of FETI-DP methods for problems in two dimensions, see Farhat, Lesoinne, Le Tallec, Pierson, and Rixen [194] and Mandel and Tezaur [341], only constraints on the degrees of freedom associated with the vertices of the substructures are enforced in each iteration, i.e., the vertex degrees of freedom belong to the primal set of variables, while all the continuity constraints

associated with the edge nodes are fully enforced only at the convergence of the iterative method. We can think of this as resulting from incisions in the mesh along the interface leaving only the subdomain vertex nodes attached; see Fig. 6.5. We thus allow multiple values for almost all of the unknowns associated with the interface. This results in a much lower computational costs than for a standard finite element problem where continuity is enforced at every point of $\Gamma_h$.

We will always work in subspaces $\widetilde{W} \subset W$ for which sufficiently many constraints are enforced so that the resulting leading diagonal block matrix of the saddle point problem, though no longer block diagonal, is strictly positive definite. Similarly, we will denote by $\widetilde{W}^h(\Omega)$ the subspace of $\prod_i W^h(\Omega_i)$ which equals $\widetilde{W}$ when restricted to the interface $\Gamma$. We also introduce two subspaces, $\widehat{W}_\Pi \subset \widetilde{W}$ and $\widetilde{W}_\Delta$, corresponding to a primal and a dual part of the space $\widetilde{W}$. They will play an important role in the description and analysis of our iterative methods and their direct sum equals $\widetilde{W}$. We note that the dual subspace $\widetilde{W}_\Delta$ will be directly associated with jumps across the interface and with the Lagrange multipliers that are introduced to eliminate these jumps. Each of our FETI-DP algorithms, in this and in the next subsection, is expressed in terms of a Schur complement $\widetilde{S}$ related to the dual space $\widetilde{W}_\Delta$. Thus, in this subsection, $\widetilde{W}$ consists of functions in $W$ that take the same value at the subdomain vertices and can be written as

$$\widetilde{W} = \widehat{W}_\Pi \oplus \widetilde{W}_\Delta. \tag{6.55}$$

Here $\widehat{W}_\Pi \subset \widehat{W}$ is the space of continuous interface functions that vanish at all nodal points on $\Gamma_h$ except at the subdomain vertices, and $\widetilde{W}_\Delta$ is the direct sum of local subspaces $\widetilde{W}_{\Delta,i}$:

$$\widetilde{W}_\Delta := \prod_{i=1}^{N} \widetilde{W}_{\Delta,i}. \tag{6.56}$$

Here $\widetilde{W}_{\Delta,i} \subset W_i$ consists of local functions on $\partial\Omega_i$ that vanish at the vertices of $\Omega_i$.

The (continuous) degrees of freedom associated with the substructure vertices and with the subspace $\widehat{W}_\Pi$ are called *primal* ($\Pi$), while those (potentially discontinuous across $\Gamma$) associated with the subspaces $\widetilde{W}_{\Delta,i}$ and with the interior of the substructure edges are called *dual* ($\Delta$).

Let $\widetilde{A}$ be the stiffness matrix, which is obtained by restricting

$$A = \operatorname{diag}\{A^{(1)}, A^{(2)}, \ldots, A^{(N)}\}$$

from $\prod_i W^h(\Omega_i)$ to $\widetilde{W}^h(\Omega)$. We note that $\widetilde{A}$ is no longer block diagonal since a coupling now exists between substructures that have a vertex in common. We can partition $\widetilde{A}$ as

$$\widetilde{A} = \begin{pmatrix} A_{II} & A_{I\Pi} & A_{I\Delta} \\ A_{I\Pi}^T & A_{\Pi\Pi} & A_{\Pi\Delta} \\ A_{I\Delta}^T & A_{\Pi\Delta}^T & A_{\Delta\Delta} \end{pmatrix}, \tag{6.57}$$

where the subscript $I$ refers to the internal degrees of freedom of the substructures, $\Pi$ to those associated with the subdomain vertices, and $\Delta$ to those of the interior of the subdomain edges; see Fig. 6.5, right. We note that $A_{II}$ and $A_{\Delta\Delta}$ are block diagonal and that each block corresponds to a single subdomain and that any nonzero of $A_{\Delta I}$ (and $A_{I\Delta}$) represents a coupling between degrees of freedoms associated with the same substructure. Similarly as for the stiffness matrix for the conforming space $\widehat{W}$, $\widetilde{A}$ can be obtained by partially assembling local contributions associated with the substructures.

The variables of the $I$ and $\Pi$ sets are then eliminated and a Schur complement, associated with the degrees of freedom of the set $\Delta$, of the interior of the edges, is obtained:

$$\widetilde{S} = A_{\Delta\Delta} - \begin{pmatrix} A_{I\Delta}^T & A_{\Pi\Delta}^T \end{pmatrix} \begin{pmatrix} A_{II} & A_{I\Pi} \\ A_{I\Pi}^T & A_{\Pi\Pi} \end{pmatrix}^{-1} \begin{pmatrix} A_{I\Delta} \\ A_{\Pi\Delta} \end{pmatrix} \tag{6.58}$$

We also obtain a reduced right hand side $\tilde{f}_\Delta$ from the load vectors associated with the individual subdomains. We denote by $u_\Delta \in \widetilde{W}_\Delta$ the vector of degrees of freedom associated with the edges.

As in Subsect. 6.3.1, we can reformulate the finite element problem, reduced to the second subspace $\widetilde{W}_\Delta$, as a minimization problem with constraints given by the requirement of continuity across all of $\Gamma$: find $u_\Delta \in \widetilde{W}$ such that

$$\begin{cases} J(u_\Delta) := \frac{1}{2}\langle \widetilde{S} u_\Delta, u_\Delta \rangle - \langle \tilde{f}_\Delta, u_\Delta \rangle \to \min \\ B_\Delta u_\Delta = 0 \end{cases} \tag{6.59}$$

The matrix $B_\Delta$ is constructed from $\{0, 1, -1\}$, in a way very similar to the matrix $B$ introduced in Sect. 6.3.1, and in such a way that the values of the solution $u$, at the nodes common to more than one subdomain, coincide when $B_\Delta u = 0$. Again these constraints are very simple and just express that the nodal values at any edge node coincide across the interface. In comparison with the FETI methods of Sect. 6.3, we can drop some of the constraints, in particular those associated with the vertex nodes which are assigned to the primal set. Since an edge node belongs to exactly two substructures, there is basically no choice for $B_\Delta$ (apart for the sign of each row) and since the only constraints are for points on the edges, no distinction need to be made between non-redundant and redundant constraints and Lagrange multipliers.

By introducing a set of Lagrange multipliers $\lambda \in V := range(B_\Delta)$, to enforce the continuity constraints, we obtain a saddle point formulation of (6.59), as in (6.27). Since $\widetilde{A}$ is symmetric, positive definite, so is $\widetilde{S}$. We can therefore eliminate the subvector $u_\Delta$, and obtain the following system for the Lagrange multipliers:

$$F\lambda = d, \tag{6.60}$$

with

$$F := B_\Delta \widetilde{S}^{-1} B_\Delta^T, \quad d := B_\Delta \widetilde{S}^{-1} \tilde{f}_\Delta.$$

Once $\lambda$ is found, we can back solve and obtain,

$$u_\Delta = \widetilde{S}^{-1}(\tilde{f}_\Delta - B_\Delta^T \lambda) \in \widetilde{W}_\Delta.$$

The values of the solution in the interior of the subdomains, $u_I$, and at the subdomain vertices, $u_\Pi$, are provided as a byproduct when solving a linear system with the block matrix given in formula (6.58).

We will now introduce a preconditioner for $F$; cf. (6.51). As in the analysis in Subsect. 6.3.3 for the case of redundant Lagrange multipliers, we introduce diagonal scaling matrices $D_\Delta^{(i)}$. Each of their diagonal elements corresponds to a Lagrange multiplier which enforces continuity between the nodal values of some $w_i \in W_i$ and $w_j \in W_j$ at some point $x \in \Gamma_h$ and it is given by $\delta_j^\dagger(x)$. We also define a scaled jump operator by

$$B_{D,\Delta} := (D_\Delta^{(1)} B_\Delta^{(1)}, \ldots, D_\Delta^{(N)} B_\Delta^{(N)}),$$

where, as before, the block $B_\Delta^{(i)}$ is obtained by extracting the columns of $B_\Delta$ associated with the local space $W_i$.

As in Subsect. 6.3.3, we solve the dual system (6.60) using the preconditioned conjugate gradient algorithm with the preconditioner

$$M^{-1} := B_{D,\Delta} S_\Delta B_{D,\Delta}^T = \sum_{i=1}^{N} D_\Delta^{(i)} B_\Delta^{(i)} S_\Delta^{(i)} B_\Delta^{(i)T} D_\Delta^{(i)}, \tag{6.61}$$

where $S_\Delta^{(i)}$ is the restriction of the local Schur complement $S^{(i)}$ to $\widetilde{W}_{\Delta,i} \subset W_i$.

The FETI-DP method is the standard preconditioned conjugate gradient algorithm for solving the preconditioned system

$$M^{-1} F \lambda = M^{-1} d$$

and can therefore be implemented as in Sect. C.5. In contrast to the one-level FETI methods, we can use an *arbitrary* initial guess $\lambda^0$. We have rewritten the algorithm in Fig. 6.6.

Most of the computational work in each step of the algorithm goes into the application of the local Schur complements $S_\Delta^{(i)}$, in the application of $M^{-1}$, and the application of the inverse of $\widetilde{S}$, in the application of $F$. The restriction $S_\Delta^{(i)}$ can in principle be found by deleting the rows and columns in $S^{(i)}$ which are related to the subdomain vertex degrees of freedom. Since in practice $S^{(i)}$ need not be formed, given $w_{\Delta,i} \in \widetilde{W}_{\Delta,i}$, we form the vector $w_i \in W_i$, which is equal to $w_{\Delta,i}$ at the interior edge nodes and which vanishes at the subdomain vertices. We then find the vector

$$S^{(i)} w_i = (A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} A_{I\Gamma}^{(i)}) w_i \tag{6.62}$$

1. Initialize: $r^0 = d - F\lambda^0$

2. Iterate $k = 1, 2, \ldots$ until convergence

Precondition: $z^{k-1} = M^{-1} r^{k-1}$

$$\begin{aligned}
\beta^k &= \langle z^{k-1}, r^{k-1} \rangle / \langle z^{k-2}, r^{k-2} \rangle \quad [\beta^1 = 0] \\
p^k &= z^{k-1} + \beta^k p^{k-1} \quad [p^1 = z^0] \\
\alpha^k &= \langle z^{k-1}, r^{k-1} \rangle / \langle p^k, F p^k \rangle \\
\lambda^k &= \lambda^{k-1} + \alpha^k p^k \\
r^k &= r^{k-1} - \alpha^k F p^k
\end{aligned}$$

**Fig. 6.6.** Implementation of the FETI-DP method as a standard preconditioned conjugate gradient methods.

and discard the entries of the vertices. The application of $S_\Delta^{(i)}$ thus involves the solution of a Dirichlet problem on each substructure.

As far as the application of $\widetilde{S}^{-1}$ to a vector $w_\Delta \in \widetilde{W}_\Delta$ is concerned, we have already shown in Sect. 4.3 that the application of an inverse of a Schur complement can be performed by solving a system with the original matrix and with a suitable right hand side. In the current case, the vector $\widetilde{S}^{-1} f_\Delta$ equals the $\Delta$-component of the vector

$$\widetilde{A}^{-1} (0 \quad 0 \quad f_\Delta^T)^T, \tag{6.63}$$

where $\widetilde{A}$ is the stiffness matrix given in (6.57). A linear system of algebraic equations is therefore solved exactly in each step of the iteration.

In order to decrease the cost of factoring $\widetilde{A}$, it is more convenient to reorder the unknowns in a different way than in (6.57) and eliminate the unknowns of type $I$ and $\Delta$ first. After this elimination, we obtain

$$\widehat{S}_\Pi = A_{\Pi\Pi} - \begin{pmatrix} A_{I\Pi}^T & A_{\Delta\Pi}^T \end{pmatrix} \begin{pmatrix} A_{II} & A_{I\Delta} \\ A_{I\Delta}^T & A_{\Delta\Delta} \end{pmatrix}^{-1} \begin{pmatrix} A_{I\Pi} \\ A_{\Delta\Pi} \end{pmatrix}. \tag{6.64}$$

We note that the $2 \times 2$ block matrix that is inverted here can be made block-diagonal, with each block associated with a single substructure. More precisely, after permuting the rows and columns of the matrix, each block is the stiffness matrix of a local problem with Neumann boundary conditions on the subdomain edges and with a zero Dirichlet condition at the vertices. Therefore, all the unknowns except those of the subdomain vertices (subscript $\Pi$) can first be eliminated at a modest expense, and in parallel across the subdomains, resulting in the Schur complement $\widehat{S}_\Pi$ for the vertex variables, which provides the coarse component of our preconditioner. In this first step, we can take full advantage of a high quality sparse Cholesky solver.

The order of $\widehat{S}_{\Pi}$ equals the number of subdomain vertices which do not belong to $\partial\Omega_D$. The matrix $\widehat{S}_{\Pi}$ is sparse. Indeed, it can be shown quite easily, as in appendix C.2, that no nonzero off-diagonal elements exist in this reduced system matrix except those that correspond to pairs of vertices which belong to the same substructure. It is clear from these remarks that each iteration step involves the solution of one Neumann and one Dirichlet problem, on each substructure, and a coarse problems involving the vertex variables and the matrix $\widehat{S}_{\Pi}$.

A similar strategy can be developed if additional optional constraints are added, as for the three dimensional methods of the next section, but we note that, obviously, the final reduced system of equations, which is directly related to the global part of the primal subspace, will be larger and therefore more expensive to handle unless we also eliminate the vertex constraints for sufficiently many or all the vertices. We will discuss several options of handling these issues in Subsect. 6.4.4 after describing and analyzing several FETI-DP algorithms.

In their paper on two-dimensional problems, Mandel and Tezaur [341] established a condition number bound of the form $C(1 + \log(H/h))^2$ for a FETI-DP method of this type when equipped with a Dirichlet preconditioner. The algorithm considered is scalable with the constant $C$ independent of the number of subregions, the subdomain diameters, as well as the mesh size $h$ of the finite element model. Mandel and Tezaur also established a similar result for a fourth-order elliptic problem in the plane. Their proof in [341], for the second order equation, uses linear algebra arguments and a lemma from a classical paper by Bramble, Pasciak, and Schatz [72, Lemma 3.5].

The same algorithm, Algorithm A, with only primal vertex constraints, can also be defined for the three dimensional case but it does not always perform well; see Farhat et al. [194, sect. 5]. This is undoubtedly related to the poor bound for the vertex-based iterative substructuring methods; see Sect. 5.4.1 or [178, Sect. 6.1]. However, we will be able to establish results as strong as in the previous sections of this chapter for three other algorithms, Algorithms B, C, and D; see Theorems 6.35 and 6.38.

We end this section with a comparison of one-level and dual-primal FETI methods. The latter class of algorithm, which has been developed more recently, and presents certain advantages:

- FETI-DP algorithms do not require the characterization of the kernels of local Neumann problems. Indeed, the enforcement of the additional constraints in each iteration always makes the local problems nonsingular and at the same time provides an underlying coarse global problem. As already pointed out in Remark 6.7, the characterization of these null spaces, required for the one-level FETI methods, may not be a trivial task for some more complicated problems.

- The same FETI-DP algorithms and codes can be employed for problems with and without nontrivial local kernels; this is not the case for one-level FETI methods (see Remark 6.7).
- FETI-DP methods do not require the introduction of a scaling matrix $Q$, which enters in the construction of a coarse solver for one-level FETI algorithms.
- One-level FETI methods are projected conjugate gradient algorithms that cannot start from an arbitrary initial guess; see Fig. 6.4. This initial guess can be far from the exact solution. In contrast, FETI-DP methods are standard preconditioned conjugate algorithms and can therefore employ an arbitrary initial guess $\lambda^0$; see Fig. 6.6.

### 6.4.2 A Family of FETI-DP Algorithms in Three Dimensions

As already mentioned, the algorithm of the previous subsection can also be defined for three-dimensional problems in a straightforward way. However, it does not always provide a satisfactory rate of convergence; see Subsect. 6.4.5. We will therefore modify it in such a way that different and often larger coarse problems are obtained. We will present five algorithms, denoted by A, B, C, D, and E, respectively, and will analyze Algorithms B, C, and D, in full detail. An analysis of Algorithm E can be found in [293].

A close look at the previous section suggests that the FETI-DP method is completely specified by the conforming subspace $\widehat{W}_\Pi$ in (6.55). More precisely, a set of interface primal degrees of freedom ($\Pi$) in the conforming space $\widehat{W}$ uniquely defines a subspace $\widehat{W}_\Pi \subset \widehat{W}$ as the span of the basis functions in $\widehat{W}$ associated with these degrees of freedom. In that subsection, we have chosen the nodal basis functions of the subdomain vertices. Here we will also consider averages on subdomain edges and faces; see Fig. 6.7.

The nonconforming spaces $\widetilde{W}_\Delta$ and $\widetilde{W}$ are specified by (6.56) and (6.55). More precisely, the local spaces $\widetilde{W}_{\Delta,i}$ are defined as the subspaces of the $W_i$ consisting of functions for which the primal degrees of freedom (a combination of nodal values at some vertices and averages over certain edges and faces) vanish. Thus, the space $\widetilde{W}$ consists of functions in $W$ for which the values associated with the degrees of freedom of the set $\Pi$ are the same regardless of which component of the product space $W$ is used in their calculation. We note that the larger $\widehat{W}_\Pi$, the smaller $\widetilde{W}_\Delta$ and $\widetilde{W}$. In the limit where the set $\Pi$ consists of all the nodal values on $\Gamma$, $\widehat{W}_\Pi$ coincides with $\widehat{W}$ and $\widetilde{W}_\Delta$ is empty.

Once the interface space $\widetilde{W}$ is specified, the global space $\widetilde{W}^h(\Omega)$ is determined as in the previous subsection. The degrees of freedom in $\widetilde{W}^h(\Omega)$ are again divided into interior ($I$), primal ($\Pi$), and dual ($\Delta$) sets. The stiffness matrix $\widetilde{A}$, defined as the restriction of $A$ to $\widetilde{W}^h(\Omega)$, has the same block structure as in (6.57), and can be obtained by partially subassembling the matrices of the subdomains. We arrive at a reduced problem by eliminating the primal variables associated with the interior nodes ($I$), the vertex nodes designated

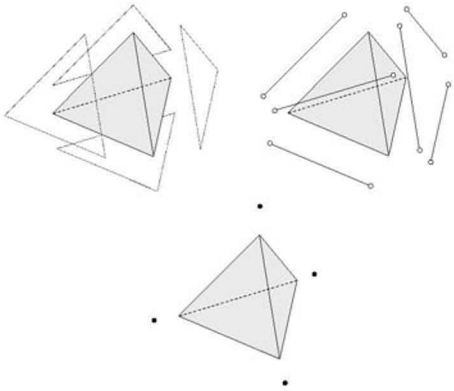

**Fig. 6.7.** Potential primal degrees of freedom for FETI-DP in three dimensions: face and edge averages, and vertex values.

as primal, as well as the variables (or a set of special Lagrange multipliers) related to the remaining primal constraints ($\Pi$). The resulting equation for the dual part of the solution, $u_\Delta$, can be written as a constrained minimization problem as in (6.59). The Schur complement $\widetilde{S}$ will be positive definite since we will always introduce enough constraints, in terms of the primal subspace $\widetilde{W}_\Pi$, to make the restriction of $A$ to $\widetilde{W}^h(\Omega)$ invertible.

As before, a matrix $B_\Delta$ is constructed from $\{0, 1, -1\}$ and the constraints $B_\Delta u_\Delta = 0$ express that the nodal values at any node coincide across the interface $\Gamma$. As in the previous subsection, we do not need any constraints at primal vertices but we will otherwise work with a fully redundant set of Lagrange multipliers for the edges as in Subsect. 6.3.3. We recall that, by assumption, every edge belongs to more than two substructures and we note that if we have a constraint on the average over a face, work with standard nodal basis functions, and use one constraint per node, we already have a redundancy. The same observation is equally valid for any edge with a primal constraint.

We now introduce Lagrange multipliers $\lambda \in V := range(B_\Delta)$, where the range is taken over functions in $\widetilde{W}_\Delta$; since any continuous function is in the null space of $B_\Delta$, the range may as well be taken over all of $\widetilde{W}$. We will in fact redefine $B_\Delta$, in comparison with the previous subsection, and view it as an operator on all of $\Gamma_h$. The variables $u_\Delta$ are then eliminated and an equation for $\lambda$ is obtained which involve the operator $F$, as in (6.60). The operator $F$ will depend on the choice of the subspaces $\widetilde{W}_\Pi$ and $\widetilde{W}_\Delta$ and we will denote the operators of the resulting linear systems by $F_A$, $F_B$, $F_C$, $F_D$, and $F_E$, respectively.

A preconditioner is defined similarly as in the previous subsection: the scaling matrices $D_\Delta^{(i)}$ are defined as before and the preconditioner $M^{-1}$ is given in terms of the block diagonal matrix $S$ and the matrix $B_{D,\Delta}$ by

$$M^{-1} := B_{D,\Delta} S B_{D,\Delta}^T = \sum_{i=1}^{N} D_\Delta^{(i)} B_\Delta^{(i)} S^{(i)} B_\Delta^{(i)T} D_\Delta^{(i)}.$$

We note that the product giving $M^{-1}$ is well defined since $B_{D,\Delta}$ is now an operator on all of $\Gamma_h$ and the computation is not affected in any way; see Sect. 6.4.4 for details. As before, multiplication by $S$ amounts to solving a Dirichlet problem on each subdomain. Just as in Subsect. 6.3.1, we should make sure that we return to $V = range(B_\Delta)$ after applying the preconditioner. We do so by premultiplying $M^{-1}$ by $P$ the $\ell_2$-orthogonal projection from $range(B_{D,\Delta})$ onto $range(B_\Delta)$. Thus, $P$ removes a component of $kernel(B_\Delta^T)$. We maintain symmetry by also multiplying $M^{-1}$ from the right by the same operator. In contrast with the one-level FETI methods of Sect. 6.3, the inclusion of this operator $P$ will make no real difference; examining the algorithm in Fig. 6.6, we find that while the Lagrange multipliers will be different, the underlying elements in the space $\widetilde{W}$ will be the same. We note that the projection on the right will always be applied to elements in $range(B_\Delta)$; it will therefore make no difference. We note, in addition, that $kernel(F) = kernel(B_\Delta^T)$; therefore $FM^{-1} = FPM^{-1}$.

The definition of $M^{-1}$ clearly depends on the choice of the subspaces $\widehat{W}_\Pi$ and $\widetilde{W}_\Delta$ for the different algorithms. The resulting preconditioners are denoted by $M_A^{-1}$, $M_B^{-1}$, $M_C^{-1}$, $M_D^{-1}$, and $M_E^{-1}$, respectively. In Lemma 6.33, we will establish that $M^{-1}$ is always invertible whenever $\widetilde{S}$ is.

We note that the matrix $\widetilde{A}$ and its blocks, as given in (6.57), are seldom formed in practice. Algorithmically, the Schur complements $\widetilde{S}$ and the $S^{(i)}$ are only needed in terms of products of $\widetilde{S}^{-1}$ and the $S^{(i)}$ and vectors. When the primal space $\widehat{W}_\Pi$ also involves averages over edges or faces, the explicit construction of the individual blocks would require a change of basis to accommodate these averages, and to separate the primal and dual part of the space. This is not desirable since it can lead to additional fill in. In preparation for the analysis, in Subsect. 6.4.3, and the discussion of the practical implementation, in Subsect. 6.4.4, we provide the following lemma with a variational definition of $\widetilde{S}$ without employing a new basis. We recall that the block diagonal matrix $S$ is defined in terms of the local Schur complements $S^{(i)}$; cf. Sect. 6.3.1.

**Lemma 6.22** *Let the spaces $\widehat{W}_\Pi$, $\widetilde{W}_\Delta$, and $\widetilde{W}$, and the Schur complement $\widetilde{S}$ be defined as above. Then, $\widetilde{S}$ satisfies*

$$\langle \widetilde{S} u_\Delta, u_\Delta \rangle = \min \langle Sw, w \rangle, \quad u_\Delta \in \widetilde{W}_\Delta, \tag{6.65}$$

*where the minimum is taken over $w = w_\Pi + u_\Delta \in \widetilde{W}$, with $w_\Pi \in \widehat{W}_\Pi$.*

*Proof.* We consider the restriction of $S$ to the space $\widetilde{W}$ and realize that $\widetilde{S}$ may be obtained by using a step of block Cholesky factorization that eliminates the degrees of freedom of type $\Pi$. An argument as in the proof of Lemma 4.9 concludes the proof. □

We note that any Schur complement of a positive semi-definite, symmetric matrix is always associated with a variational problem such as (6.65). Previously, we have often considered discrete harmonic extensions of values given on the interface $\Gamma$ (or $\partial\Omega_i$); here we compute an element in the primal subspace as a function of an element of the dual space $\widetilde{W}_\Delta$. In Subsect. 6.4.4, we will provide practical procedures for applying $\widetilde{S}^{-1}$ to vectors.

We are now ready to define our first FETI-DP algorithm in terms of a pair of subspaces. The definitions of our algorithms will employ the characteristic finite element functions $\theta_\mathcal{F}$ of a face $\mathcal{F}$, $\theta_\mathcal{E}$ of an edge $\mathcal{E}$, and $\theta_\mathcal{V}$ of a vertex $\mathcal{V}$. These functions were introduced in Sect. 4.6 and they were featured in the proof of Lemma 6.3. We also recall that we identify trace functions defined on the subdomain boundaries with the corresponding discrete harmonic extensions into the substructures.

**Algorithm 6.23 (Algorithm A)** *The primal subspace, $\widehat{W}_\Pi$, is spanned by the discrete harmonic vertex nodal finite element basis functions $\theta_{\mathcal{V}^{i\ell}}$. The local subspace $\widetilde{W}_{\Delta,i}$ is defined as the subspace of $W_i$ of elements which vanish at the subdomain vertices, i.e., by*

$$\widetilde{W}_{\Delta,i} := \{u \in W_i : u(\mathcal{V}^{i\ell}) = 0, \ \mathcal{V}^{i\ell} \in \partial\Omega_i\}.$$

*Hence, $\widetilde{W} = \widetilde{W}_A$ is the subspace of $W$ of functions that are continuous at the subdomain vertices.*

As we have already pointed out, Algorithm A is not competitive because of a poor estimate of the condition number; see also Remark 6.39. Our first algorithm for which a good condition number estimate has been obtained is Algorithm B; we note that it has a quite large primal subspace and that it is therefore quite unlikely to be competitive with one-level FETI and Neumann-Neumann algorithms in spite of a good bound of the condition number.

**Algorithm 6.24 (Algorithm B)** *The primal subspace, $\widehat{W}_\Pi$, is spanned by the vertex nodal finite element basis functions $\theta_{\mathcal{V}^{i\ell}}$ and the cutoff functions $\theta_{\mathcal{E}^{ik}}$ and $\theta_{\mathcal{F}^{ij}}$ of all the edges and faces, respectively, of the interface. The local subspace $\widetilde{W}_{\Delta,i}$ is defined as the subspace of $W_i$ where the values at the subdomain vertices vanish together with the averages $\overline{u}_{\mathcal{E}^{ik}}$ and $\overline{u}_{\mathcal{F}^{ij}}$, i.e., by*

$$\widetilde{W}_{\Delta,i} := \{u \in W_i : u(\mathcal{V}^{i\ell}) = 0, \overline{u}_{\mathcal{E}^{ik}} = 0, \overline{u}_{\mathcal{F}^{ij}} = 0, \ \mathcal{V}^{i\ell}, \mathcal{E}^{ik}, \mathcal{F}^{ij} \subset \partial\Omega_i\}.$$

*Here*

$$\overline{u}_{\mathcal{E}^{ik}} = \frac{\int_{\mathcal{E}^{ik}} u ds}{\int_{\mathcal{E}^{ik}} 1 ds} \quad \text{and} \quad \overline{u}_{\mathcal{F}^{ij}} = \frac{\int_{\mathcal{F}^{ij}} u dx}{\int_{\mathcal{F}^{ij}} 1 dx}. \tag{6.66}$$

*Hence, $\widetilde{W} = \widetilde{W}_B$ is the subspace of $W$ of functions that are continuous at the subdomain vertices and have the same values of $\overline{u}_{\mathcal{E}^{ik}}$ and $\overline{u}_{\mathcal{F}^{ij}}$ independently of which component of $u \in \widetilde{W}_B$ is used in the evaluation of these averages.*

It turns out that the constraints on the face averages can be dropped while maintaining just as satisfactory a bound on the condition number of the algorithm.

**Algorithm 6.25 (Algorithm C)** *The primal subspace, $\widehat{W}_\Pi$, is spanned by the vertex nodal finite element basis functions $\theta_{\mathcal{V}^{i\ell}}$ and the cutoff functions $\theta_{\mathcal{E}^{ik}}$ of all the edges of $\Gamma$. The local subspace $\widetilde{W}_{\Delta,i}$ is defined as the subspace of $W_i$ where the values at the subdomain vertices vanish together with the edge averages $\overline{u}_{\mathcal{E}^{ik}}$, i.e., by*

$$\widetilde{W}_{\Delta,i} := \{u \in W_i : u(\mathcal{V}^{i\ell}) = 0, \overline{u}_{\mathcal{E}^{ik}} = 0,\ \mathcal{V}^{i\ell}, \mathcal{E}^{ik} \subset \partial\Omega_i\}.$$

*Hence, $\widetilde{W} = \widetilde{W}_C$ is the subspace of $W$ of functions that are continuous at the subdomain vertices and have common averages $\overline{u}_{\mathcal{E}^{ik}}$ for all the edges.*

The number of degrees of freedom of the corresponding primal subspace $\widehat{W}_\Pi$ is therefore equal to the sum of the number of vertices and the number of edges; this $\widehat{W}_\Pi$ is of lower dimension than the primal space of Algorithm B.

Since we are able to show as strong a result for Algorithm C as for Algorithm B, it is natural to attempt to eliminate even more constraints, i.e., further decrease the primal subspace $\widehat{W}_\Pi$ while preserving the fast convergence of the FETI-DP method. This will lead to the introduction of Algorithm D; in the case of benign coefficients, the only constraints necessary are one primal edge for each face. After, quite selectively, adding edge and vertex constraints, such as those of Algorithm C, we are able to prove as strong a result as for Algorithms B and C for any choice of the $\rho_i$. As in our design of good scalings for Neumann-Neumann and one-level FETI methods, our selection of primal edges and vertices are suggested very directly by our attempts to prove strong theoretical results, in particular, results which hold for all possible positive values of the coefficients $\rho_i$.

The number of constraints enforced in all the steps of the iterations of Algorithms B and C is substantially larger than when only the vertex constraints are satisfied as in Algorithm A, but we are still able to work with a uniformly bounded number of such constraints for each substructure. In order to put this in perspective, we consider Algorithms B and C in the very regular case of cubic substructures. There are then seven global variables for each interior substructure in the case of Algorithm B since there are eight vertices, each shared by eight cubes, twelve edges, each shared by four, and six faces each shared by a pair of substructures. The count for Algorithm C is four while for Algorithm D the count can be as low as three quarters; if the coefficients are benign, we need only select one edge for each face of $\Gamma$ and such a primal edge can be shared by two faces. We note that the counts would be different, relative to the number of substructures, e.g., in the case of tetrahedral subregions.

Our requirements on a minimal set of primal constraints, which we have found necessary to give a complete proof of a good bound for Algorithm D, will be given in Assumption 6.27. We note that there are basically two requirements for the bound for the operator $P_D$ in Lemma 6.36. The first is that for each face there is a primal edge. This ensures that on each face, all functions in $\widetilde{W}_\Delta$ have vanishing mean value on at least one edge.The second is that, for each pair of substructures that have only an edge or a vertex in common, it is possible to move from one to the other moving through primal edges and substructures for which the coefficient $\rho$ is not much smaller than the minimum for the two subregions under consideration. We finally note that vertex constraints might not be necessary as long as there are enough primal edges. However, in practice, some of them might be employed in order to make the Schur complement $\widetilde{S}$ invertible even in the absence of the edge constraints; see further Subsect. 6.4.4.

Before proceeding, we give a more precise definition. Let $\Omega_i$ and $\Omega_j$ be two substructures that have an edge $\mathcal{E}^{ij}$ but not a face in common. An *acceptable edge path* is a path from $\Omega_i$ to $\Omega_j$, possibly through several other subdomains, $\Omega_k$, which all have the edge $\mathcal{E}^{ij}$ in common and the coefficients of which satisfy

$$TOL_D * \rho_k \geq \min(\rho_i, \rho_j) \tag{6.67}$$

for some chosen tolerance $TOL_D$; this parameter will appear in our bounds of the condition number of the algorithm. The path can only pass from one subdomain to another through an edge designated as primal. Clearly, if the edge $\mathcal{E}^{ij}$ is primal, an acceptable edge path between $\Omega_i$ and $\Omega_j$ exists consisting of only the two substructures. Consequently, we can always create an acceptable edge path for any edge by simply designating the edge as primal. We also note that this issue will not arise for an edge in the interior of the region which is common to only three subdomains, if we assume that every face has at least one primal edge. In this case, any pair of the three subdomains will have a face in common.

We next consider a pair of substructures which has a vertex $\mathcal{V}^{i\ell}$ but not a face or an edge in common. We then assume that $\mathcal{V}^{i\ell}$ is either a primal vertex or that we have an acceptable edge path of the same nature as above, except that we can be more lenient and only insist on

$$TOL_D * \rho_k \geq (h_k/H_k) \min(\rho_i, \rho_j). \tag{6.68}$$

We summarize:

**Definition 6.26 (Acceptable edge path).** *Let $\Omega_i$ and $\Omega_j$ be two substructures that have an edge $\mathcal{E}^{ij}$ but not a face in common (or a vertex $\mathcal{V}^{ij}$ but not an edge in common). An acceptable edge path between $\Omega_i$ and $\Omega_j$ is a sequence of substructures*

$$\{\Omega_i = \Omega_1, \Omega_2, \ldots, \Omega_k, \ldots, \Omega_M = \Omega_j\},$$

*such that:*

1. *All $\Omega_k$ in the path have the edge $\mathcal{E}^{ij}$ (resp. the vertex $\mathcal{V}^{ij}$) in common.*
2. *For $k = 1, \ldots, M-1$, the substructures $\Omega_k$ and $\Omega_{k+1}$ share a primal edge.*
3. *All $\Omega_k$ in the path satisfy (6.67) for the case of a shared edge (or (6.68) for the case of a shared vertex.)*

We stress the fact that adjacent substructures in an acceptable edge path do not need to share a face.

**Assumption 6.27 (Algorithm D)**

1. *For each face, we need at least one edge, which is part of the boundary of the face, to be designated as primal, i.e., we require that the edge averages $\overline{u}_{\mathcal{E}^{ik}}$ are the same whichever component of $u \in \widetilde{W}_D$ is used to evaluate the average.*
2. *In addition, for all pairs of substructures $\Omega_i$ and $\Omega_j$, which have an edge but not a face in common, or a vertex but not an edge in common, there exists an* acceptable edge path *between the two subdomains.*

We note that we could allow our edge paths to stray somewhat further away from the edge $\mathcal{E}^{ij}$, or the vertex $\mathcal{V}^{il}$. In fact, a careful examination of the proof of Lemma 6.36 would reveal that alternative, more liberal rules concerning the paths could be adopted.

If the coefficients of the equation do not vary significantly, there is always an acceptable edge path and only Assumption 6.27.1 is required. Otherwise, several different strategies can be developed for selecting primal edges and vertices. One strategy would be to start by selecting one primal edge per face and then try to find acceptable edge paths between any pair of subdomains which have an edge but not a face in common and where the edge has not yet been designated as primal. The principal candidates for inclusion in an enlarged set of primal edges are those for which a large $TOL_D$ would be required; we could, e.g., start by choosing those edges associated with the greatest variability of the coefficients as reflected by the potentially largest values of $TOL_D$. Similarly, we could select as primal vertices those with the worst values of $TOL_D$ as in (6.68). We could alternatively begin by first selecting as primal those edges with greatest variability in the coefficients and then, in a second phase, add primal edges to meet the requirement that there is a primal edge which is part of the boundary of each face of $\Gamma$. Whatever strategy is chosen, the selection of an edge for the primal set can depend on the set of edges already selected. Additional strategies are conceivable where edges, previously selected, are deselected and replaced by others. We can also exercise an option of designating some additional vertices of the substructures as primal; this might not be strictly necessary but if constraints are enforced at enough vertices the resulting matrix can be made invertible even without any edge constraints. As previously noted, this can be an advantage in the implementation of the method; cf. Subsect. 6.4.4.

**Algorithm 6.28 (Algorithm D)** *The primal subspace $\widehat{W}_\Pi$, is defined in terms of primal constraints associated with a subset of the edges and vertices of the interface which satisfies Assumption 6.27.*

We note that any choice of primal edges and vertices will be associated with a maximum value of the parameter $TOL_D$, which will appear in the bound of the condition number of the preconditioned operator; see Theorem 6.38. We also again note that in the case of benign coefficients and hexagonal substructures, we can have as few as three edge constraints per subdomain, and hence a very small primal space since each edge is shared by four substructures. We also note that we are free to add any other vertex, edge, or face constraints to our definition of the primal space; the bounds on the condition numbers will only improve. If all edges and vertices are made primal, we have returned to Algorithm C.

Our final algorithm, Algorithm E, can be derived from Algorithm B by eliminating the edge constraints; recall that Algorithm C was derived from Algorithm B by dropping the face constraints.

**Algorithm 6.29 (Algorithm E)** *The primal subspace, $\widehat{W}_\Pi$, is spanned by the vertex nodal finite element basis functions $\theta_{\mathcal{V}^{il}}$ and the cutoff functions $\theta_{\mathcal{F}^{ik}}$ of all the faces of $\Gamma$. The local subspace $\widetilde{W}_{\Delta,i}$ is defined as the subspace of $W_i$ where the values at the subdomain vertices vanish together with the face averages $\overline{u}_{\mathcal{F}^{ik}}$, i.e., by*

$$\widetilde{W}_{\Delta,i} := \{u \in W_i : u(\mathcal{V}^{il}) = 0, \overline{u}_{\mathcal{F}^{ik}} = 0,\ \mathcal{V}^{il}, \mathcal{F}^{ik} \subset \partial\Omega_i\}.$$

*Hence, $\widetilde{W} = \widetilde{W}_E$ is the subspace of $W$ of functions that are continuous at the subdomain vertices and have common averages $\overline{u}_{\mathcal{F}^{ik}}$ for all the faces.*

This algorithm was introduced and analyzed in [293]. In order to succeed with the analysis, a concept of *acceptable face path* was introduced. For each pair of subdomains $\Omega_i$ and $\Omega_j$, which have an edge in common, we must find a path through neighboring subdomains $\Omega_k$, such that

$$TOL_E * \rho_k \geq \min(\rho_i, \rho_j) \tag{6.69}$$

with an acceptable tolerance $TOL_E$. The path will pass from one substructure to another only through faces common to pairs of neighboring substructures. It is easy to select values of the $\rho_i$ for which $TOL_E$ is arbitrarily large; see Subsect. 6.4.5. In contrast to Algorithm D, we are not able to lower the value of $TOL_E$ by introducing additional constraints if we are unwilling to use edge constraints. We will provide experimental evidence in Subsect. 6.4.5 which indeed suggests that face constraints are less powerful than edge constraints for problems with very large jumps in the coefficients across the interface. However, we also note that the condition (6.69) is trivially satisfied with $TOL_E = 1$ for an interior edge which is common to just three

We will not provide an analysis of this algorithm and we instead refer to [293]; see also Remark 6.40.

### 6.4.3 Analysis of Three FETI-DP Algorithms

We recall that there is a central, common part of the analysis of the algorithms of Sect. 6.2 and 6.3 and that technically the proof of the central Lemma 6.3 is similar to work in Chap. 5. Here some additional ideas are required, in particular, in the analysis of Algorithm D, but we will also find that several algebraic arguments have much in common with those used for one-level FETI methods. We note that Assumption 4.3.2 is not necessary for FETI-DP algorithms. As in [341], the two different Schur complements, $\widetilde{S}$ and $S$ will play an important role in the analysis of the dual-primal iterative algorithms. We also recall that $\widetilde{S}$ represents a global problem while $S$ does not.

We recall that $V := range(B_\Delta)$ is the space of Lagrange multipliers, where the range is taken over elements in $\widetilde{W}_\Delta$ or, equivalently, $\widetilde{W}$. As in Subsect. 6.3.3, we introduce a projection

$$P_\Delta := B_{D,\Delta}^T B_\Delta.$$

As already mentioned, we will work with a fully redundant set of Lagrange multipliers. The proof of Lemma 6.16 can be carried over to the current case as well and a simple computation shows that $P_\Delta$ preserves the jump of any function $u_\Delta \in \widetilde{W}_\Delta$, i.e.,

$$B_\Delta P_\Delta u_\Delta = B_\Delta u_\Delta.$$

We also have $P_\Delta u = 0$, if $u \in \widehat{W}$. As before, we introduce an additional related operator $E_\Delta = I - P_\Delta$. We have $E_\Delta u_\Delta \in \widehat{W}$, for $u_\Delta \in \widetilde{W}_\Delta$. We can then borrow formula (6.42); for $x \in \Gamma_h$ and $w_\Delta \in \widetilde{W}_\Delta$, we have

$$(P_\Delta w_\Delta(x))_i \;=\; \sum_{j \in \mathcal{N}_x} \delta_j^\dagger(x)(w_{\Delta,i}(x) - w_{\Delta,j}(x)), \quad x \in \partial\Omega_{i,h}. \tag{6.70}$$

We note again that the coefficients in this expression are constant on the set of the nodal points of each face and edge of $\partial\Omega_i$, and that this formula is independent of the particular choice of $B_\Delta$.

In our proof of the lower bound for the preconditioned operators in Theorems 6.35 and 6.38, we will use the following lemma and corollary; the corollary replaces Lemma 6.8 which was used in the proof of Theorems 6.15 and 6.21.

**Lemma 6.30** *For any $\mu \in V$, there exists a $w_\Delta \in \widetilde{W}_\Delta$, such that $\mu = B_\Delta w_\Delta$ and $E_\Delta w_\Delta \in \widehat{W}_\Pi$. This result holds for all pairs of spaces which define the algorithms of this section.*

*Proof.* We first note that for any $\mu \in V$, there are many elements of $u_\Delta \in \widetilde{W}_\Delta$ such that $\mu = B_\Delta u_\Delta$. Given any such $u_\Delta$, we then write it as

$$u_\Delta = P_\Delta u_\Delta + E_\Delta u_\Delta.$$

$E_\Delta u_\Delta$ is an element in $\widehat{W}$ but not necessarily an element of its subspace $\widehat{W}_\Pi$. If $u_\Delta$ vanishes at a vertex, then formula (6.70) ensures that $P_\Delta u_\Delta$ (and

therefore $E_\Delta u_\Delta$) also will vanish at that vertex. Let now $\mathcal{E}^{ik}$ be a primal edge. Since $u_\Delta \in \widetilde{W}_\Delta$, the mean value of $u_\Delta$ vanishes on $\mathcal{E}^{ik}$ and

$$\overline{(P_\Delta u_\Delta)}_{\mathcal{E}^{ik}} + \overline{(E_\Delta u_\Delta)}_{\mathcal{E}^{ik}} = 0.$$

Similar considerations hold for a primal face. We can then obtain an element with the right properties by replacing $E_\Delta u_\Delta$ by its averages over the primal edges and faces. More precisely, we define $\widehat{w}_\Pi \in \widehat{W}_\Pi$ by

$$\widehat{w}_\Pi = \sum_{\mathcal{E}^{ik}} \overline{(E_\Delta u_\Delta)}_{\mathcal{E}^{ik}} \theta_{\mathcal{E}^{ik}} + \sum_{\mathcal{F}^{ij}} \overline{(E_\Delta u_\Delta)}_{\mathcal{F}^{ij}} \theta_{\mathcal{F}^{ij}},$$

where the sums are only taken over primal edges and faces, respectively. We then set $w_\Delta = P_\Delta u_\Delta + \widehat{w}_\Pi = P_\Delta w_\Delta + \widehat{w}_\Pi$. By construction, $w_\Delta \in \widetilde{W}_\Delta$, and, since $\widehat{w}_\Pi$ is continuous,

$$B_\Delta w_\Delta = B_\Delta P_\Delta u_\Delta = B_\Delta u_\Delta = \mu.$$

The proof is concluded by noting that $E_D w_\Delta = \widehat{w}_\Pi \in \widehat{W}_\Pi$. □

The following corollary will be used directly in the proof of the lower bounds for the FETI-DP operators.

**Corollary 6.31** *Let $w_\Delta$ satisfy the condition of Lemma 6.30. Then,*

$$|w_\Delta|_{\widetilde{S}} \leq |P_\Delta w_\Delta|_S.$$

*Proof.* Since $E_\Delta w_\Delta \in \widehat{W}_\Pi$, the variational formulation of the Schur complement $\widetilde{S}$, given in (6.65), implies that

$$|w_\Delta|^2_{\widetilde{S}} \leq |w_\Delta - E_\Delta w_\Delta|^2_S = |P_\Delta w_\Delta|^2_S.$$

□

*Remark 6.32.* Stronger statements than those in Lemma 6.30 and Corollary 6.31 can be made for Algorithms A and C. Indeed, from formula (6.70), we find that if $u_\Delta \in \widetilde{W}_\Delta$ then $P_\Delta u_\Delta$ vanishes at all substructure vertices, since $u_\Delta$ does. Therefore, $P_\Delta u_\Delta \in \widetilde{W}_\Delta$ for Algorithm A. In addition, for Algorithm C, formula (6.70) ensures that $P_\Delta u_\Delta$ has vanishing averages over all the edges and therefore $P_\Delta u_\Delta \in \widetilde{W}_\Delta$. This shows that, in the proof of Lemma 6.30 the averages $\overline{(E_\Delta u_\Delta)}_{\mathcal{E}^{ik}}$ all vanish and we can therefore choose $w_\Delta = P_\Delta u_\Delta$, which ensures $E_\Delta w_\Delta = 0$. We then have equality in Corollary 6.31. The same results hold for Algorithms B and D, in case the averages over edges and faces in (6.66) are constructed using only the nodal values in the interior of the edges and faces. Or, in other words, the conclusion follows in case each primal constraint employs values at the nodes $x$ where the functions $\delta_i^\dagger(x)$ take the same value; see formula (6.70).

We can now prove the invertibility of the FETI-DP operators and of the preconditioners. The key point here is that there are enough continuity constraints to make a piecewise constant function that vanish on a part of $\partial\Omega$ vanish in the whole $\Omega$.

**Lemma 6.33** *The operators $\widetilde{S}$ and $M^{-1}$ are invertible.*

*Proof.* We first consider $\widetilde{S}$ and an element $w_\Delta \in \widetilde{W}_\Delta$, such that $\widetilde{S}w_\Delta = 0$. If $w \in \widetilde{W}$ is the minimizing element of Lemma 6.22, then $Sw = 0$ and $w$ is therefore equal to a constant $c_i$ on each substructure $\Omega_i$. Clearly, $c_i = 0$ if $\Omega_i$ touches $\partial\Omega_D$, which we have assumed to be nonempty.

For Algorithms A, B, C, and E, where every vertex is primal, we can then consider all the substructures that share a vertex with $\Omega_i$ and, since $w$ must be continuous at the vertices, deduce that $w$ also vanishes in these substructures. Finally, since $\Omega$ is connected, we can visit all the substructures moving through vertices and deduce that $w = 0$ and therefore $w_\Delta = 0$.

For Algorithm D, it is enough to consider primal edges. Because of Assumption 6.27.1, for every substructure that shares a face with $\Omega_i$, $w$ has the same average over the corresponding primal edge and must therefore vanish on this substructure. We can now visit all the substructures moving through common faces and deduce that $w$ and therefore $w_\Delta$ must vanish everywhere.

We now consider a $\mu \in V$, such that $M^{-1}\mu = 0$, and the $w_\Delta \in \widetilde{W}_\Delta$ of Lemma 6.30. Then,

$$0 = \langle M^{-1}\mu, \mu\rangle = \langle M^{-1}B_\Delta w_\Delta, B_\Delta w_\Delta\rangle = \langle SP_\Delta w_\Delta, P_\Delta w_\Delta\rangle.$$

This implies $S(P_\Delta w_\Delta) = 0$ and therefore

$$w_\Delta = P_\Delta w_\Delta + E_\Delta w_\Delta = c + \widehat{w}_\Pi,$$

with $c$ equal to a constant $c_i$ on each substructure $\Omega_i$ and a continuous $\widehat{w}_\Pi \in \widehat{W}_\Pi$. We now consider Algorithms A, B, C, and E, where each vertex is primal. We note that $w_\Delta$ vanishes at the substructure vertices and, since $\widehat{w}_\Pi$ is continuous, $c$ must also be continuous at the vertices. The same argument as before then ensures that $c = 0$ and thus

$$\mu = B_\Delta w_\Delta = B_\Delta P_\Delta w_\Delta = B_\Delta c = 0.$$

A similar argument, that employs the primal edges of Assumption 6.27.1, provides the result for Algorithms D. □

We note that for Algorithm D, edge constraints are enough in order to ensure invertibility. However, we have already noted that it can be computationally advantageous in practice to add vertex constraints in a such a way that invertibility is guaranteed only by these constraints; cf. Sect. 6.4.4.

We continue by analyzing Algorithm B and begin by proving the following core estimate.

**Lemma 6.34 (Algorithm B)** *For all $w \in \widetilde{W}_B$, we have*

$$|P_\Delta w|_S^2 \leq C\,(1+\log(H/h))^2 |w|_S,$$

*where $C > 0$ is independent of $h, H$, the $\rho_i$, and $\gamma$.*

*Proof.* We use formula (6.70) and set

$$v_i(x) := (P_\Delta w(x))_i = \sum_{j \in \mathcal{N}_x} \delta_j^\dagger(x)(w_i(x) - w_j(x)), \quad x \in \partial\Omega_{i,h}. \tag{6.71}$$

We then have to estimate

$$|P_\Delta w|_S^2 = \sum_{i=1}^N |v_i|_{S^{(i)}}^2.$$

We can therefore focus on the estimate of the contribution from a single subdomain $\Omega_i$.

We model our proof on Lemma 6.3 and [289, Lemmas 4.7, 5.4] but note that the arguments need to be modified to some extent. Since the coefficients in the sum in (6.71) are constant on each face and edge of $\Omega_i$, we decompose $v_i$ into terms associated with single faces and edges. We note that we only have contributions from faces and edges since all elements in $\widetilde{W}_B$ are continuous at the vertices. Here, in contrast to the proof Theorem 6.15, we do not need to use Assumption 6.1.

We now cut the function $v_i$ using the functions $\theta_\mathcal{F}$ and $\theta_\mathcal{E}$ and write it as a sum of terms which vanish at all the interface nodes outside individual faces and edges; cf., e.g., [178, 184, 182]. We then have, since the $v_i$ vanish at the subdomain vertices,

$$v_i = \sum_{\mathcal{F} \subset \partial\Omega_i} I^h(\theta_\mathcal{F} v_i) + \sum_{\mathcal{E} \subset \partial\Omega_i} I^h(\theta_\mathcal{E} v_i).$$

**Face Terms.** We find that the face $\mathcal{F}$, shared by the substructures $\Omega_i$ and $\Omega_j$, contributes

$$I^h(\theta_\mathcal{F} \delta_j^\dagger (w_i - w_j)),$$

and we have to estimate its $H_{00}^{1/2}(\mathcal{F})$-norm.

We note that $\delta_j^\dagger$ is constant on $\mathcal{F}_h$ and that $w$ has common face averages, i.e., $\overline{w}_{i,\mathcal{F}} = \overline{w}_{j,\mathcal{F}}$. Using inequality (6.19), these observations, and Lemma 4.26, we obtain,

$$\begin{aligned}
&\rho_i \|I^h(\theta_\mathcal{F}\delta_j^\dagger(w_i - w_j))\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&= \rho_i \|I^h(\theta_\mathcal{F}\delta_j^\dagger((w_i - \overline{w}_{i,\mathcal{F}}) - (w_j - \overline{w}_{j,\mathcal{F}})))\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&\leq C\,(1+\log(H_i/h_i))^2 \min(\rho_i,\rho_j)\left(|w_i|^2_{H^{1/2}(\mathcal{F})} + |w_j|^2_{H^{1/2}(\mathcal{F})}\right) \\
&\leq C\,(1+\log(H_i/h_i))^2 \left(\rho_i |w_i|^2_{H^{1/2}(\partial\Omega_i)} + \rho_j |w_j|^2_{H^{1/2}(\partial\Omega_j)}\right).
\end{aligned} \tag{6.72}$$

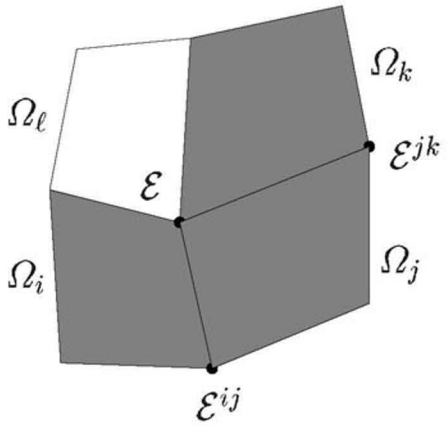


**Fig. 6.8.** Section of a configuration with an edge $\mathcal{E}$ shared by four substructures. An acceptable edge path $\{\Omega_i, \Omega_j, \Omega_k\}$, which employs the primal edges $\mathcal{E}^{ij}$ and $\mathcal{E}^{jk}$ is also shown for the substructures $\Omega_i$ and $\Omega_k$.

We note that, by assumption, $H_j$ and $H_i$ are comparable and so are $h_j$ and $h_i$, since the triangulations of $\Omega_i$ and $\Omega_j$ are locally quasi uniform and match across the interface $\Gamma$. By using Lemmas 4.9 and 4.10, we have therefore proven

$$|I^h(\theta_{\mathcal{F}} v_i)|^2_{S^{(i)}} \leq C(1+\log(H/h))^2(|w_i|^2_{S^{(i)}} + |w_j|^2_{S^{(j)}}). \tag{6.73}$$

**Edge Terms.** We next consider the edge contributions and only consider, for simplicity, the case of an edge $\mathcal{E}$ common to four subdomains $\Omega_i, \Omega_j, \Omega_k$, and $\Omega_\ell$; see Fig. 6.8. We also recall that all the edges are primal and therefore $w_i$, $w_j$, $w_k$, and $w_\ell$ have all the same average over $\mathcal{E}$. There are three terms in the estimate of the contribution of $\Omega_i$ to $|P_\Delta w|^2_S$, namely

$$\begin{aligned} &\rho_i \, \|I^h(\delta_j^\dagger \, \theta_{\mathcal{E}}(w_i - w_j))\|^2_{H^{1/2}(\partial\Omega_i)} \\ &+ \rho_i \, \|I^h(\delta_k^\dagger \, \theta_{\mathcal{E}}(w_i - w_k))\|^2_{H^{1/2}(\partial\Omega_i)} \\ &+ + \rho_i \, \|I^h(\delta_\ell^\dagger \, \theta_{\mathcal{E}}(w_i - w_\ell))\|^2_{H^{1/2}(\partial\Omega_i)}. \end{aligned} \tag{6.74}$$

We recall that $\delta_j^\dagger$, $\delta_k^\dagger$, and $\delta_\ell^\dagger$ are constant on $\mathcal{E}_h$ and that the four averages $\overline{w}_{i,\mathcal{E}}$, $\overline{w}_{k,\mathcal{E}}$, $\overline{w}_{j,\mathcal{E}}$, and $\overline{w}_{\ell,\mathcal{E}}$ are all equal.

We first consider the second term in detail assuming that $\Omega_i$ shares a face with each of $\Omega_j$ and $\Omega_\ell$, but only an edge with $\Omega_k$. Using formula (6.19), Lemmas 4.19 and 4.17, we find

$$
\begin{aligned}
&\rho_i \, \|I^h(\delta_k^\dagger \theta_{\mathcal{E}}(w_i - w_k))\|^2_{H^{1/2}(\partial\Omega_i)} \\
&\leq C\rho_i \|I^h(\delta_k^\dagger(\theta_{\mathcal{E}}(w_i - \overline{w}_{i,\mathcal{E}}) - \theta_{\mathcal{E}}(w_k - \overline{w}_{k,\mathcal{E}})))\|^2_{L^2(\mathcal{E})} \\
&\leq 2\left(\rho_i \|I^h(\theta_{\mathcal{E}}(w_i - \overline{w}_{i,\mathcal{E}}))\|^2_{L^2(\mathcal{E})} + \rho_k \|I^h(\theta_{\mathcal{E}}(w_k - \overline{w}_{k,\mathcal{E}}))\|^2_{L^2(\mathcal{E})}\right) \\
&\leq C\left(\rho_i \|w_i - \overline{w}_{i,\mathcal{E}}\|^2_{L^2(\mathcal{E})} + \rho_k \|w_k - \overline{w}_{k,\mathcal{E}}\|^2_{L^2(\mathcal{E})}\right) \\
&\leq C(1 + \log(H/h))\left(\rho_i |w_i|^2_{H^{1/2}(\mathcal{F}^i)} + \rho_k |w_k|^2_{H^{1/2}(\mathcal{F}^k)}\right) \\
&\leq C(1 + \log(H/h))\left(|w_i|^2_{S^{(i)}} + |w_k|^2_{S^{(k)}}\right),
\end{aligned}
\tag{6.75}
$$

with $\mathcal{F}^i$ a face of $\Omega_i$ and $\mathcal{F}^k$ a face of $\Omega_k$, which have the edge $\mathcal{E}$ in common. We note that since $\Omega_i$ and $\Omega_j$, as well as $\Omega_i$ and $\Omega_\ell$, have a face in common, the argument given above could be modified for the first and third edge contributions; they can be reduced to estimates for face terms directly. We have therefore proven

$$
|I^h(\theta_{\mathcal{E}} v_i)|^2_{S^{(i)}} \leq C(1 + \log(H/h)) \sum_k |w_k|^2_{S^{(k)}}, \tag{6.76}
$$

where the sum is taken over all the substructures that share the edge $\mathcal{E}$.

□

We now prove our condition number estimate for Algorithm B, which only depends polylogarithmically on the dimension of the subproblems.

**Theorem 6.35 (Algorithm B)** *The preconditioner $M_B$ satisfies*

$$
\langle M_B \lambda, \lambda \rangle \leq \langle F_B \lambda, \lambda \rangle \leq C(1 + \log(H/h))^2 \langle M_B \lambda, \lambda \rangle, \quad \lambda \in V. \tag{6.77}
$$

*Here $C$ is independent of $h, H, \gamma$, and the values of the $\rho_i$.*

*Proof.* We will estimate the smallest eigenvalue $\lambda_{min}(M_B^{-1} F_B)$ from below and the largest eigenvalue $\lambda_{max}(M_B^{-1} F_B)$ from above.

**Lower Bound:**

As in the analysis of the one-level FETI methods, see the proof of Theorem 6.15, we can use the following formula, which follows from the definition of $F$:

$$
\langle F_B \lambda, \lambda \rangle = \sup_{0 \neq w_\Delta \in \widetilde{W}_\Delta} \frac{\langle \lambda, B_\Delta w_\Delta \rangle^2}{|w_\Delta|^2_{\widetilde{S}}}. \tag{6.78}
$$

Let $\mu \in V$ be arbitrary. It then follows from Corollary 6.31 that there exists a $w_\Delta \in \widetilde{W}_\Delta$ with $\mu = B_\Delta w_\Delta$ and such that $|w_\Delta|_{\widetilde{S}} \leq |P_\Delta w_\Delta|_S$. We obtain,

$$
\langle F_B \lambda, \lambda \rangle \geq \frac{\langle \lambda, B_\Delta w_\Delta \rangle^2}{|w_\Delta|^2_{\widetilde{S}}} \geq \frac{\langle \lambda, B_\Delta w_\Delta \rangle^2}{|P_\Delta w_\Delta|^2_S} = \frac{\langle \lambda, \mu \rangle^2}{|B^T_{D,\Delta} \mu|^2_S} = \frac{\langle \lambda, \mu \rangle^2}{\langle M_B^{-1} \mu, \mu \rangle}.
$$

The left inequality of (6.77) follows by choosing $\mu := M_B \lambda$.

**Upper Bound:**

We first note that the supremum in formula (6.78) can be replaced by a supremum over the space $\widetilde{W}$. Indeed, to every $w_\Delta \in \widetilde{W}_\Delta$ we associate the minimizing element $w \in \widetilde{W}$ of Lemma 6.22, such that,

$$|w_\Delta|_{\widetilde{S}} = |w|_S.$$

In addition, since $w_\Delta$ and $w$ differ by a continuous element, we have $B_\Delta w_\Delta = B_\Delta w$. Using then Lemma 6.34, we obtain for $\lambda$ in $V$,

$$\begin{aligned}
\langle F_B\lambda,\lambda\rangle &= \sup_{0\neq w\in\widetilde{W}} \frac{\langle \lambda, B_\Delta w\rangle^2}{|w|_S^2} \\
&\leq C\,(1+\log(H/h))^2 \sup_{w\in\widetilde{W}} \frac{\langle \lambda, B_\Delta w\rangle^2}{|P_\Delta w|_S^2} \\
&= C\,(1+\log(H/h))^2 \sup_{w\in\widetilde{W}} \frac{\langle \lambda, B_\Delta w\rangle^2}{\langle M_B^{-1}B_\Delta w, B_\Delta w\rangle} \\
&= C\,(1+\log(H/h))^2 \sup_{\mu\in V} \frac{\langle \lambda,\mu\rangle^2}{\langle M_B^{-1}\mu,\mu\rangle} = C\,(1+\log(H/h))^2\langle M_B\lambda,\lambda\rangle.
\end{aligned}$$

□

We now turn to the analysis of Algorithms C and D.

**Lemma 6.36 (Algorithms C, D)** *For all $w \in \widetilde{W}_C$, we have,*

$$|P_\Delta w|_S^2 \leq C\,(1+\log(H/h))^2|w|_S^2.$$

*For all $w \in \widetilde{W}_D$, we have,*

$$|P_\Delta w|_S^2 \leq C\,\max(1,TOL_D)\,(1+\log(H/h))^2|w|_S^2.$$

*In both cases, $C > 0$ is independent of $h, H$, the $\rho_i$, and $\gamma$.*

*Proof.* We can proceed as in the proof of Lemma 6.34; we will use the same notation and only discuss details that are technically different. We note that in Algorithm D not all vertices are necessarily primal and that we therefore have to estimate terms of $P_\Delta w(x)$ related to the vertices which are not primal.

We cut the function $v_i$ using the functions $\theta_\mathcal{F}$, $\theta_\mathcal{E}$, and $\theta_\mathcal{V}$ and write it as a sum of terms which vanish at all the interface nodes outside individual faces, edges, and vertices, respectively. We then have

$$v_i = \sum_{\mathcal{F}\subset\partial\Omega_i} I^h(\theta_\mathcal{F} v_i) + \sum_{\mathcal{E}\subset\partial\Omega_i} I^h(\theta_\mathcal{E} v_i) + \sum_{\mathcal{V}\in\partial\Omega_i} \theta_\mathcal{V} v_i(\mathcal{V}),$$

where the last sum is taken over all the vertices that are not primal.

**Face Terms.** As in the proof of Lemma 6.34, we find that the face $\mathcal{F}$, shared by $\Omega_i$ and $\Omega_j$, contributes

$$I^h(\theta_{\mathcal{F}}\delta_j^\dagger(w_i - w_j)),$$

and we have to estimate its $H_{00}^{1/2}(\mathcal{F})$-norm. We recall that $\delta_j^\dagger$ is constant on $\mathcal{F}_h$ but that $w$ in general has different face averages $\overline{w}_{i,\mathcal{F}}$ and $\overline{w}_{j,\mathcal{F}}$, since we now do not work with primal faces. Using inequality (6.19), we obtain,

$$\begin{aligned}
&\rho_i \|I^h(\theta_{\mathcal{F}}\delta_j^\dagger(w_i - w_j))\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&= \rho_i \|I^h(\theta_{\mathcal{F}}\delta_j^\dagger((w_i - \overline{w}_{i,\mathcal{F}}) - (w_j - \overline{w}_{j,\mathcal{F}}) + (\overline{w}_{i,\mathcal{F}} - \overline{w}_{j,\mathcal{F}})))\|^2_{H_{00}^{1/2}(\mathcal{F})} \\
&\leq 2\min(\rho_i,\rho_j)\left( \|I^h(\theta_{\mathcal{F}}((w_i - \overline{w}_{i,\mathcal{F}}) - (w_j - \overline{w}_{j,\mathcal{F}})))\|^2_{H_{00}^{1/2}(\mathcal{F})} + \right. \\
&\qquad \left. + \|(\overline{w}_{i,\mathcal{F}} - \overline{w}_{j,\mathcal{F}})\theta_{\mathcal{F}}\|^2_{H_{00}^{1/2}(\mathcal{F})} \right).
\end{aligned} \tag{6.79}$$

The first term on the right hand side can be estimated as in (6.72) by

$$C\,(1 + \log(H/h))^2 \left(\rho_i |w_i|^2_{H^{1/2}(\mathcal{F})} + \rho_j |w_j|^2_{H^{1/2}(\mathcal{F})}\right),$$

as desired, by applying Lemma 4.26.

There remains to estimate $\|(\overline{w}_{i,\mathcal{F}} - \overline{w}_{j,\mathcal{F}})\theta_{\mathcal{F}}\|^2_{H_{00}^{1/2}(\mathcal{F})}$. We recall that for Algorithm C every edge is primal (cf. Algorithm 6.25) and that for Algorithm D every face has at least one primal edge that belongs to its boundary (cf. Assumption 6.27.1). Let $\mathcal{E} \subset \partial\mathcal{F}$ be a designated, primal edge. Since $\overline{w}_{i,\mathcal{E}} = \overline{w}_{j,\mathcal{E}}$, we then have

$$|\overline{w}_{i,\mathcal{F}} - \overline{w}_{j,\mathcal{F}}|^2 \leq 2\left(|\overline{w}_{i,\mathcal{E}} - \overline{w}_{i,\mathcal{F}}|^2 + |\overline{w}_{j,\mathcal{E}} - \overline{w}_{j,\mathcal{F}}|^2\right).$$

It is sufficient to consider the first term on the right hand side since the second can be dealt with in exactly the same way. Using Lemma 4.30, we find

$$|\overline{w}_{i,\mathcal{E}} - \overline{w}_{i,\mathcal{F}}|^2 = |\overline{(w - \overline{w}_{i,\mathcal{F}})}_{i,\mathcal{E}}|^2 \leq C/H_i \|w_i - \overline{w}_{i,\mathcal{F}}\|^2_{L^2(\mathcal{E})},$$

and, by using Lemma 4.17 and the Poincaré inequality in Lemma A.17,

$$|\overline{w}_{i,\mathcal{E}} - \overline{w}_{i,\mathcal{F}}|^2 \leq C/H_i(1 + \log(H/h))|w_i - \overline{w}_{i,\mathcal{F}}|^2_{H^{1/2}(\mathcal{F})}. \tag{6.80}$$

Combining the last inequality with the bound for $\theta_{\mathcal{F}}$ in Lemma 4.26 yields

$$\|(\overline{w}_{i,\mathcal{F}} - \overline{w}_{j,\mathcal{F}})\theta_{\mathcal{F}}\|^2_{H_{00}^{1/2}(\mathcal{F})} \leq C\,(1 + \log(H/h))^2 \left(|w_i|^2_{H^{1/2}(\mathcal{F})} + |w_j|^2_{H^{1/2}(\mathcal{F})}\right).$$

The remainder of the proof of the result for Algorithm C, involving bounds for edge contributions on primal edges, can be carried out as in the proof of Lemma 6.34; cf. (6.75).

However, for Algorithm D, we need to do some further work since some edges or vertices may not be primal.

**Edge Terms.** Proceeding as in the proof of Lemma 6.34, we can estimate the contributions of the edges of $\Omega_i$ to the energy of $v_i$ in terms of $L^2$-norms over the edges. We consider in particular the configuration in Fig. 6.8, where four substructures $\Omega_i$, $\Omega_j$, $\Omega_k$, and $\Omega_\ell$ share an edge $\mathcal{E}$ and $\Omega_i$ shares a face with each of $\Omega_j$ and $\Omega_\ell$, but only an edge with $\Omega_k$. We need to estimate the three contributions in (6.74). We first consider the second term in detail. If we have a trivial, acceptable edge path, i.e., the common edge is designated as primal, we can proceed exactly as in (6.75). Otherwise assume that we have the non-trivial, acceptable edge path $\{\Omega_i, \Omega_j, \Omega_k\}$, via the edges $\mathcal{E}^{ij}$ and $\mathcal{E}^{jk}$; in general the acceptable edge path could be more complicated but such a case could be analyzed similarly. We now note, in particular, that the definition of an acceptable edge path ensures that

$$\overline{w}_{i,\mathcal{E}^{ij}} = \overline{w}_{j,\mathcal{E}^{ij}}, \quad \overline{w}_{j,\mathcal{E}^{jk}} = \overline{w}_{k,\mathcal{E}^{jk}}$$

Using Lemma 4.19, the second term in (6.74) can be estimated by

$$\begin{aligned}
&\rho_i \left\| \delta_k^\dagger I^h(\theta_{\mathcal{E}}(w_i - w_k)) \right\|^2_{L^2(\mathcal{E})} \\
&= \rho_i \left\| \delta_k^\dagger \left( I^h(\theta_{\mathcal{E}}(w_i - \overline{w}_{i,\mathcal{E}^{ij}})) + \theta_{\mathcal{E}}(\overline{w}_{j,\mathcal{E}^{ij}} - \overline{w}_{j,\mathcal{E}^{jk}}) \right.\right. \\
&\qquad \left.\left. - I^h(\theta_{\mathcal{E}}(w_k - \overline{w}_{k,\mathcal{E}^{jk}})) \right) \right\|^2_{L^2(\mathcal{E})} \qquad (6.81) \\
&\leq C \min(\rho_i, \rho_k) \left( \left\| I^h(\theta_{\mathcal{E}}(w_i - \overline{w}_{i,\mathcal{E}^{ij}})) \right\|^2_{L^2(\mathcal{E})} + H_j \left| \overline{w}_{j,\mathcal{E}^{ij}} - \overline{w}_{j,\mathcal{E}^{jk}} \right|^2 \right. \\
&\qquad \left. + \left\| I^h(\theta_{\mathcal{E}}(w_k - \overline{w}_{k,\mathcal{E}^{jk}})) \right\|^2_{L^2(\mathcal{E})} \right),
\end{aligned}$$

where we used the fact that $0 \leq \theta_{\mathcal{E}}(x) \leq 1$ on $\mathcal{E}$. The first and third terms of the last expression can be estimated as before in (6.75). For the second, we have to use $TOL_D * \rho_j \geq \min(\rho_i, \rho_k)$. We note that the edges $\mathcal{E}^{ij}$ and $\mathcal{E}^{jk}$, which belong to $\partial\Omega_j$, may belong to different faces. We employ the average of $w_j$ over $\mathcal{W}^j$ and write

$$\left| \overline{w}_{j,\mathcal{E}^{ij}} - \overline{w}_{j,\mathcal{E}^{jk}} \right|^2 \leq 2 \left| \overline{w}_{j,\mathcal{E}^{ij}} - \overline{w}_{j,\mathcal{W}^j} \right|^2 + 2 \left| \overline{w}_{j,\mathcal{E}^{jk}} - \overline{w}_{j,\mathcal{W}^j} \right|^2.$$

Each of the two terms can then be estimated as in (6.80), using Lemma 4.17. For the first, we find:

$$\begin{aligned}
\left| \overline{w}_{j,\mathcal{E}^{ij}} - \overline{w}_{j,\mathcal{W}^j} \right|^2 &= \left| \overline{(w - \overline{w}_{j,\mathcal{W}^j})}_{j,\mathcal{E}^{ij}} \right|^2 \\
&\leq C/H_j \| w_j - \overline{w}_{j,\mathcal{W}^j} \|^2_{L^2(\mathcal{E}^{ij})} \\
&\leq C/H_j \| w_j - \overline{w}_{j,\mathcal{W}^j} \|^2_{L^2(\mathcal{W}^j)} \\
&\leq C/H_j (1 + \log(H/h)) |w_j|^2_{H^{1/2}(\partial\Omega_j)}.
\end{aligned}$$

We finally obtain,

$$\rho_i \|\delta_k^\dagger I^h(\theta_{\mathcal{E}}(w_i - w_k))\|^2_{L^2(\mathcal{E})} \leq C(1+\log(H/h)) \left(\rho_i |w_i|^2_{H^{1/2}(\partial\Omega_i)} + \right.$$
$$\left. + \rho_k |w_k|^2_{H^{1/2}(\partial\Omega_k)} + TOL_D * \rho_j |w_j|^2_{H^{1/2}(\partial\Omega_j)}\right).$$

Since $\Omega_i$ and $\Omega_j$, as well as $\Omega_i$ and $\Omega_\ell$, have a face in common, the argument given above could be simplified for the first and third edge contributions, see (6.74); they can be reduced to estimates of face terms.

**Vertex Terms.** Finally, we consider the terms resulting from the vertices. Let $\mathcal{V} \in \partial\Omega_i$ be a vertex that is not primal. We have, according to (6.71) and (6.19),

$$\begin{aligned}\rho_i |\theta_{\mathcal{V}} v_i(\mathcal{V})|^2_{H^{1/2}(\partial\Omega_i)} &\leq C \sum_{\ell \in \mathcal{N}_{\mathcal{V}}} \rho_i \, \delta_\ell^\dagger(\mathcal{V})^2 \, |\theta_{\mathcal{V}}|^2_{H^{1/2}(\partial\Omega_i)} \, |w_i(\mathcal{V}) - w_\ell(\mathcal{V})|^2 \\ &\leq C \sum_{\ell \in \mathcal{N}_{\mathcal{V}}} \min(\rho_i, \rho_\ell) h_i |w_i(\mathcal{V}) - w_\ell(\mathcal{V})|^2,\end{aligned}$$

where we have used the bound $|\theta_{\mathcal{V}}|^2_{H^{1/2}(\partial\Omega_i)} \leq C h_i$ which can be obtained from an inverse inequality. We now consider each term in the sum separately. Let $\Omega_\ell$ share the vertex $\mathcal{V}$ with $\Omega_i$. and assume that we have, an acceptable edge path $\{\Omega_i, \Omega_j, \Omega_\ell\}$ via the edges $\mathcal{E}^{ij}$ and $\mathcal{E}^{j\ell}$ meeting the condition

$$TOL_D * \rho_j \geq \frac{h_j}{H_j} \min(\rho_i, \rho_\ell). \tag{6.82}$$

We can proceed as in the analysis of the edge terms and obtain

$$\begin{aligned}&\min(\rho_i, \rho_\ell) h_i |w_i(\mathcal{V}) - w_\ell(\mathcal{V})|^2 \\ &\leq 3 \min(\rho_i, \rho_\ell) h_i \left(|w_i(\mathcal{V}) - \overline{w}_{i,\mathcal{E}^{ij}}|^2 + |\overline{w}_{j,\mathcal{E}^{ij}} - \overline{w}_{j,\mathcal{E}^{j\ell}}|^2 + |w_\ell(\mathcal{V}) - \overline{w}_{\ell,\mathcal{E}^{j\ell}}|^2\right).\end{aligned}$$

It is sufficient to estimate the first term on the last line; the third term can be treated in exactly the same way, and the second term can be estimated as above with the only difference being an additional factor $h_j/H_j$ which is accounted for in (6.82). Using Lemmas B.5 and Lemma 4.17, and estimating $|\overline{w}_{i,\mathcal{E}^{ij}}|$ as before, we obtain

$$\begin{aligned}h_i |(w_i(\mathcal{V}) - \overline{w}_{i,\mathcal{E}^{ij}}|^2 = h_i |(w_i - \overline{w}_{i,\mathcal{E}^{ij}})(\mathcal{V})|^2 &\leq \|w_i - \overline{w}_{i,\mathcal{E}^{ij}}\|^2_{L^2(\mathcal{E}^{ij})} \\ &\leq C(1+\log(H_i/h_i)) |w_i|^2_{H^{1/2}(\partial\Omega_i)}.\end{aligned}$$

Using (6.82), we finally obtain

$$\begin{aligned}\min(\rho_i, \rho_\ell) h_i |w_i(\mathcal{V}) - w_\ell(\mathcal{V})|^2 \leq C(1+\log(H/h)) &\left(\rho_i |w_i|^2_{H^{1/2}(\partial\Omega_i)} + \right. \\ &\left. \rho_\ell |w_\ell|^2_{H^{1/2}(\partial\Omega_l)} + TOL_D * \rho_j |w_j|^2_{H^{1/2}(\partial\Omega_j)}\right).\end{aligned}$$

□

*Remark 6.37.* As already noted in Sect. 6.4.2, when we defined Algorithm D, a closer look at the previous proof shows that more general paths can be considered than those in Definition 6.26. However, we note that if there are many long paths, through many subdomains, the constant in the estimate of the norm of $P_\Delta u$ will necessarily be large since there can be many contributions of the form $\rho_j |w_j|^2_{H^{1/2}(\partial\Omega_j)}$ from the same subdomain. In addition, we note that if the coefficient $\rho$ varies moderately, conditions (6.67) and (6.68) are always satisfied and only Assumption 6.27.1 is required.

We can now prove our condition number estimates for Algorithms C and D, which are as strong as those of Theorem 6.35. The proof can be carried out exactly as for Theorem 6.35, using Lemma 6.36 instead of Lemma 6.34.

**Theorem 6.38 (Algorithms C, D)** *The preconditioner $M_C$ satisfies*

$$\langle M_C\lambda,\lambda\rangle \leq \langle F_C\lambda,\lambda\rangle \leq C(1+\log(H/h))^2\langle M_C\lambda,\lambda\rangle, \quad \lambda\in V.$$

*Similarly, the preconditioner $M_D$ satisfies*

$$\langle M_D\lambda,\lambda\rangle \leq \langle F_D\lambda,\lambda\rangle \leq C\,\max(1,TOL_D)\,(1+\log(H/h))^2\langle M_D\lambda,\lambda\rangle, \quad \lambda\in V.$$

*Here $C$ is independent of $h, H, \gamma$, and the values of the $\rho_i$.*

*Remark 6.39 (Algorithm A).* By using similar tools, in particular, Lemma 4.12, we can prove the weaker bound

$$\langle M_A\lambda,\lambda\rangle \leq \langle F_A\lambda,\lambda\rangle \leq C(H/h)(1+\log(H/h))^2\langle M_A\lambda,\lambda\rangle, \quad \lambda\in V.$$

The following result is established in [293].

*Remark 6.40 (Algorithm E).* By using similar tools, we can also prove, for $\lambda \in V$,

$$\langle F_E\lambda,\lambda\rangle \geq \langle M_E\lambda,\lambda\rangle$$
$$\langle F_E\lambda,\lambda\rangle \leq C\max((1+\log(H/h))^2, TOL_E*(1+\log(H/h)))\langle M_E\lambda,\lambda\rangle,$$

where the tolerance is that of (6.69).

### 6.4.4 Implementation of FETI-DP Methods

We review an approach given in Farhat, Lesoinne, and Pierson in [195]. We then assume that we have chosen a sufficient number of primal vertices so that the stiffness matrix, which results from $A$ by partial assembly at the primal vertices, is invertible even without any additional primal constraints. An algorithm for selecting such a set of vertex constraints is given in Lesoinne [313].

We recall that in two dimensions, such a set of vertex constraints is sufficient to obtain a fast and scalable algorithm but that in three dimensions,

we should choose a primal space, which also ensures that certain edge and/or face averages are the same across the interface.

The approach is implemented by introducing additional optional Lagrange multipliers directly related to a set of constraints of the form

$$Q_\Delta B_\Delta u_\Delta = \sum_{i=1}^{N} Q_\Delta B_\Delta^{(i)} u_\Delta^{(i)} = 0. \tag{6.83}$$

Here $Q_\Delta$ is a rectangular matrix with one row for each optional constraint; each row provides a linear combination of the rows of $B_\Delta u_\Delta$ with the weights chosen so as to assure that an edge or face average will have common values across the interface. Thus, if (6.83) is satisfied, all the primal edge and/or face conditions will be satisfied. These optional constraints are fully enforced in every step of the iteration by a separate set of Lagrange multipliers.

We now assume that the unknowns are ordered such that the interior and dual variables come first, grouped together in blocks with respect to the subdomains and equipped with the subscript $r$, and that the primal vertices, equipped with the subscript $c$, are ordered last. We note that the matrix $\widetilde{A}$ is partially assembled with respect to the primal vertices; recall that we introduced a matrix $\widetilde{A}$ already in Subsect. 6.4.1 in the simpler case when there are no optional constraints. We now have

$$\widetilde{A} := \begin{bmatrix} A_{rr} & \widetilde{A}_{cr}^T & Q_r^T \\ \widetilde{A}_{cr} & \widetilde{A}_{cc} & O \\ Q_r & O & O \end{bmatrix},$$

where

$$A_{rr} := \begin{bmatrix} A_{rr}^{(1)} & & O \\ & \ddots & \\ O & & A_{rr}^{(N)} \end{bmatrix}, \quad A_{rr}^{(i)} := \begin{bmatrix} A_{II}^{(i)} & A_{\Delta I}^{(i)T} \\ A_{\Delta I}^{(i)} & A_{\Delta\Delta}^{(i)} \end{bmatrix}, \quad \widetilde{A}_{cr}^T := \begin{bmatrix} A_{cr}^{(1)T} R_c^{(1)} \\ \vdots \\ A_{cr}^{(N)T} R_c^{(N)} \end{bmatrix},$$

$$\widetilde{A}_{cc} = \sum_{i=1}^{N} {R_c^{(i)}}^T A_{cc}^{(i)} R_c^{(i)}, \quad Q_r := [Q_r^{(1)} \dots Q_r^{(N)}], \text{ and } Q_r^{(i)} := [O \;\; Q_\Delta B_\Delta^{(i)}].$$

Here, the ${R_c^{(i)}}^T$ are the matrices which perform the partial assembly at the relevant primal vertices and $\widetilde{A}_{cc}$ is the submatrix which is assembled at the primal vertices. The resulting leading two by two block of $\widetilde{A}$ is non singular since, by assumption, we have chosen a large enough set of primal vertices. Using the notation

$$B_r := [O \;\; B_\Delta]$$

and reintroducing the Lagrange multipliers $\lambda \in range(B_\Delta)$, we can reformulate the original finite element problem as follows

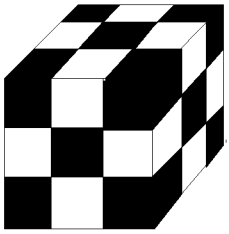

**Fig. 6.9.** Model domain decomposed into cubes with discontinuous diffusion coefficients $\rho_i = 1$ and $10^4$.

$$\begin{bmatrix} A_{rr} & \widetilde{A}_{cr}^T & Q_r^T & B_r^T \\ \widetilde{A}_{cr} & \widetilde{A}_{cc} & O & O \\ Q_r & O & O & O \\ B_r & O & O & O \end{bmatrix} \begin{bmatrix} u_r \\ u_c \\ \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} f_r \\ f_c \\ 0 \\ 0 \end{bmatrix} \tag{6.84}$$

The Schur complement $\widetilde{S}$ is obtained as the leading block of the two by two matrix that remains after that $u_I$, $u_c$, and $\mu$ have been eliminated. When computing $\widetilde{S}^{-1} f_\Delta$, we should take advantage of ordering given in (6.84) and eliminate all of $u_r$ prior to assembling and factoring a Schur complement with respect to $u_c$ and $\mu$; see (6.63) and the discussion in Sect. 4.3. This will allow us to carry out the matrix vector multiplication of $F = B_\Delta \widetilde{S}^{-1} B_\Delta^T$ with any given vector at a reasonable cost.

We note that alternative methods, without any additional Lagrange multipliers, are discussed in Klawonn and Widlund [290].

### 6.4.5 Computational Results

This subsection is based on a conference paper [286] and on experimental work by and a PETSc code developed by Oliver Rheinbach; for information on the PETSc system, see [32]. We have applied the FETI-DP algorithms A, B, C, and E to the model problem (4.3), where $\Omega := (0,1)^3$ is the unit cube. We decompose the unit cube into $N \times N \times N$ cubic subdomains with sidelength $H := 1/N$. The diffusion coefficients $\rho_i$ alternate between 1 and $10^4$ and are distributed in a three-dimensional checkerboard pattern; cf. Fig. 6.9. On the front, left, and bottom part, homogeneous Dirichlet boundary conditions are applied while on all the remaining parts of the boundary, homogeneous Neumann boundary conditions are used. The coefficients are constant on each subdomain and (4.3) is discretized by conforming trilinear elements with cubic finite elements with edge length $h$. We use the preconditioned conjugate gradient method with a zero initial guess. The stopping criterion is the relative reduction of the initial residual by $10^{-7}$ in the Euclidean norm. In order to analyze the numerical scalability of our algorithms, we have carried out two different types of experiments. In our first set of runs, we kept the subdomain problem size, i.e., $H/h$ fixed and increased the number of subdomains and

**Table 6.1.** Algorithm A - Constant $H/h$

| Subdomains | Dof/Subdom. | Dof | Iterations | $\lambda_{\min}$ | $\lambda_{\max}$ |
|---|---|---|---|---|---|
| 8 | 1000 | 6,859 | 9 | 1.00035 | 11.5539 |
| 27 | 1000 | 21,952 | 14 | 1.00051 | 28.8335 |
| 64 | 1000 | 50,653 | 19 | 1.00361 | 25.0130 |
| 125 | 1000 | 97,336 | 22 | 1.00283 | 28.8335 |
| 216 | 1000 | 166,375 | 24 | 1.00231 | 25.0127 |
| 343 | 1000 | 262,144 | 26 | 1.00188 | 28.8335 |
| 512 | 1000 | 389,017 | 25 | 1.00161 | 25.0127 |
| 729 | 1000 | 551,368 | 26 | 1.00138 | 28.8335 |
| 1000 | 1000 | 753,571 | 24 | 1.00125 | 25.0127 |

thus the overall problem size; cf. Tables 6.1, 6.2, 6.3, and 6.4. Our second series of experiments is carried out with a fixed number (216) of subdomains and an increasing $H/h$ resulting in an increased value of $1/h$; cf. Tables 6.5 and 6.6 and Fig. 6.10. From both set of runs, we see that our computational results support the theoretical condition number estimates. However, for Algorithm E, the growth of the condition number appears to be linear rather than polylogarithmic in $H/h$. We note that for this problem, the bound given in Remark 6.40 is basically meaningless since $TOL_E = 10^4$. Experiments for an isotropic material, i.e., without any jumps in the coefficients show the same polylogarithmic growth for Algorithm E as for Algorithms B and C. We also note that a variant of Algorithm E, which works with averages over the closure of the faces, has been used extensively and that there is experimental evidence that this gives a faster rate of convergence than Algorithm E. In a third set of experiments, we have tested our algorithms for parallel scalability. We consider a decomposition into 216 subdomains with 13,824 degrees of freedom for each subdomain with an overall problem size of 2,685,619 degrees of freedom; cf. Table 6.7.

The experiments in Tables 6.1, 6.2, 6.3, and 6.4 were carried out on two dual Athlon MP 2200+ PCs with 2 GByte memory each. The experiments in Tables 6.5, 6.6, and 6.7 were computed on the 350 node Linux cluster Jazz at the Argonne National Laboratory. Each node is a 2.4 GHz Pentium Xeon where half of the nodes each has 2 GByte memory and the other half each 1 GByte.

The experiments show that all algorithms have a good parallel scalability for our model problem. For this problem and the number of degrees of freedom considered, the CPU times are not significantly different, although Algorithm C is always slightly faster. To decide which method is the best, more extensive testing with different model problems and geometries is needed.

**Table 6.2.** Algorithm B - Constant $H/h$

| Subdomains | Dof/Subdom. | Dof | Iterations | $\lambda_{\min}$ | $\lambda_{\max}$ |
|---|---|---|---|---|---|
| 8 | 1000 | 6,859 | 7 | 1.00085 | 1.47091 |
| 27 | 1000 | 21,952 | 8 | 1.00049 | 1.55036 |
| 64 | 1000 | 50,653 | 8 | 1.00025 | 1.47011 |
| 125 | 1000 | 97,336 | 8 | 1.00022 | 1.55036 |
| 216 | 1000 | 166,375 | 8 | 1.00013 | 1.46995 |
| 343 | 1000 | 262,144 | 8 | 1.00013 | 1.55036 |
| 512 | 1000 | 389,017 | 8 | 1.00009 | 1.46989 |
| 729 | 1000 | 551,368 | 8 | 1.00010 | 1.55036 |
| 1000 | 1000 | 753,571 | 7 | 1.00014 | 1.46985 |

**Table 6.3.** Algorithm C - Constant $H/h$

| Subdomains | Dof/Subdom. | Dof | Iterations | $\lambda_{\min}$ | $\lambda_{\max}$ |
|---|---|---|---|---|---|
| 8 | 1000 | 6,859 | 8 | 1.00030 | 1.61492 |
| 27 | 1000 | 21,952 | 9 | 1.00040 | 2.06800 |
| 64 | 1000 | 50,653 | 9 | 1.00020 | 1.93210 |
| 125 | 1000 | 97,336 | 10 | 1.00012 | 2.06875 |
| 216 | 1000 | 166,375 | 9 | 1.00009 | 1.93192 |
| 343 | 1000 | 262,144 | 10 | 1.00008 | 2.06875 |
| 512 | 1000 | 389,017 | 9 | 1.00006 | 1.93210 |
| 729 | 1000 | 551,368 | 10 | 1.00005 | 2.06875 |
| 1000 | 1000 | 753,571 | 9 | 1.00005 | 1.93210 |

**Table 6.4.** Algorithm E - Constant $H/h$

| Subdomains | Dof/Subdom. | Dof | Iterations | $\lambda_{\min}$ | $\lambda_{\max}$ |
|---|---|---|---|---|---|
| 8 | 1000 | 6,859 | 8 | 1.00102 | 11.4671 |
| 27 | 1000 | 21,952 | 10 | 1.00185 | 16.2107 |
| 64 | 1000 | 50,653 | 14 | 1.00129 | 16.2191 |
| 125 | 1000 | 97,336 | 16 | 1.00113 | 16.2246 |
| 216 | 1000 | 166,375 | 19 | 1.00089 | 16.2281 |
| 343 | 1000 | 262,144 | 19 | 1.00079 | 16.2304 |
| 512 | 1000 | 389,017 | 20 | 1.00067 | 16.2319 |
| 729 | 1000 | 551,368 | 20 | 1.00060 | 16.2329 |
| 1000 | 1000 | 753,571 | 20 | 1.00054 | 16.2335 |

**Table 6.5.** Algorithms A and C - Constant $H$

| Subdomains | H/h | Dof | Algorithm A | | | Algorithm C | | |
|---|---|---|---|---|---|---|---|---|
| | | | Iter | $\lambda_{\min}$ | $\lambda_{\max}$ | Iter | $\lambda_{\min}$ | $\lambda_{\max}$ |
| 216 | 4 | 6,859 | 14 | 1.00018 | 4.20279 | 6 | 1.00001 | 1.28960 |
| 216 | 8 | 79,507 | 22 | 1.00147 | 16.7662 | 8 | 1.00029 | 1.75693 |
| 216 | 12 | 300,763 | 27 | 1.00306 | 34.0512 | 10 | 1.00010 | 2.08459 |
| 216 | 16 | 753,571 | 31 | 1.00371 | 53.9590 | 11 | 1.00017 | 2.34317 |
| 216 | 20 | 1,520,875 | 32 | 1.00519 | 75.7574 | 11 | 1.00024 | 2.55999 |
| 216 | 24 | 2,685,619 | 34 | 1.00651 | 99.0372 | 12 | 1.00029 | 2.74869 |
| 216 | 28 | 4,330,747 | 36 | 1.00660 | 123.530 | 12 | 1.00035 | 2.91716 |
| 216 | 32 | 6,539,203 | 36 | 1.00677 | 149.054 | 13 | 1.00034 | 3.07033 |

**Table 6.6.** Algorithms B and E - Constant $H$

| Subdomains | H/h | Dof | Algorithm B | | | Algorithm E | | |
|---|---|---|---|---|---|---|---|---|
| | | | Iter | $\lambda_{\min}$ | $\lambda_{\max}$ | Iter | $\lambda_{\min}$ | $\lambda_{\max}$ |
| 216 | 4 | 6,859 | 5 | 1.01252 | 1.06768 | 13 | 1.00006 | 4.19816 |
| 216 | 8 | 79,507 | 7 | 1.00052 | 1.31862 | 19 | 1.00044 | 12.1453 |
| 216 | 12 | 300,763 | 8 | 1.00021 | 1.62065 | 22 | 1.00058 | 20.3391 |
| 216 | 16 | 753,571 | 10 | 1.00021 | 1.90164 | 23 | 1.00054 | 28.5889 |
| 216 | 20 | 1,520,875 | 10 | 1.00033 | 2.14742 | 23 | 1.00066 | 36.8711 |
| 216 | 24 | 2,685,619 | 11 | 1.00040 | 2.36688 | 25 | 1.00062 | 45.1044 |
| 216 | 28 | 4,330,747 | 12 | 1.00040 | 2.61352 | 24 | 1.00081 | 53.3703 |
| 216 | 32 | 6,539,203 | 12 | 1.00046 | 2.80160 | 24 | 1.00097 | 61.5779 |

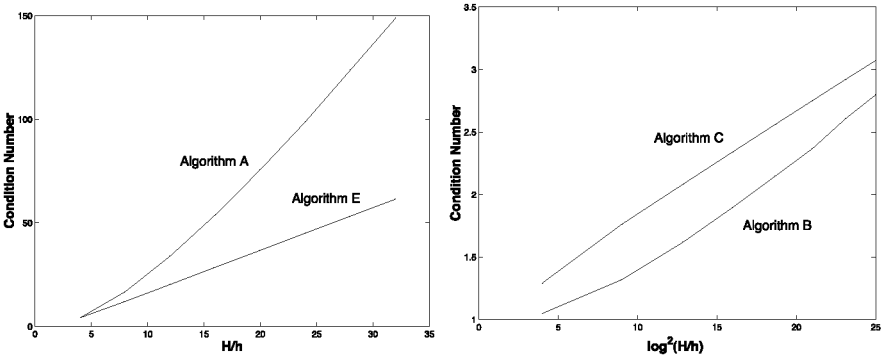


**Fig. 6.10.** Condition number as a function of $H/h$ for Algorithms A and E (*left*) and as a function of $\log(H/h)^2$ for Algorithms B and C (*right*).

**Table 6.7.** Parallel scalability - Algorithms A, B, C, and E with 216 subdomains, 13,824 dof per subdomain for a total of 2,685,619 dof.

| | Algorithm | | | |
|---|---|---|---|---|
| Processors | A | B | C | E |
| 27 | 223s | 207s | 205s | 216s |
| 54 | 113s | 106s | 106s | 110s |
| 108 | 57.0s | 54.2s | 53.8s | 55.4s |
| 216 | 29.1s | 28.9s | 27.2s | 29.1s |

# 7

# Spectral Element Methods

## 7.1 Introduction

Spectral element approximations were first introduced by Patera [374] and then for the Legendre case, which is more directly relevant to our work, by Maday and Patera [330]. These methods have been used extensively, in particular, in large scale simulations of incompressible fluid flow; see, e.g., Deville, Fischer, and Mund [160]. In this chapter, we will demonstrate that the methods and theory developed in Chap. 3, 5, and 6 can be extended to the spectral element case; we will also discuss extensions of the theory to $p$ and $hp$ methods in section 7.5. For a short introduction to spectral element methods, see appendix B.2, and for a detailed introduction to spectral methods, which form a basis for spectral elements, see [46, 48].

In this chapter, we will primarily consider three-dimensional problems. The corresponding two-dimensional algorithms can also be defined and analyzed using similar tools. We will use a conforming, shape-regular triangulation $\mathcal{T} = \{\Omega_i,\ i = 1, \ldots, N\}$ of our region $\Omega$. Each element $\Omega_i$ is the image of the reference cube $\hat{\Omega} = (-1, 1)^3$ under an affine mapping. We will denote the maximum diameter of $\mathcal{T}$ by $H$. Given a polynomial degree $k \geq 1$, the discrete space $V^k$ consists of continuous, piecewise $\mathbb{Q}_k$ functions.

Throughout we will work with the nodal basis functions of the spectral elements as introduced in appendix B.2. We recall that the nodal basis functions are defined with respect to the set of Gauss-Lobatto-Legendre (GLL) points and that their restriction to an element are constructed from the basic Lagrange interpolation functions for the points in question. We also recall that the integrals used in the variational formulation of our elliptic problems are replaced by Gauss-Lobatto-Legendre quadrature.

We will consider the second order problem

$$a(u, v) := \int_\Omega \rho \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx, \quad v \in H_0^1(\Omega), \tag{7.1}$$

where $\rho(x) \geq \rho_{\min} > 0$. Different assumptions on $\rho$ will be made in the sections on overlapping and iterative substructuring methods. Because of the equivalence (B.15) and Corollary C.2, good preconditioners for spectral element approximations with exact integration are also good for those that employ quadrature. In the following, we will therefore make no distinction between the bilinear forms $a(\cdot,\cdot)$ and $a_Q(\cdot,\cdot)$, and use the notation $a(\cdot,\cdot)$. Similarly, we will make no distinction between the corresponding stiffness matrices and use the common notation $A = A_k$ and write the linear system as

$$A_k u = f.$$

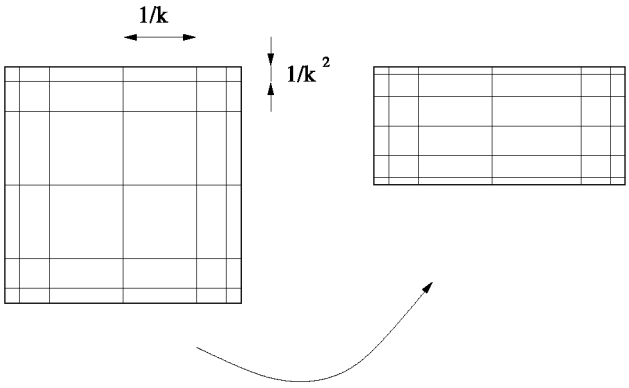


**Fig. 7.1.** Two-dimensional GLL mesh on the reference element $\hat{\Omega}$ and on an affinely mapped element $\Omega_i \in \mathcal{T}$. The dimensions of the thinnest element are $O(1/k)$ and $O(1/k^2)$ and its aspect ratio $O(k)$.

An important tool for the development and analysis of preconditioners for spectral element approximations is the finite element mesh $\hat{\mathcal{T}}_k = \mathcal{T}_k(\hat{\Omega})$ obtained by introducing planes through the GLL nodes which are parallel to the element faces. It consists of parallelepipeds and will be referred to as the GLL mesh in what follows. The mesh of a mapped element $\Omega_i \in \mathcal{T}$ is obtained from that of the reference cube $\hat{\Omega}$ and the affine mapping associated with $\Omega_i$ and it is denoted by $\mathcal{T}_i = \mathcal{T}_k(\Omega_i)$. The global GLL conforming mesh, built from the local $\mathcal{T}_i$, is denoted by $\mathcal{T}_k(\Omega)$; see Figures 7.1 and 7.2 for two-dimensional examples. We note that these families of meshes are not shape regular. Indeed, the elements in $\hat{\mathcal{T}}_k$ have dimensions ranging from $\sim 1/k^2$ to $\sim 1/k$ and the aspect ratios of some of the elements, in particular those next to the middle of an edge or face, grow in proportion to the degree $k$; cf. Figures 7.1 and 7.3.

The corresponding finite element spaces of piecewise trilinear functions on $\hat{\mathcal{T}}_k$, $\mathcal{T}_i$, and $\mathcal{T}_k(\Omega)$, are denoted by $\hat{V}^h = V^h(\hat{\Omega})$, $V^h(\Omega_i)$, and $V^h = V^h(\Omega)$, respectively. The stiffness matrix obtained by approximating the bilinear form $a(\cdot,\cdot)$ in $V^h$ is denoted by $A_h$. There is a one-to-one correspondence between

$V^h$ and $V^k$ given by

$$I^k : V^h \to V^k, \quad I^h : V^k \to V^h,$$

where $I^k$ and $I^h$ are the nodal interpolation operators onto $V^k$ and $V^h$, respectively. We use the notation $u_h \in V^h$ and $u_k \in V^k$ to distinguish between the piecewise trilinear and the piecewise polynomial functions.

A basic problem in our efforts to extend results from the lower order finite element case to spectral finite elements stems from the fact that the GLL mesh $\hat{\mathcal{T}}_k$ fails to be uniformly shape regular. More precisely, if inverse estimates as in Lemma B.27 were to be employed, the diameter $h$ in the denominator must be chosen as a typical minimum length and this could spoil some of our estimates. With the help of Lemma 7.1, given below, we will develop our theory by working with the first order finite element space $V^h$ and we will start by establishing that $A_h$ is spectrally equivalent to the spectral element matrix $A_k$. The idea of preconditioning with these sparse matrices goes back to Orszag [365] and a complete theory, which we will develop in Sect. 7.2, was apparently first given by Canuto [120]. These preconditioners are often associated with Deville and Mund, see [162, 161], who have done a great deal to develop and popularize these techniques; we will therefore call this special finite element model the Deville-Mund problem. For other early work on these methods, see also Canuto and Quarteroni [122]. Our strategy is then, in view of Corollary C.2 (a good preconditioner of a good preconditioner remains a good preconditioner), to reexamine the work in Chap. 3, 5, and 6 and show that all our results can be carried over directly to the Deville-Mund problem. While we will focus on scalar elliptic problems and, in particular, on the model problem (4.3) when working on iterative substructuring methods, we note that our results can be extended to elliptic systems, as in Chap. 8, to the same extent as for lower order finite element methods.

The approach selected here was first used systematically and successfully in the thesis of Casarin [126]; see also [127]. Casarin's work on overlapping Schwarz methods was inspired by a master of science thesis by Pahl [369]. There is an alternative route to the results on iterative substructuring methods which was taken by Pavarino and Widlund in work that preceded Casarin's; see in particular [381, 382]. For further work along those lines and work on specific algorithms, see [378, 383, 384, 385]. We also note that all that work followed even older work by Pavarino on overlapping Schwarz methods with generous element-wide overlap; see [375, 376, 377]. There has also been work on more general $p$ and $hp$ version finite element approximations, employing different types of bases, by Babuška et al. [31], Bică [52, 53] Casarin and Sherwin [417], Guo and Cao [243, 244, 245, 246, 247, 123, 248, 124], Pavarino, Warburton, and Hesthaven [380, 459], Korneev, Langer, and Xanthis [295]. We refer to Sect. 7.5 for further comments and references on extensions to $p$ and $hp$ approximations.

A fundamental (and trivial) result for trilinear functions on general parallelepipeds is given in the following lemma; cf. [126, Lem. 3.3.1]. A key point is that the constants in the bounds are independent of the diameter and the aspect ratio of the element. Given an element $K = (0, h_1) \times (0, h_2) \times (0, h_3)$ and a coordinate direction $x_i$, let $a$, $b$, $c$, and $d$ be the vertices of a face of $K$ perpendicular to this direction, and let $a'$, $b'$, $c'$, and $d'$ be the corresponding points on the parallel face.

**Lemma 7.1** *Let $K = (0, h_1) \times (0, h_2) \times (0, h_3)$ and $a$, $b$, $c$, and $d$ be the vertices of a face of $K$ perpendicular to the $x_i$ direction. Then, there are constants $c > 0$ and $C$, which are independent of $h_1$, $h_2$, and $h_3$, such that, if $u \in \mathbb{Q}_1(K)$,*

$$
\begin{aligned}
&c\|u\|^2_{L^2(K)} \le h_1 h_2 h_3 \sum_{x=a,b,c,d} (u(x)^2 + u(x')^2) \le C\|u\|^2_{L^2(K)},\\
&c\|\partial_{x_i} u\|^2_{L^2(K)} \le (h_1 h_2 h_3 / h_i^2) \sum_{x=a,b,c,d} (u(x) - u(x'))^2 \le C\|\partial_{x_i} u\|^2_{L^2(K)},\\
&c\|\partial_{x_i} u\|^2_{L^\infty(K)} \le h_i^{-2} \sum_{x=a,b,c,d} (u(x) - u(x'))^2 \le C\|\partial_{x_i} u\|^2_{L^\infty(K)}.
\end{aligned}
$$

We note that corresponding estimates for one dimension,

$$
\begin{aligned}
&\|u\|^2_{L^2(I)} \le h/2(u(a)^2 + u(a')^2) \le 3\|u\|^2_{L^2(I)},\\
&\|u_x\|^2 = h^{-1}(u(a) - u(a'))^2,
\end{aligned}
\tag{7.2}
$$

can be derived very easily.

## 7.2 Deville-Mund Preconditioners

In view of the central role of this preconditioner for our development of our theory, we will provide a full theory except for a result due to Bernardi and Maday [47, Corollary 4.6], which is also given in appendix B.12. We will follow Canuto [120] quite closely. Most of our work will concern one-dimensional problems and continuous, piecewise linear functions $\phi_h$, with nodes at the $GLL(k)$ points $\xi_i, 0 \le i \le k$, and polynomials of degree $k$ denoted by $\phi_k$. With $\Lambda := (-1, 1)$, we consider the polynomial interpolation operator $I^k : C(\Lambda) \to \mathbb{P}_k(\Lambda)$. We will compare the $L^2$- and $H^1$-norms of two functions $\phi_h$ and $\phi_k$ which share the same values on the GLL mesh, i.e., $\phi_k = I^k \phi_h$ and $\phi_h = I^h \phi_k$.

The proof that the Deville-Mund preconditioner is optimal for two and three dimensions can be reduced to a result on the interval $\Lambda$, namely,

**Lemma 7.2** *There exist constants $c > 0$ and $C$, which are independent of the degree $k$, such that*

$$
c\|\phi_k\|^2_{L^2(\Lambda)} \le \|\phi_h\|^2_{L^2(\Lambda)} \le C\|\phi_k\|^2_{L^2(\Lambda)}, \tag{7.3}
$$

*and*

$$
c|\phi_k|^2_{H^1(\Lambda)} \le |\phi_h|^2_{H^1(\Lambda)} \le |\phi_k|^2_{H^1(\Lambda)}. \tag{7.4}
$$

*Proof.* Let

$$\Lambda_j := (\xi_j, \xi_{j+1}), \quad h_j := \xi_{j+1} - \xi_j, \qquad 0 \le j \le k-1.$$

By direct calculation, it is easy to show that, as in (7.2),

$$\|\phi_h\|^2_{L^2(\Lambda)} \ge \frac{h_0}{6}\phi(\xi_0)^2 + \frac{h_{k-1}}{6}\phi(\xi_k)^2 + \sum_{j=1}^{k-1}\frac{h_{j-1}+h_j}{6}\phi(\xi_j)^2$$

$$\|\phi_h\|^2_{L^2(\Lambda)} \le \frac{h_0}{2}\phi(\xi_0)^2 + \frac{h_{k-1}}{2}\phi(\xi_k)^2 + \sum_{j=1}^{k-1}\frac{h_{j-1}+h_j}{2}\phi(\xi_j)^2.$$

By using formulas (B.5) and (B.6), we find

$$(1/C)h_0 \le w_0 \le Ch_0$$

$$(1/C)h_{k-1} \le w_k \le Ch_{k-1}$$

$$(1/C)(h_{j-1}+h_j) \le w_j \le C(h_{j-1}+h_j), \quad 1 \le j \le k-1;$$

see Canuto [120] for details. The inequalities of (7.3) then follow by using formula (B.8).

The lower bound of formula (7.4) follows directly from a result of Bernardi and Maday given in Lemma B.12, the fact that $I^k$ reproduces constants, and the Poincaré inequality in Lemma A.13. For the upper bound, let us consider $|\phi_h|_{H^1(\Lambda_j)}$. Since $\phi_h$ is a linear function, it is harmonic in $\Lambda_j$. $\phi_k$ and $\phi_h$ also share the same values at the endpoints of $\Lambda_j$ and the upper bound follows immediately by adding over $j$. □

What remains is to generalize Lemma 7.2 to two and three dimensions. The $L^2$-bound follows easily by using the one-dimensional bound for one variable at a time. Similarly, we can estimate the $L^2$-norm of a first derivative of $u_h$ in terms of the norm of the derivative of $u_k$ by using all the inequalities of Lemma 7.2, one by one. For an affinely mapped element, the result follows by a standard scaling argument. We obtain:

**Theorem 7.3** *Let $u_h$ be a piecewise trilinear finite element function on $\mathcal{T}_k(\hat{\Omega})$ and let $u_k$ be the $\mathbb{Q}_k$ polynomial which takes on the same values as $u_h$ at the GLL nodes of $\hat{\Omega}$. Then, there exist constants $c > 0$ and $C$, which are independent of the degree $k$, such that*

$$c\|u_k\|^2_{L^2(\hat{\Omega})} \le \|u_h\|^2_{L^2(\hat{\Omega})} \le C\|u_k\|^2_{L^2(\hat{\Omega})}$$

$$c|u_k|^2_{H^1(\hat{\Omega})} \le |u_h|^2_{H^1(\hat{\Omega})} \le C|u_k|^2_{H^1(\hat{\Omega})}.$$

*A similar result holds for an affinely mapped element $\Omega_i$ and for the domain $\Omega$, with constants that only depend on the aspect ratio of $\Omega_i$ and of the mesh $\mathcal{T}$, respectively.*

**Corollary 7.4** *The spaces $V^k$ and $V^h$ are isomorphic and the matrices $A_k$ and $A_h$ are spectrally equivalent.*

## 7.3 Two-Level Overlapping Schwarz Methods

In this section, we will follow the analysis first given by Casarin [127, 126] quite closely. We now assume that the coefficient $\rho$ in (7.1) is constant; as in Chap. 3, there are no strong results for arbitrary variations of the coefficients. Two-level overlapping Schwarz methods can be developed straightforwardly for the Deville-Mund problem. To simplify our analysis, as well as the practice, we will use the underlying spectral element mesh $\mathcal{T}$ as our coarse mesh. We define our coarse space, $V_0 = V^H$, as a first (or low order) finite element space on this mesh. It then follows immediately that $V_0$ is contained in $V^k$ and also in $V^h$ in the linear case. $R_0^T : V_0 \to V^k$, again denotes the natural extension operator from the coarse to the fine mesh and $A_H$ the stiffness matrix for the coarse problem on $V_0$.

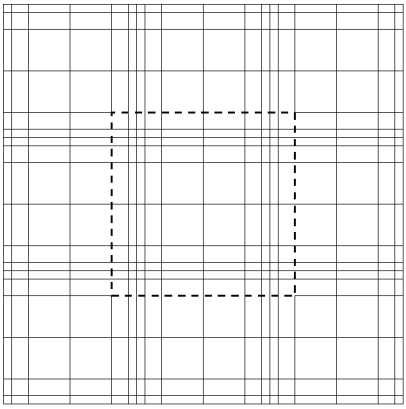

**Fig. 7.2.** Overlapping subdomain in two dimensions.

The local spaces are associated with an overlapping partition of $\Omega$. Here, we will construct the overlapping subdomains $\Omega_i'$ from the individual spectral elements $\Omega_i$ by adding one or several layers of Gauss-Lobatto points from the neighboring elements. We refer to Figure 7.2 for a two-dimensional example. We assume a small overlap and a finite covering as in Assumptions 3.1 and 3.2. We denote by $\mathcal{T}_i'$ the GLL mesh obtained from the GLL nodes in $\Omega_i'$. The local space $V_i = V^h(\Omega_i')$ is the space of piecewise trilinear functions on $\mathcal{T}_i'$ that vanish on $\partial\Omega_i'$ and $R_i^T : V^h(\Omega_i') \to V^k(\Omega)$, is the interpolation operator that extends such a local finite element function by zero to the rest of $\Omega$, and then interpolates into the spectral element space $V^k$ at the GLL nodes. In terms of degrees of freedom, $R_i^T$ simply takes a vector of degrees of freedom in the interior of $\Omega_i'$ and returns a global vector with zeros at the GLL nodes in $\Omega \setminus \Omega_i'$. The local finite element stiffness matrix on $V_i$ is denoted by $A_{h,i}$. With these definitions, a two-level additive preconditioner can be defined by

$$\hat{A}_h^{-1} := R_0^T A_H^{-1} R_0 + \sum_{i=1}^{N} R_i^T A_{h,i}^{-1} R_i,$$

and the preconditioned operator by $P_{ad} = \hat{A}_h^{-1} A_k$.

We note that $\hat{A}_h^{-1}$ is exactly the two-level preconditioner for the finite element matrix $A_h$ defined in Chap. 3. Indeed, while the spaces $V^h$ and $V^k$ are different, they give rise to the same space of degrees of freedom and it is easy to show that the extension matrices $R_0^T$ and $R_i^T$ are the same for the two spaces. According to Corollaries C.2 and 7.4, it is then enough to bound the condition number of the two-level preconditioned finite element operator $\hat{A}_h^{-1} A_h$.

We will now examine the proofs of Sect. 3.6 and, in particular, the technical tools of Sect. 3.5. We find that there are no problems with the coarse global space; we can use Lemma 3.6 directly and we do not need to use Lemma 3.8 since $V_0 \subset V^h$. We find that Lemma 3.9 relies on an inverse inequality; we will soon see how to overcome this restriction and still prove a strong result in Lemma 7.5, using an alternative argument. Lemma 3.10, on the other hand, holds for all of $H^1$ and it therefore presents no problems. We also find that Lemma 3.11, which relies on a coloring argument for the overlapping partition, remains valid. We finally turn to the remaining task which is to prove a counterpart of Lemma 3.12. Here we find that Lemma 3.9 is used to derive the estimate of formula (3.20) but that the rest of the proof remains valid without any assumptions on the aspect ratios of the mesh elements. Following Casarin, we formulate and prove the following lemma.

**Lemma 7.5** *Let $K = (0, h_1) \times (0, h_2) \times (0, h_3)$, $\theta \in W^{1,\infty}(K)$, with $0 \leq \theta \leq 1$, and $u$ be a trilinear function. Then, there is a uniformly bounded constant $C$, independent in particular of $K$ and its aspect ratio, such that*

$$|I^h(\theta u)|^2_{H^1(K)} \leq C(|u|^2_{H^1(K)} + \|\nabla\theta\|^2_{L^\infty(K)} \|u\|^2_{L^2(K)}).$$

*A similar bound holds for an affinely mapped element $K$, with a constant that depends only on the angles of $K$.*

*Proof.* By using Lemma 7.1, we have, for the direction $x_i$,

$$\|\partial_{x_i} I^h(\theta u)\|^2_{L^2(K)} \leq (h_1 h_2 h_3 / h_i^2) \sum_{x=a,b,c,d} (\theta(x)u(x) - \theta(x')u(x'))^2.$$

A term in the sum above can be bounded by

$$2\theta(x')^2 (u(x) - u(x'))^2 + 2u(x)^2 (\theta(x) - \theta(x'))^2$$

and another application of Lemma 7.1 concludes the proof for an axi-parallel element. The case of an affinely mapped element can be treated using a standard scaling argument. □

*Remark 7.6.* A uniform bound as in Lemma 7.5 can also be established, by a simple computation, for any tetrahedron with a right angle. By examining this proof, we also see that is not possible to get a uniform bound for all, arbitrarily distorted, tetrahedra. We also note that alternative Deville-Mund preconditioners can be developed by replacing the piecewise trilinear finite element functions by piecewise linears after subdividing each GLL element into six tetrahedra. There is clear numerical evidence that such a preconditioner is more powerful than the one based on trilinears.

The proof of Lemma 3.12 can now easily be modified and we have established the following direct counterpart of Theorem 3.13.

**Theorem 7.7** *In case exact solvers are employed on all subspaces, the condition number of the additive Schwarz operator satisfies*

$$\kappa(P_{ad}) \leq C\left(1 + \frac{H}{\delta}\right),$$

*where $C$ depends on $N^c$, but is otherwise independent of $k$, $H$, and $\delta$.*

*Remark 7.8.* A different two-level preconditioner can be defined without using local GLL meshes and local spaces of piecewise trilinear functions. Following Casarin [126, Sect. 3.5], the local finite element spaces $V^h(\Omega_i')$ can be replaced by spectral element spaces consisting of piecewise polynomial functions with vanishing values at all the GLL points in $\Omega \setminus \Omega_i'$. The extension matrices $R_i^T$ remain the same while the local matrices $A_{h,i}$ are now replaced by $A_{k,i} = R_i A_k R_i^T$. We note that the support of these local functions is not contained in $\Omega_i'$. The same condition number bound as in Theorem 7.7 can be proven in this case. We note that the resulting matrices are dense but that if the subdomains are rectangular the resulting linear systems can be solved efficiently by exploiting the tensor product structure of the problems to develop fast diagonalization methods; see Deville, Fischer, and Mund [160, Chap. 4] and also Couzy and Deville [144, 145] and Lynch, Rice, and Thomas [326].

## 7.4 Iterative Substructuring Methods

Iterative substructuring methods can be devised for spectral element approximations as well. The mesh $\mathcal{T}$ provides a nonoverlapping partition into shape-regular subdomains of maximum diameter $H$, obtained from affine mappings of the reference cube $\hat{\Omega}$. We use the same notations as for low order finite element approximations, introduced in Sect. 4.2. Thus, the interface $\Gamma$ is the union of the interior subdomain faces, edges (all regarded as open sets), and vertices. The faces of $\Omega_i$ are denoted by $\mathcal{F}^{ij}$, its edges by $\mathcal{E}^{i\ell}$, its vertices by $\mathcal{V}^{i\ell}$, and its wire basket by $\mathcal{W}^i$. We will also use notation with one or even without any superscript. The sets $\Gamma_h$, $\partial\Omega_{i,h}$, $\mathcal{F}_h^{ij}$, $\mathcal{E}_h^{i\ell}$, and $\mathcal{W}_h^i$ now consists

of GLL nodes. We note that a vertex $\mathcal{V}^{il}$ is always a node of the local GLL mesh $\mathcal{T}_i$.

As before, we assume that all possible large jumps of the coefficient $\rho(x)$ in (7.1) are aligned with the subdomain boundaries

In the case of a region of diameter $H_i$, we use norms with different relative weights obtained by a simple dilation argument as in formulas (4.4) and (4.5). To every substructure, we associate a local space $V^k(\Omega_i) = \mathbb{Q}_k(\Omega_i)$ and a local bilinear form $a_i(\cdot,\cdot)$ as in (4.3).

We can now proceed as in Sect. 4.3 and order the global degrees of freedom relative to interior GLL nodes of the subdomains first (subscript $I$), followed by those on the interface $\Gamma$ (subscript $\Gamma$). For local functions, an analogous splitting can be considered. The contributions to the stiffness matrix $A = A_k$ and the right hand side $f$ can be formed one subdomain at a time and we obtain a linear system as in (4.7). The unknowns in the interior of the substructures are then eliminated and a system with a Schur complement matrix $S = S_k$ is obtained, as in (4.10) and (4.11):

$$Su_\Gamma = \tilde{f}_\Gamma.$$

The contributions to the Schur complement matrix can also be formed one subdomain at a time, using the local Schur complements $S^{(i)} = S_k^{(i)}$. The same is true for the right hand side $\tilde{f}_\Gamma$.

Discrete harmonic extensions can also be defined in this case: a local function $u_i \in V^k(\Omega_i)$, represented by the local vector $u^{(i)}$, is said to be discrete harmonic if (4.14) holds. As in the case of continuous, piecewise linear functions, $u^{(i)} =: \mathcal{H}(u_\Gamma^{(i)})$ is completely defined by its nodal values on $\partial\Omega_i$. A function $u \in V^k$ is piecewise discrete harmonic if its restriction $u_i$ to $\Omega_i$ is discrete harmonic; the function $u =: \mathcal{H}(u)$ is completely defined by its nodal values on $\Gamma$.

We will work with the inner products defined by the Schur complements:

$$\begin{aligned} s(u_\Gamma, v_\Gamma) &:= u_\Gamma^T S v_\Gamma, \\ s_i(u^{(i)}, v^{(i)}) &:= {u^{(i)}}^T S^{(i)} v^{(i)}; \end{aligned}$$

cf. (4.15) The minimum property of discrete harmonic extensions in Lemma 4.9 remains valid in this case and we can either work with functions defined on the interface $\Gamma$ (and with the scalar product $s(\cdot,\cdot)$) or with the corresponding discrete harmonic extensions (with the scalar product $a(\cdot,\cdot)$). The equivalence between the $H^{1/2}$- and $H^1$-norms on a substructure for discrete harmonic extensions given in Lemma 4.10 remains valid for spectral approximations. Indeed, the proof involves the construction of a stable polynomial extension, in a way similar as for finite element functions in Lemma 4.6. Stable extensions can indeed be constructed for polynomials; see, e.g., [328, 45, 39].

We have the following result for the condition number of the Schur complement matrix. The proof can be found, e.g., in [352]; see in particular Theorem

2.1. We are not aware of any result for three dimensional approximations and the generalization of the two-dimensional proof in [352] does not appear to be trivial.

**Lemma 7.9** *Assume that $\rho_i = 1$, $i = 1, \ldots, N$, that $\Omega \subset \mathbb{R}^2$, and that the mesh $\mathcal{T}$ is quasi uniform. Then, there exists a constant $C$, such that*

$$\kappa(S_k) \leq C \frac{k}{H^2}.$$

We will again consider the local and global GLL meshes and the corresponding spaces of piecewise trilinear functions. Unknowns interior to the subdomains can be eliminated and a system involving the Schur complement matrix $S_h$ is obtained. We will denote by $S_h^{(i)}$ the local Schur complement associated with $\Omega_i$. As for the original stiffness matrices $A_h$ and $A_k$, $S_h$ is an optimal preconditioner for $S_k$.

**Lemma 7.10** *The matrices $S_k^{(i)}$ and $S_h^{(i)}$, and $S_k$ and $S_h$ are spectrally equivalent.*

*Proof.* We need only show the equivalence of the local Schur complements. Bounds for the global ones can then be obtained by subassembling as in (4.10). Let $u_\Gamma^{(i)}$ be a vector of nodal degrees of freedom on $\partial\Omega_i$, and let $u_k$ and $u_h$ be the corresponding discrete harmonic extensions. We obtain

$$u_\Gamma^{(i)T} S_k^{(i)} u_\Gamma^{(i)} = a_i(u_k, u_k) \leq a_i(I^k u_h, I^k u_h) \leq C a_i(u_h, u_h) = C u_\Gamma^{(i)T} S_h^{(i)} u_\Gamma^{(i)},$$

where we use Lemma 4.9 for the first and last equalities and for the first inequality, and Theorem 7.3 for the second inequality. A lower bound can be found in a similar way. □

### 7.4.1 Technical Tools

We recall that the basic tools for the analysis of the iterative substructuring methods developed in Chap. 5 and 6 were collected in Sect. 4.6. We will therefore first systematically consider the extent by which these auxiliary results can be extended to the spectral element case. We again note that because of the large and growing aspect ratios of some of the elements of the Gauss-Lobatto mesh, we have to avoid using inverse inequalities, i.e., we should make sure that any argument based on such an inequality can be replaced. In addition, as in Casarin [127, Sect. 3.3], once we have stability results for these finite element functions, they translate into corresponding results for polynomials, by using Lemma 7.2 and Theorem 7.3. We will only consider the modifications of the proofs required for the particular anisotropic GLL mesh and refer to Sect. 4.6 for the rest of the proofs. Finally, we only give proofs for cubic substructures. The general case of affinely mapped elements can easily be dealt by a scaling argument.

The following lemma is an analog of Lemmas 4.16 and 4.17. We note that the latter employs Lemma 4.15 which relies on an inverse inequality.

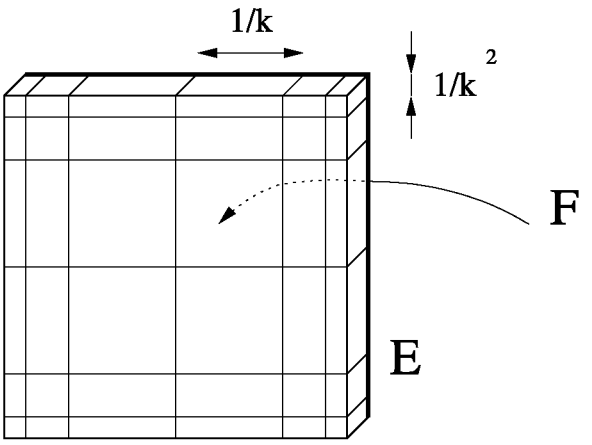


**Fig. 7.3.** Three-dimensional GLL mesh on the reference element $\hat{\Omega}$ close to a face.

**Lemma 7.11** *Let $\bar{u}_{\mathcal{E}^l}$ be the average value of $u$ over $\mathcal{E}^l$, an edge of a face $\mathcal{F}^j$ of $\Omega_i$. Then, there exists a constant $C$, such that*

$$||u||^2_{L^2(\mathcal{E}^l)} \leq C(1+\log(k))||u||^2_{H^1(\Omega_i)},$$
$$||u||^2_{L^2(\mathcal{E}^l)} \leq C(1+\log(k))||u||^2_{H^{1/2}(\mathcal{F}^j)},$$

*and*

$$||u-\bar{u}_{\mathcal{E}^l}||^2_{L^2(\mathcal{E}^l)} \leq C(1+\log(k))|u|^2_{H^1(\Omega_i)},$$
$$||u-\bar{u}_{\mathcal{E}^l}||^2_{L^2(\mathcal{E}^l)} \leq C(1+\log(k))|u|^2_{H^{1/2}(\mathcal{F}^j)},$$

*for $u \in V^k(\Omega_i)$. Similar bounds also hold for the wire basket $\mathcal{W}^i$ of the subdomain $\Omega_i$ if the edge $\mathcal{E}^l$ and the face $\mathcal{F}^j$ are replaced by $\mathcal{W}^i$ and the boundary $\partial\Omega_i$, respectively.*

*Proof.* Let $u_h = I^h u$. We refine the GLL mesh $\mathcal{T}_i$ creating a quasi-uniform mesh which has a mesh size on the order of $1/k^2$ on the unit cube; see Figure 7.3. For the first two inequalities, we first consider the case of $H = 1$. The function $u_h$ can also be represented on the finer quasi-uniform mesh and we can then employ Lemmas 4.16 and 4.17 and obtain bounds for $u_h$ with $1/h$ replaced by $k^2$. Lemma 7.2 and Theorem 7.3 then give the bounds for $u$. The resulting inequalities are equally valid for any $H$ as can be seen by a simple scaling argument. Therefore the factor $\log(H/h)$ is replaced by $\log(k)$ in all the appropriate inequalities. The inequalities involving the seminorms can then be proven by using a Poincaré inequality. □

We next turn to an analog of Lemma 4.19:

**Lemma 7.12** *Let $\vartheta_{\mathcal{E}^j} \in V^h(\Omega_i)$ be the finite element function that equals one at the nodes of an edge $\mathcal{E}^j$ of $\Omega_i$ and which vanishes on the rest of the nodes in $\Omega_{i,h}$. Then,*

$$|\mathcal{H}_i(\vartheta_{\mathcal{E}^j}u)|^2_{H^1(\Omega_i)} \leq C||I^k(\vartheta_{\mathcal{E}^j}u)||^2_{L^2(\mathcal{E}^j)},$$
$$|\mathcal{H}_i(\theta_{\mathcal{E}^j}u)|^2_{H^1(\Omega_i)} \leq C||u||^2_{L^2(\mathcal{E}^j)}.$$

*Similar bounds also hold for the wire basket $\mathcal{W}^i$ of the subdomain $\Omega_i$.*

*Proof.* The proof is very similar to that of Lemma 4.19. We estimate the energy of the zero extension from the edge. We consider an element $K$, of dimensions $h_x$, $h_y$, and $h_z$, in the GLL mesh $\mathcal{T}_i$ that touches an edge $\mathcal{E}$; see Figure 7.3. We assume that $\mathcal{E}$ is parallel to the $z$ direction. Then, $h_x = h_y \leq h_z$.

The nodal values of $I^h(\vartheta_{\mathcal{E}^j}u)$ on $K$ are 0, 0, 0, 0, 0, 0, $\vartheta_{\mathcal{E}^j}(a)u(a)$, $\vartheta_{\mathcal{E}^j}(b)u(b)$, with $a$ and $b$ the vertices on $\mathcal{E}^j$. Using Lemma 7.1 and the inequalities in (7.6), we can easily show that

$$|I^h(\vartheta_{\mathcal{E}^j}u)|^2_{H^1(K)} \leq Ch_z\left((\vartheta_{\mathcal{E}^j}(a)u(a))^2 + (\vartheta_{\mathcal{E}^j}(b)u(b))^2\right).$$

Using Lemma 7.1 and that $0 \leq \vartheta_{\mathcal{E}^j} \leq 1$, the right hand side can be bounded both by $C\int_a^b (I^h(\vartheta_{\mathcal{E}^j}u))^2 dz$ and by $C\int_a^b u^2 dz$. Summing over the elements and using the minimum property of discrete harmonic extensions and Theorem 7.3 concludes the proof. □

Given a face $\mathcal{F}^j$ of a substructure $\Omega_i$, we recall that for first order finite element approximations, we have constructed two functions $\theta_{\mathcal{F}^j}$ and $\vartheta_{\mathcal{F}^j}$ in Lemma 4.25. They have the same values on $\partial\Omega_i$ and vanish at all the nodes of $\partial\Omega_{i,h}$ except at those of $\mathcal{F}^j$ where they take the value one. $\theta_{\mathcal{F}^j}$ is discrete harmonic while an explicit construction of the interior values of $\vartheta_{\mathcal{F}^j}$ is given in the proof of Lemma 4.25. Both sets of functions provide partitions of unity at all the nodal points in $\overline{\Omega_i}$ except for those on the wire basket.

In the proofs of Lemmas 4.25 and 4.24, the bound for $||\theta_{\mathcal{F}^j}||_{L^2(\Omega_i)}$ depends on an inverse inequality. The resulting estimate is used only once in Chap. 5 and 6, namely in the proof of formula (5.11) in the proof of Theorem 5.13. We have already noted, in Remark 5.14, that we can modify that argument and use the functions $\vartheta_{\mathcal{F}^j}$ instead of the $\theta_{\mathcal{F}^j}$. The proofs of the $H^1$ bounds in Lemmas 4.25 and 4.24 can also easily be modified. We proceed as follows:

**Lemma 7.13** *Given a face $\mathcal{F}^j$ of $\Omega_i$, there exists a finite element function $\vartheta_{\mathcal{F}^j} \in V^h(\Omega_i)$, that is equal to one at the nodal points of $\mathcal{F}^j_h$ and zero on $\Gamma_{i,h} \setminus \mathcal{F}^j_h$, such that*

$$\begin{aligned} &\sum_{\mathcal{F}^j \subset \Gamma_i} \vartheta_{\mathcal{F}^j}(x) = 1, \quad x \in (\Omega_{i,h} \cup \partial\Omega_{i,h}) \setminus \mathcal{W}^i_h, \\ &0 \leq \vartheta_{\mathcal{F}^j} \leq 1, \\ &|\nabla\vartheta_{\mathcal{F}^j}(x)| \leq C/r(x), \end{aligned} \tag{7.5}$$

*for $x$ in every element that does not touch $\mathcal{W}^i$ and $r = r(x)$ the distance to $\mathcal{W}^i$.*

*Proof.* We construct $\vartheta_{\mathcal{F}^j}$ as in the proof of Lemma 4.25, in particular, by interpolation on the GLL mesh $\mathcal{T}_i$. A careful look at that proof reveals that the properties of $\vartheta_{\mathcal{F}^j}$ are independent of the finite element mesh in $\Omega_i$ and therefore remain valid in the GLL case. □

**Lemma 7.14** *Let $\vartheta_{\mathcal{F}^j}$ be the functions in Lemma 7.13, where $\mathcal{F}^j$ is a face of the substructure $\Omega_i$. Then, for every $x \in \Omega_{i,h} \cup \Gamma_{i,h}$ that is not on $\mathcal{W}^i$,*

$$\sum_j I^k(\vartheta_{\mathcal{F}^j} u)(x) = \sum_j I^h(\vartheta_{\mathcal{F}^j} u)(x) = u(x), \quad u \in V^k$$

*and*

$$|I^k(\vartheta_{\mathcal{F}^j} u)|^2_{H^1(\Omega_i)} \leq C(1+\log(k))^2 ||u||^2_{H^1(\Omega_i)},$$
$$|I^k(\vartheta_{\mathcal{F}^j} u)|^2_{H^{1/2}(\partial\Omega_i)} \leq C(1+\log(k))^2 ||u||^2_{H^{1/2}(\partial\Omega_i)},$$
$$|I^k(\vartheta_{\mathcal{F}^j}(u-\bar{u}_{\mathcal{F}^j})|^2_{H^{1/2}(\partial\Omega_i)} \leq C(1+\log(k))^2 |u|^2_{H^{1/2}(\partial\Omega_i)}.$$

*Proof.* The first equality follows directly from Lemma 7.13.

For the first inequality, we only consider the case of a substructure of unit diameter. The more general case can be dealt with using a scaling argument. Because of Theorem 7.3, it is enough to estimate the energy of the finite element function $I^h(\vartheta_{\mathcal{F}^j} u)$.

We recall that the proof of the corresponding result for low order approximations in Lemma 4.24 consists of two independent parts. In the first, we estimated the contributions from elements that do not touch the wire basket and this can be done in exactly the same way in this case, using Lemma 7.13, and noting that the elements that touch the wire basket have a minimum mesh size on the order of $1/k^2$. We obtain the same bound where the ratio $H/h$ is replaced by $k$; see, in particular, formula (4.21). In a second part, we estimated contributions from elements that touch the wire basket. We need to modify this part of the proof since some of the elements that touch the wire basket have large aspect ratios; cf. Figure 7.3. We follow an argument in [126, Lem. 3.3.7].

We consider an element $K$, of dimensions $h_x$, $h_y$, and $h_z$, in the GLL mesh $\mathcal{T}_i$ that touches an edge $\mathcal{E}$; see Figure 7.3. We assume that $\mathcal{E}$ is parallel to the $z$ direction. In fact, for the GLL mesh we have

$$h_x = h_y \leq h_z. \tag{7.6}$$

We note that $h_x = h_y$ are on the order of $1/k^2$, while $h_z$ ranges from $C/k^2$ to $C/k$.

The nodal values of $I^h(\vartheta_{\mathcal{F}^j} u)$ on $K$ are 0, 0, 0, 0, $u(a)$, $u(b)$, $\vartheta_{\mathcal{F}^j}(c)u(c)$, and $\vartheta_{\mathcal{F}^j}(d)u(d)$, with $a$ and $b$ vertices on $\mathcal{F}^j$ and $c$ and $d$ vertices in the interior of $\Omega_i$. Using Lemma 7.1 and the inequalities in (7.6), we can easily show that

$$\begin{aligned} |I^h(\vartheta_{\mathcal{F}^j} u|^2_{H^1(K)} &\leq C h_z \left(u(a)^2 + u(b)^2 + u(c)^2 + u(d)^2\right) \\ &\leq C \left( \int_a^b u^2 dz + \int_c^d u^2 dz \right), \end{aligned}$$

where we have also used the fact that $\vartheta_{\mathcal{F}^j}$ has values between zero and one. Summing over the element $K$ and using Lemma 7.11 give

$$\sum_K |I^h(\vartheta_{\mathcal{F}^j} u)|^2_{H^1(K)} \leq C\,(1+\log(k))\,\|u\|^2_{H^1(\Omega_i)},$$

where the sum is taken over the elements in $\mathcal{T}_i$ that touch $\mathcal{W}^i$. This concludes the proof of the first inequality. The two remaining bounds can be proven as in Lemma 4.26, using properties of discrete harmonic extensions and the Poincaré inequality in Lemma A.17. □

We note that the proof of the previous lemma is not valid for general anisotropic meshes since it relies on the property of the GLL mesh, given in (7.6), which ensures that long elements are aligned with the subdomain edges.

**Lemma 7.15** *Let $\vartheta_{\mathcal{F}^j}$ be the functions in Lemma 7.13, where $\mathcal{F}^j$ is a face of the substructure $\Omega_i$. Then,*

$$|\vartheta_{\mathcal{F}^j}|^2_{H^1(\Omega_i)} \leq C(1+\log(k))H_i,$$
$$\|\vartheta_{\mathcal{F}^j}\|^2_{L^2(\Omega_i)} \leq CH_i^3,$$
$$|\vartheta_{\mathcal{F}^j}|^2_{H^{1/2}(\partial\Omega_i)} \leq C(1+\log(k))/H_i.$$

*If $\theta_{\mathcal{F}^j} \in V^k(\Omega_i)$ is discrete harmonic and vanishes at the nodes in $\partial\Omega_{i,h}$ except at those in $\mathcal{F}_h^j$ where it is equal to one, then*

$$|\theta_{\mathcal{F}^j}|^2_{H^1(\Omega_i)} \leq C(1+\log(k))H_i.$$

*Proof.* The proof of the first bound can be carried out as in Lemma 4.25 using the modifications developed in the proof of Lemma 7.14. The second follows directly from the fact that $\vartheta_{\mathcal{F}^j}$ has values in $[0,1]$; cf. Lemma 7.13 and the third from a trace theorem. Finally, the last inequality is a direct consequence of the minimum properties of discrete harmonic extensions. □

### 7.4.2 Algorithms and Condition Number Bounds

It is clear that every lemma in Sect. 4.6 now has a counterpart for spectral element functions. We can immediately find algorithms for spectral element approximations corresponding to those in Chap. 5 and 6, by observing that while for $h$ approximations we work with vectors of nodal values at the mesh nodes, here we work with vectors of nodal values at the GLL points. In particular, degrees of freedom and, equivalently, GLL nodes can be partitioned into internal, face, edge, and vertex nodes. Indeed, written in matrix form, the algorithms and operators are exactly the same.

Here we only consider two algorithms in more detail but refer to Chap. 5 and 6 for others and for a more thorough treatment.

We first consider the wire basket method in Sect. 5.4.2. Here we work with the space of piecewise discrete harmonic extensions $\tilde{V}^k \subset V^k$. Since the

values of these functions are uniquely determined by their value on $\Gamma$ and since Lemma 4.9 holds, we could equivalently work with functions defined on $\Gamma$. We note that

$$S : \tilde{V}^k \longrightarrow \tilde{V}^k.$$

We employ the preconditioned Schur complement

$$B^{-1}S = R_0^T \tilde{S}_{\mathcal{W}\mathcal{W}}^{-1} R_0 S + \sum_i R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} S.$$

The coarse component consists of the solution of a coarse problem in the subspace $\tilde{V}_{\mathcal{W}}^k \subset \tilde{V}^k$, a subspace which can be defined as the range of the interpolation operator $I_{\mathcal{W}}^k : \tilde{V}^k \to \tilde{V}_{\mathcal{W}}^k$, given by

$$I_{\mathcal{W}}^k u := \sum_{x_j \in \mathcal{W}_h} u(x_j)\theta_j + \sum_{\mathcal{F}^j \subset \Gamma} \bar{u}_{\partial\mathcal{F}^j} \theta_{\mathcal{F}^j}. \tag{7.7}$$

The first sum is extended to all GLL nodes on the wire basket $\mathcal{W}$, while the second to all subdomain faces on $\Gamma$. For a node $x_j$, $\theta_j$ is the discrete harmonic function that takes the value one at $x_j$ and vanishes at all other nodes on $\Gamma$, while, for a face $\mathcal{F}^j$, the discrete harmonic function $\theta_{\mathcal{F}^j}$ is defined in Lemma 7.15. As usual, the extension $R_0^T : \tilde{V}_{\mathcal{W}}^k \to \tilde{V}^k$, is the natural interpolation operator from the subspace to the global space. The approximate solver $\tilde{S}_{\mathcal{W}\mathcal{W}}^{-1}$ is defined by the bilinear form

$$\tilde{a}_0^{\mathcal{W}}(u,u) = \sum_i (1+\log(k))\rho_i \min_{\eta_i} \|u - \eta_i\|^2_{L^2(\mathcal{W}^i)};$$

cf. Equation (5.7). We note that, because of Lemma 7.3, the $L^2$-norm can be replaced by a discrete norm obtained using a quadrature formula based on the GLL nodes.

The local components of the preconditioner are associated with single faces. They are built on subspaces of discrete harmonic functions that vanish everywhere on $\Gamma_h$ except at the nodes in $\mathcal{F}_h^j$. Exact solvers are employed and the local bilinear forms are therefore the restrictions of $s(\cdot,\cdot)$ to these subspaces. A method for solving these local problems was given in Equation (5.3) and one for solving the coarse problem in Equation (5.9).

Using the same arguments as for Theorem 5.11, we can prove

**Theorem 7.16 (Wire basket method)** *There is a constant dependent of the aspect ratio of the mesh $\mathcal{T}$, but otherwise independent of the diameters of the elements in $\mathcal{T}$, the polynomial degree $k$, and the coefficients $\rho_i$, such that*

$$\kappa(B^{-1}S) \leq C\,(1+\log(k))^2.$$

We note that the above result was originally proven in [381, 382] using different techniques.

Similarly, all the other algorithms of Chap. 5 can be generalized to spectral element approximations on shape-regular meshes. In particular, we note that for the local edge and vertex components of the preconditioners of Sect. 5.2, the bilinear forms for edge and vertex components are

$$\tilde{a}_{\mathcal{E}^i}(u,u) = \rho_{\mathcal{E}^i}||u||^2_{L^2(\mathcal{E}^i)}, \quad \tilde{a}_{\mathcal{V}^j}(u,u) = \rho_{\mathcal{V}^j}||u||^2_{L^2(\mathcal{W})}.$$

In case the $L^2$-norms are replaced by discrete ones obtained by using a quadrature formula based on GLL points, the edge and vertex bilinear forms give rise to diagonal matrices. Similar considerations apply to the face based algorithms in Sect. 5.4.3.

**Theorem 7.17** *The results of Theorems 5.2, 5.5, 5.7, 5.12, and 5.16 are valid after replacing $H/h$ by $k$.*

We now turn to the Neumann-Neumann and FETI algorithms of Chap. 6. They can also be defined and analyzed in a straightforward way for spectral element approximations. Roughly speaking, the definitions of the various operators are the same, provided that the nodal interpolation operator $I^h$ is replaced by $I^k$. Here we only consider the balancing Neumann-Neumann method in detail.

As in Sect. 6.2, we introduce a trace space $W_i, i = 1, \ldots, N$, for each $\partial\Omega_i$, consisting of restrictions to the subdomain boundary of spectral element functions in $V^k(\Omega_i)$, and an associated product space $W := \prod_{i=1}^N W_i$. The subspace of functions which are continuous across the interface $\Gamma$ is denoted by $\widehat{W}$. The contributions $S^{(i)}$ to the Schur complement $S$ from the individual subdomains are computed as before; each of them is directly related to the component $W_i$ of the product space $W$. The weighted counting functions $\delta_i$ are defined as in (6.1):

$$\delta_i(x) = \frac{\sum_{j \in \mathcal{N}_x} \rho_j^\gamma(x)}{\rho_i^\gamma(x)}, \quad x \in \partial\Omega_{i,h},$$

where $\mathcal{N}_x$ is the set of subdomain indices to which the node $x$ belongs. The values at the GLL nodes in $\partial\Omega_{i,h}$ are then interpolated in order to obtain a spectral element function in $W_i$. The pseudoinverses $\delta_i^\dagger$ are then defined as in formula (6.2). An averaging operator similar to that of (6.4) can then be defined by

$$E_D u = \sum_{i=1}^N R_i^T I^k(\delta_i^\dagger u_i),$$

where $u_i \in W_i$ is the component of $u \in W$ associated with the subregion $\Omega_i$ and $R_i^T$ an interpolation operator as in Sect. 6.2.

The coarse space $W_0 \subset \widehat{W}$ is given by

$$W_0 = \text{span}\,\{R_i^T \delta_i^\dagger, \quad \partial\Omega_i \cap \partial\Omega_D = \emptyset\}$$

and, as for finite element approximations, we are free to extend this space by, e.g., adding basis functions associated with the boundary subdomains. The local space $W_i$ is simply the individual component of the product space and the bilinear form for $W_i$ is given by

$$\tilde{s}_i(u,v) = a_i(\mathcal{H}_i(\delta_i u), \mathcal{H}_i(\delta_i v));$$

cf. (6.6). Projection-like operators $P_i = R_i^T \tilde{P}_i$ are then defined with $\tilde{P}_i$ given by equation (6.7). In matrix form, we have

$$P_i = R_i^T \tilde{P}_i = R_i^T D^{(i)} S^{(i)\dagger} D^{(i)} R_i S,$$

where $D^{(i)}$ is a diagonal matrix that corresponds to the multiplication of a vector by the values of $\delta_i^\dagger$ and $S^{(i)\dagger}$ is a pseudoinverse of $S^{(i)}$ which provides a solution of a local Neumann problem. Here, we choose the solution that ensures $S^{(i)\dagger} u_i \in \mathit{Range}(S^{(i)})$, but note that other choices are possible; cf. (6.8).

The hybrid preconditioned operator $P_{hyl}$ is then defined in (6.10). We refer to Sect. 6.2 for details and some practical aspects.

The key to the theory is a stability property for the operator $E_D$ in Lemma 6.3. It relies on a Poincaré type inequality for certain functions; see Lemma 6.2. Here, we have

**Lemma 7.18** *Let $u \in \mathit{Range}(S^{(i)})$. Then,*

$$\|u\|^2_{L^2(\partial\Omega_i)} \leq C H_i \, |u|^2_{H^{1/2}(\partial\Omega_i)}, \quad i = 1, \ldots, N,$$

*with a constant that is independent of $u$, $H_i$, and $k$.*

*Proof.* The proof is very similar to that of Lemma 6.2 and we only consider the case of a substructure of unit diameter. Here, if $\Omega_i$ is a floating subdomain, we have

$$l_i(u) := k^{-3} \, (\mathbf{1}^T u) = 0 \quad \Longleftrightarrow \quad u \in \mathit{Range}(S^{(i)}),$$

where $\mathbf{1} \in W_i$ is a vector of ones; cf. (6.13). By using the lower bound in (B.7) for the weights of the quadrature formula on the GLL nodes and the equivalence in formula B.8, $l_i(\cdot)$ is a bounded linear functional in $W_i$:

$$\begin{aligned} |l_i(u)|^2 &\leq k^{-6} \textstyle\sum_{j,l} (w_j w_l)^{-1} \sum_{j,l} w_j w_l u(\xi_{jl})^2 \\ &\leq C k^{-6} \left( \textstyle\sum_{j,l} k^4 \right) \|u\|^2_{L^2(\partial\Omega_i)} \leq C \|u\|^2_{L^2(\partial\Omega_i)}, \end{aligned}$$

where the sums are taken over the GLL nodes $\xi_{jl}$ on $\partial\Omega_i$. Since the constant $C$ is independent of $k$, $l_i(\cdot)$ can be extended to the whole of $H^{1/2}(\partial\Omega_i)$ and the proof can be concluded as in Lemma 6.2. □

The previous lemma and the technical tools of the previous section enable us to prove a stability estimate for $E_D$ as in Lemma 6.3, with the ratio $H/h$ replaced by $k$. The same arguments as in Subsect. 6.2.3 then provides the following result.

**Theorem 7.19 (Balancing Neumann-Neumann method)** *The hybrid Schwarz method defined by the operator (6.10), and the spaces and bilinear forms of this subsection, satisfies*

$$s(u,u) \leq s(P_{hy1}u,u) \leq C(1+\log(k))^2 s(u,u),$$

*where the constant $C$ is independent not only of the polynomial degree and the number of substructures, but also of the values of the coefficients $\rho_i$.*

We note that a purely additive version of Neumann-Neumann methods for spectral element approximations was originally analyzed in [378] using different tools. One-level and dual-primal FETI methods can be developed and analyzed as in Chap. 6. We have:

**Theorem 7.20** *The results of Theorems 6.4, 6.15, 6.21, 6.35, and 6.38 are valid after replacing $H/h$ by $k$.*

## 7.5 Remarks on $p$ and $hp$ Approximations

### 7.5.1 More General $p$ Approximations

In the previous sections, we have extended overlapping and iterative substructuring methods to spectral element approximations. The development and analysis relied on the fact that degrees of freedom are *localized* and associated with GLL nodes. In addition, the use of GLL meshes, the spectral equivalences in Lemma 7.2, and the equivalence results for trilinear functions in Lemma 7.1 allowed us to prove Lemma 7.5 and the Sobolev type inequalities of Sect. 7.4.1.

Spectral elements basis functions present certain advantages but they are not the only choice for higher order approximations. They, indeed, provide a natural quadrature formula for the evaluation of the integrals. The use of quadrature does not spoil the accuracy of the solution and provides discrete norms that are equivalent to the original ones. These are all properties that are not guaranteed by other types of bases.

However, spectral element bases also have some disadvantages. They require tensor product meshes and are therefore not well suitable for triangles or tetrahedra. In addition, they are not hierarchical and are not particularly well suited for adaptive procedures; if a better approximation is needed and the polynomial degree $k$ is increased, the GLL nodes change and completely new basis functions need to be constructed. Other basis functions are more suitable for adaptive $p$ and $hp$ finite elements. They are hierarchical and allow different polynomial degrees for interior, face, and edge functions. We refer, e.g., to [415] for additional details.

Basis functions currently used in $p$ and $hp$ codes generally satisfy the following assumption:

**Assumption 7.21** *The shape functions on an element $K \in \mathcal{T}$ can be partitioned into* internal, face, edge, *and* vertex *functions.*

1. *An* internal *shape function vanishes on $\partial K$ and outside $K$.*
2. *A* face *function, associated with a face $\mathcal{F}$, vanishes on $\partial\mathcal{F}$ and on the rest of the interface.*
3. *An* edge *function, associated with an edge, vanishes on the rest of the interface except for that edge and the faces that share $\mathcal{E}$. In particular, it vanishes at the end points of $\mathcal{E}$.*
4. *A* vertex *shape function is a basis functions of the low order space $\mathbb{P}_1(K)$, for tetrahedral meshes, or $\mathbb{Q}_1(K)$, for meshes consisting of mapped cubes.*

In two dimensions, we only have interior, edge, and vertex functions.

Even though basis functions are still associated with the element vertices, edges, faces, and the interior, they are not in general associated with nodes, except for the vertex functions. For tensor product meshes, they are often chosen as tensor products of one dimensional functions

$$\left\{ \frac{1-x}{2}, \frac{1+x}{2}; \int_{-1}^{x} L_j(t)dt, \quad j = 1, \ldots, k-1 \right\} \subset \mathbb{P}_k(-1,1),$$

where $L_j$ is the Legendre polynomial of degree $j$.

We first consider overlapping methods. Since basis functions are not localized and are not associated with single nodes, we cannot in general have the great flexibility in the choice of the overlapping partition into subdomains possible for spectral elements; see Sect. 7.3. Indeed, overlapping subdomains now need to be unions of *entire* elements. This was indeed the method proposed by Pavarino in [375, 376]. There, any subdomain is taken as the union of the elements that have a subdomain vertex in common. The overlap is thus of order $H$, and a uniform bound is obtained for the corresponding preconditioner. However, for high polynomial degree, this procedure forces us to take subdomains with a large number of degrees of freedom in three dimensions and thus may result in very large local problems.

We now turn our attention to iterative substructuring methods. In two dimensions, two-level preconditioners can be devised in a straightforward way and satisfactory bounds can be obtained for *all* choices of basis functions that satisfy Assumption 7.21. After eliminating the internal degrees of freedom, we obtain a Schur complement system involving degrees of freedom associated with edges and vertices:

$$S = \begin{pmatrix} S_{\mathcal{E}\mathcal{E}} & S_{\mathcal{E}\mathcal{V}} \\ S_{\mathcal{E}\mathcal{V}}^T & S_{\mathcal{V}\mathcal{V}} \end{pmatrix};$$

see Sect. 5.3 and 5.4.1, and in particular Remark 5.4. A preconditioner is then obtained by removing the couplings between vertices and edges and between all pairs of edges:

$$B^{-1} = \sum_{\mathcal{E}^i} R_{\mathcal{E}^i}^T S_{\mathcal{E}^i\mathcal{E}^i}^{-1} R_{\mathcal{E}^i} + R_{\mathcal{V}}^T S_{\mathcal{V}\mathcal{V}}^{-1} R_{\mathcal{V}} = \sum_{\mathcal{E}^i} R_{\mathcal{E}^i}^T S_{\mathcal{E}^i\mathcal{E}^i}^{-1} R_{\mathcal{E}^i} + R_H^T S_H^{-1} R_H,$$

We note that vertex functions are the linear or bilinear shape functions associated with the vertices. They therefore span a coarse space $V_H$ and no change of basis is necessary; see Sect. 5.4.1.

For the analysis of this method, we refer to Babuška, Craig, Mandel, and Pitkäranta [31]. We recall that the analysis of iterative substructuring methods basically relies on three tools: the first provides stability estimates for functions associated with the wire basket or single edges (single vertices in two dimensions); cf. Lemmas 4.16 and 4.19 (or 4.15 in two dimensions). In particular, we recall that the three-dimensional result in Lemma 4.16 is a consequence of the two-dimensional inverse inequality in Lemma 4.15. The second tool is a decomposition result for components associated with single faces (edges in two dimensions); cf. Lemma 4.24. The third is the construction of a stable discrete extension of trace functions from the boundaries of the substructures; cf. Lemmas 4.10 and 4.9.

These tools are also available for two-dimensional polynomial functions and are given in [31] for triangular and quadrilateral meshes. For simplicity, we only consider quadrilateral meshes. Similarly to Lemma 4.15, for every substructure $\Omega_i$, we have that

$$\|u\|^2_{L^\infty(\Omega_i)} \leq C\,(1+\log(k))\|u\|^2_{H^1(\Omega_i)}, \quad u \in \mathbb{Q}_k(\Omega_i); \tag{7.8}$$

see [31, Th. 6.2]. The result for edge components corresponding to Lemma 4.24 is replaced by the inequality

$$\|u\|^2_{H^{1/2}_{00}(\mathcal{E})} \leq |u|^2_{H^{1/2}(\mathcal{E})} + C\,(1+\log(k))\|u\|^2_{L^\infty(\mathcal{E})}, \quad u \in \mathbb{Q}_k(\Omega_i), \tag{7.9}$$

which is valid for functions $u \in \mathbb{Q}_k(\Omega_i)$ that vanish on $\partial\Omega \setminus \mathcal{E}$, with $\mathcal{E}$ an edge of $\Omega_i$; see [31, Th. 6.6]. We refer to appendix A.2 for the definition of the trace spaces and norms; see in particular Remark A.9. Stable polynomial extensions can be found in [31, Sect. 7].

The three results mentioned above allow us to prove that the condition number of the two level preconditioner satisfies

$$\kappa(B^{-1}S) \leq C\,(1+\log(k))^2,$$

for *any* choice of basis functions that satisfy Assumption 7.21; see [31, Sect. 3] for details.

The situation is however much less favorable in three dimensions. Here, we can partition the Schur complements with respect to the wire basket degrees of freedom (associated with edges and vertices) and with face degrees of freedom:

$$S = \begin{pmatrix} S_{\mathcal{FF}} & S_{\mathcal{FW}} \\ S^T_{\mathcal{FW}} & S_{\mathcal{WW}} \end{pmatrix},$$

as in Sect. 5.4.2. A wire basket preconditioner is obtained by removing the couplings between the wire basket and the faces and between all pairs of faces.

After a change of basis for the wire basket basis functions and the introduction of an inexact solver, we obtain

$$B^{-1} = R_0^T \bar{S}_{\mathcal{W}\mathcal{W}}^{-1} R_0 + \sum_i R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i}.$$

However, unless a spectral element basis is employed, the condition number of $B^{-1}S$ may be very large. In [417], numerical tests are performed on a single tetrahedral element and a particular choice of basis functions. The condition number of $B^{-1}S$ is shown to be of the order of $k^{5/2}$ and is primarily due to the strong coupling between wire basket and face components.

We note that this strong coupling prevents us from finding Sobolev type inequalities as in Sect. 4.6 with good quality constants. Consequently all the algorithms in Chap. 5 and 6 are bound to show unfavorable condition numbers. In particular, we know from Assumption 2.2 and the work in Chap. 5 that if we try to develop effective primal iterative substructuring methods, it is necessary work with subspaces which allow for a stable decomposition of arbitrary elements of our space $V$. In fact, the failure to do so will necessarily lead to small eigenvalues of the operator $P_{ad}$; cf. Lemma 2.5. It is therefore natural, as originally suggested by Mandel [333], to change the basis to ensure a weaker coupling between wire basket and face components and between different face components.

These ideas were explored theoretically and experimentally in the thesis of Bică [52, 53]. He principally worked on extending the wire basket method to $p-$method finite elements using individual tetrahedral elements as subdomains. The goal was to develop effective methods for which the condition number can be bounded by $C(1+\log(p))^2$ with $C$ independent of the number of elements as well as the degree $p$ of the piecewise polynomials.

We recall that in Subsect. 7.4.1 of this chapter, we systematically developed technical tools similar to those of Sect. 4.4 and 4.6. For the spectral element case, there is complete success and all results on iterative substructuring methods previously developed for lower order finite elements are therefore equally valid for spectral elements. For the $p-$method there are many new technical difficulties. However, an extension theorem for polynomials, similar to that of Lemma 4.10, has been developed for three dimensions by Muñoz-Sola [355] by working with a particular extension operator. That work followed earlier similar work in two dimensions by Babuška, Craig, Mandel, and Pitkäranta [31] and Maday [328], which includes a bound in $H^{1/2}(T)$ of a polynomial extension in terms of the $L^2-$norm over the boundary of the triangle $T$. Another classical tool is an inequality of Markov; see Rivlin [397]. Bică's work, which uses all these tools, led to a $C(1+\log(p))^4$ bound for one of his wire basket algorithms; he also presents experimental evidence that $C(1+\log(p))^2$ should be possible.

While, generally, the development of efficient iterative substructuring preconditioners for general $p$ approximations remains less advanced than for lower

order and spectral elements, there has been additional work that deserves attention. Casarin and Sherwin [417] successfully explored a change of basis for a different higher order method on tetrahedra. We also wish to mention Guo and Cao [123, 248] and Korneev and Jensen [296], Korneev, Langer, and Xanthis [295].

### 7.5.2 Extensions to *hp* Approximations

In *hp* finite element approximations higher accuracy of the discrete solution is achieved by refining the mesh and/or by increasing the polynomial degree. We introduce a finite element triangulation $\mathcal{T} = \mathcal{T}_h$, consisting of triangles or tetrahedra, or affinely-mapped squares or cubes, and of maximum diameter $h$. Piecewise polynomial functions of degree $k$ are then employed. We note that different polynomial degrees may also be chosen, on different elements, as well as on different edges and faces; cf. Assumption 7.21. In addition certain types of nonconforming meshes are possible. We refer to, e.g., [415] for details. The same basis functions discussed previously may be employed as well as spectral element basis functions.

In recent years, there has been some work on the extension of some domain decomposition preconditioners to *hp* approximations. Here, we mention Ainsworth [11], Guo and Cao [243, 244, 245, 246], Oden, Patra, and Feng [364], Le Tallec and Patra [310], Korneev, Flaherty, Oden and Fish [294], Ainsworth and Sherwin [12], Bauer and Patra [38], Pardo and Demkowicz [371], Korneev, Langer, and Xanthis [295], and Toselli and Vasseur [448]. Additional material can be found in the references therein. Here, we only briefly describe the two-dimensional algorithm by Guo and Cao in [244] as an example.

We consider a nonoverlapping partition into subdomains and assume that the substructures are unions of element of the fine mesh $\mathcal{T}_h$. As usual, we also assume that they are unions of elements of a coarse mesh $\mathcal{T}_H$ that is contained in $\mathcal{T}_h$.

We eliminate the degrees of freedom internal to the substructures and obtain a modified system involving the Schur complement $S$ for the degrees of freedom on the interface $\Gamma$ between the substructures. We note that Assumption 7.21 ensures that the vertex basis functions span the low order space of piecewise linear functions on the fine mesh $\mathcal{T}_h$. The idea in [244] is to use the edge-based preconditioner of Remark 5.4 for low order $h$ approximations. This consists of a coarse solve on a low order coarse space on $\mathcal{T}_H$ and local solvers associated with single subdomain edges acting on piecewise linear functions that are supported on single edges. In order to obtain a decomposition of the *hp* space, higher order components associated with single edges $e$ of the fine mesh must also be accounted for; we use local spaces associated with single fine edges on $\Gamma$ and define a preconditioner by

$$B^{-1} = R_H^T S_H^{(h)-1} R_H + \sum_{\mathcal{E}^i \subset \Gamma} R_{\mathcal{E}^i}^T S_{\mathcal{E}^i \mathcal{E}^i}^{(h)-1} R_{\mathcal{E}^i} + \sum_{\mathcal{E}^i \subset \Gamma} \sum_{e \subset \mathcal{E}^i} R_e^T S_{ee}^{(k)-1} R_e.$$

The first term involves the solution of a coarse problem on a piecewise linear space on $\mathcal{T}_H$. The second involves the solution of a Dirichlet problem for piecewise linear elements on unions of two substructures that share an edge, while the third involves the solution of Dirichlet problems for $p$ finite elements on unions of the two fine elements in $\mathcal{T}_h$ that share an edge $e \subset \Gamma$. The results for low order finite elements in Sect. 4.6 and a generalization of (7.8) to $hp$ finite element functions in [244, Lem. 4.2] makes it possible to prove that the condition number of the additive operator satisfies:

$$\kappa(B^{-1}S) \leq C\,(1 + \log(Hk/h))^2.$$

We note that $(H/h)^n$ provides a measure of the number of elements of a substructure, while $k^n$ is a measure of the degrees of freedom associated with a single element. Therefore, $(Hk/h)^n$ measures the number of degrees of freedom associated with a single substructure. We refer to [244] for a detailed analysis and additional comments.

Preconditioning three-dimensional $hp$ approximations presents the same problems and difficulty as for $p$ approximations that employ basis functions not associated with GLL nodes. In conclusion, we mention the work in [444, 445, 446] where certain Neumann-Neumann and FETI preconditioners are extended to some $hp$ approximations in two and three dimensions on anisotropic meshes. If basis functions on GLL nodes are employed and the subdomains are suitably chosen, the condition numbers retain the logarithmic dependence in the polynomial degree and they remain independent of arbitrarily large aspect ratios of the mesh.

# 8

# Linear Elasticity

## 8.1 Introduction

The equilibrium of linear elastic material is described by a system of elliptic equations. We refer to appendix A.6.2 for an introduction. In this chapter, we only consider three-dimensional problems. With $\Omega \subset \mathbb{R}^3$ a polyhedral domain and with $\partial\Omega_D$ a subset of positive measure of its boundary $\partial\Omega$, we introduce the space

$$H_0^1(\Omega, \partial\Omega_D)^3 := \{\mathbf{v} \in H^1(\Omega)^3 : \mathbf{v}|_{\partial\Omega_D} = 0\}.$$

We next consider the problem of finding $\mathbf{u} \in H_0^1(\Omega, \partial\Omega_D)^3$, such that

$$a(\mathbf{u}, \mathbf{v}) := \int_\Omega 2\mu(x)\epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\ dx + \int_\Omega \lambda(x)\mathrm{div}\mathbf{u}\ \mathrm{div}\mathbf{v}\ dx\ = <\mathbf{F}, \mathbf{v}>, \quad (8.1)$$

for all $\mathbf{v} \in H_0^1(\Omega, \partial\Omega_D)^3$. Here $\epsilon_{ij}(\mathbf{u}) = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})$ is the linearized strain tensor, $\lambda$ and $\mu$ are the Lamé parameters, which are positive functions, and the inner products are defined as

$$\epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) = \sum_{i,j=1}^{3} \epsilon_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{v}),$$

$$<\mathbf{F}, \mathbf{v}> = \int_\Omega \sum_{i=1}^{3} f_i v_i\ dx + \int_{\partial\Omega_N} \sum_{i=1}^{3} g_i v_i\ ds,$$

where $\partial\Omega_N = \partial\Omega \setminus \partial\Omega_D$.

We assume that the Poisson ratio $\nu = \frac{\lambda}{2(\lambda+\mu)}$ is bounded away from 1/2. As remarked in appendix A.6.2, in this case, we have

$$\lambda = \frac{2\nu}{1-2\nu}\mu \leq C\mu. \quad (8.2)$$

The elliptic system (8.1) has much in common with our scalar elliptic model problem discussed in the previous chapters. Thus, it can be discretized

effectively by low order finite element methods. Methods based on mapped quadrilaterals and parallelepipeds are in particular frequent use. Let $\mathbf{V}^h$ be the corresponding finite element space of vector functions that vanish on $\partial\Omega_D$. We note that the resulting stiffness matrices are often very much more ill-conditioned than those of scalar elliptic problem and that simple iterative methods such as Jacobi-conjugate gradient methods are utterly inefficient.

The basic theory, in the constant coefficient case, for this problem is summarized in appendix A.6.2; see also Ciarlet [137] for a detailed treatment of nonlinear and linear elasticity. While the null space of a scalar elliptic problem consists of constants, we now have a six-dimensional null space of rigid body modes. This is reflected in the discussion of pure Neumann problems on floating subdomains where the data has to satisfy the compatibility condition given in (A.38) and the solution is defined only up to a rigid body mode. This will have important consequences when we redesign our overlapping as well as our iterative substructuring methods.

The space of rigid body modes, $\mathcal{RB}$, is spanned by the $\mathbf{r}_j$: three translations

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{r}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{r}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{8.3}$$

and three rotations

$$\mathbf{r}_4 = \begin{bmatrix} 0 \\ x_3 \\ -x_2 \end{bmatrix}, \quad \mathbf{r}_5 = \begin{bmatrix} x_3 \\ 0 \\ -x_1 \end{bmatrix}, \quad \mathbf{r}_6 = \begin{bmatrix} x_2 \\ -x_1 \\ 0 \end{bmatrix}. \tag{8.4}$$

This is a good basis if we consider a region centered at the origin; in other cases, it is natural to replace $x_i$ in the rotations $\mathbf{r}_4, \mathbf{r}_5, \mathbf{r}_6$ by $x_i - x_i^0$ where $x^0$ is a point in a relevant region. We note that it is easy to show that both the divergence and the linearized strain tensor of all rigid body modes vanish. The rigid body modes are all linear functions and they thus are contained in the polynomial space $(\mathbb{P}_1)^3$ and indeed in any conforming finite element space that we might consider for Equation (8.1). An important consequence of this fact is that the null space of the stiffness matrix $A^{(i)}$ of any floating subdomain $\Omega_i$ is equal to $\mathcal{RB}$. Similarly, since the rigid body modes have no elastic energy, the Schur complement $S^{(i)}$ of such a subregion has the same null space. We recall that in order to construct a scalable algorithm, we must include the entire null space of the operator in a coarse, global space; see Mandel [332] or Smith, Bjørstad, and Gropp, [424, p. 132] for a discussion of this *null space property*.

In separate sections, we will consider two-level overlapping Schwarz methods and a wire basket based iterative substructuring method, first considered in Pavarino and Widlund [383] in the spectral element as well as lower order finite element case. Some modifications are required of the algorithms, developed for the scalar case, to deal with the richer null space. We will

also show how balancing Neumann-Neumann and several FETI methods can be extended to compressible linear elasticity. The extension of the balancing and one-level FETI algorithms and theory turns out to be routine while the redesign of the dual-primal FETI methods is more challenging; new results have been obtained only recently by Klawonn and Widlund [290, 291]. All our results closely parallel those for the scalar elliptic problems of the previous chapters.

## 8.2 A Two-Level Overlapping Method

We will now examine Chap. 3 to see what changes of the algorithms and theory will be required. We will primarily consider the algorithm of Sect. 3.6. We will also adopt the same assumptions.

We introduce an overlapping partition into subdomains $\{\Omega_i'\}$, as in Sect. 3.2, satisfying Assumptions 3.1 and 3.2. The local problems are then defined just as in Sect. 3.2 in terms of this overlapping decomposition after replacing the bilinear form by that of problem (8.1).

We next introduce a shape-regular, but not necessarily quasi-uniform, coarse mesh $\mathcal{T}_H$. We assume that Assumption 3.5 holds and recall that in Sect. 3.3 the coarse space $V^H$ is chosen to be a continuous, piecewise linear, or trilinear finite element space on a $\mathcal{T}_H$. This space will serve equally well for each of the components of the elasticity problem; all rigid body modes can be represented exactly in the resulting space.

As for the technical tools given in Sect. 3.2 and 3.5, only the construction of the quasi-interpolant $\tilde{I}^H$ needs to be modified so as to reproduce the entire null space $\mathcal{RB}$. As already pointed out in Remark 3.7, this can be accomplished straightforwardly by a simple change in the definition of the values at the vertices of the coarse mesh to make the new operator $\tilde{I}^H$ reproduce all linear functions. More precisely, with the same definitions as in Sect. 3.5, let $y$ be a node of the coarse mesh $\mathcal{T}_H$, and $\omega_y$ be the union of the elements in $\mathcal{T}_H$ that share $y$. Let in addition $\pi_y : L^2(\omega_y) \to \mathbb{P}_1(\omega_y)$, be the $L^2$-projection onto the space of *linear* functions on $\omega_y$. For every vertex $y$, we define

$$(\tilde{I}^H u)(y) = \begin{cases} 0, & y \in \partial\Omega, \\ (\pi_y u)(y) & \text{otherwise}, \end{cases}$$

It is a simple matter to check that, if $u$ is a linear function on $\omega_K$, with $K \in \mathcal{T}_H$, then $\tilde{I}^H u = u$ on $K$. Here, $\omega_K$ is the union of $K$ and the elements, the boundaries of which have a nonempty intersection with $\partial K$. We use the same notation for the operator acting on vector-valued functions, which is obtained by applying $\tilde{I}^H$ component by component.

If $\mathbf{H}(\omega_K) = H^1(\omega_K)^3/\mathcal{RB}$ is the quotient space introduced in appendix A.6.2, then by using similar arguments as in Lemma 3.6, we obtain:

**Lemma 8.1** *Let $\mathcal{T}_H$ be shape regular and let $\tilde{I}^H$ be the modified quasi-interpolant of Remark 3.7 discussed above. Then, there exists a constant $C$ such that*

$$\|\tilde{I}^H \mathbf{u} - \mathbf{u}\|_{L^2(K)} \leq C H_K |\mathbf{u}|_{\mathbf{H}(\omega_K)}, \tag{8.5}$$

$$|\tilde{I}^H \mathbf{u}|_{\mathbf{H}(\omega_K)} \leq C |\mathbf{u}|_{\mathbf{H}(\omega_K)}. \tag{8.6}$$

It is now straightforward to modify the proof of Theorem 3.13 to obtain:

**Theorem 8.2** *In case exact solvers are employed on all subspaces, the condition number of the additive Schwarz operator for the linear elasticity problem satisfies*

$$\kappa(P_{ad}) \leq C \left(1 + \frac{H}{\delta}\right),$$

*where $C$ depends on $N^c$, the number of colors of the decomposition $\{\Omega_i'\}$, but is otherwise independent of $h$, $H$, and $\delta$. However, we cannot guarantee that the estimate is independent of variations in the Lamé parameters $\mu$ and $\lambda$.*

We will not discuss the modification necessary for the methods of Sect. 3.10. We only note that the coarse spaces need to be enriched and that therefore the point of departure would be a set of new coarse basis functions for the coarse space. We could take a set $\{\Phi_i \mid 1 \leq i \leq N\}$, developed for the scalar case and replace them by $\{I^h(\Phi_i \mathbf{r}_k) \mid 1 \leq i \leq N, 1 \leq k \leq 6\}$, where the $\mathbf{r_k}$ are the rigid body modes, given in (8.3) and (8.4), and $I^h$ is the interpolation operator which maps into the finite element space selected for the problem (8.1).

## 8.3 Iterative Substructuring Methods

We now turn to iterative substructuring methods. As for the overlapping methods considered in the previous section, we will find that the redesign of the algorithm for the problem at hand is very directly related to the larger null space of our system of elliptic equations. Otherwise, much of the framework of the discrete problem can remain the same. Thus, exactly as in Sect. 4.2, the given region $\Omega$ in three dimensional space is partitioned into nonoverlapping subdomains $\Omega_i$ which define an interface $\Gamma$, which consists of vertices, edges, and faces. As before, the wire basket $\mathcal{W}$ is the union of the edges and vertices of $\Gamma$.

As for the scalar elliptic problems, we consider a model problem where the coefficients, the Lamé parameters, are constant in each subdomain:

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \sum_{i=1}^{N} a_i(\mathbf{u}, \mathbf{v}) \\ &= \sum_{i=1}^{N} \left( 2\mu_i \int_{\Omega_i} \epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) \, dx + \lambda_i \int_{\Omega_i} \mathrm{div}\mathbf{u} \, \mathrm{div}\mathbf{v} \, dx \right). \end{aligned} \tag{8.7}$$

We recall that, by assumption, the substructures are polyhedral domains and unions of elements of a shape-regular fine mesh. We denote the standard finite element space of continuous, piecewise linear vector-valued functions on $\Omega_i$ by $\mathbf{V}^h(\Omega_i)$ and always assume that these functions vanish on $\partial\Omega_D$. We recall that the triangulation of each subdomain is assumed to be quasi uniform.

Since a substructure may share a single edge or face with $\partial\Omega_D$ (see Assumption 4.3), we need to characterize the kernels of local Neumann problems.

**Lemma 8.3** *The subspace of vectors* $\mathbf{u} \in \mathbf{V}^h(\Omega_i)$, *such that*

$$a_i(\mathbf{u}, \mathbf{u}) = 0,$$

*consists of:*

1. *the null vector if* $\partial\Omega_i \cap \partial\Omega_D$ *contains at least three points that are not aligned and thus, in particular, if the intersection contains one or several faces;*
2. *a one-dimensional space of rotations if* $\partial\Omega_i \cap \partial\Omega_D$ *contains two points on a straight line and no points outside this line, thus, in particular, if it consists of an edge;*
3. *the whole space* $\mathcal{RB}$ *if* $\partial\Omega_i \cap \partial\Omega_D$ *is empty.*

To simplify our discussion, we will assume that no subdomain has a boundary which intersects $\partial\Omega_D$ in just one or few points, or along an edge; in fact, we assume that such an intersection contains a whole subdomain face. We denote by $\mathcal{RB}_i$ the null space defined for the substructure $\Omega_i$ in the previous lemma. Because of the Korn inequalities in Lemma A.39 and the compressibility condition (8.2), we can equivalently work with $H^1$-seminorms, scaled only with the Lamé coefficient $\mu_i$, for certain local subspaces.

**Lemma 8.4** *There exist positive constants* $c$ *and* $C$, *which depend only on the shape of* $\Omega_i$ *and on the local Poisson's ratio* $\nu_i$, *such that*

$$c\mu_i|\mathbf{u}|^2_{H^1(\Omega_i)^3} \leq a_i(\mathbf{u}, \mathbf{u}) \leq C\mu_i|\mathbf{u}|^2_{H^1(\Omega_i)^3},$$

*for* $\mathbf{u}$ *in the quotient space* $\mathbf{V}^h(\Omega_i)/\mathcal{RB}_i$.

We note that in case $\Omega_i$ intersects $\partial\Omega$ in only an edge an additional logarithmic factor in $H/h$ is likely to arise in the constants, as in Lemma 4.21 and the discussion that follows. To simplify our discussion, we have excluded such cases.

## 8.4 A Wire Basket Based Method

In this section, we will describe a wire basket based method for the equations of linear elasticity. This algorithm is based on Algorithm 5.10, which is due

to Smith [420]. This method was first developed for the equations of elasticity in Pavarino and Widlund [383].

In our discussion of Algorithm 5.10, we began by introducing an operator $\hat{R}_{\mathcal{W}}$ the transpose of which maps values on the wire basket onto the faces and a closely related interpolation operator $I^h_{\mathcal{W}}$ defined in (5.6). As in the scalar case, the values obtained at the nodes of a face will depend only on the values on its boundary. This will make the elements of the range of this operator, which will serve as the coarse space of our Schwarz method, continuous across the interface $\Gamma$. While this extension operator is exceptionally simple in the scalar case, a more elaborate construction is required for the system of linear elasticity.

### 8.4.1 An Extension from the Interface

We will need new extension operators, $\mathcal{E}_i$, based on solving Dirichlet problems of the elastic problem on the individual substructures. We will call them the elastic extension operators. As in Sect. 4.4, such a function, which is defined in $\Omega_i$, is the solution of

$$A^{(i)}_{II}\mathbf{u}^{(i)}_I + A^{(i)}_{I\Gamma}\mathbf{u}^{(i)}_\Gamma = 0.$$

Such a solution can also be regarded as the element $\mathcal{E}_i(\mathbf{u}_\Gamma)$ which equals $\mathbf{u}_\Gamma$ on $\partial\Omega_i$ and which is orthogonal, in the $a_i(\cdot,\cdot)$–inner product, to the space $\mathbf{V}^h \cap (H^1_0(\Omega_i, \partial\Omega_i \cap \Gamma))^3$. We can also characterize this solution variationally as in Lemma 4.9.

### 8.4.2 An Extension from the Wire Basket

In the construction of our algorithm, we need an extension of finite element functions given on the wire basket to the faces. This is a local operation and we can therefore restrict our attention to the reference element $(-1,1)^3$; we will only consider the case of hexagonal substructures. A preliminary operator $\tilde{I}^h_{\mathcal{W}}$ is simply given by the restriction of $\mathbf{u}$ to the wire basket, i.e., by

$$\tilde{I}^h_{\mathcal{W}}\mathbf{u} = \sum_{x_k \in \mathcal{W}_h} \mathbf{u}(x_k)\phi_k.$$

This operator works component by component; we will use the same notation when it is applied to individual components and vector valued functions. Since the resulting functions vanish at all face mesh points, it is clear that this operator will not preserve any of the rigid body modes.

We next consider the difference between each of the $\mathbf{r}_j$ and $\tilde{I}^h_{\mathcal{W}}\mathbf{r}_j$. They can all be expressed in terms of four scalar functions, defined by

$$\mathcal{G}^0 = 1 - \tilde{I}^h_{\mathcal{W}}1, \quad \mathcal{G}^1 = x_1 - \tilde{I}^h_{\mathcal{W}}x_1, \quad \mathcal{G}^2 = x_2 - \tilde{I}^h_{\mathcal{W}}x_2, \quad \mathcal{G}^3 = x_3 - \tilde{I}^h_{\mathcal{W}}x_3.$$

These four functions vanish on the wire basket and each of them can be split into six face terms,

$$\mathcal{G}^0 = \sum_{k=1}^{6} \mathcal{G}_k^0, \qquad \mathcal{G}^1 = \sum_{k=1}^{6} \mathcal{G}_k^1, \qquad \mathcal{G}^2 = \sum_{k=1}^{6} \mathcal{G}_k^2, \qquad \mathcal{G}^3 = \sum_{k=1}^{6} \mathcal{G}_k^3.$$

Here, the $\mathcal{G}_k^j, j = 0, 1, 2, 3$, vanish on all faces except on $\mathcal{F}^k$. For each component $u_i$ of $\mathbf{u}$, we define a new extension $I_{\mathcal{W}}^h u_i$ from the wire basket to the interface as follows: on a face $\mathcal{F}^k$, for which the two relevant variables are $x_1$ and $x_2$, the restriction of $I_{\mathcal{W}}^h u_i$ to $\mathcal{F}^k$ has the form

$$I_{\mathcal{W}}^h u_i = \tilde{I}_{\mathcal{W}}^h u_i + a_{k,i}\mathcal{G}_k^0 + b_{k,i}^1 \mathcal{G}_k^1 + b_{k,i}^2 \mathcal{G}_k^2. \tag{8.8}$$

The weights $a_{k,i}, b_{k,i}^1$, and $b_{k,i}^2$ are chosen so that $I_{\mathcal{W}}^h$ preserves the functions $1, x_1, x_2, x_3$, and therefore the rigid body modes. The weights are given by

$$a_{k,i} = \frac{\int_{\partial\mathcal{F}^k} u_i ds}{\int_{\partial\mathcal{F}^k} 1 ds}; \qquad b_{k,i}^j = \frac{\int_{\partial\mathcal{F}^k} u_i x_j ds}{\int_{\partial\mathcal{F}^k} x_j^2 ds}, \quad j = 1, 2.$$

We note that on each face only three correction terms come into play; see (8.8).

A simple computation shows that, the new extension operator reproduces all linear functions, on each face, and therefore also all the rigid body modes. If, e.g., $u_i = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3$, we find that on the face $\mathcal{F}^k = \{x_3 = 1\}$, that

$$a_{k,i} = \frac{1}{8}(c_0 + c_1 x_1 + c_2 x_2 + c_3, 1)_{\partial\mathcal{F}^k} = c_0 + c_3,$$

$$b_{k,i}^1 = \frac{3}{16}(c_0 + c_1 x_1 + c_2 x_2 + c_3, x_1)_{\partial\mathcal{F}^k} = c_1,$$

$$b_{k,i}^2 = \frac{3}{16}(c_0 + c_1 x_1 + c_2 x_2 + c_3, x_2)_{\partial\mathcal{F}^k} = c_2,$$

as required.

For a vector valued displacement $\mathbf{u}$, the extension operator can now be defined as the discrete elastic extension of the vector with components given by (8.8) and by similar expressions for the other faces, i.e., by

$$I_{\mathcal{W}}^h \mathbf{u} = \mathcal{E}(I_{\mathcal{W}}^h u_1, I_{\mathcal{W}}^h u_2, I_{\mathcal{W}}^h u_3).$$

Any rigid body mode $\mathbf{r}$ is reproduced inside each element, i.e., $\mathcal{E}_i(\mathbf{r}) = \mathbf{r}$. This follows from the minimium property of the elastic extension and the fact that $a(\mathbf{r}, \mathbf{r}) = 0$. Therefore, $I_{\mathcal{W}}^h \mathbf{r} = \mathbf{r}, \quad \mathbf{r} \in \mathcal{RB}$.

The extension operator $I_{\mathcal{W}}^h$ defines a change of basis in $\mathbf{V}_\Gamma$; the face basis functions are unchanged, but the wire basket basis functions are transformed according to (8.8).

### 8.4.3 A Wire Basket Preconditioner for Linear Elasticity

We can now describe our wire basket preconditioner for linear elasticity problems in matrix form. We proceed as in Subsect. 5.4.1 and 5.4.2, to which we refer for more details.

We perform a change of basis in the space spanned by the wire basket functions in order to satisfy the null space property, i.e., in order to ensure that the null space of the local contribution $\widehat{S}^{(i)}$ to the preconditioner is the space of rigid body modes $\mathcal{RB}$. This can be accomplished by using the extension operator $I_{\mathcal{W}}^h$ defined by (8.8), since $I_{\mathcal{W}}^h$ reproduces all rigid body modes. In matrix form, this change of basis from the new basis to the original nodal basis is represented locally by the matrix

$$\begin{bmatrix} I_{\mathcal{FF}}^{(i)} & \hat{R}_{\mathcal{W}}^{(i)T} \\ 0 & I_{\mathcal{WW}}^{(i)} \end{bmatrix},$$

where the diagonal blocks are identity matrices of the appropriate order. Then, $S^{(i)}$ is transformed into

$$\begin{bmatrix} I_{\mathcal{FF}}^{(i)} & 0 \\ \hat{R}_{\mathcal{W}}^{(i)} & I_{\mathcal{WW}}^{(i)} \end{bmatrix} \begin{bmatrix} S_{\mathcal{FF}}^{(i)} & S_{\mathcal{FW}}^{(i)} \\ S_{\mathcal{FW}}^{(i)T} & S_{\mathcal{WW}}^{(i)} \end{bmatrix} \begin{bmatrix} I_{\mathcal{FF}}^{(i)} & \hat{R}_{\mathcal{W}}^{(i)T} \\ 0 & I_{\mathcal{WW}}^{(i)} \end{bmatrix} = \begin{bmatrix} S_{\mathcal{FF}}^{(i)} & \text{nonzero} \\ \text{nonzero} & \widetilde{S}_{\mathcal{WW}}^{(i)} \end{bmatrix}.$$

The local preconditioner $\widehat{S}^{(i)}$ is constructed by

a) eliminating the coupling between faces and the wire basket in the transformed matrix;

b) eliminating the coupling between all pairs of faces, i.e., by replacing $S_{\mathcal{FF}}^{(i)}$ by its block-diagonal part $\widehat{S}_{\mathcal{FF}}^{(i)}$;

c) replacing the wire basket block $\widetilde{S}_{\mathcal{WW}}^{(i)}$ by a simpler matrix $\widehat{S}_{\mathcal{WW}}^{(i)}$: replace $\widetilde{S}_{\mathcal{WW}}^{(i)}$ by a rank-six perturbation of the identity. On the reference element, it is given by

$$\widehat{S}_{\mathcal{WW}}^{(i)} = (1 + \log(H/h))\mu_i (I^{(i)} - \sum_{j=1}^{6} \frac{\mathbf{r}_j \mathbf{r}_j^T}{\mathbf{r}_j^T \mathbf{r}_j}). \tag{8.9}$$

(This corresponds to using a cheaper, approximate solver for the wire basket variables.)

d) returning to the original basis:

$$\widehat{S}^{(i)} = \begin{bmatrix} I_{\mathcal{FF}}^{(i)} & 0 \\ -\hat{R}_{\mathcal{W}}^{(i)} & I_{\mathcal{WW}}^{(i)} \end{bmatrix} \begin{bmatrix} \widehat{S}_{\mathcal{FF}}^{(i)} & 0 \\ 0 & \widehat{S}_{\mathcal{WW}}^{(i)} \end{bmatrix} \begin{bmatrix} I_{\mathcal{FF}}^{(i)} & -\hat{R}_{\mathcal{W}}^{(i)T} \\ 0 & I_{\mathcal{WW}}^{(i)} \end{bmatrix}. \tag{8.10}$$

The action of $\hat{R}_{\mathcal{W}}^{(j)T}$ and $\hat{R}_{\mathcal{W}}^{(i)T}$ on a face shared by two elements $\Omega_j$ and $\Omega_i$ is the same, because the extension of any function defined on the wire basket to a face, using the operator $I_{\mathcal{W}}^h$, is determined only by the values on the boundary of that face. Therefore the preconditioner can be obtained by subassembly. We obtain,

$$\widehat{S} = \begin{bmatrix} I_{\mathcal{FF}} & 0 \\ -\hat{R}_{\mathcal{W}} & I_{\mathcal{WW}} \end{bmatrix} \begin{bmatrix} \widehat{S}_{\mathcal{FF}} & 0 \\ 0 & \widehat{S}_{\mathcal{WW}} \end{bmatrix} \begin{bmatrix} I_{\mathcal{FF}} & -\hat{R}_{\mathcal{W}}^T \\ 0 & I_{\mathcal{WW}} \end{bmatrix},$$

and

$$\widehat{S}^{-1}S = R_0^T \widehat{S}_{\mathcal{WW}}^{-1} R_0 S + \sum_k R_{\mathcal{F}^k}{}^T \widehat{S}_{\mathcal{F}^k\mathcal{F}^k}^{-1} R_{\mathcal{F}^k} S,$$

where $R_0 = (\hat{R}_{\mathcal{W}}, I_{\mathcal{WW}})$; see Subsect. 5.4.2 or Dryja, Smith, and Widlund [178].

We define the usual orthogonal projections $P_{\mathcal{F}^k}$ onto the face spaces and a projection-like operator $P_0$, defined by a bilinear form $\tilde{a}_0(\cdot,\cdot)$ consistent with (8.9), onto the wire basket space, and obtain the additive Schwarz operator

$$P = P_0 + \sum_{\mathcal{F}^k} P_{\mathcal{F}^k}.$$

**Theorem 8.5** *The condition number of the iteration operator $P$ is bounded by*

$$\kappa(P) \leq C(1 + \log(H/h))^2,$$

*where $C$ is a constant independent not only of the mesh size and the number of substructures, but also of the values of the Lamé parameters $\mu_i$ and $\lambda_i$ of our model problem (8.7) provided that the Poisson ratios $\nu_i$ are uniformly bounded away from* $1/2$.

The proof follows, step by step, that of Theorem 5.11; for full details, see [383].

## 8.5 Neumann-Neumann and FETI Methods

In this section, we will first consider a balancing Neumann-Neumann and one-level FETI methods previously discussed in Sect. 6.2 and 6.3, respectively. As in the previous sections of this chapter, the changes of the algorithms and analysis are very directly related to the larger null space of the elliptic system (8.1). We recall that the key to the analysis of these two families of algorithms for scalar elliptic problems is Lemma 6.3; such a result also holds for the problem at hand and we will show that only a few changes of the algorithms are required. In contrast, the FETI-DP algorithm considered in the final subsection of this chapter will require more extensive modifications of the algorithms. As in Chap. 6, we will change notations and denote our finite element space of continuous, piecewise linear functions by $\mathbf{W}^h$, etc.

### 8.5.1 A Neumann-Neumann Algorithm for Linear Elasticity

As in Sect. 6.2, given a finite element approximation of the elliptic system (8.1) and a partition of the region $\Omega$ into substructures, we can introduce a finite

element trace space $\mathbf{W}_i, i = 1, \ldots, N$, for each $\partial\Omega_i$ and an associated product space $\mathbf{W} := \prod_{i=1}^{N} \mathbf{W}_i$. The subspace of vector valued functions which are continuous across the interface $\Gamma$ is denoted by $\widehat{\mathbf{W}}$. A stiffness matrix, $A$, and a Schur complement, $S$, are then introduced. The contributions $S^{(i)}$ to $S$ from the individual subdomains are computed as before; each of them is directly related to the component $\mathbf{W}_i$ of the product space $\mathbf{W}$. The weighted counting functions $\delta_i$ that are appropriate for our elasticity problem are based on the Lamé parameter $\mu_i$ of the subdomains, which replace the $\rho_i$ in the definition (6.1):

$$\delta_i(x) = \frac{\sum_{j \in \mathcal{N}_x} \mu_j^\gamma(x)}{\mu_i^\gamma(x)}, \quad x \in \partial\Omega_{i,h} \cap \Gamma_h.$$

As in Lemma 8.4 only one of the Lamé parameters enters this definition. The pseudoinverses $\delta_i^\dagger$ are then defined as in formula (6.2). An operator similar to that of (6.4) can then be defined by

$$E_D \mathbf{u} = \sum_{i=1}^{N} R_i^T I^h(\delta_i^\dagger \mathbf{u}_i), \tag{8.11}$$

where $\mathbf{u}_i \in \mathbf{W}_i$ is the component of $\mathbf{u} \in \mathbf{W}$ associated with the subregion $\Omega_i$ and $R_i^T$ an interpolation operator as in Sect. 6.2.

The larger null space must be accommodated when we redefine the subspaces and bilinear forms for the balancing Neumann-Neumann algorithm of Sect. 6.2. We first modify the coarse space; cf. (6.5). The new $\mathbf{W}_0 \subset \widehat{\mathbf{W}}$ is given by

$$\mathbf{W}_0 = \text{span}\,\{R_i^T I^h(\delta_i^\dagger \mathbf{r}_k), \quad \partial\Omega_i \cap \partial\Omega_D = \emptyset, \quad k = 1, \ldots, 6\}; \tag{8.12}$$

we note that this space contains the entire space of rigid body modes. As in the scalar case, we are free to enrich this space, e.g., by adding basis functions associated with the boundary subdomains. The local space $\mathbf{W}_i$ is simply the individual component of the product space $\mathbf{W}$ and the bilinear form for $\mathbf{W}_i$ is given by

$$\tilde{s}_i(\mathbf{u}, \mathbf{v}) \;=\; a_i(\mathcal{E}_i(\delta_i \mathbf{u}), \mathcal{E}_i(\delta_i \mathbf{v})),$$

where $\mathcal{E}_i$ are the discrete elastic extension operators introduced in Subsect. 8.4.1; cf. also (6.6). Projection-like operators $P_i = R_i^T \tilde{P}_i$ are then defined with $\tilde{P}_i$ given by equation (6.7). Finally, the constraint (6.9) is replaced by the requirement that $I^h(\delta_i \tilde{P}_i \mathbf{u})$ belongs to $range(S^{(i)})$, a space of codimension six.

The key to the theory is a lemma that replaces Lemma 6.3:

**Lemma 8.6** *Let $E_D$ be the operator defined in formula (8.11). Then,*

$$|E_D \mathbf{w}|_S^2 \leq C(1 + \log(H/h))^2 |\mathbf{w}|_S^2, \quad \mathbf{w} \in range(S). \tag{8.13}$$

The proof of this lemma requires no new ideas: using Lemma 8.4, we can estimate the individual components of $E_D\mathbf{w}$ using the same tools as in the scalar elliptic case. Similarly, no new ideas are required to establish:

**Theorem 8.7** *The hybrid Schwarz method defined by the operator (6.10), and the spaces and bilinear forms of this subsection, satisfies*

$$s(\mathbf{u},\mathbf{u}) \;\leq\; s(P_{hyl}\mathbf{u},\mathbf{u}) \;\leq\; C(1+\log(H/h))^2 s(\mathbf{u},\mathbf{u}),$$

*where $C$ is independent not only of the mesh size and the number of substructures, but also of the values of the Lamé parameters $\mu_i$ and $\lambda_i$ of our model problem (8.7) provided that the Poisson ratios $\nu_i$ are uniformly bounded away from* $1/2$.

### 8.5.2 One-Level FETI Algorithms for Linear Elasticity

These FETI methods can easily be extended to problems of elasticity. Examining Sect. 6.3, we find that the matrix $R$ of vectors of elements of *kernel*$(S)$, defined in (6.28), must be expanded to account for all the rigid body modes of the floating subdomains, characterized by Lemma 8.3; cf. Remark 6.6. As in the scalar elliptic case, Lemma 8.6 plays a central role in the theory for one-level FETI methods. All other changes are completely routine. In the fully redundant case, e.g., we obtain a result which is completely analogous to Theorem 6.21:

**Theorem 8.8** *The preconditioner $\widehat{M}_r$, defined as in (6.51), and with $Q = \widehat{M}_r^{-1}$, satisfies*

$$\langle \widehat{M}_r\lambda,\lambda\rangle \leq \langle F_r\lambda,\lambda\rangle \leq C(1+\log(H/h))^2\langle \widehat{M}_r\lambda,\lambda\rangle, \quad \lambda\in V.$$

*Here $C$ is a constant independent of $h, H, \gamma$, and the values of the Lamé parameters $\mu_i$ and $\lambda_i$ of our model problem (8.7) provided that the Poisson ratios $\nu_i$ are uniformly bounded away from* $1/2$.

### 8.5.3 FETI-DP Algorithms for Linear Elasticity

We first recall that this family of algorithms was first introduced by Farhat, Lesoinne, and Pierson [195]; their versions of the algorithm have been used extensively to solve very large and difficult problems. As in Sect. 6.4, we will now try to identify subspaces $\widetilde{\mathbf{W}}$ with a sufficiently rich set of primal constraints that will allow us to prove satisfactory bounds on the condition numbers of our FETI-DP algorithms. We recall that for scalar elliptic problems and Algorithm D, we always make sure that the boundary of each face of the interface $\Gamma$ contains at least one primal edge. In addition, it is sometimes necessary to make some vertices and additional edges primal in order to obtain bounds

that are not strongly dependent on the values of the coefficients in the different subdomains. In order to understand what might be appropriate in the present context, we will first consider two special cases.

In the first, we assume that we have two subdomains made of the same material, which have a face in common and are surrounded by subdomains made of a material with much smaller Lamé parameters. Such a problem will clearly have six low energy modes related to the rigid body modes of the union of the two special subdomains. Since six rigid body modes are associated with each substructure, any preconditioner that has less than six $(12-6)$ linearly independent primal constraints across that face will have at least seven low energy modes and is likely to be poor.

This observation suggests that we might introduce six linearly independent constraint across each face of the interface. Given that point constraints have been shown to be ineffective in Sect. 6.4, we will choose the constraints in terms of averages (and first order moments) over edges of the components of the displacement vector. One can show that for a rectangular face, averages from three (or four) different edges are needed to obtain the required set of six linearly independent constraints. In fact, what matters is that the set of constraints control all rigid body modes in the sense that if the functionals which represent the constraints all vanish for a rigid body mode, then the rigid body must vanish. A face equipped with such a full set of primal constraints is call *fully primal.*

In the second case, we again consider two subdomains surrounded by subdomains with much smaller Lamé parameters. We now assume that the two special subdomains share only an edge. In this case, there are seven low energy modes of the finite element model corresponding to the same rigid body modes as before and an additional one. The new mode corresponds to a relative rotation of the two subdomains around their common edge. We conclude that in such a case, we should introduce five $(12-7)$ linearly independent primal constraints for the special edge since without any constraints, we would have twelve low energy modes. Such edges will be called *fully primal* in our discussion.

Three of the five constraints of a fully primal edge $\mathcal{E}^{ik}$ are given in terms of the averages over the edge of the three displacement components. We note that one rigid body mode, corresponding to a rotation around the edge vanishes on the edge. The additional two constraints, which completes our full set of constraints for this fully primal edge, are given by the $L^2(\mathcal{E}^{ik})$–inner product of the other two rotational rigid body modes with the vector of displacements.

We can now design an algorithm similar to Algorithm C of the scalar case by insisting that all edges of the interface are fully primal. We note that all faces then are also fully primal. We also make all vertices fully primal. We can then establish that this algorithm satisfies the same strong bound as the first bound in Theorem 6.38.

**Theorem 8.9 (Algorithm C for elasticity)** *The preconditioner $M_C$ satisfies*

$$\langle M_C\lambda,\lambda\rangle \leq \langle F_C\lambda,\lambda\rangle \leq C(1+\log(H/h))^2\langle M_C\lambda,\lambda\rangle, \quad \lambda \in V.$$

*Here, the constant $C$ is independent of $h, H, \gamma$, and the values of the Lamé parameters $\mu_i$ and $\lambda_i$ of our model problem (8.7) provided that the Poisson ratios $\nu_i$ are uniformly bounded away from* $1/2$.

A natural question now arises if a counterpart to Algorithm D can be developed. This is the main topic of a forthcoming paper by Klawonn and Widlund [290]; see also the conference paper [291]. In these papers, the acceptable edge paths of Sect. 6.4 are replaced by paths that can pass from one subdomain to one of its neighbors either through a fully primal face or a fully primal edge. If such paths exist for every edge and vertex of the interface, and with a modest tolerance as in formulas (6.67) and (6.68), we can again establish as satisfactory a bound as the second one in Theorem 6.38.

We also note that it is not necessary to make an edge fully primal if it is common to no more than three subdomains, which each shares a fully primal face with the other two. Current research is focused on finding effective and even smaller sets of primal constraints for this and more general configurations.

It is interesting to consider the constraints that so far have been employed in practice for some large scale FETI-DP codes; see, e.g., [195]. First, one normally chooses enough vertex constraints in order to make the global stiffness matrix invertible, as in the scalar case; see Subsect. 6.4.4 and, in particular, Lesoinne [313]. In addition, three average constraints over the closure of each face are typically employed and, occasionally, three additional first order moments. All these optional constraints are usually enforced by using an additional set of Lagrange multipliers as outlined in Subsect. 6.4.4; for an alternative approach see [290]. We also note that if only vertex constraints are used, the iteration counts will often exceed several hundreds.

# 9

# Preconditioners for Saddle Point Problems

## 9.1 Introduction

In this chapter, we consider preconditioners for several indefinite problems of the form

$$\mathcal{A}\mathcal{U} := \begin{pmatrix} A & B^T \\ B & -t^2 C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} =: \mathcal{F}, \tag{9.1}$$

which arise from the approximation of saddle point problems. The theory for continuous and discrete problems of this kind is reviewed in appendices A.7.2 and B.4, respectively. We assume that the matrix $A$ is symmetric, positive semidefinite and that $C$ is symmetric, nonnegative definite. The real number $t$ is small and may vanish. Systems with $t = 0$ arises from the Stokes problem (see appendices A.7.2, B.4.1, and B.4.2) and from flows in porous media (see appendices A.7.2 and B.4.3). Almost incompressible elasticity (see appendices A.7.2, B.4.1, and B.4.2) and certain stabilized formulations of the Stokes problem (see, e.g., [460, 418]) give rise to problems with $t > 0$; in this case, good preconditioners should be robust with respect to arbitrarily small values of $t$. We assume that $A$, $B$, and $C$ are operators representing the related bilinear forms and that the algebraic problem (9.1) arises from a variational problem: find $u \in V$ and $p \in Q$, such that

$$\begin{aligned} a(u,v) + b(v,p) &= F(v), \quad v \in V, \\ b(u,q) - t^2 c(p,q) &= G(q), \quad q \in Q. \end{aligned} \tag{9.2}$$

For simplicity, we employ a notation without superscripts for the finite dimensional spaces $V$ and $Q$ which are finite or spectral element spaces; see appendix B.4. We employ the scalar products $(\cdot,\cdot)_V$ and $(\cdot,\cdot)_Q$, and the corresponding induced norms $\|\cdot\|_V$ and $\|\cdot\|_Q$, respectively. These scalar products are assumed to be independent of the discretization parameters. In the following, we will also work with vectors $\mathcal{U} \in X$ in the product space $X := V \times Q$ and we will use the norm given by $\|\mathcal{U}\|_X^2 := \|u\|_V^2 + \|p\|_Q^2$ for $\mathcal{U} = \{u, p\}$.

We make the following assumptions on the bilinear forms and the spaces; cf. appendix B.4.

(i) $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are continuous, i.e.,

$$\begin{aligned} |u^T A v| = |a(u,v)| &\le \alpha_1 \|u\|_V \|v\|_V, \quad u,v \in V \\ |u^T B^T p| = |p^T B u| = |b(u,p)| &\le \alpha_2 \|u\|_V \|p\|_Q, \quad u \in V, \quad p \in Q. \end{aligned} \tag{9.3}$$

(ii) $a(\cdot,\cdot)$ is coercive on $Z$, i.e.,

$$u^T A u = a(u,u) \ge \beta \|u\|_V^2, \quad u \in Z, \tag{9.4}$$

where $Z \subset V$ is defined as

$$Z = kernel(B) = \{v \in V \mid \quad b(v,q) = 0,\ q \in Q\}.$$

(iii) $b(\cdot,\cdot)$ satisfies an inf-sup condition, i.e.,

$$\inf_{0 \ne q \in Q} \sup_{0 \ne v \in V} \frac{q^T B v}{\|v\|_V \|q\|_Q} = \inf_{0 \ne q \in Q} \sup_{0 \ne v \in V} \frac{b(v,q)}{\|v\|_V \|q\|_Q} \ge \gamma > 0. \tag{9.5}$$

(iv) $c(\cdot,\cdot)$ is symmetric, nonnegative, and continuous, i.e.,

$$p^T C p = c(p,p) \ge 0, \qquad |q^T C p| = |c(p,q)| \le \delta \|p\|_Q \|q\|_Q, \quad p,q \in Q. \tag{9.6}$$

In some cases, a stronger assumption on the coercivity of $a(\cdot,\cdot)$ holds:

$$u^T A u = a(u,u) \ge \beta \|u\|_V^2, \quad u \in V. \tag{9.7}$$

We note that the matrix $A$ may still be positive definite even if condition (9.7) does not hold. This is the case for flows in porous media where $A$ is a scaled mass matrix.

We note that the continuity properties (9.3) and (9.6) ensure the boundedness of the operator $\mathcal{A}$, which is equivalent to the sup-sup condition

$$\sup_{\mathcal{V} \in X} \sup_{\mathcal{U} \in X} \frac{\mathcal{U}^T \mathcal{A} \mathcal{V}}{\|\mathcal{U}\|_X \|\mathcal{V}\|_X} \le C_1, \tag{9.8}$$

where $C_1$ depends only on $\alpha_1$, $\alpha_2$, and $\delta$, and is independent of $t \in [0,1]$. In addition, conditions (9.3–9.6) imply the invertibility of $\mathcal{A}$, which is equivalent to the inf-sup condition

$$\inf_{\mathcal{V} \in X} \sup_{\mathcal{U} \in X} \frac{\mathcal{U}^T \mathcal{A} \mathcal{V}}{\|\mathcal{U}\|_X \|\mathcal{V}\|_X} \ge C_2 > 0, \tag{9.9}$$

where $C_2$ depends only on $\alpha_1$, $\alpha_2$, $\beta$, $\gamma$, and $\delta$, and is independent of $t$. For additional details see, e.g., [30, 94, 95].

One of the oldest methods for the solution of the algebraic problem (9.1) is the Uzawa algorithm, see [24], which, in case $A$ is invertible amounts to eliminating the variable $u$ and then solving a Schur complement system, with the positive definite matrix

$$S := t^2 C + S_0 = t^2 C + B A^{-1} B^T,$$

for $p$. However, this algorithm requires the exact application of $A^{-1}$ which, in general, is quite expensive. A possible remedy is then to use an inexact Uzawa algorithm where an iterative method is employed instead of the exact inverse of $A$; see [35, 78, 186, 403]. Another strategy is to work with a block preconditioner without eliminating the variable $u$; see, e.g., [406, 405, 270, 460, 418, 282, 279, 281, 461, 280] and Sect. 9.2. Still another strategy, which has been employed for flows in porous media, consists in introducing an additional Lagrange multiplier that enforces the continuity of $u$ across the element boundaries of the triangulation. In this case, $A$ becomes block diagonal and can be inverted inexpensively. A Schur complement system is then obtained for $p$ and the new Lagrange multiplier. Such algorithms will be presented in Sect. 9.3. A different remedy is based on the observation that the variable $u$ satisfies a positive definite problem in case $t = 0$. Thus, let $u = u_g + w$, with $u_g \in V$ such that $Bu_g = g$ and $w \in Z$. Therefore, $w$ satisfies

$$a(w, v) = F(v) - a(u_g, v), \quad v \in Z;$$

see condition (9.4). Certain methods solve this problem directly; particular care is then needed in order to make sure that the iterates belong to the subspace $Z$. Several methods in Sect. 9.4 and 9.3 employ these ideas; see, e.g., [310, 12, 346, 347, 189].

We note that saddle point problems also arise from certain approximations on nonmatching grids (e.g., the mortar method) and the fictitious domain methods. Fictitious domain techniques may also be employed for the preconditioning of standard approximations, where a finite element problem on an unstructured mesh and a complicated domain is replaced by a problem on a larger but simpler domain and on a uniform mesh, for which fast solution methods may be employed. We will not present preconditioners for these problems and of this type in this monograph but refer the interested reader to, e.g., [298, 4, 3, 6, 129, 303, 302, 426, 104, 469, 387, 55] and the references therein, for mortar approximations, and to [26, 299, 224, 225, 300, 263], for fictitious domain techniques and related topics. Nor will we discuss preconditioners for stabilized formulations of Stokes or elasticity problems for which an inf-sup condition does not hold; see, e.g., [460, 418].

Before proceeding, we recall some properties of the Schur complements $S_0$ and $S$. Their condition numbers can be related to the properties of the bilinear form $b(\cdot, \cdot)$; see, e.g., [95, Sect. II.3.2] and the references therein for more details. See also [418, Sect. 1].

**Lemma 9.1** *Let $A$ be invertible. We have, for $p \in Q$,*

$$p^T S p \geq p^T S_0 p \geq \frac{\gamma^2}{\alpha_1} \|p\|_Q^2.$$

*If the stronger coercivity property (9.7) holds, then*

$$p^T S_0 p \leq (\alpha_2^2/\beta) \|p\|_Q^2,$$

*and*

$$p^T S p \leq ((\alpha_2^2/\beta) + t^2\delta) \|p\|_Q^2 \leq ((\alpha_2^2/\beta) + \delta) \|p\|_Q^2.$$

*Proof.* We first note that $S_0$ and $S$ are symmetric and positive semidefinite. In order to find the lower bound, it is enough to consider $S_0$, since the bilinear form $c(\cdot,\cdot)$ is positive semidefinite. Since $A$ is positive semidefinite and invertible, it is positive definite and we can define the scalar product

$$(u,v)_{A^{-1}} := v^T A^{-1} u, \quad u,v \in V,$$

and the corresponding induced norm.

Using a property of Hilbert norms, the continuity of $a(\cdot,\cdot)$ in (9.3), and the inf-sup condition (9.5), we find, for $q \in Q$,

$$\begin{aligned} q^T S_0 q = q^T B A^{-1} B^T q = \|B^T q\|_{A^{-1}}^2 &= \sup_{0 \neq v \in V} \frac{(v, B^T q)_{A^{-1}}^2}{\|v\|_{A^{-1}}^2} \\ &= \sup_{0 \neq v \in V} \frac{(q^T B A^{-1} v)^2}{v^T A^{-1} v} = \sup_{0 \neq \tilde{v} \in V} \frac{(q^T B \tilde{v})^2}{\tilde{v}^T A \tilde{v}} \\ &\geq \frac{1}{\alpha_1} \sup_{0 \neq \tilde{v} \in V} \frac{(q^T B \tilde{v})^2}{\|\tilde{v}\|_V^2} \geq \frac{\gamma^2}{\alpha_1} \|q\|_Q^2, \end{aligned}$$

which proves the lower bound.

In order to prove the upper bound for $S_0$, it is enough to consider the continuity of $b(\cdot,\cdot)$ in (9.3) for $q \in Q$ and $u = A^{-1} B^T q$, and the coercivity of $a(\cdot,\cdot)$ in (9.7):

$$\begin{aligned} q^T S_0 q = q^T B (A^{-1} B^T q) &\leq \alpha_2 \|q\|_Q \|A^{-1} B^T q\|_V \\ &\leq \alpha_2 \beta^{-1/2} \|q\|_Q (q^T B A^{-1} A A^{-1} B^T q)^{1/2} = \alpha_2 \beta^{-1/2} \|q\|_Q (q^T S_0 q)^{1/2}, \end{aligned}$$

which proves the upper bound for $S_0$. That of $S$ is then a consequence of the continuity of $c(\cdot,\cdot)$ given in (9.6). □

We note that the constants in the previous lemma are independent of the small parameter $t$. This result highlights why Uzawa's method, which employs the Schur complement $S$ are potentially attractive. While the condition number of the original system $\mathcal{A}$ usually depends on discretization parameters, that of $S$ often does not. (However, we recall that for certain spectral

element approximations the inf-sup constant $\gamma$ may depend on the polynomial degree; see appendix B.4.2.) Lemma 9.1 may be summarized, by saying that the quadratic forms with the Schur complements $S_0$ and $S$ are spectrally equivalent to the scalar product in $Q$, which typically is a subspace of $L^2$ and therefore to a mass matrix. We note that this is true only if the stronger coercivity condition (9.7) holds, as for the Stokes problem and almost incompressible elasticity. For flows in porous media, only the weaker property (9.4) holds; we will see in Sect. 9.3, that $S$ is equivalent to an operator related to a Laplacian. Finally, in case the bilinear form $c(\cdot,\cdot)$ is not continuous, we can obtain the same results by employing the modified norm $\|p\|_Q + t\, c(p,p)^{1/2}$ in $Q$; see, e.g., [279, 280].

## 9.2 Block Preconditioners

We consider the system (9.1) and will look for symmetric, block diagonal preconditioners of the form

$$\hat{\mathcal{B}} = \begin{pmatrix} \hat{A} & 0 \\ 0 & \hat{C} \end{pmatrix}, \tag{9.10}$$

with $\hat{A}$ and $\hat{C}$ suitable symmetric, positive definite matrices. Since $\mathcal{A}$ is symmetric and $\hat{\mathcal{B}}$ is positive definite, we can employ the conjugate residual algorithm (CR) or MINRES; see appendix C.6.2. In this case, convergence is determined by the condition number of the preconditioned operator $\hat{\mathcal{B}}^{-1}\mathcal{A}$,

$$\kappa = \kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) = \frac{\mu_{max}}{\mu_{min}} = \frac{\max\{|\lambda| :\ \lambda \in \sigma(\hat{\mathcal{B}}^{-1}\mathcal{A})\}}{\min\{|\lambda| :\ \lambda \in \sigma(\hat{\mathcal{B}}^{-1}\mathcal{A})\}};$$

see Lemma C.12.

A first strategy is given in [279, 280]. It is very general and does not require any additional assumptions beyond the usual continuity, coercivity, and inf-sup conditions. The proofs are similar to those of Lemmas 1 and 4 in [279].

Since $\hat{\mathcal{B}}^{-1}\mathcal{A}$ and $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ have the same spectrum, the sup conditions (9.8) and (9.9) suggest that bounds on the spectrum of a suitable preconditioned operator may be obtained by rescaling the vectors $\mathcal{V}$ and $\mathcal{U}$ in these conditions in such a way that Euclidean norms are obtained in the denominators. This leads to the definition

$$\mathcal{U}^T \mathcal{B} \mathcal{U} := \|\mathcal{U}\|_X^2,$$

and to the matrices $H_V$ and $H_Q$, defined as the representations of the scalar products in $V$ and $Q$, respectively:

$$u^T H_V u := \|u\|_V^2, \quad u \in V; \qquad p^T H_Q p := \|p\|_Q^2, \quad p \in Q.$$

**Lemma 9.2** *Let*

$$\mathcal{B} = \begin{pmatrix} H_V & 0 \\ 0 & H_Q \end{pmatrix}.$$

*Then,*

$$\sup_{\mathcal{V}\in X}\sup_{\mathcal{U}\in X}\frac{\mathcal{U}^T\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}\mathcal{V}}{\|\mathcal{U}\|_2\,\|\mathcal{V}\|_2} \le C_1$$

$$\inf_{\mathcal{V}\in X}\sup_{\mathcal{U}\in X}\frac{\mathcal{U}^T\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}\mathcal{V}}{\|\mathcal{U}\|_2\,\|\mathcal{V}\|_2} \ge C_2 > 0,$$

*with $C_1$ and $C_2$ the* same *constants as in conditions (9.8) and (9.9). Therefore,*

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) \le \frac{C_1}{C_2}.$$

*Proof.* Since $\|\mathcal{U}\|_X^2 = \mathcal{U}^T\mathcal{B}\mathcal{U}$, it is enough to make the substitutions

$$\mathcal{U}' = \mathcal{B}^{1/2}\mathcal{U}, \qquad \mathcal{V}' = \mathcal{B}^{1/2}\mathcal{V}$$

in conditions (9.8) and (9.9). □

We are then led to the following strategy: *a good block diagonal preconditioner for the mixed system $\mathcal{A}$ is obtained if we employ good preconditioners for the operators representing the scalar products in $V$ and $Q$.* We therefore assume that

$$\begin{aligned} a_1\|u\|_V^2 \le u^T\hat{A}u \le a_2\|u\|_V^2, \quad u \in V, \\ c_1\|p\|_Q^2 \le p^T\hat{C}p \le c_2\|p\|_Q^2, \quad p \in Q. \end{aligned} \tag{9.11}$$

**Theorem 9.3** *Let the conditions in (9.11) hold. Then,*

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \le \frac{C_1}{C_2}\frac{\max\{a_2,c_2\}}{\min\{a_1,c_1\}},$$

*with $C_1$ and $C_2$ the* same *constants as in conditions (9.8) and (9.9).*

*Proof.* We will look for similar sup-sup and inf-sup conditions as in Lemma 9.2. We first note that from the definition of $\hat{\mathcal{B}}$ and the inequalities in (9.11), we have, for $\mathcal{U} = \{u,p\} \in X$,

$$\min\{a_1,c_1\}\,\|\mathcal{U}\|_X^2 \le \mathcal{U}^T\hat{\mathcal{B}}\mathcal{U} = u^T\hat{A}u + p^T\hat{C}p \le \max\{a_2,c_2\}\,\|\mathcal{U}\|_X^2.$$

We can then prove that

$$\sup_{\mathcal{V}\in X}\sup_{\mathcal{U}\in X}\frac{\mathcal{U}^T\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\mathcal{V}}{\|\mathcal{U}\|_2\,\|\mathcal{V}\|_2} \le C_1\,\max\{a_2,c_2\}$$

using the substitutions $\mathcal{U}' = \hat{\mathcal{B}}^{1/2}\mathcal{U}$ and $\mathcal{V}' = \hat{\mathcal{B}}^{1/2}\mathcal{V}$, and the sup-sup condition (9.8). An inf-sup bound can be found similarly. □

We stress that the condition number depends on the spectral bounds in (9.11) and the parameters $\alpha_1$, $\alpha_2$, $\beta$, $\gamma$, and $\delta$ in conditions (9.3–9.6), but is uniform in $t \in [0,1]$.

We will now discuss an alternative way of devising a block preconditioner, which, to the best of our knowledge, was proposed earlier than the previous one in [270]. We note that a similar strategy was already proposed in [406, 405]. It works under somewhat more restrictive assumptions, since we will now require that the block $A$ is invertible, and this may lead to a different choice of the blocks $\hat{A}$ and $\hat{C}$. We arrive at this idea after reminding ourselves that an Uzawa algorithm requires the exact solution of a linear system with the matrix $A$; here we will work with preconditioners for $A$ and $S$; see also [70] for similar and related ideas. For the case $t = 0$, we will now obtain bounds that are independent of the constants in conditions (9.3)-(9.6).

We will use the following fundamental result; see [270, Sect. 5.1]. Our proof follows [282, Lem. 1].

**Lemma 9.4** *Let $A$ be invertible and $t = 0$. If*

$$\mathcal{B} = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix},$$

*then, $\mathcal{B}^{-1}\mathcal{A}$ has exactly three distinct eigenvalues:*

$$\frac{1-\sqrt{5}}{2}, \quad 1, \quad \frac{1+\sqrt{5}}{2}.$$

*Therefore,*

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) = \frac{\sqrt{5}+1}{\sqrt{5}-1}.$$

*Proof.* To find the spectrum of $\mathcal{B}^{-1}\mathcal{A}$, we can equivalently work with the matrix

$$\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2} = \begin{pmatrix} I & \tilde{B}^T \\ \tilde{B} & 0 \end{pmatrix}, \quad \tilde{B} = S^{-1/2}BA^{-1/2}.$$

Since $S = S_0$ and therefore $\tilde{B}\tilde{B}^T = I$, a simple computation shows that the eigenvalues of $\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}$ that are different from 1 are solutions of

$$\lambda^2 - \lambda - 1 = 0.$$

The spectrum of $\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}$ is therefore

$$\{\frac{1-\sqrt{5}}{2}, \quad 1, \quad \frac{1+\sqrt{5}}{2}\}.$$

□

We stress that in the previous lemma the condition number is independent of all the continuity, coercivity, and stability constants of the bilinear forms.

*Remark 9.5.* A similar result holds in case $t > 0$. Indeed, a simple calculation shows that the eigenvalues of $\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}$ that are different from one, are solutions of

$$\lambda^2 - \lambda - 1 - (1-\lambda)t^2\mu = 0,$$

with $\mu$ an eigenvalue of $S^{-1/2}CS^{-1/2}$.

Even though the Schur complement $S$ and the block $A$ are often difficult to invert, this result suggests the following strategy: *a good preconditioner for the mixed system $\mathcal{A}$ can be obtained if good preconditioners for the $A$ block and for the Schur complement $S$ are employed.*

The assumptions that we make are therefore that $\hat{A}$ and $\hat{C}$ are good preconditioners for $A$ and $S$, respectively, i.e., there exist strictly positive constants, such that,

$$\begin{aligned} a_1 u^T A u \le u^T \hat{A} u \le a_2 u^T A u, \quad u \in V, \\ c_1 p^T S p \le p^T \hat{C} p \le c_2 p^T S p, \quad p \in Q. \end{aligned} \tag{9.12}$$

The following result can be proven in the same way as Theorem 9.3. More precise bounds for the spectrum of the preconditioned operator are provided in [270, Th. 5.1].

**Theorem 9.6** *If $A$ is invertible and $\mathcal{B}$ is defined as in Lemma 9.4, then the condition number of the preconditioned operator $\hat{\mathcal{B}}^{-1}\mathcal{A}$ satisfies, for $t > 0$,*

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \le \kappa(\mathcal{B}^{-1}\mathcal{A}) \frac{\max\{a_2, c_2\}}{\min\{a_1, c_1\}}.$$

*In case $t = 0$, we have*

$$\kappa(\hat{\mathcal{B}}^{-1}\hat{\mathcal{A}}) \le \left(\frac{\sqrt{5}+1}{\sqrt{5}-1}\right) \frac{\max\{a_2, c_2\}}{\min\{a_1, c_1\}}.$$

We note that the constants in this result are now independent of the parameters $\alpha_1$, $\alpha_2$, $\beta$, $\gamma$, and $\delta$ in conditions (9.3–9.6). In addition, in case the weaker coercivity property (9.4) holds, this second approach may lead to a different preconditioning strategy; see the case of flows in porous media in Sect. 9.3, where $A$ is spectrally equivalent to a mass matrix.

Block triangular preconditioners can also be employed. In this case, the conjugate residual method must be replaced by another Krylov space method, such as GMRES. See, e.g., [70, 280, 281, 287]. In our discussion, we will follow Klawonn [281]. When these methods are used, the block diagonal preconditioner given in (9.10) is replaced by

$$\hat{\mathcal{B}} = \begin{pmatrix} \hat{A} & B^T \\ 0 & -\hat{C} \end{pmatrix}, \tag{9.13}$$

or by its transpose, with $\hat{A}$ and $\hat{C}$ again suitable symmetric, positive definite matrices. Restricting the discussion to the upper triangular preconditioner,

one can establish spectral bounds for the preconditioned operator $\mathcal{A}\mathcal{B}^{-1}$ similar to those for the block diagonal case. One can also show that, if we choose to make $A - \hat{A}$ positive definite, then multiplying by the positive definite matrix

$$\mathcal{H} = \begin{pmatrix} A - \hat{A} & 0 \\ 0 & \hat{C} \end{pmatrix}, \tag{9.14}$$

from the right will make $\mathcal{A}\mathcal{B}^{-1}\mathcal{H}$ symmetric, positive definite. This makes it possible to show a strong bound on the decay of the error in the preconditioned GMRES algorithm that resembles the bound for the conjugate gradient method rather than that of the conjugate residual method; see [281, Theorem 3.7].

## 9.3 Flows in Porous Media

We consider saddle point approximations of second order diffusive problems and assume that the boundary of a bounded Lipschitz domain $\Omega$ is partitioned into two sets $\partial\Omega_D$ and $\partial\Omega_N$.

In order to be consistent with the notations and the abstract framework employed in the previous sections, we will not employ bold letters for velocities. The letters $u$, $v$, $w$,... will be employed for primal variables (velocities) and are understood to be vectors, while the letters $p$, $q$, $r$,... are employed for the dual variables (pressures) and represent scalar functions. We also use the notation $n$ for the outward unit normal of a domain.

We consider the following problem for a velocity $u$ and a pressure $p$:

$$\begin{aligned} u &= -\mathcal{K}\nabla p && \text{in } \Omega, \\ \nabla \cdot u &= f && \text{in } \Omega, \\ u \cdot n &= g_D && \text{on } \partial\Omega_{\mathrm{D}}, \\ p &= g_N && \text{on } \partial\Omega_{\mathrm{N}}, \end{aligned} \tag{9.15}$$

where $\mathcal{K} = [\mathcal{K}_{ij}]$ is a symmetric, uniformly positive definite and bounded diffusive matrix, $f \in L^2(\Omega)$, $g_D \in H_{00}^{-1/2}(\partial\Omega_D)$, and $g_N \in H^{1/2}(\partial\Omega_N)$. We will also refer to this system as Darcy's problem in the following. In case $\partial\Omega_N = \emptyset$, we require the compatibility condition

$$\int_\Omega f \, dx + \int_{\partial\Omega} g_D \, ds = 0 \tag{9.16}$$

and the pressure $p$ is determined up to an additive constant.

The variational formulation for problem (9.15) and some finite element approximations are reviewed in appendices A.7.2 and B.4.3, respectively. We recall that we obtain a saddle point problem of the form (9.2) with $t = 0$ by defining the bilinear and linear forms by

$$a(u,v) = \int_\Omega (\mathcal{K}^{-1}u)\cdot v\,dx,$$
$$b(u,p) = -\int_\Omega \nabla\cdot u\,p\,dx,$$
$$G(p) = -\int_\Omega f\,p\,dx,$$
$$F(v) = -\int_{\partial\Omega_N} g_N\,(v\cdot n)\,ds.$$

We consider a shape-regular triangulation $\mathcal{T}_h$ of $\Omega$, of maximum diameter $h$. The appropriate finite element spaces $V$ and $Q$ are conforming spaces in $H(\mathrm{div};\Omega)$ and $L^2(\Omega)$, respectively. Here, we only recall that $V$ may be chosen as a Raviart-Thomas finite element space consisting of vector functions with a continuous normal component across the interfaces between the elements of a mesh, while $Q$ may consist of discontinuous scalar functions:

$$V = RT_k^h(\Omega), \quad Q = Q_{k-1}^h(\Omega), \qquad k \geq 1.$$

We refer to appendices B.3.1 and B.4.3 for more details. Other spaces conforming in $H(\mathrm{div};\Omega)$ may also be employed; cf., e.g., [95, Sect. III.3]. The solution $u$ satisfies $u\cdot n = g_D$ on $\partial\Omega_D$, and, in addition, the space $V$ consists of vectors with a vanishing normal component on $\partial\Omega_D$. In case $\partial\Omega_N = \emptyset$, then $V$ consists of vectors with a vanishing normal component on the whole of $\partial\Omega$ and $p$ is defined up to an additive constant. Uniqueness can be ensured by imposing the condition that the mean value of $p$ over $\Omega$ vanishes, i.e., by requiring additionally that the functions in $Q$ have a vanishing mean value.

Finite element approximations give rise to an algebraic problem of the form (9.1), with $t = 0$. We recall that the weaker coercivity property (9.4) holds in this case, and that the matrix $A$ remains symmetric and positive definite. The Schur complement $BA^{-1}B^T$ can therefore be formed but only the lower bound in Lemma 9.1 holds. We also recall that the subspace $Z = \mathit{kernel}(B) \subset V$ is the kernel of the divergence operator defined in $V$, while the range of the divergence equals $Q$; cf. appendix B.3.4 and in particular Lemma B.24.

We note that the classification into Dirichlet and Neumann conditions may be confusing when we devise iterative substructuring methods of, e.g., Neumann-Neumann or FETI type for this problem. Indeed, Neumann (natural) boundary conditions on $\partial\Omega_N$ for the mixed problem arise from Dirichlet (essential) conditions for the underlying second order equation, and, conversely, Dirichlet (essential) conditions on $\partial\Omega_D$ for the primal variable of the mixed problem arise from Neumann (natural) conditions for the second order equation. For this reason, Neumann problems for the mixed system are always well posed, while Dirichlet problems require the compatibility condition (9.16) and provide a pressure that is only defined up to an additive constant. We also point out that a Dirichlet condition for the mixed system provides the value of $u\cdot n$ (the appropriate trace for the primal variable $u \in H(\mathrm{div};\Omega)$) on

the boundary, while a Neumann condition imposes the value of the *flux* $pn$, which is in general not defined for pressures $p \in L^2(\Omega)$.

Block diagonal preconditioners of type (9.10) can be employed with the conjugate residual algorithm. The two strategies considered in Sect. 9.2 give rise to different types of preconditioners. The first is considered in Lemma 9.2 and Theorem 9.3; according to condition (9.11), the blocks $\hat{A}$ and $\hat{C}$ need to be good preconditioners for the matrices representing the scalar products in $H(\mathrm{div}\,;\Omega)$ and $L^2(\Omega)$, respectively. For the first block, we refer to Chap. 10, where many preconditioners are presented, while for the second block it is enough to take the block diagonal mass matrix for a finite element space consisting of discontinuous functions.

The second strategy in Sect. 9.2 is considered in Lemma 9.4 and Theorem 9.6. According to condition (9.12), it is enough to have good preconditioners for the $A$ block and for the Schur complement $BA^{-1}B^T$. The block $A$ is spectrally equivalent to a mass matrix and it is therefore uniformly well conditioned. On the other hand, we will see that $BA^{-1}B^T$ is spectrally equivalent to suitable approximations of the Laplace operator; in case $\mathcal{K}$ does not vary very much across $\Omega$, it was proven in [462] that $BA^{-1}B^T$ is spectrally equivalent to a finite difference approximation of a Laplacian. Many efficient preconditioners can then be employed. However, many practical problems exhibit large jumps in $\mathcal{K}$. We will present several successful strategies of constructing good preconditioners for this Schur complement in the next few sections.

We will first consider some iterative substructuring methods; they were indeed the first to be studied. We will then present some more recent work and some overlapping preconditioners.

### 9.3.1 Iterative Substructuring Methods

In this section, we derive some simple iterative substructuring methods in the same way as we did for the Laplace equation in Chap. 1; see, in particular, Sect. 1.1, 1.2, and 1.3.

To the best of our knowledge, two of the first iterative substructuring methods were presented in Glowinski and Wheeler [227] (some older work of the same authors can also be found in the references therein). In that work, two methods are presented; they are unpreconditioned conjugate gradient algorithms for certain interface variables. The important issue of singular local Dirichlet problems was already addressed there. We also note that an analog of harmonic discrete Stokes extensions was also employed; cf. Sect. 9.4.2.

Here, we present the two methods in [227] for a very simple situation. We derive equations for interface variables in a fashion similar to those of Chap. 1. We then present simple Neumann-Neumann and Dirichlet-Dirichlet algorithms. The latter was successfully generalized in [147, 148]. The same framework applies to the Stokes problem, for which a Neumann-Neumann method is presented in Sect. 9.4.2. However, there can be differences if there are more than two subdomains.

We consider a partition of $\Omega$ into two subdomains $\Omega_1$ and $\Omega_2$, with a common interface $\Gamma$, as in Sect. 1.1 and Fig. 1.1. Here, we assume that $\partial\Omega_D = \emptyset$, in order to avoid complications with possibly singular local Dirichlet problems. Problem (9.15) is equivalent to the following transmission problem:

$$\begin{aligned}
u_1 &= -\mathcal{K}\nabla p_1 && \text{in } \Omega_1,\\
\nabla\cdot u_1 &= f && \text{in } \Omega_1,\\
p_1 &= g_N && \text{on } \partial\Omega_1 \cap \partial\Omega,\\
u_1\cdot n_1 &= -u_2\cdot n_2 && \text{on } \Gamma,\\
-p_1\, n_1 &= p_2\, n_2 && \text{on } \Gamma,\\
u_2 &= -\mathcal{K}\nabla p_2 && \text{in } \Omega_2,\\
\nabla\cdot u_2 &= f && \text{in } \Omega_2,\\
p_2 &= g_N && \text{on } \partial\Omega_2 \cap \partial\Omega,
\end{aligned} \tag{9.17}$$

We note that this problem can be obtained by variational arguments but also directly from (1.2).

We now rewrite the algebraic problem (9.1) in block form, by further partitioning the degrees of freedom, as in Sect. 1.2. Here the appropriate partition is into velocity degrees of freedom internal to the two subdomains, $u_I^{(1)}$ and $u_I^{(2)}$, respectively, pressure degrees of freedom associated with one substructure, $p^{(1)}$ and $p^{(2)}$, respectively, and velocity degrees of freedom on the interface $\Gamma$, $u_\Gamma$. We recall that pressures are discontinuous and therefore any pressure degree of freedom is associated with only one substructure. This corresponds to a partition of finite element functions into velocities supported on single subdomains and with vanishing normal component on $\Gamma$, pressures supported on single subdomains, and velocities with vanishing internal degrees of freedom. Since the degrees of freedom for the velocity are normal components along the faces (edges in two dimensions) of the triangulation, we assume that the degrees of freedom in $u_\Gamma$ are values of $u\cdot n_1 = -u\cdot n_2$.

The discrete problem (9.1) can then be written as

$$\left(\begin{array}{cc|cc|c}
A_{II}^{(1)} & B_I^{(1)T} & 0 & 0 & A_{I\Gamma}^{(1)}\\
B_I^{(1)} & 0 & 0 & 0 & B_\Gamma^{(1)}\\
\hline
0 & 0 & A_{II}^{(2)} & B_I^{(2)T} & A_{I\Gamma}^{(2)}\\
0 & 0 & B_I^{(2)} & 0 & B_\Gamma^{(2)}\\
\hline
A_{\Gamma I}^{(1)} & B_\Gamma^{(1)T} & A_{\Gamma I}^{(2)} & B_\Gamma^{(2)T} & A_{\Gamma\Gamma}
\end{array}\right)
\left(\begin{array}{c} u_I^{(1)}\\ p^{(1)}\\ u_I^{(2)}\\ p^{(2)}\\ u_\Gamma \end{array}\right)
=
\left(\begin{array}{c} f_I^{(1)}\\ g^{(1)}\\ f_I^{(2)}\\ g^{(2)}\\ f_\Gamma \end{array}\right), \tag{9.18}$$

where the interface blocks can be found by subassembling local contributions:

$$A_{\Gamma\Gamma} = A_{\Gamma\Gamma}^{(1)} + A_{\Gamma\Gamma}^{(2)}, \qquad f_\Gamma = f_\Gamma^{(1)} + f_\Gamma^{(2)}.$$

We note that $f_\Gamma^{(1)} = f_\Gamma^{(2)} = 0$ but that we keep these terms so that the same formulas can be employed for the Stokes problem.

In order to write the coupled problem in matrix form, equivalent to the original system (9.18), we need to employ a definition of the fluxes $p_1 n_1$ and $p_2 n_2$ using duality; cf. equation (1.7). Thus, if $\phi_j$ is a velocity basis function relative to a degree of freedom on $\Gamma$, we have, using (9.17),

$$\begin{aligned}\int_\Gamma (-p_i\, n_i)\cdot\phi_j\, ds &= \int_{\partial\Omega_i\setminus\Gamma} (p_i\, n_i)\cdot\phi_j\, ds - \int_{\Omega_i} \nabla\cdot(p_i\,\phi_j)\, dx \\ &= \int_{\partial\Omega_i\setminus\Gamma} g_N\,\phi_j\cdot n_i\, ds + \int_{\Omega_i} (\mathcal{K}^{-1}u_i)\cdot\phi_j\, dx - \int_{\Omega_i} p_i\,\nabla\cdot\phi_j\, dx.\end{aligned}$$

An approximation $\lambda^{(i)}$ of the functional representing the flux $(-p_i\, n_i)$ can be found by replacing the exact solution $\{u_i, p_i\}$ with its finite element approximation. Letting $j$ run over the velocity degrees of freedom on $\Gamma$ and using the local matrices, we have

$$\lambda^{(i)} = A_{\Gamma I}^{(i)} u_I^{(i)} + A_{\Gamma\Gamma}^{(i)} u_\Gamma^{(i)} + {B_\Gamma^{(i)}}^T p^{(i)} - f_\Gamma^{(i)}, \tag{9.19}$$

with $u_\Gamma^{(1)}$ and $u_\Gamma^{(2)}$ local vectors of degrees of freedom, which are values of $u_1\cdot n_1$ and $u_2\cdot n_1$, respectively.

As for the Laplace problem, this approximation coincides with the residual corresponding to the degrees of freedom on $\Gamma$ of a Darcy problem with a Neumann condition on $\Gamma$. Using this flux, we can write the transmission problem (9.17) as

$$\begin{aligned}&\begin{pmatrix} A_{II}^{(1)} & {B_I^{(1)}}^T \\ B_I^{(1)} & 0 \end{pmatrix}\begin{pmatrix} u_I^{(1)} \\ p^{(1)} \end{pmatrix} + \begin{pmatrix} A_{I\Gamma}^{(1)} \\ {B_I^{(1)}}^T \end{pmatrix} u_\Gamma^{(1)} = \begin{pmatrix} f_I^{(1)} \\ g^{(1)} \end{pmatrix}, \\ &u_\Gamma^{(1)} = u_\Gamma^{(2)} = u_\Gamma, \\ &A_{\Gamma I}^{(1)} u_I^{(1)} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^{(1)} + {B_\Gamma^{(1)}}^T p^{(1)} - f_\Gamma^{(1)} \\ &= -(A_{\Gamma I}^{(2)} u_I^{(2)} + A_{\Gamma\Gamma}^{(2)} u_\Gamma^{(2)} + {B_\Gamma^{(2)}}^T p^{(2)} - f_\Gamma^{(2)}) = \lambda_\Gamma \\ &\begin{pmatrix} A_{II}^{(2)} & {B_I^{(2)}}^T \\ B_I^{(2)} & 0 \end{pmatrix}\begin{pmatrix} u_I^{(2)} \\ p^{(2)} \end{pmatrix} + \begin{pmatrix} A_{I\Gamma}^{(2)} \\ {B_I^{(2)}}^T \end{pmatrix} u_\Gamma^{(2)} = \begin{pmatrix} f_I^{(2)} \\ g^{(2)} \end{pmatrix}.\end{aligned} \tag{9.20}$$

This coupled problem can be obtained directly from the algebraic system (9.18) by introducing the two interface vectors of degrees of freedom $u_\Gamma^{(1)}$ and $u_\Gamma^{(2)}$. As for the Laplace problem in Sect. 1.2, the transmission condition involving the fluxes coincides with the last equation in (9.18). We note that the first and last equations are discretizations of local problems with Dirichlet conditions $u_\Gamma^{(i)}$ on the interface $\Gamma$ and that they are well posed since we assume that only Neumann conditions are given on $\partial\Omega$.

We can now obtain equations for the interface variable $u_\Gamma$ and the flux $\lambda_\Gamma$ by proceeding in the same way as in Sect. 1.3.1 and 1.3.2.

A Schur complement system can be obtained from (9.18) by eliminating the interior variables using the first four equations (see also the first and fourth equations in the coupled system (9.20)):

$$S_\Gamma u_\Gamma = g_\Gamma.$$

We note that the Schur complement $S_\Gamma$ can be constructed by subassembling local contributions

$$S_\Gamma = S_\Gamma^{(1)} + S_\Gamma^{(2)},$$

where each contribution involves the solution of a local Darcy problem with Dirichlet conditions; cf. the first and fourth equations in (9.20). This equation is the unpreconditioned problem of Method 1 in [227]. The Schur complement $S_\Gamma$ is symmetric and positive definite; see [227, Th. 3.1].

An equation for the flux $\lambda_\Gamma$ can be obtained as in Sect. 1.3.2 by finding local velocities and pressures from the coupled system (9.20):

$$\begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} & B_I^{(i)T} \\ A_{\Gamma I}^{(1)} & A_{\Gamma\Gamma}^{(i)} & B_\Gamma^{(1)T} \\ B_I^{(i)} & B_\Gamma^{(i)} & 0 \end{pmatrix} \begin{pmatrix} u_I^{(i)} \\ u_\Gamma^{(i)} \\ p^{(i)} \end{pmatrix} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} \pm \lambda_\Gamma \\ g^{(i)} \end{pmatrix}, \tag{9.21}$$

and then using the interface variables $u_\Gamma^{(i)}$ in the second equation of (9.20). We obtain the system

$$F_\Gamma \lambda_\Gamma = d_\Gamma, \qquad F_\Gamma = S_\Gamma^{(1)-1} + S_\Gamma^{(2)-1}.$$

We note that the application of each $S_\Gamma^{(i)-1}$ involves the solution of a local Darcy problem with Neumann conditions on $\Gamma$; cf. (9.21). This equation is the unpreconditioned problem of Method 2 in [227]; see, in particular, equation (7.21). The matrix $F_\Gamma$ is also symmetric and positive definite.

Preconditioned Neumann-Neumann and Dirichlet-Dirichlet methods can then be obtained as in Sect. 1.3.4 and 1.3.5, respectively. Dirichlet-Neumann algorithms are also possible. A Neumann-Neumann algorithm corresponds to the preconditioned equation

$$F_\Gamma S_\Gamma u_\Gamma = F_\Gamma g_\Gamma.$$

This can be obtained as in Sect. 1.3.4 by starting from an initial guess $u_\Gamma^0$, first solving Dirichlet problems on each $\Omega_i$ with data $u_\Gamma^0$ on $\Gamma$, and then a problem on each subdomain, with Neumann data on $\Gamma$ chosen as the difference of fluxes of the solutions of the two Dirichlet problems. The values on $\Gamma$ of the solutions of these Neumann problems are then employed to correct the initial $u_\Gamma^0$ and find the new iterate.

A preconditioned method for the flux equation is given by

$$S_\Gamma F_\Gamma \lambda_\Gamma = S_\Gamma d_\Gamma. \tag{9.22}$$

This method was studied in [147, 148] for the more general case of many subdomains and jumps of the diffusion matrix; it is presented in Sect. 9.3.3. Here, we only mention that it can be obtained as in Sect. 1.3.5, by starting from an initial guess $\lambda_\Gamma^0$ of the flux, first solving two Neumann problems on the subdomains $\Omega_i$ with data $\lambda_\Gamma^0$ on $\Gamma$, and then a problem on each subdomain, with Dirichlet data, on $\Gamma$, chosen as the difference of the traces, on $\Gamma$, of the solutions of the two Neumann problems. The values on $\Gamma$ of the fluxes of the solutions of these Dirichlet problems are then employed to correct the initial $\lambda_\Gamma^0$ and find the new iterate.

We point out that in case there are Dirichlet conditions on $\partial\Omega$ or if there are floating subdomains, the local mixed Dirichlet problems may be singular. Local pressures are then not uniquely determined. They can be chosen in such a way that scalability is ensured. This issue was already addressed in [227] for the unpreconditioned Schur complement system. For the preconditioned flux equation this was considered in [147, 148].

To provide a comparison, we will now consider the Stokes system, with discontinuous finite elements employed for the pressure; see Sect. 9.4 for additional algorithms. We consider the two subdomain configuration of Fig. 1.1, already employed in this section and Dirichlet conditions on $\partial\Omega$. An equivalent coupled Stokes problem is

$$\begin{aligned}
-\nu\Delta u_1 + \nabla p_1 &= f && \text{in } \Omega_1,\\
\nabla \cdot u_1 &= 0 && \text{in } \Omega_1,\\
u_1 &= g_D && \text{on } \partial\Omega_1 \cap \partial\Omega,\\
u_1 &= u_2 && \text{on } \Gamma,\\
\nu\nabla u_1 \cdot n_1 - p_1 n_1 &= -(\nu\nabla u_2 \cdot n_2 - p_2 n_2) && \text{on } \Gamma,\\
-\nu\Delta u_2 + \nabla p_2 &= f && \text{in } \Omega_2,\\
\nabla \cdot u_2 &= 0 && \text{in } \Omega_2,\\
u_2 &= g_D && \text{on } \partial\Omega_2 \cap \partial\Omega;
\end{aligned} \tag{9.23}$$

see [392, Sect. 5.5]. We recall that the velocity $u$ is continuous and with nodal degrees of freedom on the mesh.

The degrees of freedom can be partitioned exactly as before in equation (9.18). We note, in particular, that the subvector $u_\Gamma$ consists of values of the velocity on the interface $\Gamma$. The *same* system as in (9.18) is obtained for the Stokes problem. Approximations of the fluxes

$$\lambda^{(i)} \sim \nu\nabla u_i \cdot n_i - p_i n_i$$

can be found as before by using duality and exactly the same formula (9.19) is obtained. The algebraic coupled system corresponding to (9.23) can therefore be written as in (9.20). Preconditioned Neumann-Neumann and FETI method can be devised in exactly the same way. The more general case of many subdomains have been considered in [386, 229, 230, 316, 314, 314] and some results are presented in Sect. 9.4.2. We note that there is a difference

from Darcy's problem when many subdomains are considered. For Stokes problems Neumann boundary conditions give rise to singular problems (as for the Laplace equation), while for Darcy flows, Dirichlet problems can be singular. The singularity of these problems can be exploited to ensure scalability of algorithms.

### 9.3.2 Hybrid-Mixed Formulations and Spectral Equivalencies with Crouzeix-Raviart Approximations

As already mentioned, even though the matrix $A$ is spectrally equivalent to a mass matrix, forming the Schur complement $BA^{-1}B^T$ of (9.1) can be expensive. A possible remedy is to work with vector functions that have a discontinuous normal component across the interelement boundaries and impose continuity using Lagrange multipliers. In this case, the matrix $A$ becomes block diagonal and can be easily inverted. The new Schur complement now acts on the pressures and the new Lagrange multipliers (which also has the meaning of a pressure). We refer to [20] for details and additional results. Our presentation is largely based on [412, Sect. 3.4]. In particular, we will see that there is an equivalence between mixed formulations of second order problems and some nonconforming approximations that employ Crouzeix-Raviart finite elements. Many preconditioning strategies construct preconditioners for this new Schur complement.

For simplicity, we assume here that $\partial\Omega_D = \emptyset$, $g_N = 0$, and that $\mathcal{T}_h$ consists of tetrahedra (triangles in two dimensions). We consider the case of $k = 1$. We employ the set of faces of the elements of the triangulation $\mathcal{T}_h$ that do not lie on $\partial\Omega$, which we denote by $\mathcal{F}_h$, with the understanding that in two dimensions it consists of edges. We will refer to $\mathcal{F}_h$ as the interface between the elements.

We first define a velocity space consisting of discontinuous vectors:

$$\widetilde{V} = \prod_{K \in \mathcal{T}_h} RT_1^h(K);$$

the pressure space remains the same. We then need the space of normal traces of vectors in $\widetilde{V}$ on the interface $\mathcal{F}_h$:

$$W = \prod_{F \in \mathcal{F}_h} \mathbb{P}_0(F).$$

We assume that functions in $W$ are also defined on $\partial\Omega$ by extending them by zero. We then consider the problem of finding $u \in \widetilde{V}$, $p \in Q$, $\lambda \in W$, such that,

$$\begin{aligned}
&\int_\Omega (\mathcal{K}^{-1}u)\cdot v\,dx - \sum_{K\in\mathcal{T}_h}\Big(\int_K \nabla\cdot v\,p\,dx - \int_{\partial K}(v\cdot n_K)\lambda\,ds\Big) = 0,\ v\in\widetilde{V}\\
&-\sum_{K\in\mathcal{T}_h}\int_K \nabla\cdot u\,q\,dx = G(q),\quad q\in Q\\
&\sum_{K\in\mathcal{T}_h}\int_{\partial K}(u\cdot n_K)\mu\,ds = 0,\quad \mu\in W,
\end{aligned} \tag{9.24}$$

with $n_K$ the outward unit normal to the element $K$ and $G(\cdot)$ defined as before. We note that the last condition means that the solution $u$ has a continuous normal component across the element boundaries. By considering the differential problem on an element $K$, we immediately see that $\lambda$ represents the value of the pressure on the interface $\mathcal{F}_h$.

Problem (9.24) is equivalent to the original one. The bilinear forms considered satisfy the continuity and coercivity conditions that ensure well-posedness. The new problem gives rise to the algebraic problem

$$\begin{pmatrix} \tilde{A} & \tilde{B}_1^T & \tilde{B}_2^T \\ \tilde{B}_1 & 0 & 0 \\ \tilde{B}_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{h} \\ 0 \end{pmatrix}, \tag{9.25}$$

where the couple $\{p,\lambda\}$ determines a pressure. Values inside the elements $K\in\mathcal{T}_h$ are given by $p$ (one for each element), while values on the element interface are given by $\lambda$ (one for each face $F\in\mathcal{F}_h$). We can then employ one pressure variable:

$$\eta = \begin{pmatrix} p \\ \lambda \end{pmatrix},\quad \tilde{B} = \begin{pmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{pmatrix},\quad \tilde{g} = \begin{pmatrix} \tilde{h} \\ 0 \end{pmatrix}$$

and write

$$\begin{pmatrix} \tilde{A} & \tilde{B}^T \\ \tilde{B} & 0 \end{pmatrix} \begin{pmatrix} u \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{g} \end{pmatrix}, \tag{9.26}$$

with $u\in\widetilde{V}$ and $\eta\in Q\times W$.

The matrix $\tilde{A}$ is now block diagonal, with each block corresponding to an element $K\in\mathcal{T}_h$. It can easily be inverted and an equation for $\eta$ ($p$ and $\lambda$) is obtained:

$$\tilde{S}\eta = \tilde{g}, \tag{9.27}$$

with the Schur complement

$$\tilde{S} = \tilde{B}\tilde{A}^{-1}\tilde{B}^T = \begin{pmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{pmatrix} \tilde{A}^{-1} \begin{pmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{pmatrix}^T.$$

We note that $\tilde{S}$ is symmetric and Lemma 9.1 can be employed to show that it positive definite. In addition, $\tilde{S}$ can be constructed by subassembling local contributions from each element; this can easily be seen by partitioning the

columns of $\tilde{B}$ into contributions acting on single elements. We can indeed find the bilinear form corresponding to $\tilde{S}$. By considering the original problem (9.24) and eliminating $u$, we find that the bilinear form corresponding to $\tilde{S}$ is

$$\eta^T \tilde{S} \zeta = a_{\tilde{S}}(\eta, \zeta) = \sum_{K \in \mathcal{T}_h} \int_K (\mathcal{K}^{-1} \nabla_K \eta) \cdot \nabla_K \zeta \, dx, \tag{9.28}$$

where the discrete operator $\nabla_K : Q \times W \to \widetilde{V}$, is defined on the element $K$ and for $\eta = \{p, \lambda\}$ by

$$\int_K (\mathcal{K}^{-1} \nabla_K \eta) \cdot v \, dx = - \int_K p \nabla \cdot v \, dx + \int_{\partial K} \lambda (v \cdot n_K) \, ds, \quad v \in \widetilde{V}.$$

We next introduce the finite element space

$$P = P_{CR}(\mathcal{T}_h) \oplus P_B(\mathcal{T}_h),$$

with $P_{CR}(\mathcal{T}_h)$ the space of Crouzeix-Raviart finite elements on $\mathcal{T}_h$ that vanish on $\partial\Omega$, see [149], and $P_B(\mathcal{T}_h)$ the space of quartic bubbles. In particular, $P_{CR}(\mathcal{T}_h)$ consists of piecewise linear functions that are continuous only at the barycenters $b_F$ of the faces $F \in \mathcal{F}_h$; the degrees of freedom are chosen as the values at these barycenters (one per face). The space $P_B(\mathcal{T}_h)$ is spanned by the quartic polynomials that vanish on $\mathcal{F}_h$ (there is then one degree of freedom for each element); for this space we can choose a basis consisting of bubbles that take the value one at the barycenters $b_K$ of the elements $K \in \mathcal{T}_h$. However, the degrees of freedom need not be associated with these barycenters.

The spaces $P$ and $Q \times W$ have the same dimension. We introduce an interpolation operator $\Pi : P \to Q \times W$, as $\Pi\psi = \{\Pi_{\mathcal{T}}\psi, \Pi_{\mathcal{F}}\psi\}$, with $L^2$-projections defined, for each element $K \in \mathcal{T}_h$ and face $F \in \mathcal{F}_h$, by

$$\Pi_{\mathcal{T}}\psi = |K|^{-1} \int_K \psi \, dx = \psi(b_K), \quad \text{on } K,$$
$$\Pi_{\mathcal{F}}\psi = |F|^{-1} \int_F \psi \, dS = \psi(b_F), \quad \text{on } F.$$

We see immediately that $\Pi$ is an isomorphism. If $\eta = \{p, \lambda\} \in Q \times W$ is the solution of the system (9.24), we define a new pressure by $\psi = \Pi^{-1}\{p, \lambda\}$. The new couple $\{u, \psi\}$ satisfies

$$\begin{pmatrix} \tilde{A} & \tilde{B}^T \Pi \\ \Pi^T \tilde{B} & 0 \end{pmatrix} \begin{pmatrix} u \\ \psi \end{pmatrix} = \begin{pmatrix} 0 \\ \Pi^T \tilde{g} \end{pmatrix}. \tag{9.29}$$

If we then eliminate the velocity $u$, we obtain an equation for $\psi \in P$ which involves the modified Schur complement

$$S_P = \Pi^T \tilde{S} \Pi : P \longrightarrow P.$$

The bilinear form corresponding to $S_P$ can be written in terms of integrals over the elements and a suitable $L^2$-projection over $\widetilde{V}$; see [20] for details. The key point is now that $S_P$ is spectrally equivalent to the bilinear form which approximates the Laplacian over the space $P$:

$$\psi^T A_P \phi := a_P(\psi, \phi) = \sum_{K \in \mathcal{T}_h} \int_K (\mathcal{K} \nabla \psi) \cdot \nabla \phi \, dx, \quad \psi, \phi \in P.$$

In addition, the Schur complements $\tilde{S}$ and $S_P$ are also spectrally equivalent for our choice of basis. We summarize these findings in the following lemma. We note that the proof relies on local arguments only and that the result therefore is valid for meshes that are not quasi uniform. The first bound is a direct consequence of the definition of $S_P$ and the choice of a basis for the space $P$. A proof, which employs techniques in [84], can be found in [412, Lem. 3.2].

**Lemma 9.7** *Let the triangulation $\mathcal{T}_h$ be shape regular. Then, there are constants, independent of $h$ and the diffusive tensor $\mathcal{K}$, such that*

$$c_1 \psi^T \tilde{S} \psi \leq \psi^T S_P \psi \leq C_2 \psi^T \tilde{S} \psi,$$
$$c_3 \psi^T A_P \psi \leq \psi^T S_P \psi \leq \psi^T A_P \psi.$$

The lemma can be summarized by saying that in order to find a good preconditioner for the Schur complement $\tilde{S}$ it is enough to find a good preconditioner for the stiffness matrix for Crouzeix-Raviart finite elements augmented by bubble functions.

The last step that remains to perform is to note that for some special cases the stiffness matrix $A_P$ is block diagonal or that it is spectrally equivalent to a block diagonal matrix, with blocks corresponding to the two subspaces $P_{CR}(\mathcal{T}_h)$ and $P_B(\mathcal{T}_h)$.

**Lemma 9.8** *Assume that $\mathcal{T}_h$ is shape regular.*

1. *Let $\mathcal{K} = \rho I$, with $\rho$ constant on each element $K \in \mathcal{T}_h$ and $I$ the identity matrix. Then, $c_3 = 1$ in Lemma 9.7 and $A_P = \operatorname{diag}\{A_{CR}, A_B\}$, where $A_{CR}$ is the stiffness matrix for the Crouzeix-Raviart finite element space $P_{CR}(\mathcal{T}_h)$ and $A_B$ is block diagonal, with each block corresponding to one element.*
2. *Let $\mathcal{K}$ be smooth and a small perturbation of a constant matrix on each element. Then, there exist constants independent of $h$ and $\mathcal{K}$, such that,*

$$c_4 \, a_P(\psi, \psi) \leq a_P(\psi_1, \psi_1) + a_P(\psi_2, \psi_2) \leq C_5 \, a_P(\psi, \psi), \quad \psi = \psi_1 + \psi_2,$$

*for every $\psi_1 \in P_{CR}(\mathcal{T}_h)$ and $\psi_2 \in P_B(\mathcal{T}_h)$.*

We have therefore reduced the problem of preconditioning $\tilde{S}$ from the hybrid-mixed system (9.24) to that of preconditioning nonconforming approximations of Crouzeix-Raviart type. Domain decomposition preconditioners have been proposed for the latter. Here, we mention Sarkis [410, 412, 413],

where overlapping methods are considered together with some choices of coarse spaces that provide condition numbers that are independent of possibly large jumps of $\mathcal{K}$, and Cowsar [147, 146], where some overlapping and iterative substructuring methods are proposed. We will consider the balancing Neumann-Neumann method of Cowsar, Mandel, and Wheeler [148] in more detail in the next section. We also mention Brenner [83], Braess and Verfürth [68], Oswald [366, 367], Vassilevski and Wang [457], Sarkis [411, 412], for some multilevel preconditioners.

### 9.3.3 A Balancing Neumann-Neumann Method

We now consider a method that was proposed and analyzed in [148]. It provides a preconditioner for Method II in [227].

We first present the algorithm as originally introduced in [148]. The balancing Neumann-Neumann algorithm of Sect. 6.2 is applied to the pressure equation (9.27), which involves the bilinear form $a_{\tilde{S}}(\cdot,\cdot)$ in (9.28). Given a nonoverlapping partition into subdomains, it is in fact possible to employ the algorithm of Sect. 6.2 without any essential modification since the degrees of freedom of the pressure space $Q \times W$ can be partitioned into those internal to the subdomains and those on the interface. In addition, the matrix corresponding to the bilinear form can be constructed by subassembling contributions from single subdomains and the local kernels consist of constant functions. Solutions of local Dirichlet problems for the pressure are used to eliminate degrees of freedom internal to the subdomains and we then use solutions of local Neumann problems also for the pressure. We will finally discuss ways of calculating the local operators.

We consider a partition of $\Omega$ into nonoverlapping subdomains, as in Sect. 4.2. For simplicity, we only consider three-dimensional problems; the results carry over to the two-dimensional case. In particular, we assume that $\Omega$ is decomposed into nonoverlapping subdomains (substructures) $\Omega_i, i = 1, \ldots, N$, each of which is the union of shape-regular elements with the finite element nodes on the boundaries of neighboring subdomains matching across the interface. The interface $\Gamma$ is the union of the interior subdomain boundaries. As in the previous section, we employ the set of faces $\mathcal{F}_h$ of the elements of the triangulation $\mathcal{T}_h$ that do not lie on $\partial\Omega$. The letter $F$ will always denote a face in $\mathcal{F}_h$.

Here, we rename the matrix corresponding to $a_{\tilde{S}}(\cdot,\cdot)$ $A_{\tilde{S}}$ and rewrite the system (9.27) as

$$A_{\tilde{S}}\,\eta = \tilde{g}.$$

Degrees of freedom can be partitioned into those internal to the subdomains (corresponding to those in $Q$ and those in $W$ that do not lie on $\Gamma$) and those on the interface (those in $W$ that lie on $\Gamma$):

$$A_{\tilde{S}} = \begin{pmatrix} A_{II} & A_{I\Gamma} \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{pmatrix}.$$

The matrix $A_{\tilde{S}}$ has the structure considered in Chap. 4, where we introduced iterative substructuring methods for conforming approximations of second order, scalar problems; see, in particular, equation (4.7). The matrix $A_{II}$ is block diagonal and a Schur complement system involving the unknowns on the interface $\Gamma$ alone can be obtained, after a step of block Gaussian elimination:

$$S = A_{\Gamma\Gamma} - A_{\Gamma I} {A_{II}}^{-1} A_{I\Gamma} : W_\Gamma \to W_\Gamma,$$

with $W_\Gamma$ the space of traces of functions in $Q \times W$ on $\Gamma$; cf. Sect. 4.3. We note that $W_\Gamma$ consists of discontinuous, piecewise constant functions on $\Gamma$.

The Schur complement $S$, as well as the matrix $A_{\tilde{S}}$, can be constructed by subassembly. We define $W_i$ as the space of traces of functions in $Q \times W$ on $\partial\Omega_i \setminus \partial\Omega$ and let $R_i^T : W_i \to W_\Gamma$, be the extension by zero to the rest of $\Gamma$. We have:

$$S = \sum_{i=1}^{N} R_i^T S^{(i)} R_i,$$

with $S^{(i)}$ the local Schur complements constructed from local stiffness matrices; cf. equation (4.10). The local stiffness matrices are given by formula (9.28) by restricting the sum to the elements contained in one subdomain. A formula for the application of the $S^{(i)}$ will be given below.

We now consider exactly the same hybrid algorithm as in Sect. 6.2, as defined in equation (6.10); the various components are defined in Sect. 6.2.2. The local components involve pseudoinverses of the $S^{(i)}$, for which a formula is given below. The scaling functions $\delta_i^\dagger$ are defined by (6.2), with the assumption that $\mathcal{K} = \rho_i I$ on $\Omega_i$. We note that exactly two terms appear in the sum in (6.2) since degrees of freedom in $W_\Gamma$ are associated with faces in $\mathcal{F}_h$. A coarse space of minimal dimension is the span of the functions $\delta_i^\dagger$ associated with floating subdomains. We note that the only subdomains that are not floating are those that share a face with $\partial\Omega$. We refer to Sect. 6.2 and [148] for further details.

The logarithmic bound for the condition number of the preconditioned operator in Theorem 6.4 remains valid in this case. The proof is given in [148]; we note that it relies on an equivalence of the bilinear form $a_{\tilde{S}}(\cdot,\cdot)$ with one defined on a conforming finite element space.

We now consider the problem of applying the local Schur complements and their pseudoinverses to a vector. The formulas, already given in [148], involve solutions of local saddle point problems and will allow us to relate this algorithm to others that have been proposed over the years. We first note that in practice, it is only necessary to use Lagrange multipliers on the interface $\Gamma$ in order to enforce the continuity of normal velocities; cf. equation (9.24). The Schur complement $S$ remains unchanged.

**Application of the Local Schur Complements $S^{(i)}$.** Given $\lambda_i \in W_i$, we want to calculate

$$\mu_i = S^{(i)}\lambda_i \in W_i.$$

We first need to recall that the Schur complement $S$ (and therefore $S^{(i)}$) has been obtained from the original mixed system (9.24) after repeatedly eliminating variables. In particular, we first eliminated the velocity $u$ and obtained equation (9.27) for $p$ and $\lambda$. We then eliminated the degrees of freedom internal to the subdomains from $\tilde{S} = A_{\tilde{S}}$ and obtained an equation involving $S$. The Schur complements $S$ or $S^{(i)}$ can be then constructed from the blocks of the mixed systems of the form (9.24).

We now fix a subdomain $\Omega_i$, and consider the local spaces

$$RT_1^h(\Omega_i), \qquad Q_0^h(\Omega_i),$$

where $RT_1^h(\Omega_i)$ is the conforming Raviart-Thomas finite element space and $Q_0^h(\Omega_i)$ the space of piecewise constant functions on $\Omega_i$. We consider the following saddle point *Neumann* problem: given $\lambda_i \in W_i$, find $u_i \in RT_1^h(\Omega_i)$, $p_i \in Q_0^h(\Omega_i)$, such that,

$$\begin{aligned} \int_{\Omega_i} (\mathcal{K}^{-1}u_i)\cdot v_i\,dx - \int_{\Omega_i} \nabla\cdot v_i\, p_i\,dx &= -\int_{\partial\Omega_i\cap\Gamma} \lambda_i (v_i\cdot n_i)\,ds,\ v_i \in RT_1^h(\Omega_i) \\ -\int_{\Omega_i} \nabla\cdot u_i\, q_i\,dx &= 0, \quad q_i \in Q_0^h(\Omega_i), \end{aligned}$$

which can be written in matrix form as

$$\begin{pmatrix} A^{(i)} & {B_1^{(i)}}^T \\ B_1^{(i)} & 0 \end{pmatrix} \begin{pmatrix} u_i \\ p_i \end{pmatrix} = \begin{pmatrix} -{B_2^{(i)}}^T\lambda_i \\ 0 \end{pmatrix}.$$

It can easily be shown that

$$\mu_i = -B_2^{(i)}u_i.$$

We have therefore shown that

$$S^{(i)} = \begin{pmatrix} B_2^{(i)} & 0 \end{pmatrix} \begin{pmatrix} A^{(i)} & {B_1^{(i)}}^T \\ B_1^{(i)} & 0 \end{pmatrix}^{-1} \begin{pmatrix} B_2^{(i)} & 0 \end{pmatrix}^T; \tag{9.30}$$

cf. (9.25). We note that the application of $S^{(i)}$ requires the solution of a local mixed Neumann problem, which is always invertible, with datum given by the flux $\lambda_i$.

**Application of the Pseudoinverses of the Schur Complements $S^{(i)}$.** Given $\mu_i \in W_i$, we now want to calculate

$$\lambda_i = {S^{(i)}}^\dagger \mu_i \in W_i.$$

As we have already remarked many times in this monograph, inverses of Schur complements can be applied by employing inverses of the original matrices;

see Sect. 4.3. Looking at the expression for $S^{(i)}$ in (9.30), it is easy to find the $3 \times 3$ mixed system that needs to be considered. The $\lambda_i$ to be computed is the third component of

$$\begin{pmatrix} u_i \\ p_i \\ \lambda_i \end{pmatrix} = - \begin{pmatrix} A^{(i)} & B_1^{(i)^T} & B_2^{(i)^T} \\ B_1^{(i)} & 0 & 0 \\ B_2^{(i)} & 0 & 0 \end{pmatrix}^{\dagger} \begin{pmatrix} 0 \\ 0 \\ \mu_i \end{pmatrix}.$$

We write the equation for the vector on the left-hand side in variational form in order to recognize what type of problem it satisfies. We start by considering the spaces $RT_1^h(\Omega_i)$ and $Q_0^h(\Omega_i)$, employed for the application of $S^{(i)}$, and $W_i$, the space of normal components of local vectors on the subdomain boundary, previously defined. We then look for $u_i \in RT_1^h(\Omega_i)$, $p_i \in Q_0^h(\Omega_i)$, and $\lambda_i \in W_i$, such that,

$$\begin{aligned} &\int_{\Omega_i} (\mathcal{K}^{-1} u_i) \cdot v_i \, dx - \int_{\Omega_i} \nabla \cdot v_i \, p_i \, dx + \int_{\partial\Omega_i \cap \Gamma} (v_i \cdot n_i) \lambda_i \, ds = 0, \\ &- \int_{\Omega_i} \nabla \cdot u_i \, q_i \, dx = 0, \\ &\int_{\partial\Omega_i \cap \Gamma} (u_i \cdot n_i) \psi_i \, ds = - \int_{\partial\Omega_i \cap \Gamma} \mu_i \psi_i \, ds, \end{aligned}$$

for $v_i \in RT_{1;0}^h(\Omega_i)$, $q_i \in Q_0^h(\Omega_i)$, and $\psi_i \in W_i$. We note that the last condition means that $u_i \cdot n_i = \mu_i$ on $\partial\Omega_i \cap \Gamma$. We can therefore solve the following Dirichlet problem: find $u_i \in RT_1^h(\Omega_i)$ and $p_i \in Q_0^h(\Omega_i)$, such that $u_i \cdot n_i = -\mu_i$ on $\partial\Omega_i \cap \Gamma$ and

$$\begin{aligned} &\int_{\Omega_i} (\mathcal{K}^{-1} u_i) \cdot v_i \, dx - \int_{\Omega_i} \nabla \cdot v_i \, p_i \, dx = 0, \quad v_i \in RT_{1;0}^h(\Omega_i), \\ &- \int_{\Omega_i} \nabla \cdot u_i \, q_i \, dx \qquad\qquad\qquad\qquad = 0, \quad q_i \in Q(\Omega_i), \end{aligned}$$

with $RT_{1;0}^h(\Omega_i)$ denoting the subspace of $RT_1^h(\Omega_i)$ of vectors with vanishing normal component on $\partial\Omega_i \cap \Gamma$. We note that $u_i$ provides a *discrete harmonic Darcy extension* of the normal component $\mu_i$; see Sect. 9.4.2 for a discrete Stokes extension.

We have already noted in Sect. 9.3 that this problem, in general, is not solvable on a floating subdomain unless a compatibility condition is satisfied; see (9.16). In particular, we need

$$\int_{\partial\Omega_i} \mu_i \, ds = 0.$$

This property is ensured by the fact that we apply local solvers after applying a coarse solver, i.e., to a balanced vector, which ensures that the required compatibility condition is satisfied.

Once we have a solution of the Dirichlet problem, we choose any $\lambda_i$ satisfying the first equation of the $3 \times 3$ mixed system:

$$\int_{\partial\Omega_i \cap \Gamma} \lambda_i (v_i \cdot n_i)\, ds = -\int_{\Omega_i} (\mathcal{K}^{-1} u_i) \cdot v_i\, dx + \int_{\Omega_i} \nabla \cdot v_i\, p_i\, dx, \quad v_i \in RT_1^h(\Omega_i),$$

or, equivalently,

$$B_2^{(i)^T} \lambda_i = -A^{(i)} u_i - B_1^{(i)^T} p_i.$$

More details can be found in [227] and [148].

The attentive reader, who has followed us this far, should certainly has noted some similarities with some other methods presented in this monograph. We recall that in the previous section, we reformulated the mixed Darcy system by employing Lagrange multipliers in order to enforce continuity of the velocity across the elements. The equation considered by the balancing Neumann-Neumann of this section is related to these Lagrange multipliers, which have the meaning of a pressure on the interface $\Gamma$. A similar procedure is employed in Chap. 6 for the FETI methods, where preconditioned equations for Lagrange multipliers are considered; cf. Sect. 6.3, in particular.

In addition, formula (9.30), which gives the local component of the unpreconditioned operator $S$, resembles that of the FETI operator $F$ in (6.30). Indeed, we have the inverse of a local Neumann matrix, pre- and post-multiplied by a matrix related to the enforcement of continuity across the subdomain boundaries; cf. the matrix $B_2^{(i)}$ in (9.30). Moreover, the application of the local components of the preconditioner involves the solution of local Dirichlet problems, just as in the Dirichlet preconditioner for FETI methods in formula (6.35).

It can be shown that the unpreconditioned operator $S$ of this section coincides with the operator $F_\Gamma$ in Sect. 9.3.1 for the case of two subdomains and if basis functions for the velocity spaces $RT_1^h(\Omega_i)$ and the trace space $W_\Gamma$ are chosen so that the matrices $B_2^{(i)}$ become identities. In addition, the preconditioned operator of the balancing Neumann-Neumann method of this section coincides with the operator $S_\Gamma F_\Gamma$ of Sect. 9.3.1.

Far from trying to force any one method to fit into a specific class, we have added these remarks in order to point out some connections between seemingly different methods that have been proposed over the years and to show that they rely on similar important ideas that have permeated domain decomposition research over the years. Balancing Neumann-Neumann and FETI methods for the Stokes system have been developed later and some of them are presented in Sect. 9.4.2. Generalizations of balancing Neumann-Neumann methods for Darcy problems on nonmatching grids have also been proposed in [387]. Iterative substructuring methods for Raviart-Thomas approximations are presented in Sect. 10.2.

### 9.3.4 Overlapping Methods

We now briefly consider two overlapping methods that were originally proposed by Ewing and Wang [189] and Mathew [345, 346, 347]. The two methods both reduce the Darcy problem to a symmetric, positive definite problem for a divergence free vector. They next employ a two level overlapping method in a space of functions which are divergence free in a finite element sense.

We consider the mixed system (9.1) and assume, for simplicity, that $\partial\Omega_N = \emptyset$, $g_D = 0$, and $k = 1$. We obtain,

$$\begin{aligned} Au + B^T p &= F, \\ Bu \qquad &= G. \end{aligned} \tag{9.31}$$

We recall that $u \in V = RT^h_{1;0}(\Omega)$, the space of Raviart-Thomas vectors on the mesh $\mathcal{T}_h$ with vanishing normal component on the boundary, that $p \in Q = Q^h_{0;0}(\Omega)$, the space of piecewise constant functions with vanishing mean value on $\Omega$, and that a compatibility condition needs to be satisfied; cf. (9.16).

We next consider the coarse mesh $\mathcal{T}_H$, which provides a nonoverlapping partition into substructures, $\mathcal{T}_H = \{\Omega_i,\ 1 \le i \le N\}$, and we then extend each substructure $\Omega_i$ to a larger subregion $\Omega_i'$, such that $\partial\Omega_i'$ does not cut through any fine element. We assume that these new subregions are shape regular and that Assumptions 3.1 (small overlap) and 3.2 (finite covering) hold. We also set $\Omega_0' = \Omega$. We refer to Sect. 3.2 for details.

The first step of the algorithm consists in finding a velocity $u^* \in V$ that satisfies the second equation of (9.31):

$$Bu^* = G,$$

and writing

$$u = u^* + w.$$

The equation for $w \in V$ then becomes

$$\begin{aligned} Aw + B^T p &= F - Au^* = \tilde{f}, \\ Bw \qquad &= 0. \end{aligned}$$

The new velocity is therefore divergence free in the sense that

$$w \in Z = kernel(B) = \{v \in V \mid \quad \nabla \cdot v = 0\}.$$

We therefore have a symmetric, positive definite problem: find $w \in Z$, such that

$$a(w, v) = \tilde{f}(v), \quad v \in Z, \tag{9.32}$$

with the appropriate right-hand side, and an overlapping preconditioner, as in Chap. 3.

**Step 1.** We consider the problem of finding $u^*$. We introduce the Raviart-Thomas space on the coarse mesh $\mathcal{T}_H$, $V_0 = RT^H_{1;0}(\Omega)$, of vectors with a vanishing normal component on $\partial\Omega$, and the space, $Q_0 = Q^H_{0;0}(\Omega)$ of constant pressures on each coarse element with vanishing mean value. Let $R_0^T : V_0 \times Q_0 :\to V \times Q$, be the natural interpolation operator from the subspace. Let

$$\begin{pmatrix} A_0 & B_0^T \\ B_0 & 0 \end{pmatrix} = R_0 \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} R_0^T,$$

be the matrix for the mixed problem on the coarse spaces. We then find that

$$\begin{pmatrix} u_0^* \\ p_0^* \end{pmatrix} = R_0^T \begin{pmatrix} A_0 & B_0^T \\ B_0 & 0 \end{pmatrix}^{-1} R_0 \begin{pmatrix} F \\ G \end{pmatrix}.$$

We note that the residual $G - Bu_0^*$ has vanishing mean value on each coarse element $\Omega_i$, since, for $q_0 \in Q_0$,

$$q_0^T(G - Bu_0^*) = q_0^T(G - B_0 u_0^*) = 0.$$

We can therefore consider local Dirichlet problems. We define the local spaces by

$$V_i = RT^h_{1;0}(\Omega_i) = RT^h_1(\Omega_i) \cap H_0(\mathrm{div}\,;\Omega_i), \qquad Q_i = Q^h_{0;0}(\Omega_i),$$

where the velocities have vanishing normal component on the subdomain boundaries and the pressures have vanishing mean values. We also employ the natural interpolation operators $R_i^T : V_i \times Q_i :\to V \times Q$, and the local matrices

$$\begin{pmatrix} A_i & B_i^T \\ B_i & 0 \end{pmatrix} = R_i \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} R_i^T.$$

We note that these local matrices represent local Dirichlet problems on the substructures and may therefore not be invertible. However, the following local vectors

$$\begin{pmatrix} u_i^* \\ p_i^* \end{pmatrix} = R_i^T \begin{pmatrix} A_i & B_i^T \\ B_i & 0 \end{pmatrix}^{\dagger} R_i \begin{pmatrix} F - Au_0^* \\ G - Bu_0^* \end{pmatrix}$$

are well defined since the residual $G - Bu_0^*$ has a vanishing mean value on $\Omega_i$ and the solution $u_i^*$ should have a vanishing normal component along $\partial\Omega_i$; cf. (9.16). It can easily be verified that

$$u^* = u_0^* + \sum_{i=1}^{N} u_i^*$$

is the desired vector. We note that the $A_i$, can be replaced by different, simpler matrices for $1 \le i \le N$, since we only need to provide a velocity $u^*$ that satisfies the second equation of (9.31). The procedure given here can be found in [345, 346]; see [189] for a similar construction.

**Step 2.** We now consider the preconditioning of the positive definite problem (9.32). The space $Z$ can be fully characterized using Lemma B.23 in three dimensions and the decomposition (B.32) in two dimensions. Two dimensional problems are considered in [189]. We have

$$Z = \mathbf{curl}\, S, \quad S \subset H^1(\Omega),$$

with $S$ the space of continuous piecewise linear functions. We can therefore look for $w = \mathbf{curl}\,\psi$, $\psi \in S$, such that,

$$a(\mathbf{curl}\,\psi, \mathbf{curl}\,\phi) = \int_\Omega (\mathcal{K}^{-1}\mathbf{curl}\,\psi) \cdot \mathbf{curl}\,\phi\, dx = \tilde{f}(\mathbf{curl}\,\phi), \quad \phi \in S.$$

The relevant bilinear form becomes the usual grad-grad form arising from the Laplacian after a rotation of the axes, and the two-level overlapping preconditioner in Chap. 3 can be employed. Additive and multiplicative methods are considered in [189]. Optimal convergence is proven for the case of generous overlap.

The three-dimensional case is more complicated since $Z$ consists of the curls of Nédélec finite element functions. In addition, the vector potential that provides the velocity $w$ is defined only up to the gradient of a continuous function. This approach has been explored in [118].

Another approach is employed in [345, 346, 347], where coarse mixed problems and local mixed problems on overlapping subdomains are employed. We do not go into details but refer to [346] for the precise algorithm. An analysis in two dimensions and for the case of generous overlap can also be found in that paper.

## 9.4 The Stokes Problem and Almost Incompressible Elasticity

Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. We recall that the Stokes problem corresponds to the choices $t = 0$ and that

$$a(u, v) = \nu \int_\Omega \nabla u \cdot \nabla v\, dx,$$
$$b(u, p) = -\int_\Omega \nabla \cdot u\, p\, dx,$$

while almost incompressible elasticity results in a $t > 0$ and the forms

$$a(u, v) = 2\mu \int_\Omega \epsilon(u) : \epsilon(v)\, dx,$$
$$b(u, p) = -\int_\Omega \nabla \cdot u\, p\, dx,$$
$$c(p, q) = \int_\Omega p\, q\, dx.$$

see appendices A.7.2, B.4.1, and B.4.2. In the following, we will refer generically to these problems as *Stokes problems*. In our discussion, we restrict ourselves to the case of homogeneous Dirichlet conditions, i.e., a zero velocity or displacement $u$ on $\partial\Omega$. We employ a finite or spectral element space $V \subset H_0^1(\Omega)^n$, consisting of continuous, piecewise polynomial functions, that vanish on $\partial\Omega$, on a given mesh. For the pressure the space $Q \subset L_0^2(\Omega)$ consists of discontinuous polynomial functions with a vanishing mean value in $\Omega$. There are of course also well-known mixed finite element approximations with continuous pressure spaces, such as the Taylor-Hood methods. However, the study of domain decomposition methods for such cases appear to introduce a number of additional technical difficulties; see, e.g., [284, 119] for some methods with continuous pressures.

We note that the stronger coercivity property (9.7) holds in this case, because of the Friedrichs' and Korn's inequalities in Lemmas A.14 and A.39, respectively. For this reason, we choose

$$\|u\|_V^2 = a(u,u) = u^T A u.$$

Since the bilinear form $c(\cdot,\cdot)$ is the $L^2$-scalar product, we also set

$$\|q\|_Q^2 = c(q,q) = q^T C q.$$

For this particular case, we have

$$\alpha_1 = \beta = \delta = 1,$$

in (9.3), (9.7), and (9.6). We note that spectral bounds for the Schur complement $S$ can be found since Lemma 9.1 can be applied. In particular, for $t = 0$,

$$\gamma^2\, q^T C q \le q^T S_0 q \le \alpha_2^2\, q^T C q, \quad q \in Q. \tag{9.33}$$

We will only consider block preconditioners and iterative substructuring algorithms in the following sections. We note that successful overlapping preconditioners have also been considered by Gervasio [222], Fischer [203], Klawonn and Pavarino [283, 284, 285], Rønquist [404], Fischer, Miller, and Tufo [204], and Pavarino [379]

### 9.4.1 Block Preconditioners

We consider the system (9.1) and look for symmetric, block preconditioners of the form

$$\hat{\mathcal{B}} = \begin{pmatrix} \hat{A} & 0 \\ 0 & \hat{C} \end{pmatrix}, \tag{9.34}$$

with $\hat{A}$ and $\hat{C}$ positive definite, symmetric matrices. We denote the special case with $\hat{A} = A$ and $\hat{C} = C$ by $\mathcal{B}$. Since $\mathcal{A}$ is symmetric and $\hat{\mathcal{B}}$ is positive

definite, we can employ the CR algorithm and the theory developed in Sect. 9.2.

The two strategies presented in Sect. 9.2 give rise to the same conclusion: Theorem 9.3 can be applied if $\hat{A}$ and $\hat{C}$ are good preconditioners for the operators associated with the scalar products in $V$ and $Q$ (or, equivalently, for $A$ and $C$). Here, we instead consider the particular analysis in [282] which provides explicit bounds for the condition number in terms of the stability constants of the discrete Stokes problem. We also refer to [406, 405, 270, 460, 418, 461] for alternative analyses and results.

We assume that there exist strictly positive constants, such that,

$$\begin{aligned} a_1\|u\|_V^2 = a_1 u^T A u \le u^T \hat{A} u \le a_2 u^T A u = a_2\|u\|_V^2, \quad u \in V, \\ c_1\|p\|_Q^2 = c_1 p^T C p \le p^T \hat{C} p \le c_2 p^T C p = c_2\|p\|_Q^2, \quad p \in Q. \end{aligned} \tag{9.35}$$

The next lemma gives explicit bounds; a proof follows those of Lemmas 1 and 2 and Theorems 3 and 4 in [282].

**Theorem 9.9** *Let $0 \le t \le 1$. The condition number of the preconditioned operator $\mathcal{B}^{-1}\mathcal{A}$ satisfies*

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) \le \frac{1/2 + \sqrt{\alpha_2 + 1/4}}{-1/2 + \sqrt{\gamma + 1/4}},$$

*with $\alpha_2$ and $\gamma$ the continuity and inf-sup constants of the bilinear form $b(\cdot,\cdot)$. In addition,*

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \le \kappa(\mathcal{B}^{-1}\mathcal{A}) \frac{\max\{a_2, c_2\}}{\min\{a_1, c_1\}} \le \frac{1/2 + \sqrt{\alpha_2 + 1/4}}{-1/2 + \sqrt{\gamma + 1/4}} \left(\frac{\max\{a_2, c_2\}}{\min\{a_1, c_1\}}\right).$$

*Proof.* We will use the notation

$$\begin{aligned} \lambda_{max}(L) &= \max\{\lambda : \lambda \in \sigma(L)\}, \quad \mu_{max}(L) = \max\{|\lambda| : \lambda \in \sigma(L)\}, \\ \lambda_{min}(L) &= \min\{\lambda : \lambda \in \sigma(L)\}, \quad \mu_{min}(L) = \min\{|\lambda| : \lambda \in \sigma(L)\}, \end{aligned}$$

for a generic matrix $L$ and its spectrum $\sigma(L)$. The spectral radius $\rho(L)$ satisfies

$$\rho(L) = \mu_{max}(L) = \|L\|_2.$$

**Step 1.** We consider the operator $\mathcal{B}^{-1}\mathcal{A}$ and show that

$$\begin{aligned} \mu_{max}(\mathcal{B}^{-1}\mathcal{A}) &= \frac{1-t^2}{2} + \sqrt{\lambda_{max}(C^{-1/2} S_0 C^{-1/2}) + \left(\frac{1+t^2}{2}\right)^2}, \\ \mu_{min}(\mathcal{B}^{-1}\mathcal{A}) &= -\frac{1-t^2}{2} + \sqrt{\lambda_{min}(C^{-1/2} S_0 C^{-1/2}) + \left(\frac{1+t^2}{2}\right)^2}. \end{aligned} \tag{9.36}$$

In order to do so, we employ the matrix

$$\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2} = \begin{pmatrix} I & \tilde{B}^T \\ \tilde{B} & -t^2 I \end{pmatrix}, \qquad \tilde{B} = C^{-1/2} B A^{-1/2}.$$

Setting up the eigenvalue problem for the eigenvalues $\lambda$ and the eigenvectors $\mathcal{U} = \{u, p\}$, we see that the eigenvalues that are different from 1, satisfy

$$\tilde{B}\tilde{B}^T p = C^{-1/2} S_0 C^{-1/2} p = (\lambda - 1)(t^2 + \lambda)p,$$

and therefore

$$(\lambda - 1)(t^2 + \lambda) = z, \qquad z \in \sigma(C^{-1/2} S_0 C^{-1/2}).$$

Finding $\lambda$ yields (9.36).

**Step 2.** Finding the extremal values for $0 \leq t \leq 1$ then gives

$$\begin{aligned} \mu_{max}(\mathcal{B}^{-1}\mathcal{A}) &\leq \frac{1}{2} + \sqrt{\lambda_{max}(C^{-1/2} S_0 C^{-1/2}) + \frac{1}{4}} \\ \mu_{min}(\mathcal{B}^{-1}\mathcal{A}) &\geq -\frac{1}{2} + \sqrt{\lambda_{min}(C^{-1/2} S_0 C^{-1/2}) + \frac{1}{4}}. \end{aligned} \tag{9.37}$$

**Step 3.** We note that the Schur complement $S_0$ is symmetric and positive definite. Upper and lower bounds for its extremal eigenvalues are provided by Lemma 9.1 and are given in (9.33) for our particular case. This proves the bound for $\kappa(\mathcal{B}^{-1}\mathcal{A})$.

**Step 4.** In order to prove the second bound, we write

$$\begin{aligned} \mu_{max}(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}) = \rho(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}) &= \|\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\|_2 \\ &= \sup_{0 \neq \mathcal{U} \in X} \frac{\mathcal{U}^T \hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\mathcal{U}}{\mathcal{U}^T\mathcal{U}} \\ &\leq \left( \sup_{0 \neq \mathcal{U} \in X} \frac{\mathcal{U}^T\mathcal{A}\mathcal{U}}{\mathcal{U}^T\mathcal{B}\mathcal{U}} \right) \left( \sup_{0 \neq \mathcal{U} \in X} \frac{\mathcal{U}^T\mathcal{B}\mathcal{U}}{\mathcal{U}^T\hat{\mathcal{B}}\mathcal{U}} \right) \\ &\leq \mu_{max}(\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}) \ \max\{a_2, c_2\}. \end{aligned}$$

A lower bound can be found in a similar way by noting that, for a matrix $L$, $\mu_{min}(L) = 1/\rho(L^{-1})$. □

This result can be summarized as follows: in order for the block diagonal preconditioner to be effective, it is enough that the blocks $\hat{A}$ and $\hat{C}$ are good preconditioners for $A$ and $C$. We note that this result is independent

of the small parameter $t$ (and therefore holds for $t = 0$), while it may depend on the discretization parameters if the inf-sup constant $\gamma$ depends on the discretization order, as for certain spectral element approximations.

Many good preconditioners for $A$ have been developed in the previous chapters. For the vector Laplacian, we refer to Chap. 3 for overlapping preconditioners, to Chap. 5 for primal iterative substructuring methods, and to Chap. 6 for Neumann-Neumann and FETI methods. For the elasticity operator, we refer to Chap. 8, where some overlapping and nonoverlapping methods have been presented. Finally for spectral element approximations, we refer to Chap. 7.

For the block $C$, we note that this operator is spectrally equivalent to a mass matrix on a finite or spectral element space consisting of discontinuous functions. It is therefore block diagonal, with each block corresponding to one element, and it can therefore be inverted at very low cost. In addition, we recall that for spectral elements the local mass matrices are diagonal; see Chap. 7.

### 9.4.2 Iterative Substructuring Methods

In this section, we will discuss some iterative substructuring methods for the Stokes problem. As for symmetric, positive definite problems, they are based on a nonoverlapping partition into subdomains or substructures $\Omega_i, i = 1, \ldots, N$, of the computational domain $\Omega$. We will employ the notation and assumptions of Sect. 4.2. In particular, we recall that the substructures are unions of fine elements for $h$ approximations or consist of one or a few spectral elements for spectral element approximations (cf. Sect. 7.4 for the latter case). The interface $\Gamma$ is the union of the intersections of the subdomain boundaries. We refer to Sect. 4.2 for a precise definition of this interface and for additional definitions and notations.

We recall that two simple iterative substructuring preconditioned algorithms have been presented in Sect. 9.3.1 for the case of two subdomains, namely a Neumann-Neumann and a FETI method. Here we mainly focus on the generalization of the former to the case of many subdomains while providing references for some other methods at the end of this subsection. (We note that the work that is most similar to the Neumann-Neumann method that we will present is the work by Li [314, 315, 316] on FETI-DP algorithms.)

For the Stokes problem not only are the local Neumann problems singular for floating subdomains but additionally the boundary values of the local Dirichlet problems should satisfy the zero flux constraints

$$\int_{\partial\Omega_i} u \cdot n ds = 0.$$

We recall that by the divergence theorem, this is a necessary condition for a divergence-free extension. After the averaging of the solutions of the local

Neumann problems, we cannot expect that this condition will be met. However, we can effectively drop this constraint for the local Dirichlet problems by restricting the pressure spaces to have zero averages. The resulting reduced Dirichlet problems are well posed. The pressure component that are piecewise constant on the subdomains will instead be part of a coarse component of the preconditioner. For this coarse space, we also have to include a sufficiently rich velocity space to make this global problem inf-sup stable; naturally, we will build our preconditioner only from stable subspace components. We note that the solution of the coarse problem and, in particular, its pressure components will provide the necessary mechanism of global communication of information across the region $\Omega$.

We will rely on a partition of the degrees of freedom into those *interior* to the substructures and those on the *interface* $\Gamma$. The interior degrees of freedom consist of local velocities supported on single subdomains and local pressures that have vanishing mean values on the substructures. These unknowns can then be eliminated in parallel by a step of block Gaussian elimination. A Schur complement system is then obtained for the interface $\Gamma$; this will correspond to an indefinite, saddle point problem. The elimination of the pressure variables, which are constant on each subdomain, through the solution of a saddle point coarse problem, yields a symmetric, positive definite system for the interface velocities for which the conjugate gradient method can be employed.

We note that natural boundary conditions for the Stokes problems have been already given in Sect. 9.3.1:

$$\nu \nabla u_i \cdot n - p n_i = r_i, \quad \text{on} \quad \partial\Omega, \ i = 1, 2, \ldots, n. \tag{9.38}$$

They are derived by using Green's formula. In contrast with the case of Dirichlet boundary conditions, the pressure is now uniquely determined, and the pressure space should now be taken to be $L^2(\Omega)$. In this case, as for the Laplace operator, each velocity component is determined only up to a constant, and the $r_i$ must satisfy compatibility conditions.

**Schur Complement Systems.** We will now derive a Schur complement system; further details and additional formulas can be found in [386]. We proceed as in Sect. 4.3 but we note that the notation here is a little heavier since we deal with saddle point problems.

In preparation for the elimination of the interior degrees of freedom, we reorder the vector of unknowns as

$$\begin{pmatrix} u_I \\ p_I \\ u_\Gamma \\ p_0 \end{pmatrix} \quad \begin{array}{l} \text{interior velocities} \\ \text{interior pressures with zero average in each } \Omega_i \\ \text{interface velocities} \\ \text{constant pressures in each } \Omega_i. \end{array}$$

Then, after using the same permutation, the discrete Stokes system matrix can be written as

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{II} & \mathcal{A}_{\Gamma I}^T \\ \mathcal{A}_{\Gamma I} & \mathcal{A}_{\Gamma\Gamma} \end{pmatrix} = \left(\begin{array}{cc|cc} A_{II} & B_{II}^T & A_{\Gamma I}^T & 0 \\ B_{II} & 0 & B_{I\Gamma} & 0 \\ \hline A_{\Gamma I} & B_{I\Gamma}^T & A_{\Gamma\Gamma} & B_0^T \\ 0 & 0 & B_0 & 0 \end{array}\right);$$

see also (9.18) for the case of two subdomains for which constant pressures do not need to be considered.

Eliminating the interior unknowns $u_I$ and $p_I$ by static condensation, we obtain the saddle point Schur complement system

$$S \begin{pmatrix} u_\Gamma \\ p_0 \end{pmatrix} = \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix}, \tag{9.39}$$

where

$$\begin{aligned} S &= \mathcal{A}_{\Gamma\Gamma} - \mathcal{A}_{\Gamma I}\mathcal{A}_{II}^{-1}\mathcal{A}_{\Gamma I}^T \\ &= \begin{pmatrix} A_{\Gamma\Gamma} & B_0^T \\ B_0 & 0 \end{pmatrix} - \begin{pmatrix} A_{\Gamma I} & B_{I\Gamma}^T \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_{II} & B_{II}^T \\ B_{II} & 0 \end{pmatrix}^{-1} \begin{pmatrix} A_{\Gamma I}^T & 0 \\ B_{I\Gamma} & 0 \end{pmatrix} = \begin{pmatrix} S_\Gamma & B_0^T \\ B_0 & 0 \end{pmatrix}, \end{aligned}$$

and a similar formula holds for the new right hand side $(\tilde{b}^T \;\; 0)^T$; see [386]. We note that this Schur complement $S$ is not the same as the one derived in the beginning of this chapter in the context of Uzawa-type algorithms.

By using a second permutation that reorders the interior velocities and pressures subdomain by subdomain, we note that $\mathcal{A}_{II}^{-1}$ represents the solution of $N$ decoupled Stokes problems, one for each subdomain and all uniquely solvable, with Dirichlet data given on $\partial\Omega_i$:

$$\mathcal{A}_{II}^{-1} = \begin{pmatrix} \mathcal{A}_{II}^{(1)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \mathcal{A}_{II}^{(N)^{-1}} \end{pmatrix}.$$

This is the matrix associated with the discrete Stokes extension operator $\mathcal{SH}$ described below.

As for iterative substructuring methods for positive definite problems, the matrix $S$ need not be explicitly assembled since only its action on a vector is needed in a Krylov iteration. This operation, essentially, only requires the action of $\mathcal{A}_{II}^{-1}$ on a vector, i.e., the solution of $N$ decoupled Stokes problems. In other words, its action can be computed by subassembling the actions of the subdomain Schur complements $S^{(i)}$ defined for $\Omega_i$, by

$$\begin{aligned} S^{(i)} &= \mathcal{A}_{\Gamma\Gamma}^{(i)} - \mathcal{A}_{\Gamma I}^{(i)}\mathcal{A}_{II}^{(i)-1}\mathcal{A}_{\Gamma I}^{(i)^T} \\ &= \begin{pmatrix} A_{\Gamma\Gamma}^{(i)} & B_0^{(i)^T} \\ B_0^{(i)} & 0 \end{pmatrix} - \begin{pmatrix} A_{\Gamma I}^{(i)} & B_{I\Gamma}^{(i)^T} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_{II}^{(i)} & B_{II}^{(i)^T} \\ B_{II}^{(i)} & 0 \end{pmatrix}^{-1} \begin{pmatrix} A_{\Gamma I}^{(i)^T} & 0 \\ B_{I\Gamma}^{(i)} & 0 \end{pmatrix} \\ &= \begin{pmatrix} S_\Gamma^{(i)} & B_0^{(i)^T} \\ B_0^{(i)} & 0 \end{pmatrix}. \end{aligned}$$

Once $\begin{pmatrix} u_\Gamma \\ p_0 \end{pmatrix}$ is known, $\begin{pmatrix} u_I \\ p_I \end{pmatrix}$ can be found by back-substitution:

$$\begin{pmatrix} u_I \\ p_I \end{pmatrix} = \begin{pmatrix} A_{II} & B_{II}^T \\ B_{II} & 0 \end{pmatrix}^{-1} \left( \begin{pmatrix} b_I \\ 0 \end{pmatrix} - \begin{pmatrix} A_{I\Gamma}^T & 0 \\ B_{I\Gamma} & 0 \end{pmatrix} \begin{pmatrix} u_\Gamma \\ p_0 \end{pmatrix} \right).$$

We note that $u_I$ and $p_I$ are independent of $p_0$.

The substructuring procedure described here is associated with the space decomposition

$$V \times Q = \oplus_{i=1}^N V_i \times Q_i \oplus \tilde{V}_\Gamma \times Q_0 .$$

The interior spaces are defined as

$$V_i = V \cap H_0^1(\Omega_i), \quad Q_i = Q \cap L_0^2(\Omega_i).$$

As for iterative substructuring methods for positive definite problems, the interface velocity space is associated with an extension operator from the interface $\Gamma$; in addition the coarse pressures are constant in each subdomain. We have

$$\tilde{V}_\Gamma = \mathcal{SH}(V_\Gamma) = \{v \in V \mid v|_{\Omega_i} = \mathcal{SH}(v|_{\partial\Omega_i}),\ i = 1, 2, \ldots, N\},$$
$$Q_0 = \{q \in Q \mid q|_{\Omega_i} = \text{constant},\ i = 1, 2, \ldots, N\}.$$

Here $\mathcal{SH} : V_\Gamma \to V$, is the velocity component of the *discrete Stokes harmonic extension* operator that maps an interface velocity $u_\Gamma \in V_\Gamma := V|_\Gamma$ onto the solution $(\tilde{u}, \tilde{p})$ of the following homogeneous Stokes problem defined on each subdomain separately: find $\tilde{u} \in V$, with $\tilde{u} = u_\Gamma$ on each $\partial\Omega_i$, and $\tilde{p} \in \sum_{i=1}^N Q_i$ such that on $\Omega_i$

$$\begin{aligned} a(\tilde{u}, v) + b(v, \tilde{p}) &= 0 \quad v \in V_i, \\ b(\tilde{u}, q) &= 0 \quad q \in Q_i. \end{aligned}$$

The following comparison of the energy of the discrete Stokes extensions $\mathcal{SH}$ and the discrete harmonic extensions $\mathcal{H}$ (see Sect. 4.4) of each velocity component can be found in [71] and [223] for finite element discretizations and in [128] and [310] for spectral element discretizations. We note that the corresponding local bounds for individual subdomains are equally valid and that the upper bound has an elementary proof. This lemma allows us to reduce part of the proof of our main result to arguments for the scalar elliptic case.

**Lemma 9.10** *Let $u_\Gamma \in V_\Gamma$, and let $\gamma$ be the inf-sup constant of the chosen mixed finite element space $V \times Q$. Then,*

$$c\gamma^2\, a(\mathcal{SH}u_\Gamma, \mathcal{SH}u_\Gamma) \le a(\mathcal{H}u_\Gamma, \mathcal{H}u_\Gamma) \le a(\mathcal{SH}u_\Gamma, \mathcal{SH}u_\Gamma).$$

If we define an interface inner product by

$$s(u_\Gamma, v_\Gamma) = a(\mathcal{SH}(u_\Gamma), \mathcal{SH}(v_\Gamma)) = u_\Gamma^T S_\Gamma v_\Gamma,$$

and by $b_0(u_\Gamma, p_0)$ the restriction of the other bilinear form to the Stokes harmonic extensions, the variational formulation of the saddle point Schur complement problem (9.39) can be given by (see Pavarino and Widlund [384, lemma 7.1]): find $u_\Gamma \in V_\Gamma$ and $p_0 \in Q_0$ such that

$$\begin{array}{lll} s(u_\Gamma, v_\Gamma) + b_0(v_\Gamma, p_0) = \widetilde{F}(v_\Gamma) & v_\Gamma \in V_\Gamma, \\ b_0(u_\Gamma, q_0) \qquad\qquad = \quad 0 & q_0 \in Q_0 \,. \end{array} \tag{9.40}$$

Problem (9.40) is equivalent to the following positive definite problem: find $u_\Gamma \in V_{\Gamma,B}$ such that

$$s(u_\Gamma, v_\Gamma) = \widetilde{F}(v_\Gamma) \quad v_\Gamma \in V_{\Gamma,B}\,, \tag{9.41}$$

where $V_{\Gamma,B}$ is the subspace of *balanced velocities* defined by

$$\begin{aligned} V_{\Gamma,B} &= \ker B_0 = \{v_\Gamma \in V_\Gamma \mid B_0 v_\Gamma = 0\} \\ &= \{v \in V \mid v|_{\Omega_i} = \mathcal{SH}(v|_{\partial\Omega_i}),\ \textstyle\int_{\partial\Omega_i} v \cdot n = 0,\ i = 1, \ldots, N\}\,. \end{aligned}$$

In the following, we will also use the term *balanced* for the couple $V_{\Gamma,B} \times Q_0$. (We also note that in Pavarino and Widlund [386], this is referred to as the *benign* space.)

The equivalence of problems (9.40) and (9.41) follows from the fact that $V_{\Gamma,B}$ consists of the functions of $\mathcal{SH}(\widetilde{V})$ that satisfy the zero flux constraints on the subdomain boundaries. We obtain (9.40) by using Lagrange multipliers $p_0 = (p_{0,i})_{i=1}^N$ to enforce the constraints given in the definition of the balanced subspace.

**A Neumann-Neumann Preconditioner.** We will solve the saddle point Schur complement problem

$$S \begin{pmatrix} u_\Gamma \\ p_0 \end{pmatrix} = \begin{pmatrix} S_\Gamma & B_0^T \\ B_0 & 0 \end{pmatrix} \begin{pmatrix} u_\Gamma \\ p_0 \end{pmatrix} = \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} \tag{9.42}$$

by a preconditioned Krylov space method such as GMRES or PCG. The latter can be applied to this indefinite problem, because we will start and keep the iterates in the positive definite subspace $V_{\Gamma,B}$ of balanced velocities; we note that the indefinite bilinear form defined by $S$ is a norm on space $V_{\Gamma,B} \times Q_0$.

We consider the balancing Neumann-Neumann preconditioner of Pavarino and Widlund [386], which is based on the solution of a coarse Stokes problem with a few degrees of freedom per subdomain and of local Stokes problems with natural and essential boundary conditions on each subdomain. This preconditioner is of the same hybrid form that was introduced for symmetric,

positive definite problems in Sect. 6.2; see also Sect. 2.5.2. Thus, the coarse problem is treated multiplicatively, while the local problems are treated additively. The preconditioned operator has the same form as in Sect. 6.2:

$$P_{hy1} = \hat{Q}S = P_0 + (I - P_0)(\sum_{i=1}^{N} P_i)(I - P_0),$$

where $P_0 = \hat{Q}_H S$ and $P_i = \hat{Q}_i S$. The matrix form of the preconditioner is

$$\hat{Q} = P_{hy1}S^{-1} = \hat{Q}_H + (I - \hat{Q}_H S)\sum_{i=1}^{N} \hat{Q}_i(I - S\hat{Q}_H),$$

where the precise form of the coarse operator $\hat{Q}_H$ and local operators $\hat{Q}_i$ is given below. We note that $\hat{Q}$ can be written as a three-step preconditioner; see Sect. 2.6. Given a residual vector $r$ the preconditioned vector $\hat{r} = \hat{Q}r$ is given by

$$\begin{aligned} t &\leftarrow \hat{Q}_H r, \\ t &\leftarrow \sum_{i=1}^{N} \hat{Q}_i(r - St), \\ \hat{r} &\leftarrow t + \hat{Q}_H(r - St). \end{aligned}$$

We note that the first step can be left out if $r$ is a residual in the range of $S(I - P_0)$; we can show that $P_0(I - P_0) = 0$. We also note that we have returned to the balanced space after each full application of the operator $P_{hy1}$, but that prior to the application of $(I - P_0)$, we are generally outside this space; we can say that the coarse correction restores the balance. While $P_0$ is a projection on the balanced space, it is not on the larger space; this is an explanation why the stability of the coarse solver, expressed in terms of its inf-sup constant $\gamma_0$, will enter the bound of the condition number of this iterative method; see Theorem 9.11.

**Coarse Solver.** Given a residual vector $r$, the coarse term $\hat{Q}_H r$ is the solution of a coarse, global Stokes problem with a few velocity degrees of freedom and one constant pressure per subdomain $\Omega_i$:

$$\hat{Q}_H = R_H^T S_0^{-1} R_H,$$

where

$$R_H = \begin{pmatrix} L_0^T & 0 \\ 0 & I \end{pmatrix},$$

and

$$S_0 = R_H S R_H^T = \begin{pmatrix} L_0^T S_\Gamma L_0 & L_0^T B_0^T \\ B_0 L_0 & 0 \end{pmatrix}. \tag{9.43}$$

Four choices for the extension matrix $L_0$ (and therefore for the velocity coarse space span$(L_0)$) are proposed in [386]. Some of the columns of $L_0$ are defined in terms of the pseudoinverses $\delta_i^\dagger$ of the Neumann-Neumann counting functions associated with each subdomain $\Omega_i$; see (6.1). We use the function $\delta_i^\dagger$ in all or almost all of the subdomains and for each velocity component. The columns of $L_0$ can be defined by one of the following four choices:

0) the functions $\delta_i^\dagger$,
1) the $\delta_i^\dagger$ and the continuous coarse piecewise bi- or tri-linear functions;
2) the $\delta_i^\dagger$ and the continuous coarse piecewise bi- or tri-quadratic functions;
3) the $\delta_i^\dagger$ and the quadratic coarse edge/face bubble functions for the normal direction.

Choice 0) provides a quite minimal coarse velocity space and it turns out not to be inf-sup stable; it is included since it corresponds to the standard choice for the Laplacian. Choices 1), 2), and 3) are enrichments of 0). The choice 2) is uniformly inf-sup stable, i.e., its inf-sup stability constant $\gamma_0$ is independent of the number of subdomains. The inf-sup constant of 3) satisfies

$$\gamma_0^2 \geq C/(1 + \log(H/h)),$$

with the constant $C$ independent of the number of subdomains. For spectral elements, $\log(k)$ replaces $\log(H/h)$; $k$ is the degree of the polynomials which define the spectral elements. There is no full theory for the choice 1), but it has worked quite well in numerical experiments. We note that 2) is a very natural choice because it is based directly on a well-known stable discretization of the Stokes problem on the coarse finite or spectral elements. We note that 1) and 2) might not be easy to adapt to unstructured subdomain meshes produced by automatic mesh partitioners. This is one of the reasons why the choice 3) is also considered, since it can relatively easily be adapted for general unstructured subdomains. For detailed proofs, see Pavarino and Widlund [386, Lem. 5.2].

In order to avoid linearly dependent $\delta_i^\dagger$ functions, and hence a singular coarse space problem, we might have to drop all of the components of these functions for one subdomain. This depends on the coarse triangulation; on a regular hexahedral mesh in three dimensions, we should, e.g., include at most $3(N-1)$ such functions.

We refer to [386] for precise formulas for the velocity and pressure components of the coarse operator $\hat{Q}_H S$.

**Local Solvers.** The local operators $\hat{Q}_i$ are only applied to residuals of balanced velocity fields in $V_{\Gamma,B}$ and thus the second residual component will vanish. It can also be shown that the pressure components obtained in this

step of the preconditioner plays no further role when we next apply the operator $(I-P_0)$; see [386, Lem. 4.1]. Each local operator $\hat{Q}_i$ is based on the solution of a local Stokes problem on $\Omega_i$ with natural boundary conditions. This local problem is nonsingular for any subdomain $\Omega_i$ the boundary of which intersects $\partial\Omega$, but it is singular otherwise, i.e., for the *floating* subdomains; cf. Sect. 4.2. In the latter case any constant velocity belongs to the null space, while the pressure is now uniquely determined because a pressure term is present in the Neumann boundary condition; see (9.38). To avoid possible complications with singular problems, we can modify the local Stokes problems on the floating subdomains, by adding $\epsilon$ times the velocity mass matrix to the local stiffness matrix $\mathcal{A}^{(i)}$ or we can make these solutions unique by requiring that each velocity component has a zero average over $\Omega_i$; the right hand sides will always be compatible.

Given a residual vector with a first component $r_\Gamma$ and a zero second component, $\hat{Q}_i r$ is the weighted solutions of a local Stokes problem on subdomain $\Omega_i$ with a natural boundary condition on $\partial\Omega_i \setminus \partial\Omega$:

$$\hat{Q}_i r = \begin{pmatrix} R_i^T D^{(i)} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S_{\Gamma,\epsilon}^{(i)} & B_0^{(i)T} \\ B_0^{(i)} & 0 \end{pmatrix}^{-1} \begin{pmatrix} D^{(i)} R_i & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} r_\Gamma \\ 0 \end{pmatrix}; \tag{9.44}$$

see [386] for details. Here $R_i$ are $0,1$ restriction matrices mapping $r_\Gamma$ into $r_{\Gamma_i}$ and $D^{(i)}$ are diagonal matrices representing multiplication by the functions $\delta_i^\dagger$. This operator has a form quite similar to that of the Neumann-Neumann methods for positive definite problems introduced in Sect. 6.2.

We have the following main result; see [386].

**Theorem 9.11** *The balancing Neumann-Neumann operator $P_{hy1}$ is symmetric, positive definite, with respect to the scalar product defined by $S$ on the balanced subspace $V_{\Gamma,B} \times Q_0$ and*

$$\kappa(P_{hy1}) \leq C(1 + \frac{1}{\gamma_0})\frac{1}{\gamma^2}\ \eta,$$

*where*

$$\eta = \begin{cases} (1+\log(H/h))^2 & \text{for finite elements,} \\ \\ (1+\log k)^2 & \text{for spectral elements,} \end{cases}$$

*and $\gamma_0$ and $\gamma$ are the inf-sup constants of the coarse problem and the original discrete Stokes problem, respectively. The constant $C$ is independent of the number of subdomains and degrees of freedom.*

A similar Neumann-Neumann algorithm has been developed for almost incompressible linear elasticity in [230, 229]; see also Subsect. 9.4.3.

Other iterative substructuring algorithms have also been proposed in the literature. Here, we mention Bramble and Pasciak [71] Pasciak [373], Quarteroni [390], Marini and Quarteroni [344], Fischer and Rønquist [205], Casarin [126, 128], Rønquist [402], Le Tallec and Patra [310], Ainsworth and Sherwin [12], Calgaro and Laminie [119], and Li [314, 315, 316].

**Table 9.1.** Constant coefficients, $\nu = 0.499$

| **A.** Fixed number of subdomains: $N = 4 \times 4$ | | | | | |
|---|---|---|---|---|---|
| mesh size | local prob. | dofs | iter. | $\lambda_{\max}$ | time |
| 160 × 160 | 40 × 40 | 206,082 | 17 | 7.21 | 10.5 |
| 240 × 240 | 60 × 60 | 462,722 | 18 | 8.30 | 26.3 |
| 320 × 320 | 80 × 80 | 821,762 | 19 | 9.12 | 65.6 |
| 400 × 400 | 100 × 100 | 1,283,202 | 19 | 9.78 | 96.2 |
| 480 × 480 | 120 × 120 | 1,847,042 | 19 | 10.34 | 188.8 |
| **B.**Fixed local prob.: 80 × 80 elements (51,842 dofs) | | | | | |
| mesh size | $N$ | dofs | iter. | $\lambda_{\max}$ | time |
| 320 × 320 | 4 × 4 | 821,762 | 19 | 9.12 | 65.6 |
| 640 × 640 | 8 × 8 | 3,281,922 | 20 | 9.33 | 69.5 |
| 960 × 960 | 12 × 12 | 7,380,482 | 21 | 9.44 | 72.1 |
| 1280 × 1280 | 16 × 16 | 13,117,442 | 21 | 9.48 | 73.8 |
| 1600 × 1600 | 20 × 20 | 20,492,802 | 21 | 9.49 | 75.5 |

### 9.4.3 Computational Results

This subsection is based on the thesis work of Paulo Goldfeld; cf. [229]. His thesis is an extension of the work carried out jointly by Goldfeld, Pavarino, and Widlund [230] in which Neumann-Neumann methods and results very similar to those of Theorem 9.11 were developed for almost incompressible elasticity. We note that the scaling matrices $D^{(i)}$ in (9.44) will now depend on the values of the Lamé parameter $\mu$, which can vary dramatically across the interface; cf. Sect. 6.2.

Several parallel C codes, using the MPI-based PETSc library, were developed and extensively tested on *Seaborg*, an IBM SP RS/6000 of the National Energy Research Scientific Computing Center of the US Department of Energy's Office of Science, using up to 2000 of its processors, as well as on other distributed computer systems. For information on the PETSc system, see [32]. A mixed finite element method, $Q_2$ - discontinuous $P_1$, is implemented on rectangular elements for almost incompressible elasticity in two dimensions; cf. appendix A.6.2. The region $\Omega$ is a rectangle partitioned uniformly into $N$ substructures. A type 1) coarse space is used. One subdomain was assigned to each processor used in the experiments. The iterations were stopped after a reduction of $10^{-6}$ of the $L^2-$norm of the initial residual. The performance of the balancing Neumann-Neumann algorithm is first illustrated for an almost incompressible material with a Poisson ratio $\nu = 0.499$ in Table 9.1.

We note that only the estimate of the largest eigenvalue of the preconditioned operator is given; the smallest is always very close to 1. The times, given in seconds, are the total times of the computation, which also include the setup of the problem, the assembling of the matrices, and the factorization of the coarse and local matrices. We note that the condition number grows

**Table 9.2.** Heterogeneous Material

| **A.** Fixed number of subdomains: $N = 8 \times 8$ | | | | | |
|---|---|---|---|---|---|
| mesh size | local prob. | dofs | iter. | $\lambda_{\max}$ | time |
| 320 × 320 | 40 × 40 | 821,762 | 20 | 8.60 | 12.5 |
| 480 × 480 | 60 × 60 | 1,847,042 | 22 | 10.08 | 31.2 |
| 640 × 640 | 80 × 80 | 3,281,922 | 23 | 11.22 | 72.3 |
| 800 × 800 | 100 × 100 | 5,126,402 | 23 | 12.14 | 115.6 |
| **B.** Fixed local prob.: 80 × 80 elements (51,842 dofs) | | | | | |
| mesh size | $N$ | dofs | iter. | $\lambda_{\max}$ | time |
| 640 × 640 | 8 × 8 | 3,281,922 | 23 | 11.22 | 72.3 |
| 960 × 960 | 12 × 12 | 7,380,482 | 22 | 10.67 | 71.7 |
| 1280 × 1280 | 16 × 16 | 13,117,442 | 22 | 10.78 | 74.4 |
| 1600 × 1600 | 20 × 20 | 20,492,802 | 22 | 10.76 | 76.2 |
| 2560 × 2560 | 32 × 32 | 52,449,282 | 22 | 10.70 | 88.7 |
| 3520 × 3520 | 44 × 44 | 99,151,362 | 21 | 10.64 | 102.9 |

slowly with the dimension of the local problems, as predicted by the theory. It also remains almost constant when the number of subdomains increases.

Experiments for heterogeneous materials were also carried out. In the final table, the different subdomains are constructed using Lamé parameters of a rubber-like material, steel, and aluminium with Poisson ratios of 0.495, 0.275, and 0.341, respectively. (The rubber-like material is closer to incompressibility than natural rubber.) The results again show a very satisfactory performance. We also note that the largest example involve almost $10^8$ degrees of freedom. We note that while ideally, the execution time would remain constant in the **B** tables, the actual growth is modest for up to almost 2000 subdomains.

# 10

# Problems in $H(\mathrm{div}\,;\Omega)$ and $H(\mathrm{curl}\,;\Omega)$

In this chapter, we analyze some domain decomposition iterative methods for the variational problems which involve the bilinear forms

$$a_{\mathrm{div}}(\mathbf{u},\mathbf{v}) = \int_{\Omega} (a\,\mathrm{div}\,\mathbf{u}\,\mathrm{div}\,\mathbf{v} + B\,\mathbf{u}\cdot\mathbf{v})\,dx,\ \mathbf{u},\,\mathbf{v}\in H(\mathrm{div}\,;\Omega), \quad (10.1)$$

$$a_{\mathbf{curl}}(\mathbf{u},\mathbf{v}) = \int_{\Omega} (A\mathbf{curl}\,\mathbf{u}\cdot\mathbf{curl}\,\mathbf{v} + B\mathbf{u}\cdot\mathbf{v})dx,\ \mathbf{u},\mathbf{v}\in H(\mathbf{curl}\,;\Omega). \quad (10.2)$$

We refer to appendix A.6.3 for an introduction to the variational problems and appendix A.5 for a discussion of the spaces $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$. The matrices $A$ and $B$ are symmetric uniformly positive definite with entries $a_{ij}$ and $b_{ij}\in L^\infty(\Omega)$, $1\le i,j\le n$, and $a\in L^\infty(\Omega)$ is a scalar-valued positive function bounded away from zero. The two bilinear forms are thus symmetric and coercive.

Given a Lipschitz region $\Omega\subset\mathbb{R}^n$, $n=2,3$ of unit diameter, we introduce a family of conforming and shape-regular triangulations $\mathcal{T}_h$, as in appendix B.1.1. We consider conforming Raviart-Thomas and Nédélec finite element spaces; see appendix B.3. For simplicity, we only consider the case of triangulations made of triangles, for $n=2$, and of tetrahedra, for $n=3$. The results in this chapter also hold for meshes consisting of affinely mapped squares or cubes.

Given a polynomial degree $k\ge 1$, the Raviart-Thomas finite element spaces $RT^h_{k;0}(\Omega)$, which are conforming in $H_0(\mathrm{div}\,;\Omega)$, are introduced in Sect. B.3.1, while two- and three-dimensional Nédélec spaces $ND^h_{k;0}(\Omega)$, which are conforming in $H_0(\mathbf{curl}\,;\Omega)$, are presented in appendices B.3.2 and B.3.3, respectively.

For $\mathbf{f}\in L^2(\Omega)^n$, we consider the two discrete problems: find $\mathbf{u}\in RT^h_{k;0}(\Omega)$ such that

$$a_{\mathrm{div}}(\mathbf{u},\mathbf{v}) = \int_{\Omega}\mathbf{f}\cdot\mathbf{v}\,dx, \quad \mathbf{v}\in RT^h_{k;0}(\Omega), \quad (10.3)$$

and $\mathbf{u} \in ND^h_{k;0}(\Omega)$ such that

$$a_{\mathbf{curl}}(\mathbf{u},\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v}\,dx, \quad \mathbf{v} \in ND^h_{k;0}(\Omega); \tag{10.4}$$

see Sect. A.6.3. In this chapter, we set $a(\cdot,\cdot) = a_{\mathrm{div}}(\cdot,\cdot)$ or $a(\cdot,\cdot) = a_{\mathbf{curl}}(\cdot,\cdot)$, and $V = RT^h_{k;0}(\Omega)$ or $V = ND^h_{k;0}(\Omega)$, depending on the problem we consider. Similarly, $A$ is the stiffness matrix obtained by approximating the bilinear form $a(\cdot,\cdot)$ in $V$. We recall that, for $k = 1$, the degrees of freedom in $RT^h_{k;0}(\Omega)$ are associated with edges and faces of $\mathcal{T}_h$ for $n = 2$ or $n = 3$, respectively, and those in $ND^h_{k;0}(\Omega)$ with edges.

The development and analysis of domain decomposition preconditioners for finite element approximations in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$ is fairly recent. Compared to the scalar and vector second order problems in $H^1(\Omega)^n$ in Chap. 5, 6, and 8, where the relevant operator has a small null space (consisting of constants for the Laplacian, and rigid body modes for linear elasticity), here the divergence and curl operators have large null spaces; see appendix A.5.4. We recall that the null space of the divergence consists of curls of $H^1$ functions or vectors (cf. Lemmas A.25, A.27, and A.28) and that of the the curl consists of gradients of $H^1$ functions (cf. Lemmas A.23 and A.25).

For problems in $H^1(\Omega)^n$, the kernels are finite dimensional and consist of low-frequency functions that are contained in the chosen coarse space; see, e.g., the case of overlapping methods in Chap. 3, primal iterative substructuring methods in Chap. 5, Neumann-Neumann and FETI methods in Chap. 6, and linear elasticity problems in Chap. 8. However, for the divergence and curl operators the large null spaces also consist of high-frequency functions that cannot be represented by a low-dimensional coarse space. We recall that these kernels can also be characterized for Raviart-Thomas and Nédélec finite element spaces; see appendix B.3.4 and, in particular, Lemmas B.21 and B.23.

The size of the null spaces and the fact that only the normal or the tangential components of the finite element functions are continuous across the interelement boundaries make the analysis much more cumbersome. Indeed, while certain algorithms can be implemented in exactly the same way as for approximations in $H^1(\Omega)$, their analysis requires more sophisticated tools. This is the case for overlapping preconditioners, where standard coarse spaces consisting of Raviart-Thomas or Nédélec functions on a coarse mesh can be employed and local spaces are formed by selecting degrees of freedom inside overlapping subdomains, as in the case of approximations of $H^1(\Omega)$ in Chap. 3. Similarly, some iterative substructuring methods for two-dimensional problems, which employ standard coarse spaces and local spaces associated with single edges, and some Neumann-Neumann methods, which also employ standard coarse spaces, are defined in the same way as their counterparts for problems in $H^1(\Omega)$; see Sect. 5.4.1 and in particular Remark 5.4, and Sect. 6.2.

However, other methods are not directly applicable to Raviart-Thomas and Nédélec approximations. Certain subspaces need to be modified and completely new algorithms often need to be devised. This is the case for certain balancing Neumann-Neumann and FETI methods; cf. Chap. 6. In addition, many important problems remain open, as is the case of iterative substructuring methods for Nédélec approximations in three dimensions.

For Raviart-Thomas approximations, similar methods can be developed for two and three dimensions and the analysis is also quite similar. In addition, since vectors in $H(\mathbf{curl};\Omega)$ and Nédélec spaces in two dimensions can be obtained from those in $H(\mathrm{div};\Omega)$ and Raviart-Thomas spaces, respectively, by a rotation of 90 degrees (cf. Lemma A.20 and appendix B.3.2), algorithms for two-dimensional Raviart-Thomas approximations can also be defined and analyzed for Nédélec approximations. This will become clear as we proceed with our presentation.

To our knowledge, the first studies on domain decomposition preconditioners for Raviart-Thomas approximations of the bilinear form (10.1) are due to Arnold, Falk, and Winther in [21], which was followed by [22] and [23], and by work by Hiptmair [259], which was continued in [261, 260]. In [21], two-level overlapping and certain multilevel methods are proposed and analyzed for Raviart-Thomas approximations in two dimensions. In [259, 261], multilevel preconditioners are studied in three dimensions. This pioneering work introduced the fundamental idea that stable decompositions of Raviart-Thomas functions should treat the components in the kernel of the divergence operator and those in its orthogonal complement *separately*. Multilevel preconditioners for Nédélec approximations in three dimensions were then considered in [260, 23]. Overlapping methods for three-dimensional Nédélec approximations were specifically studied in [439] and a unified framework for $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$ was later given in [262]. The work mentioned above, then paved the way to the additional results on iterative substructuring methods, for which we refer to section 10.2 and the references therein.

We also mention the work of Alonso and Valli which concerns certain domain decomposition methods for various types of three-dimensional electromagnetic problems; see, e.g., [16, 17] and the references therein. Among multilevel methods, we further mention [249, 395]. Overlapping methods have also been considered for spectral element approximations in [257, 258]. Two-level and multilevel preconditioners have also been applied to and studied for time-harmonic electromagnetic fields: in this case an indefinite problem is solved; see Chap. 11. We mention the two-level method in [232] for $h$ approximations, which extends the theory for scalar problems developed in Chap. 11, and that in [371] for $hp$ approximations in two dimensions, which appears to be robust also for certain anisotropic meshes.

Some earlier results on mixed approximations of scalar diffusion problems deserve a special mention. The Laplace problem for a pressure $p$ can be reformulated in mixed form by introducing the flux $\mathbf{u} = -\mathbf{grad}\, p$ as an additional unknown; see appendices A.7.2 and B.4. With $p \in L^2(\Omega)$ and $\mathbf{u} \in H(\mathrm{div};\Omega)$,

finite element spaces consisting of discontinuous functions for $p$ and Raviart-Thomas spaces for $\mathbf{u}$ can then be employed. An equation for $p$ can be found by eliminating $\mathbf{u}$, thus giving a nonconforming approximation for the Laplacian. Some overlapping and nonoverlapping domain decomposition preconditioners have been proposed for these types of approximations; see chapter 9.

## 10.1 Overlapping Methods

In this section, we study some two-level overlapping methods for three-dimensional problems in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$. The algorithms are defined and can be analyzed in exactly the same way for two-dimensional problems. We follow the analysis in [439] for $H(\mathbf{curl};\Omega)$ in three dimensions and then adapt it to the case of $H(\mathrm{div};\Omega)$. We recall that two-dimensional problems in $H(\mathrm{div};\Omega)$ were originally studied in [21] and that a common framework was given in [262]; see Sect. 10.1.3 for additional comments.

We only consider the case of constant coefficients:

$$a_{\mathrm{div}}(\mathbf{u},\mathbf{v}) = \eta_1 \int_\Omega \mathrm{div}\,\mathbf{u}\,\mathrm{div}\,\mathbf{v}\,dx + \eta_2 \int_\Omega \mathbf{u}\cdot\mathbf{v}\,dx, \tag{10.5}$$

$$a_{\mathbf{curl}}(\mathbf{u},\mathbf{v}) = \eta_1 \int_\Omega \mathbf{curl}\,\mathbf{u}\cdot\mathbf{curl}\,\mathbf{v}\,dx + \eta_2 \int_\Omega \mathbf{u}\cdot\mathbf{v}\,dx, \tag{10.6}$$

with $\eta_1 \geq 0$ and $\eta_2 > 0$. In the more general case, our bounds can be expected to depend on the maximum and minimum of $a$ and of the eigenvalues of $B$ in (10.1), and on the maximum and minimum eigenvalues of $A$ and $B$ in (10.2).

We assume that the domain $\Omega$ is convex; we will indeed need to use the regularity result of Lemma A.54. Here, we consider less general partitions than those in Sect. 3.2 and assume that our overlapping partitions arises from a coarse mesh. More precisely, we introduce a shape-regular, quasi-uniform coarse mesh $\mathcal{T}_H$ on the domain $\Omega$, with $H$ the maximum of the diameters of its elements. In contrast to Sect. 3.3, we now assume that the fine mesh $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$:

$$\mathcal{T}_h \subset \mathcal{T}_H.$$

The local overlapping subregions $\Omega_i'$ are constructed from the elements of the coarse mesh $\mathcal{T}_H$ by extending the substructures $\Omega_i$ to larger regions in such a way that $\partial\Omega_i'$ does not cut through any fine element. We assume that these new subregions are shape-regular and that Assumptions 3.1 (small overlap) and 3.2 (finite covering) hold. We also set $\Omega_0' = \Omega$.

We summarize our geometric assumptions:

**Assumption 10.1** *The domain $\Omega$ is a convex polyhedron and the coarse mesh is quasi uniform. The overlapping partition consists of shape-regular subdomains and satisfies Assumptions 3.1 and 3.2, with overlap $\delta$ and $N_c$ colors.*

As in Chap. 3, overlapping preconditioners are built from a coarse solver and local solvers associated with the overlapping partition into subdomains. Equivalently, they are given in terms of a coarse space $V_0$ and local spaces $V_i$, $i = 1, \ldots, N$, and solvers on these subspaces. Here, we define $V_0$ as the lowest order Raviart-Thomas or Nédélec space built on the coarse triangulation $\mathcal{T}_H$:

$$V_0 = RT^H_{1;0}(\Omega) \quad \text{or} \quad V_0 = ND^H_{1;0}(\Omega).$$

Since $\mathcal{T}_h \subset \mathcal{T}_H$, $V_0$ is a subspace of $V$, and

$$R_0^T : V_0 \longrightarrow V,$$

is the natural interpolation operator from the coarse to the fine mesh.

Local spaces are associated with single subdomains; they are defined as fine Raviart-Thomas or Nédélec spaces on individual subdomains and consist of local functions with vanishing normal or tangential component on the boundary of the subdomain:

$$V_i = RT^h_{k;0}(\Omega_i') \quad \text{or} \quad V_i = ND^h_{k;0}(\Omega_i'), \qquad i = 1, \ldots, N.$$

The extension operators

$$R_i^T : V_i \longrightarrow V,$$

are defined as the extensions by zero to the rest of $\Omega$. We note that, as for nodal finite elements in section 3.2, the restriction $R_i$ consists of zeros and ones and simply extracts the degrees of freedom inside $\Omega_i'$. We have $V = \sum_i R_i^T V_i$; cf. Lemma 10.9.

We select exact solvers for the coarse and local problems. Following the notation of Sect. 2.2, we define, e.g., for $H(\mathrm{div}\,;\Omega)$

$$\tilde{a}_i(\mathbf{u}_i, \mathbf{v}_i) = \eta_1 \int_{\Omega_i'} \mathrm{div}\,\mathbf{u}_i \,\mathrm{div}\,\mathbf{v}_i \, dx + \eta_2 \int_{\Omega_i'} \mathbf{u}_i \cdot \mathbf{v}_i \, dx, \quad \mathbf{u}_i, \mathbf{v}_i \in V_i, \; 0 \le i \le N.$$

Analogous definitions hold for $H(\mathbf{curl}\,;\Omega)$.

If $\{A_i\}$ are the matrices corresponding to the local bilinear forms, a two-level additive Schwarz method gives rise to preconditioned operators of the form:

$$P_{ad} = A_{ad}^{-1} A, \quad A_{ad}^{-1} = R_0^T A_0^{-1} R_0 + \sum_{i=1}^{N} R_i^T A_i^{-1} R_i.$$

We stress the fact that for Raviart-Thomas and Nédélec approximations two-level overlapping methods can be implemented in exactly the same way as in the case of nodal elements, in terms of local and coarse solvers and restriction and extension operators. The analysis employs the abstract Schwarz theory of chapter 2 but is more involved, as already noted. Multiplicative and hybrid methods can also be considered and analyzed, and inexact solvers can also be considered; see Chap. 2 for more details.

We have the following result.

**Theorem 10.2** *Let $k \geq 1$. If exact solvers are employed on all the subspaces, then the condition number of the additive Schwarz operators for Raviart-Thomas and Nédélec approximations satisfy*

$$\kappa(P_{ad}) \leq C\left(1+\left(\frac{H}{\delta}\right)^2\right).$$

*Here, $C$ may depend on $k$, the number of colors of the partition $N_c$, or the shape-regularity constants of the coarse and the fine meshes, but is independent of $h$, $H$, $\delta$, and the coefficients $\eta_1$ and $\eta_2$.*

We note that here we obtain a quadratic growths in $H/\delta$, as opposed to the case of approximations in $H^1(\Omega)$. This bound does not appear to be sharp; see Sect. 10.1.3. The largest eigenvalue of $P_{ad}$ is bounded from above by $(N_c+1)$. This can be proven by using Assumptions 2.4, 2.3, and 3.2, and Lemma 2.10. We refer to Sect. 3.6 for additional comments.

According to Lemma 2.5, we need to find a stable decomposition in order to prove a bound for the smallest eigenvalue. This is done first for $H(\mathbf{curl};\Omega)$ and then for $H(\mathrm{div};\Omega)$ in the following subsections, after providing some technical tools.

### 10.1.1 Problems in $H(\mathbf{curl};\Omega)$

#### Technical Tools

We recall that the space $H_0(\mathbf{curl};\Omega)$ has the following decomposition, which is orthogonal both in $L^2(\Omega)$ and $H(\mathbf{curl};\Omega)$:

$$H_0(\mathbf{curl};\Omega) = \mathbf{grad}\, H_0^1(\Omega) \oplus H_0^{\perp}(\mathbf{curl};\Omega), \tag{10.7}$$

with

$$H_0^{\perp}(\mathbf{curl};\Omega) = H_0(\mathbf{curl};\Omega) \cap H(\mathrm{div}_0;\Omega);$$

see Remark A.26 in appendix A.5.4.

We recall that Nédélec finite element spaces also have discrete Helmholtz decompositions. Following (B.29), we set

$$S = V_{k;0}^h(\Omega), \quad V^{\perp} = ND_{k;0}^{h;\perp}(\Omega),$$

and obtain the orthogonal decomposition

$$V = \mathbf{grad}\, S \oplus V^{\perp}. \tag{10.8}$$

We note that while $\mathbf{grad}\, S \subset \mathbf{grad}\, H_0^1(\Omega)$, in general $V^{\perp} \not\subset H_0^{\perp}(\mathbf{curl};\Omega)$, and that both decompositions (10.7) and (10.8) are also orthogonal with respect to the bilinear form $a(\cdot,\cdot)$.

We summarize the properties of the orthogonal complements in the following lemma; see Lemma A.54 and inequalities (A.20) and (B.31).

**Lemma 10.3** *There exists a constant $C$, depending only on $\Omega$, such that*

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C \, \|\mathbf{curl}\, \mathbf{u}\|_{L^2(\Omega)^n}, \quad \mathbf{u} \in H_0^{\perp}(\mathbf{curl}\,; \Omega) \cup V^{\perp}.$$

*In addition, if $\Omega$ is convex, the space $H_0^{\perp}(\mathbf{curl}\,; \Omega)$ is continuously embedded in $H^1(\Omega)^n$:*

$$\|\mathbf{u}\|^2_{H^1(\Omega)^n} \leq C(\|\mathbf{u}\|^2_{L^2(\Omega)^n} + \|\mathbf{curl}\, \mathbf{u}\|^2_{L^2(\Omega)^n}) \leq C\|\mathbf{curl}\, \mathbf{u}\|^2_{L^2(\Omega)^n},$$

*for $\mathbf{u} \in H_0^{\perp}(\mathbf{curl}\,; \Omega)$.*

In our analysis, we will find stable decompositions of the two terms corresponding to the two subspaces in (10.8). In order to do so, we need some additional projection operators and spaces. Following Hiptmair [259, 260], we first define a finite dimensional space of more regular functions as follows: let

$$\Theta^{\perp} : \; H_0(\mathbf{curl}\,; \Omega) \longrightarrow H_0^{\perp}(\mathbf{curl}\,; \Omega),$$

be the orthogonal projection onto $H_0^{\perp}(\mathbf{curl}\,; \Omega)$. In particular, $\Theta^{\perp}\mathbf{u}$ is defined by

$$\Theta^{\perp}\mathbf{u} := \mathbf{u} - \mathbf{grad}\, q,$$

where $q \in H_0^1(\Omega)$ satisfies

$$(\mathbf{grad}\, q, \mathbf{grad}\, p)_{L^2(\Omega)^n} = (\mathbf{u}, \mathbf{grad}\, p)_{L^2(\Omega)^n}, \quad p \in H_0^1(\Omega).$$

It is readily seen that $\Theta^{\perp}$ preserves the curl and is an orthogonal projection in $L^2(\Omega)^n$ as well. We now define

$$V^{+} := \Theta^{\perp}(V^{\perp}) \subset H_0^{\perp}(\mathbf{curl}\,; \Omega).$$

The space $V^{+}$ is finite dimensional, but is not a finite element space. However, the curls of these vectors are finite element functions. In the following, we need to interpolate vectors in $V^{+}$. Error estimates can be found for these particular vectors.

**Lemma 10.4** *Let $\mathcal{T}_h$ be a shape-regular triangulation and let $\mathbf{u} \in H^1(\Omega)^n$, such that*

$$\mathbf{curl}\, \mathbf{u}_{|_K} \in \mathbb{P}_{k+1}(K)^n, \quad K \in \mathcal{T}_h.$$

*Then, the following estimate holds for $k \geq 1$ and $K \in \mathcal{T}_h$:*

$$\|\mathbf{u} - \Pi^h_{ND_k}\mathbf{u}\|_{L^2(K)^n} \leq C \, h_K |\mathbf{u}|_{H^1(K)^n}, \tag{10.9}$$

*with a constant $C$ independent of $\mathbf{u}$ and $h$.*

*Proof.* We follow [23]; see also [262, Lemma 4.3] for a similar proof. We first assume that $K$ is the reference tetrahedron. Since $\mathbf{curl}\, \mathbf{u} \in \mathbb{P}_{k+1}(K)^n$, it also belongs to $L^p(K)^3$, for $p \geq 1$, and thus Lemma B.19 can be applied.

Using the triangle inequality and the equivalence of the $L^p$- and $L^\infty$-norms on the finite dimensional space $\mathbb{P}_{k+1}(K)^n$, we find

$$\|\mathbf{u} - \Pi^h_{ND_k}\mathbf{u}\|_{L^2(K)^n} \le C(\|\mathbf{u}\|_{H^1(K)^n} + \|\mathbf{curl}\,\mathbf{u}\|_{L^\infty(K)^n}) \le C\|\mathbf{u}\|_{H^1(K)^n}.$$

The norm on the right hand side can be replaced by a seminorm using a quotient space argument. A scaling argument which uses the appropriate transformation for vector functions (cf. [359]) then proves the error bound for a general element $K$. □

We next define a projection $P^h$ onto the semicontinuous space $V^+$ by

$$\begin{aligned} &P^h : H_0(\mathbf{curl};\Omega) \longrightarrow V^+, \\ &\big(\mathbf{curl}\,(P^h\mathbf{u} - \mathbf{u}), \mathbf{curl}\,\mathbf{v}\big)_{L^2(\Omega)^n} \;=\; 0, \quad \mathbf{v} \in V^+. \end{aligned}$$

Lemma 10.3 ensures that the operator $P^h$ is well defined.

*Remark 10.5.* It can easily be checked that whenever $P^h$ is applied to a vector $\mathbf{u}^\perp \in V^\perp$, it coincides with $\Theta^\perp$. In this case, we have

$$\mathbf{curl}\,(P^h\mathbf{u}^\perp) = \mathbf{curl}\,(\Theta^\perp\mathbf{u}^\perp) = \mathbf{curl}\,\mathbf{u}^\perp.$$

The following lemma holds; see [262, Lemma 4.4].

**Lemma 10.6** *Let $\Omega$ be convex. Then, the operator $P^h$ satisfies the following error estimate*

$$\|\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n} \le C\,h\,\|\mathbf{curl}\,\mathbf{u}_h^\perp\|_{L^2(\Omega)^n}, \quad \mathbf{u}_h^\perp \in V^\perp, \tag{10.10}$$

*with $C$ independent of $h$ and $\mathbf{u}_h^\perp$*

*Proof.* A proof can be found in [262]. Here, we give a direct argument, similar to that of [262, Lemma 4.3] and originally proposed in [439, Lemma 3.3]. Let $\mathbf{u}_h^\perp \in V^\perp$. Because of Remark 10.5, $\mathbf{curl}\,(\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp) = 0$, and, since the boundary of $\Omega$ is simply connected, Lemma A.25 ensures that

$$\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp = \mathbf{grad}\,q,$$

for a $q \in H_0^1(\Omega)$. Using interpolation into the Nédélec space and (B.23) in Lemma B.20 yields

$$\begin{aligned} \mathbf{u}_h^\perp - \Pi^h_{ND_k}P^h\mathbf{u}_h^\perp &= \Pi^h_{ND_k}(\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp) \\ &= \Pi^h_{ND_k}(\mathbf{grad}\,q) = \mathbf{grad}\,q_h, \end{aligned} \tag{10.11}$$

with $q_h = I_k^h q \in S$. We can then write

$$\begin{aligned} &\|\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp\|^2_{L^2(\Omega)^n} \\ &= \big(\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp, (\mathbf{u}_h^\perp - \Pi^h_{ND_k}P^h\mathbf{u}_h^\perp) + (\Pi^h_{ND_k}P^h\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp)\big)_{L^2(\Omega)^n} \\ &= \big(\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp, \Pi^h_{ND_k}P^h\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp\big)_{L^2(\Omega)^n} \\ &\le \|\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n}\,\|P^h\mathbf{u}_h^\perp - \Pi^h_{ND_k}P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n}, \end{aligned}$$

where, for the second equality, we have used the fact that, because of (10.11), both $\mathbf{u}_h^\perp$ and $P^h\mathbf{u}_h^\perp$ are orthogonal to $\mathbf{u}_h^\perp - \Pi_{ND_k}^h P^h\mathbf{u}_h^\perp = \mathbf{grad}\, q_h$. We then obtain

$$\|\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n} \leq \|P^h\mathbf{u}_h^\perp - \Pi_{ND_k}^h P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n}. \tag{10.12}$$

It is then enough to bound the right hand side of (10.12). This can be done by using Lemma 10.4:

$$\|P^h\mathbf{u}_h^\perp - \Pi_{ND_k}^h P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n} \leq C\, h\, |P^h\mathbf{u}_h^\perp|_{H^1(\Omega)^n}.$$

The bounds and the regularity result in Lemma 10.3 finally give

$$\begin{aligned}\|P^h\mathbf{u}_h^\perp - \Pi_{ND_k}^h P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n} &\leq Ch\|\mathbf{curl}\,(P_h\mathbf{u}_h^\perp)\|_{L^2(\Omega)^n} \\ &= Ch\|\mathbf{curl}\,\mathbf{u}_h^\perp\|_{L^2(\Omega)^n}.\end{aligned}$$

□

We end this section by defining the operator

$$Q^H : L^2(\Omega)^n \longrightarrow V_0,$$

as the $L^2$ projection onto the coarse space $V_0$. The following error estimates follow directly from those in Lemma 4.14 for scalar functions.

**Lemma 10.7** *Let $\mathcal{T}_H$ be quasi uniform. Then, the following inequalities hold*

$$\begin{aligned}&\|Q^H\mathbf{u}\|_{L^2(\Omega)^n} \leq \|\mathbf{u}\|_{L^2(\Omega)^n}, \\ &\|\mathbf{curl}\,(Q^H\mathbf{u})\|_{L^2(\Omega)^n} \leq C\,|\mathbf{u}|_{H^1(\Omega)^n}, \\ &\|\mathrm{div}\,(Q^H\mathbf{u})\|_{L^2(\Omega)} \leq C\,|\mathbf{u}|_{H^1(\Omega)^n}, \\ &\|\mathbf{u} - Q^H\mathbf{u}\|_{L^2(\Omega)^n} \leq C\,H\,|\mathbf{u}|_{H^1(\Omega)^n},\end{aligned}$$

*with constants independent of* $\mathbf{u}$ *and* $H$.

As for overlapping methods for approximations in $H^1$, local components of a stable splitting are found by employing the continuous, piecewise linear partition of unity functions for the overlapping partition constructed in Lemma 3.4. We have the following lemma; cf. Lemmas 3.9 and 4.31 for analogous results.

**Lemma 10.8** *Let $\mathbf{u} \in V$ and let $\theta$ be a continuous, piecewise linear, scalar function on $\Omega$. Then, the interpolated vector $\Pi_{ND_k}^h(\theta\mathbf{u})$ belongs to $V$ and, for $K \in \mathcal{T}_h$, the following inequalities hold,*

$$\|\Pi_{ND_k}^h(\theta\mathbf{u})\|_{L^2(K)^n} \leq C\,\|\theta\mathbf{u}\|_{L^2(K)^n}, \tag{10.13}$$

$$\|\mathbf{curl}\,(\Pi_{ND_k}^h(\theta\mathbf{u}))\|_{L^2(K)^n} \leq C\,\|\mathbf{curl}\,(\theta\mathbf{u})\|_{L^2(K)^n}, \tag{10.14}$$

*with constants $C$ independent of* $\mathbf{u}$, $\theta$ *and* $h$.

*Proof.* We recall that the moments that define the interpolant $\Pi^h_{ND_k}$ involve integrals of the tangential components over edges and faces, as well as moments of $\mathbf{u}$ computed over each tetrahedron in $\mathcal{T}_h$; see appendix B.3.3.

The vector $\mathbf{u}$ has continuous tangential component across the edges and faces of the tetrahedra; since $\theta$ belongs to $C^0(\Omega)$, the vector $\theta\mathbf{u}$ also has a continuous tangential component and the moments that define $\Pi^h_{ND_k}$ are thus well defined and provide a Nédélec vector in $V$.

Let us now consider (10.13). The function $\theta\mathbf{u}$ belongs to $H^1(K)^n$ and its curl to $\mathbb{P}_{k+1}(K)^n$. Lemma 10.4 can then be applied:

$$\begin{aligned}\|\Pi^h_{ND_k}(\theta\mathbf{u})\|_{L^2(K)^n} &\le \|\theta\mathbf{u}\|_{L^2(K)^n} + \|\theta\mathbf{u} - \Pi^h_{ND_k}(\theta\mathbf{u})\|_{L^2(K)^n}\\ &\le \|\theta\mathbf{u}\|_{L^2(K)^n} + Ch_k|\theta\mathbf{u}|_{H^1(K)^n} \le C\|\theta\mathbf{u}\|_{L^2(K)^n},\end{aligned}$$

where, for the last inequality, we have used an elementary inverse inequality for polynomials in $\mathbb{P}_{k+1}$; see Lemma B.27.

Let us next consider inequality 10.14. Using the commuting property (B.25) in Lemma B.20, we find

$$\|\mathbf{curl}\,(\Pi^h_{ND_k}(\theta\mathbf{u}))\|_{L^2(K)} = \|\Pi^h_{RT_k}(\mathbf{curl}\,(\theta\mathbf{u}))\|_{L^2(K)}.$$

The proof can also be carried out, using similar arguments, for the Raviart-Thomas interpolant $\Pi^h_{RT_k}$ and the vector $\mathbf{curl}\,(\theta\mathbf{u})$. □

### A Stable Decomposition

We are now ready to prove the existence of a stable decomposition in $V$.

**Lemma 10.9** *There exists a constant $C$, independent of $h$, $H$, $\delta$, and the coefficients $\eta_1$ and $\eta_2$, such that, for $\mathbf{u} \in V$, there exists a decomposition*

$$\mathbf{u} = \sum_{i=0}^{N} R_i^T \mathbf{u}_i, \quad \{\mathbf{u}_i \in V_i,\ 0 \le i \le N\},$$

*that satisfies*

$$\sum_{i=0}^{N} \tilde{a}_i(\mathbf{u}_i, \mathbf{u}_i) \le C\left(1 + \left(\frac{H}{\delta}\right)^2\right) a(\mathbf{u}, \mathbf{u}).$$

*Proof.* Equation (10.8) ensures that $\mathbf{u}$ has the following decomposition

$$\mathbf{u} = \mathbf{grad}\, q + \mathbf{w}^{\perp}, \tag{10.15}$$

where $q \in S$ and $\mathbf{w}^{\perp} \in V^{\perp}$. We decompose $\mathbf{grad}\, q$ and $\mathbf{w}^{\perp}$ separately.

We first consider the gradient term. Using the result for continuous finite element spaces conforming in $H^1_0(\Omega)$ given in the proof of Lemma 3.12, we can find a decomposition $\{q_i\} \subset S$ of $q$, such that

$$\begin{aligned}\sum_{i=0}^{N} \tilde{a}_i(\mathbf{grad}\, q_i, \mathbf{grad}\, q_i) &= \eta_2 \sum_{i=0}^{N} |q_i|^2_{H^1(\Omega_i')} \\ &\leq C\left(1+\frac{H}{\delta}\right)\eta_2\, |q|^2_{H^1(\Omega)} \\ &= C\left(1+\frac{H}{\delta}\right)\, a(\mathbf{grad}\, q, \mathbf{grad}\, q).\end{aligned} \tag{10.16}$$

We then consider $\mathbf{w}^\perp \in V^\perp$. We will find a nonconforming approximation of $\mathbf{w}^\perp$, as described in [262], by first projecting onto the semicontinuous space $V^+$, and then onto the coarse space $V_0$, using the $L^2$ projection $Q^H$. We will next divide the remainder into a sum of functions supported on the individual subdomains $\{\Omega_i'\}$.

The first step is performed in the following way:
consider the splitting

$$\mathbf{w}^\perp = \mathbf{w}_0 + \mathbf{v}, \tag{10.17}$$

where

$$\begin{aligned}\mathbf{w}_0 &:= Q^H\left(P_h \mathbf{w}^\perp\right) \in V_0, \\ \mathbf{v} &:= \mathbf{w}^\perp - \mathbf{w}_0 \in V.\end{aligned} \tag{10.18}$$

We note that $P_h\mathbf{w}^\perp$ belongs to $H_0^\perp(\mathbf{curl}\,;\Omega)$ and thus to $H^1(\Omega)^n$ because of Lemma 10.3. We then decompose $\mathbf{v}$ as a sum of terms in the local spaces $\{V_i\}_{i=1}^N$. With $\{\theta_i\}_{i=1}^J$, the piecewise linear partition of unity relative to the covering $\{\Omega_i'\}_{i=1}^N$ constructed in Lemma 3.4, we define

$$\mathbf{w}_i := \Pi^h_{ND_k}(\theta_i \mathbf{v}) \in V_i, \quad i = 1, \ldots, N. \tag{10.19}$$

The function $\mathbf{w}^\perp$ is thus decomposed as $\mathbf{w}^\perp = \sum_{i=0}^N R_i^T \mathbf{w}_i$.

We first consider the component in the coarse space $V_0$:
for its curl, we can employ Lemmas 10.7 and 10.3, and the definition of the projection $P_h$, and write

$$\begin{aligned}\|\mathbf{curl}\, \mathbf{w}_0\|_{L^2(\Omega)^n} &\leq C|P_h\mathbf{w}^\perp|_{H^1(\Omega)^n} \leq C\|\mathbf{curl}\, P_h\mathbf{w}^\perp\|_{L^2(\Omega)^n} \\ &= C\|\mathbf{curl}\, \mathbf{w}^+\|_{L^2(\Omega)^n}.\end{aligned} \tag{10.20}$$

Using the definition of $Q^H$ and $\Theta$, Lemma 10.7, and Remark 10.5, we next find

$$\|\mathbf{w}_0\|_{L^2(\Omega)^n} \leq \|P_h\mathbf{w}^\perp\|_{L^2(\Omega)^n} = \|\Theta^\perp \mathbf{w}^\perp\|_{L^2(\Omega)^n} \leq \|\mathbf{w}^\perp\|_{L^2(\Omega)^n}. \tag{10.21}$$

For the terms $\mathbf{w}_i$, we employ (10.14), and write

$$
\begin{aligned}
&\|\mathbf{curl}\,\mathbf{w}_i\|^2_{L^2(\Omega_i')^n}\\
&\le C\,\|\mathbf{curl}\,(\theta_i\mathbf{v})\|^2_{L^2(\Omega_i')^n}\\
&\le C\,\left(\|\mathbf{grad}\,\theta_i\times\mathbf{v}\|^2_{L^2(\Omega_i')^n}+\|\theta_i\,\mathbf{curl}\,\mathbf{v}\|^2_{L^2(\Omega_i')^n}\right)\\
&\le C\,\left(\|\mathbf{grad}\,\theta_i\|^2_{L^\infty(\Omega_i')^n}\|\mathbf{v}\|^2_{L^2(\Omega_i')^n}+\|\theta_i\|^2_{L^\infty(\Omega_i')}\|\mathbf{curl}\,\mathbf{v}\|^2_{L^2(\Omega_i')^n}\right)\\
&\le C\,\left(\delta^{-2}\,\|\mathbf{v}\|^2_{L^2(\Omega_i')^n}+\|\mathbf{curl}\,\mathbf{v}\|^2_{L^2(\Omega_i')^n}\right),
\end{aligned}
$$

where we have used Lemma 3.4 and equation (3.7) for the last inequality.

Summing over the subdomains and using the finite covering property in Assumption 10.1, we find

$$
\sum_{i=1}^{N}\|\mathbf{curl}\,\mathbf{w}_i\|^2_{L^2(\Omega_i')^n}\le C\,\left(\delta^{-2}\,\|\mathbf{v}\|^2_{L^2(\Omega)^n}+\|\mathbf{curl}\,\mathbf{v}\|^2_{L^2(\Omega)^n}\right). \tag{10.22}
$$

The first term on the right hand side of (10.22) can be bounded, using Lemmas 10.6 and 10.7. We have

$$
\begin{aligned}
\|\mathbf{v}\|^2_{L^2(\Omega)^n}&\le 2\|\mathbf{w}^\perp-P_h\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}+2\|P_h\mathbf{w}^\perp-Q^HP_h\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}\\
&\le C\,\left(h^2\,\|\mathbf{curl}\,\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}+H^2\,\|\mathbf{curl}\,P_h\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}\right)\\
&\le C\,H^2\,\|\mathbf{curl}\,\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}.
\end{aligned} \tag{10.23}
$$

The second term can be bounded using (10.20), as

$$
\begin{aligned}
\|\mathbf{curl}\,\mathbf{v}\|^2_{L^2(\Omega)^n}&\le 2\left(\|\mathbf{curl}\,\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}+\|\mathbf{curl}\,\mathbf{w}_0\|^2_{L^2(\Omega)^n}\right)\\
&\le C\,\|\mathbf{curl}\,\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}.
\end{aligned} \tag{10.24}
$$

Combining (10.22), (10.23), and (10.24) yields

$$
\sum_{i=1}^{N}\|\mathbf{curl}\,\mathbf{w}_i\|^2_{L^2(\Omega)^n}\le C\,\left(1+\left(\frac{H}{\delta}\right)^2\right)\|\mathbf{curl}\,\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}. \tag{10.25}
$$

We then find bounds for the $L^2$-norms of the local terms. Lemma 10.8, the finite covering property, and (10.21) give

$$
\begin{aligned}
\sum_{i=1}^{N}\|\mathbf{w}_i\|^2_{L^2(\Omega_i')^n}&\le C\sum_{i=1}^{N}\|\theta_i(\mathbf{w}^\perp-\mathbf{w}_0)\|^2_{L^2(\Omega_i')^n}\\
&\le C\sum_{i=1}^{N}(\|\mathbf{w}^\perp\|^2_{L^2(\Omega_i')^n}+\|\mathbf{w}_0\|^2_{L^2(\Omega_i')^n})\\
&\le C(\|\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}+\|\mathbf{w}_0\|^2_{L^2(\Omega)^n})\le C\|\mathbf{w}^\perp\|^2_{L^2(\Omega)^n}.
\end{aligned} \tag{10.26}
$$

Finally, by employing (10.20), (10.21), (10.25), and (10.26), and the equivalence of the graph and the energy norm, we obtain

$$\sum_{i=0}^{N} \tilde{a}_i(\mathbf{w}_i, \mathbf{w}_i) \leq C \left(1 + \left(\frac{H}{\delta}\right)^2\right) a(\mathbf{w}^\perp, \mathbf{w}^\perp). \tag{10.27}$$

The proof is concluded by defining

$$\mathbf{u}_i := \mathbf{grad}\, q_i + \mathbf{w}_i \subset V_i, \quad i = 0, \ldots, N,$$

using the bounds for the two components in (10.16) and (10.27), and noting that the decomposition (10.15) is also orthogonal with respect to $a(\cdot,\cdot)$. □

### 10.1.2 Problems in $H(\mathrm{div}\,;\Omega)$

The technical tools in Sect. 10.1.1 and the proof of Lemma 10.9 rely on a few basic properties: a Helmholtz decomposition of the continuous and finite element spaces, a regularity result for vectors with certain vanishing traces on the boundary (see Lemmas 10.3 and A.54), and estimates for vectors in certain orthogonal complements (see Lemma 10.3). All these tools are also available for vector functions in $H(\mathrm{div}\,;\Omega)$ and $RT_k^h(\Omega)$ and a similar stable decomposition can be found in this case. Indeed, this was the approach employed in [262], where a common framework was given for the analysis of certain two-level overlapping and multilevel preconditioners for $H^1(\Omega)$, $H(\mathrm{div}\,;\Omega)$, and $H(\mathbf{curl}\,;\Omega)$. Speaking informally, if we replace the '**curl**' with the 'div' operator and '**grad**' with '**curl**', the same results and proofs as in the previous subsection hold verbatim. Here, we only review the technical tools for $H(\mathrm{div}\,;\Omega)$ and briefly sketch the proof of Lemma 10.9 for Raviart-Thomas approximations in three dimensions. The result collected in this section can be found in appendix A.5 for the continuous spaces, and in appendix B.3 for the discrete ones.

The space $H_0(\mathrm{div}\,;\Omega)$ has the following orthogonal decomposition:

$$\begin{aligned} H_0(\mathrm{div}\,;\Omega) &= \mathbf{curl}\, H_0(\mathbf{curl}\,;\Omega) \oplus H_0^\perp(\mathrm{div}\,;\Omega) \\ &= \mathbf{curl}\, H_0^\perp(\mathbf{curl}\,;\Omega) \oplus H_0^\perp(\mathrm{div}\,;\Omega), \end{aligned} \tag{10.28}$$

with

$$H_0^\perp(\mathrm{div}\,;\Omega) = H_0(\mathrm{div}\,;\Omega) \cap H(\mathbf{curl}_0;\Omega);$$

see Lemma A.27 and Remark A.30. In addition, Raviart-Thomas finite element spaces also have a discrete Helmholtz decomposition. Following (B.33), we set

$$S = ND_{k;0}^{h;\perp}(\Omega), \quad V^\perp = RT_{k;0}^{h;\perp}(\Omega),$$

and obtain the orthogonal decomposition

$$V = \mathbf{curl}\, S \oplus V^\perp. \tag{10.29}$$

We note that

$$\begin{aligned}\mathbf{curl}\, ND_{k;0}^{h;\perp}(\Omega) &= \mathbf{curl}\, ND_{k;0}^{h}(\Omega),\\ \mathbf{curl}\, H_0^{\perp}(\mathbf{curl}\,;\Omega) &= \mathbf{curl}\, H_0(\mathbf{curl}\,;\Omega).\end{aligned}$$

For a given $\mathbf{w} \in \mathbf{curl}\, H_0^{\perp}(\mathbf{curl}\,;\Omega)$, there is a unique $\mathbf{u} \in H_0^{\perp}(\mathbf{curl}\,;\Omega)$, such that $\mathbf{w} = \mathbf{curl}\,\mathbf{u}$; an analogous result holds for $\mathbf{w} \in \mathbf{curl}\, ND_{k;0}^{h;\perp}(\Omega)$. We recall, in addition, that

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C\, \|\mathbf{curl}\,\mathbf{u}\|_{L^2(\Omega)^n}, \quad \mathbf{u} \in H_0^{\perp}(\mathbf{curl}\,;\Omega) \cup S. \tag{10.30}$$

Both decompositions (10.28) and (10.29) are also orthogonal with respect to the bilinear form $a(\cdot,\cdot)$. We summarize the properties of these orthogonal complements in the following lemma.

**Lemma 10.10** *There exists a constant $C$, depending only on $\Omega$, such that*

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C\, \|\mathrm{div}\,\mathbf{u}\|_{L^2(\Omega)}, \quad \mathbf{u} \in H_0^{\perp}(\mathrm{div}\,;\Omega) \cup V^{\perp}.$$

*In addition, if $\Omega$ is convex, the space $H_0^{\perp}(\mathrm{div}\,;\Omega)$ is continuously embedded in $H^1(\Omega)^n$.*

As for the analysis in $H(\mathbf{curl}\,;\Omega)$, we can find a space of more regular functions by defining the orthogonal projection

$$\Theta^{\perp} : H_0(\mathrm{div}\,;\Omega) \longrightarrow H_0^{\perp}(\mathrm{div}\,;\Omega).$$

Here, $\Theta^{\perp}\mathbf{u}$ is defined by

$$\Theta^{\perp}\mathbf{u} := \mathbf{u} - \mathbf{curl}\,\mathbf{w},$$

where $\mathbf{w} \in H_0^{\perp}(\mathbf{curl}\,;\Omega)$ satisfies

$$(\mathbf{curl}\,\mathbf{w}, \mathbf{curl}\,\mathbf{v})_{L^2(\Omega)^n} = (\mathbf{u}, \mathbf{curl}\,\mathbf{v})_{L^2(\Omega)^n}, \quad \mathbf{v} \in H_0^{\perp}(\mathbf{curl}\,;\Omega).$$

We note that the solution $\mathbf{w}$ is uniquely defined because of (10.30).

It is readily seen that $\Theta^{\perp}$ preserves the divergence and is an orthogonal projection in $L^2(\Omega)^n$ as well. We now define

$$V^{+} := \Theta^{\perp}(V^{\perp}) \subset H_0^{\perp}(\mathrm{div}\,;\Omega).$$

The space $V^{+}$ is finite dimensional, but is not a finite element space. The divergence of these functions however are finite element functions.

We define the projection $P^h$ onto the semicontinuous space $V^{+}$ by

$$\begin{aligned}&P^h : H_0(\mathrm{div}\,;\Omega) \longrightarrow V^{+},\\ &\big(\mathrm{div}\,(P^h\mathbf{u} - \mathbf{u}), \mathrm{div}\,\mathbf{v}\big)_{L^2(\Omega)} = 0, \quad \mathbf{v} \in V^{+}.\end{aligned}$$

Lemma 10.10 ensures that the operator $P^h$ is well defined. It can also be checked that whenever $P^h$ is applied to a vector $\mathbf{u}^{\perp} \in V^{\perp}$, it coincides with $\Theta^{\perp}$. In this case, we have

$$\mathrm{div}\,(P^h\mathbf{u}^{\perp}) = \mathrm{div}\,(\Theta^{\perp}\mathbf{u}^{\perp}) = \mathrm{div}\,\mathbf{u}^{\perp}.$$

The following two lemmas can be proven in a similar way as their counterparts 10.4 and 10.6.

**Lemma 10.11** *Let $\mathcal{T}_h$ be a shape-regular triangulation and let $\mathbf{u} \in H^1(\Omega)^n$, such that*

$$\operatorname{div} \mathbf{u}_{|K} \in \mathbb{P}_{k+1}(K), \quad K \in \mathcal{T}_h.$$

*Then, the following estimate holds for $k \geq 1$ and $K \in \mathcal{T}_h$:*

$$\|\mathbf{u} - \Pi^h_{RT_k}\mathbf{u}\|_{L^2(K)^n} \leq C\, h_K |\mathbf{u}|_{H^1(K)^n}, \tag{10.31}$$

*with a constant $C$ independent of $\mathbf{u}$ and $h$.*

**Lemma 10.12** *Let $\Omega$ be convex. Then, the operator $P^h$ satisfies the following error estimate*

$$\|\mathbf{u}_h^\perp - P^h\mathbf{u}_h^\perp\|_{L^2(\Omega)^n} \leq C\, h\, \|\operatorname{div} \mathbf{u}_h^\perp\|_{L^2(\Omega)}, \quad \mathbf{u}_h^\perp \in V^\perp, \tag{10.32}$$

*with $C$ independent of $h$ and $\mathbf{u}_h^\perp$*

We finally need a result for the interpolation of the product of two finite element functions. The proof can be carried out as for Lemma 10.8.

**Lemma 10.13** *Let $\mathbf{u} \in V$ and let $\theta$ be a continuous, piecewise linear, scalar function on $\Omega$. Then, the interpolated vector $\Pi^h_{RT_k}(\theta\mathbf{u})$ belongs to $V$ and, for $K \in \mathcal{T}_h$, the following inequalities hold,*

$$\|\Pi^h_{RT_k}(\theta\mathbf{u})\|_{L^2(K)^n} \leq C\, \|\theta\mathbf{u}\|_{L^2(K)^n}, \tag{10.33}$$

$$\|\operatorname{div}(\Pi^h_{RT_k}(\theta\mathbf{u}))\|_{L^2(K)} \leq C\, \|\operatorname{div}(\theta\mathbf{u})\|_{L^2(K)}, \tag{10.34}$$

*with constants $C$ independent of $\mathbf{u}$, $\theta$ and $h$.*

We are finally able to prove the existence of a stable decomposition in $V$. Lemma 10.9 holds verbatim for the definitions and notations given in this subsection.

*Proof.* Here we only sketch the proof and only point out the parts that are different from the $H(\mathbf{curl}\,;\Omega)$ case. We consider the discrete Helmholtz decomposition (10.29) and find

$$\mathbf{u} = \mathbf{curl}\, \mathbf{w}^\perp + \mathbf{z}^\perp, \tag{10.35}$$

where $\mathbf{w}^\perp \in S$ and $\mathbf{z}^\perp \in V^\perp$, and decompose $\mathbf{curl}\, \mathbf{w}^\perp$ and $\mathbf{z}^\perp$ separately. Since $\mathbf{w}^\perp$ is a vector in the Nédélec space, we can use the decomposition already proven for $H(\mathbf{curl}\,;\Omega)$ in Lemma 10.9. More precisely, $\mathbf{w}^\perp$ belongs to the orthogonal complement $ND^{h;\perp}_{k;0}(\Omega)$ and it can thus be decomposed according to (10.17) and (10.19). Using the bounds in (10.20) and (10.25), we find

$$\sum_{i=0}^{N} \tilde{a}_i(\mathbf{curl}\, \mathbf{w}_i, \mathbf{curl}\, \mathbf{w}_i) = \eta_2 \sum_{i=0}^{N} \|\mathbf{curl}\, \mathbf{w}_i\|^2_{L^2(\Omega_i')^n}$$

$$\leq C\left(1+\left(\frac{H}{\delta}\right)^2\right)\eta_2\left\|\mathbf{curl}\,\mathbf{w}^\perp\right\|^2_{L^2(\Omega)^n}$$
$$= C\left(1+\left(\frac{H}{\delta}\right)^2\right)a(\mathbf{curl}\,\mathbf{w}^\perp,\mathbf{curl}\,\mathbf{w}^\perp).$$

The component $\mathbf{z}^\perp$ in the orthogonal complement can be decomposed as in Lemma 10.9:

$$\mathbf{z}^\perp=\sum_{i=0}^{N}R_i^T\mathbf{z}_i, \tag{10.36}$$

with

$$\begin{aligned}&\mathbf{z}_0:=Q^H\left(P_h\mathbf{z}^\perp\right)\in V_0,\\ &\mathbf{v}:=\mathbf{z}^\perp-\mathbf{z}_0\in V,\\ &\mathbf{z}_i:=\Pi^h_{RT_k}(\theta_i\mathbf{v})\in V_i,\quad i=1,\ldots,N.\end{aligned} \tag{10.37}$$

Bounds for the single terms can be found in exactly the same way as before and the stable decomposition for $\mathbf{u}$ is

$$\mathbf{u}_i:=\mathbf{curl}\,\mathbf{w}_i+\mathbf{z}_i\subset V_i,\quad i=0,\ldots,N.$$

□

### 10.1.3 Final Remarks on Overlapping Methods and Numerical Results

For the proof of Theorem 10.2, we required two quite restrictive assumptions: the domain $\Omega$ needs to be convex, to enable us to use the regularity result of Lemma A.54, and the coarse mesh is required to be quasi uniform, in order to use the bounds for the $L^2$-projection in Lemma 10.7. For approximations in $H^1(\Omega)$, more general meshes can be considered since the $L^2$-projection can be replaced by a more local operator such as the quasi-interpolant in Lemma 3.6.

Theorem 10.2 provides a quadratic bound in $H/\delta$, as opposed to the case of problems in $H^1(\Omega)$, for which a linear bound is obtained. Indeed, a quantitative result as in Lemma 3.10 is missing for vectors in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$ and a linear bound has not been found. Numerical results in [439] for three-dimensional Nédélec approximations are consistent with a *linear* bound, suggesting that the bound in Theorem 10.2 is not sharp; see also the numerical results below.

The proof that we have given of Theorem 10.2 is basically that in [439] for Nédélec approximations in three dimensions. In [262], a similar analysis was given which employs a different stable decomposition. More precisely, the $L^2$-projection in the definition of the coarse component in (10.18) and (10.18)

**Table 10.1.** Two-level additive Schwarz algorithm. Estimated condition number and number of CG iterations for a residual norm reduction of $10^{-6}$ (in parentheses), versus $n$, number of subdomains and $H/\delta$.

| | | $H/\delta$ | | | |
|---|---|---|---|---|---|
| $n$ | $n_{sub}$ | 8 | 4 | 2 | 1.33 |
| 8 | $2^3$ | - | 8.94 (19) | 8.98 (20) | 9.05 (15) |
| 16 | $2^3$ | 14.05 (21) | 8.49 (19) | - | - |
| 16 | $4^3$ | - | 8.61 (19) | 9.86 (21) | 27. 54 (28) |
| 16 | $8^3$ | - | - | 10.59 (19) | - |
| 24 | $3^3$ | 13.86 (21) | 8.45 (19) | - | - |
| 24 | $6^3$ | - | 8.43 (19) | 9.05 (19) | 21.93 (27) |
| 24 | $12^3$ | - | - | 9.39 (18) | - |
| 32 | $4^3$ | 13.02 (20) | 8.30 (18) | - | - |
| 32 | $8^3$ | - | 8.37 (19) | 8.78 (19) | 21.39 (25) |
| 40 | $5^3$ | 13.12 (20) | 8.29 (18) | - | - |
| 40 | $10^3$ | - | 8.29 (18) | 8.68 (19) | 22.28 (25) |
| 48 | $6^3$ | 12.91 (20) | 8.36 (18) | - | - |
| 48 | $12^3$ | - | 8.32 (18) | 8.64 (18) | 22.93 (24) |

is replaced by the orthogonal projection onto the complements $ND_{1;0}^{H;\perp}(\Omega)$ and $RT_{1;0}^{H;\perp}(\Omega)$, respectively. In [21], solutions of suitable mixed problems in the coarse space $V_0$ are employed for the coarse component. It can be proven that the decompositions considered in [21] and in [262] are the same, even though they are written in different ways. However, none of the methods in [439] or in [21, 262] seem to allow us to remove the restrictions in Theorem 10.2 mentioned above.

We finally present some numerical results for the two-level overlapping preconditioner of Sect. 10.1 and 10.1.1. They were originally given in [439]. We consider the Dirichlet problem (10.4) in the unit cube $\Omega$ and with uniform triangulations $\mathcal{T}_h$ and $\mathcal{T}_H$. The fine triangulation $\mathcal{T}_h$ consists of $n^3$ cubical elements, with $h = 1/n$. The number of subdomains $n_{sub}$ equals the number of coarse elements in $\mathcal{T}_H$. Low order edge elements are chosen, $k = 1$, and $\eta_1 = \eta_2 = 1$; cf. (10.6).

Table 10.1 shows the estimated condition number and the number of conjugate gradient iterations to obtain an error reduction of $10^{-6}$ of the residual, as functions of the problem size, the number of subregions, and the relative overlap. A behavior similar to that of the overlapping preconditioners for scalar problems is observed. The condition number initially decreases with $H/\delta$. For larger values of the relative overlap, the number of colors $N_c$ increases, and, consequently, the condition number increases, in accordance with our analysis. In addition, for a fixed value of the relative overlap, the condition number appears to be bounded independently of the number of subregions.

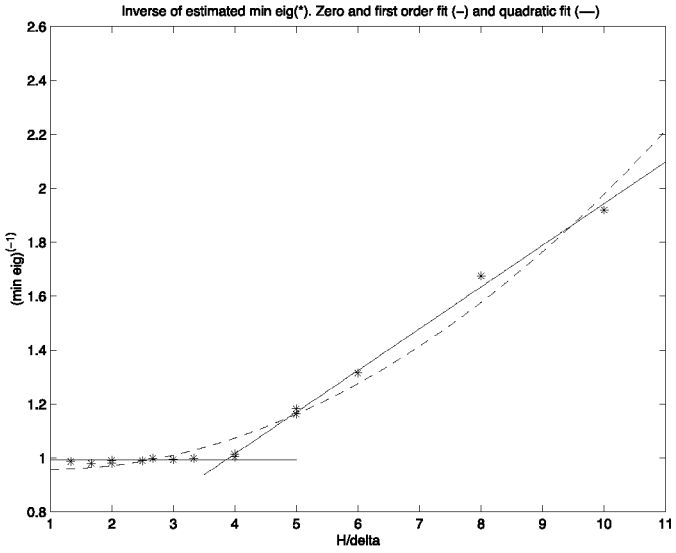


**Fig. 10.1.** Two-level method. Inverse of the minimum eigenvalue (asterisk), least-square zero (solid line), first (solid line), and second order (dashed line) fitting polynomials, versus $H/\delta$. Case for $n$ equal to 16, 18, and 20.

Fig. 10.1 shows the inverse of the smallest eigenvalue and some least-square fitting polynomials, versus $H/\delta$. We observe that the inverse of the smallest eigenvalue appears to grow linearly for $H/\delta \geq 4$ and is practically constant for $H/\delta \leq 4$. This suggests that the leading term of its asymptotic expansion is linear in $H/\delta$ and that the bound given in Theorem 10.2 is probably not sharp.

Additional results in [439] confirm that condition numbers are independent of $\eta_1$ and $\eta_2$.

## 10.2 Iterative Substructuring Methods

In this section, we present some iterative substructuring methods for problem (10.3), involving the lowest order Raviart-Thomas spaces $RT_0^h(\Omega) = RT_{1;0}^h(\Omega)$ in three dimensions and homogeneous Dirichlet conditions on $\partial\Omega$. The methods presented here are also valid for two-dimensional problems in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$; cf. Sect. 10.2.4. Throughout this section, we drop the subscripts that refer to the polynomial order in our notations, since we only consider lower-order approximation spaces. We only consider the case of meshes consisting of affinely mapped cubes: our results also hold for tetrahedral meshes.

We use the notation, introduced in Sect. 4.2, for problems in $H^1(\Omega)$. In particular, we assume that $\Omega$ is decomposed into nonoverlapping subdomains (substructures) $\Omega_i, i = 1, \ldots, N$, each of which is the union of shape-regular elements with the finite element nodes on the boundaries of neighboring subdomains matching across the interface.

The interface $\Gamma$ is the union of the interior subdomain faces, edges (all regarded as open sets), and vertices. For our methods, we only deal with subdomain faces and no further assumption needs to be made concerning the boundary edges. The faces of $\Omega_i$ are denoted by $\mathcal{F}^{ij}$, its edges by $\mathcal{E}^{ik}$, its vertices by $\mathcal{V}^{i\ell}$, and its wire basket by $\mathcal{W}^i$. We will also use notation with one or without any superscript.

Let $\mathcal{T}_i$ be a triangulation of the subdomain $\Omega_i$, of diameter $h_i$, and let $\mathcal{T} = \mathcal{T}_h$ be the geometrically conforming global mesh on $\Omega$, of diameter $h = \max\{h_i\}$; cf. appendix B.1.1. The diameter of $\Omega_i$ is $H_i$, with $H = \max\{H_i\}$. The degrees of freedom in the space $RT_0^h(\Omega)$ are associated with the faces of the fine triangulation $\mathcal{T}$. Fine faces on $\partial\Omega$, $\partial\Omega_i$, and $\Gamma$ are denoted here by $\partial\Omega_h$, $\partial\Omega_{i,h}$, and $\Gamma_h$, respectively.

We assume that all possible large jumps of the coefficients $a(x)$ and $B(x)$ are aligned with the subdomain boundaries and that on each subregion $\Omega_i$, $a(x)$ and $B(x)$ have constant values $a_i$ and $B_i$, respectively. The $\gamma_i$ and $\beta_i > 0$ are the largest and smallest eigenvalues of $B_i$:

$$\beta_i \eta^T \eta \leq \eta^T B_i \eta \leq \gamma_i \eta^T \eta, \quad \eta \in \mathbb{R}^3. \tag{10.38}$$

If we define

$$a_i(\mathbf{u}, \mathbf{v}) := \int_{\Omega_i} (a_i \operatorname{div} \mathbf{u} \operatorname{div} \mathbf{v} + B_i \, \mathbf{u} \cdot \mathbf{v}) \, dx, \qquad \mathbf{u}, \mathbf{v} \in RT^h(\Omega_i), \tag{10.39}$$

then the global bilinear form $a(\mathbf{u}, \mathbf{v}) = a_{\text{div}}(\mathbf{u}, \mathbf{v})$ can be written in terms of contributions from individual subregions as for (4.3). A similar procedure can be used for the right-hand side of (10.3).

We also assume that Assumption 4.3.1 holds: each substructure is the union of shape-regular coarse elements of a global conforming mesh $\mathcal{T}_H$ and the number of such elements forming an individual subdomain is uniformly bounded. We will employ the low-order Raviart-Thomas space $RT_0^H(\Omega) = RT_{1;0}^H(\Omega)$ on the coarse triangulation $\mathcal{T}_H$.

In the case of a region of diameter $H_i$, we use norms with relative weights obtained by a simple dilation argument:

$$\|u\|^2_{H^{1/2}(\partial\Omega_i)} = |u|^2_{H^{1/2}(\partial\Omega_i)} + \frac{1}{H_i} \|u\|^2_{L^2(\partial\Omega_i)}.$$

For $H^{-1/2}(\partial\Omega_i)$, we employ the dual norm

$$\|u\|_{H^{-1/2}(\partial\Omega_i)} = \sup_{\substack{\phi \in H^{1/2}(\partial\Omega_i) \\ \phi \neq 0}} \frac{\langle u, \phi \rangle}{\|\phi\|_{H^{1/2}(\partial\Omega_i)}},$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $H^{-1/2}(\partial\Omega_i)$ and $H^{1/2}(\partial\Omega_i)$.

We will work with finite element functions defined on the boundaries $\partial\Omega_i$ and the interface $\Gamma$. Indeed, the normal component of a vector $\mathbf{u} \in RT^h(\Omega_i)$ belongs to $Q_i = Q_i^h$, the space of functions that are constant on each fine face $f \in \partial\Omega_{i,h}$ and that vanish on $\partial\Omega \cap \partial\Omega_i$. We define a subspace thereof by

$$Q_{i;0} = Q_{i;0}^h := \left\{ \mu \in Q_i^h \mid \int_{\partial\Omega_i} \mu \, ds = 0 \right\}.$$

Once a normal vector $\mathbf{n}$ to $\Gamma$ is defined, the normal component of a vector $\mathbf{u} \in RT_0^h(\Omega)$ belongs to the space of functions that are constant on each fine face $f \in \Gamma_h$ and that vanish on $\partial\Omega \cap \partial\Omega_i$ and, from now on, we denote this by $Q = Q^h$.

We finally introduce a space of traces of coarse functions: let $Q^H$ be the space of functions $\mu$ defined on $\Gamma$, such that for each substructure $\Omega_i$ and each face $\mathcal{F}^{ij}$, $\mu$ is constant on $\mathcal{F}^{ij}$. We note that $Q^H$ is the space of normal traces on $\Gamma$ of vectors in $RT_0^H(\Omega)$.

We introduce the following convention: given a Raviart-Thomas vector $\mathbf{u} \in RT_0^H(\Omega)$, we denote by $u$ the column vector of its degrees of freedom. This makes sense since degrees of freedom in $RT_0^H(\Omega)$ are given in terms of (scalar) normal components over the faces of $\mathcal{T}$. In addition, this ensures that many of the results in Chap. 4, 5, and 6 hold verbatim. For degrees of freedom associated with a substructure $\Omega_i$, we use the notations $u^{(i)}$ or, equivalently, $u_i$. In addition, for the same vector $\mathbf{u}$ we also denote its normal component on the boundary $\Gamma$ by $u$ and on $\partial\Omega_i$, by $u^{(i)}$ or $u_i$, along fixed unit normal vectors ($\mathbf{n}$ and $\mathbf{n}_i$, respectively).

We can now proceed as in Sect. 4.3:
we number the degrees of freedom relative to faces interior to the subdomains first (subscript $I$), followed by those on the interface $\Gamma$ (subscript $\Gamma$). The contributions to the stiffness matrix $A$ and the right hand side $f$ can be formed one subdomain at a time. We obtain the same linear system as in (4.7). The unknowns in the interior of the substructures are eliminated and a Schur complement system, involving the degrees of freedom on $\Gamma$, is obtained, as in (4.10) and (4.11):

$$Su_\Gamma = \tilde{f}_\Gamma.$$

The contributions to the Schur complement matrix $S : Q \to Q$, can also be formed one subdomain at a time, using the local Schur complements $S^{(j)} : Q_j \to Q_j$. A similar procedure applies to the right hand side $\tilde{f}_\Gamma$. We note that, since no degrees of freedom are associated with subdomain edges and vertices, according to the representations (4.12), (4.13), and (5.1), Schur complements only have blocks associated with single faces: $S^{(j)} = S^{(j)}_{\mathcal{F}\mathcal{F}}$ and $S = S_{\mathcal{F}\mathcal{F}}$.

Discrete harmonic extensions can also be defined in this case: a local vector $\mathbf{u}_i$, determined by the local degrees of freedom $u^{(i)}$, is said to be discrete harmonic on $\Omega_i$ if (4.14) holds. This corresponds to the continuous problem

$$\begin{aligned} -\mathbf{grad}\,(a\,\mathrm{div}\,\mathbf{u}) + B\mathbf{u} &= 0 && \text{in } \Omega_i, \\ \mathbf{u}\cdot\mathbf{n}_i &= u_\Gamma^{(i)} && \text{on } \partial\Omega_i \setminus \partial\Omega, \\ \mathbf{u}\cdot\mathbf{n}_i &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega. \end{aligned}$$

As for continuous, piecewise polynomial functions, $u^{(i)} =: \mathcal{H}(u_\Gamma^{(i)})$, for $u_\Gamma^{(i)} \in Q_i$, or, equivalently, $\mathbf{u}_i = \mathcal{H}(u_\Gamma^{(i)})$, is completely defined by its normal component on $\partial\Omega_i$ and is orthogonal, in the $a_i(\cdot,\cdot)$-inner product, to the space $RT_0^h(\Omega_i)$. A vector $\mathbf{u}$ is piecewise discrete harmonic if its restriction, $\mathbf{u}_i$, to $\Omega_i$ is discrete harmonic; the vector $\mathbf{u} =: \mathcal{H}(u)$ is completely defined by its normal component $u$ on $\Gamma$.

We will work with the inner products defined by the Schur complements (see (4.15)):

$$\begin{aligned} s(u,v) &:= u^T S v, \quad u, v \in Q, \\ s_i(u_i,v_i) &:= u_i^T S^{(i)} v_i, \quad u_i, v_i \in Q_i. \end{aligned}$$

The minimization property of discrete harmonic extensions in Lemma 4.9 remains valid in this case and we can equivalently work with functions defined on the interface $\Gamma$ (in the spaces $Q_i$ and $Q$ with the scalar product $s(\cdot,\cdot)$) and the corresponding discrete harmonic extensions (with scalar product $a(\cdot,\cdot)$). We note however that we cannot, in general, work with norms of local discrete harmonic extensions and traces on the subdomains as in Lemma 4.10. A trace theorem is indeed given in Lemma A.19 and it provides bounds that are independent of the diameter of the substructure: for $\mathbf{u} \in H_0(\mathrm{div}\,;\Omega_i)$, a scaling argument gives

$$\begin{aligned} \|u\|^2_{H^{-1/2}(\partial\Omega_i)} &= \|\mathbf{u}\cdot\mathbf{n}_i\|^2_{H^{-1/2}(\partial\Omega_i)} \\ &\le C(H_i^2\,\|\mathrm{div}\,\mathbf{u}\|^2_{L^2(\Omega_i)} + \|\mathbf{u}\|^2_{L^2(\Omega_i)}). \end{aligned} \tag{10.40}$$

However, harmonic extensions that ensure stability constants that are independent of the diameter of the subdomain can in general be found only for special functions; see Lemma 10.19 below.

We first need to introduce some technical tools. We will only present the most relevant proofs and refer to the references for the others.

### 10.2.1 Technical Tools

#### Raviart-Thomas Interpolants onto Standard Coarse Spaces

We consider the interpolation operator $\Pi_{RT}^H = \Pi_{RT_1}^H$ onto $RT_0^H(\Omega)$, which, we recall, is defined in terms of the degrees of freedom

$$\lambda_{\mathcal{F}}(\Pi_{RT}^H \mathbf{u}) := |\mathcal{F}|^{-1} \int_{\mathcal{F}} \mathbf{u}\cdot\mathbf{n}\,ds,$$

with $\mathcal{F}$ a face of $\mathcal{T}_H$; see appendix B.3.1. The interpolant $\Pi_{RT}^H$ is logarithmically stable in the $\|\cdot\|_{\mathrm{div}}$-norm, in three dimensions, in contrast to the nodal

interpolant on continuous finite element spaces, which has a norm which grows algebraically with $H/h$; cf. Lemma 4.12. We note that for the case at hand the best bound for the $L^2$-norm alone involves a factor of $H/h$; this can easily be seen by considering an element $\mathbf{u}$, for which all the interior degrees of freedom vanish. The proof of the following result was originally given in [470].

**Lemma 10.14** *There exists a constant $C$, which depends only on the aspect ratios of $K \in \mathcal{T}_H$ and of the elements of $\mathcal{T}_h$, such that, for $\mathbf{u} \in RT_0^h(\Omega)$,*

$$\begin{aligned} &\|\mathrm{div}\,(\Pi_{RT}^H \mathbf{u})\|_{L^2(K)}^2 \le \|\mathrm{div}\,\mathbf{u}\|_{L^2(K)}^2, \\ &\|\Pi_{RT}^H \mathbf{u}\|_{L^2(K)}^2 \le C\,(1+\log(H/h))\left(\|\mathbf{u}\|_{L^2(K)}^2 + H_K^2 \|\mathrm{div}\,\mathbf{u}\|_{L^2(K)}^2\right). \end{aligned} \tag{10.41}$$

*Proof.* The commuting diagram property (B.27) ensures that

$$\left(\mathrm{div}\,(\Pi_{RT}^H \mathbf{u})\right)_{|K} = \left(\Pi_Q^H (\mathrm{div}\,\mathbf{u})\right)_{|K}, \tag{10.42}$$

where $\Pi_Q^H$ is the $L^2$-projection onto the space of constants on $K$. The first inequality in (10.41) follows immediately.

The proof of the second inequality of (10.41) uses Green's formula, Lemma B.17, and the *continuous*, piecewise linear partition of unity functions $\{\theta_{\mathcal{F}}\}$, associated with the faces $\{\mathcal{F}\}$ of $K$, constructed in Lemma 4.25.

We consider a face $\mathcal{F} \subset \partial K$, and note that it is partitioned into nonoverlapping faces $f \in \mathcal{F}_h$. We number these faces so that $f_i$, $1 \le i \le n_{\mathcal{F}}$, are the faces that have at least one vertex on the boundary $\partial\mathcal{F}$; see, e.g., Fig. 5.1. We note that since, by assumption, the triangulation of the face $\mathcal{F}$ is quasi uniform, $n_{\mathcal{F}} \le C(H_K/h)$. For a face $f_i \in \mathcal{F}_h$, let $k_i \subset K$ be the element of the fine triangulation to which $f_i$ belongs.

We apply Lemma B.17 to the coarse triangulation $\mathcal{T}_H$; it is then sufficient to bound $\lambda_{\mathcal{F}}(\Pi_{RT}^H \mathbf{u})$, for each face $\mathcal{F} \subset \partial K$. We can write

$$\begin{aligned} |\mathcal{F}|\,\lambda_{\mathcal{F}}(\Pi_{RT}^H \mathbf{u}) &= \int_{\mathcal{F}} (\mathbf{u}\cdot\mathbf{n})\,ds \\ &= \int_{\mathcal{F}} (1-\theta_{\mathcal{F}})(\mathbf{u}\cdot\mathbf{n})\,ds + \int_{\partial K} \theta_{\mathcal{F}}(\mathbf{u}\cdot\mathbf{n})\,ds =: I + II. \end{aligned} \tag{10.43}$$

The first term on the right hand side is an integral over the fine faces $f_i$ that touch $\partial\mathcal{F}$, where $\mathbf{u}\cdot\mathbf{n}$ is constant and equal to $\lambda_{f_i}(\mathbf{u})$. Since $\theta_{\mathcal{F}}$ has values between zero and one, applying Lemma B.17 to the fine triangulation $\mathcal{T}$ gives

$$\begin{aligned} |I| \le C \sum_{i=1}^{n_{\mathcal{F}}} |\lambda_{f_i}(\mathbf{u})|\,|f_i| &\le \left(\sum_{i=1}^{n_{\mathcal{F}}} \|\mathbf{u}\|_{L^2(k_i)}^2\right)^{1/2} \left(\sum_{i=1}^{n_{\mathcal{F}}} h\right)^{1/2} \\ &\le C H_K^{1/2} \|\mathbf{u}\|_{L^2(K)}. \end{aligned} \tag{10.44}$$

For the second term in (10.43), we use Green's formula and Lemma 4.25:

$$
\begin{aligned}
|II| &\leq \int_K |\mathrm{div}\,(\theta_{\mathcal{F}}\mathbf{u})|\,dx = \int_K |\theta_{\mathcal{F}}\mathrm{div}\,\mathbf{u} + \mathbf{grad}\,\theta_{\mathcal{F}}\cdot\mathbf{u}|\,dx \\
&\leq CH_K^{3/2}\|\mathrm{div}\,\mathbf{u}\|_{L^2(K)} + C(H_K(1+\log(H/h)))^{1/2}\|\mathbf{u}\|_{L^2(K)}.
\end{aligned}
\tag{10.45}
$$

Combining (10.43), (10.44), and (10.45) yields

$$
|\mathcal{F}|^2\,\lambda_{\mathcal{F}}(\Pi_{RT}^H\mathbf{u})^2 \leq C\left(H_K^3\|\mathrm{div}\,\mathbf{u}\|_{L^2(K)}^2 + H_K(1+\log(H/h))\|\mathbf{u}\|_{L^2(K)}^2\right).
$$

The proof is concluded by using Lemma B.17 and summing over the faces $\mathcal{F}\subset\partial K$. □

In our analysis, we will also employ the operator $\tilde{\Pi}_{RT}^H$, which provides the discrete harmonic function with the same normal component as $\Pi_{RT}^H$ on $\Gamma$. More precisely, $\mathbf{v} = \tilde{\Pi}_{RT}^H\mathbf{u}$ is obtained by first applying the coarse interpolation operator $\Pi_{RT}^H$ to $\mathbf{u}$ and then extending its normal component on $\Gamma$ harmonically inside the substructures. We note that $\tilde{\Pi}_{RT}^H\mathbf{u}$ and $\Pi_{RT}^H\mathbf{u}$ have the same normal component $v\in Q^H$ on $\Gamma$ and bounds for the energy norm of the former can be obtained from Lemma 10.14 and the minimizing property of discrete harmonic extensions. We will also employ local operators $\tilde{\Pi}_{RT}^{H_i}$, $i=1,\ldots,N$, for which the previous procedure is applied on single substructures.

### Equivalent Trace Norms

We introduce equivalent norms for the finite-dimensional trace spaces $Q$, $Q_i$, and $Q^H$. We refer to [470, Lem. 2.2] for a proof which employs a Poincaré inequality in $H^{1/2}(\partial\Omega_i)$.

**Lemma 10.15** *There exists a constant $c>0$, which is independent of the diameter of $\Omega_i$, $i=1,\ldots,N$, such that for each $\mu\in H^{-1/2}(\partial\Omega_i)$ with $\langle\mu,1\rangle = 0$*

$$
c\sup_{\substack{\phi\in H^{1/2}(\partial\Omega_i)\\ \phi\neq\mathrm{const}}}\frac{\langle\mu,\phi\rangle}{|\phi|_{H^{1/2}(\partial\Omega_i)}} \leq \|\mu\|_{H^{-1/2}(\partial\Omega_i)} \leq \sup_{\substack{\phi\in H^{1/2}(\partial\Omega_i)\\ \phi\neq\mathrm{const}}}\frac{\langle\mu,\phi\rangle}{|\phi|_{H^{1/2}(\partial\Omega_i)}}.
$$

For finite element functions in $Q^h$ (and thus in $Q^H$), the supremum in the definition of the $H^{-1/2}$-norm can be taken over a finite-dimensional space. It is not enough to replace $H^{1/2}(\partial\Omega_i)$ with the space of continuous, piecewise bilinear functions

$$
V_i = V_i^h := \{\phi\in C^0(\partial\Omega_i)\mid \phi_{|_f}\in\mathbb{Q}_1(f),\ f\in\mathcal{F}_h^{ij},\ \mathcal{F}^{ij}\subset\partial\Omega_i\},
$$

but a larger space needs to be employed. We introduce the space of quadratic bubbles

$$
B_i = B_i^h := \{\phi\in C^0(\partial\Omega_i)\mid \phi_{|_f} = \alpha_f\varphi_1\,\varphi_2\,\varphi_3\,\varphi_4,\ f\in\mathcal{F}_h^{ij},\ \mathcal{F}^{ij}\subset\partial\Omega_i\},
$$

where $\varphi_i$, $1 \le i \le 4$, are the nodal basis functions that span $\mathbb{Q}_1(f)$ on the face $f$. The support of any bubble basis function is exactly one element. We then define

$$\tilde{V}_i = \tilde{V}_i^h := V_i + B_i.$$

We have the following property; see [470, Lem. 2.3] for a proof.

**Lemma 10.16** *For $\phi \in \tilde{V}_i$, there exist a unique decomposition $\phi = \phi_V + \phi_B$, with $\phi_V \in V_i$ and $\phi_B \in B_i$. In addition, there exists a constant $C$, which depends only on the aspect ratio of $\Omega_i$, such that*

$$|\phi|_{H^{1/2}(\partial\Omega_i)} \le |\phi_V|_{H^{1/2}(\partial\Omega_i)} + |\phi_B|_{H^{1/2}(\partial\Omega_i)} \le C|\phi|_{H^{1/2}(\partial\Omega_i)}.$$

The proof of the following result can be found in [470, Lem. 4.2].

**Lemma 10.17** *There exist constants, $c$ and $C$, such that, for $\mu \in Q_i$,*

$$c \sup_{\phi\in\tilde{V}_i} \frac{\langle\mu,\phi\rangle}{\|\phi\|_{H^{1/2}(\partial\Omega_i)}} \le \|\mu\|_{H^{-1/2}(\partial\Omega_i)} \le C \sup_{\phi\in\tilde{V}_i} \frac{\langle\mu,\phi\rangle}{\|\phi\|_{H^{1/2}(\partial\Omega_i)}}. \tag{10.46}$$

*Furthermore, if $\langle\mu,1\rangle = 0$, then the $H^{1/2}$-norm in (10.46) can be replaced by the seminorm and the supremum can be taken over the nonconstant functions $\phi$.*

The following lemma compares norms of trace functions on substructures that share a face.

**Lemma 10.18** *Let $\Omega_i$ and $\Omega_j$ be two substructures with a common face $\mathcal{F}$. Let $\mu_{\mathcal{F}}$ be a function in $L^2(\partial\Omega_i \cup \partial\Omega_j)$, that vanishes outside $\mathcal{F}$. Then, there is a constant $C$, that depends only on the aspect ratios of $\Omega_i$ and $\Omega_j$, such that*

$$\|\mu_{\mathcal{F}}\|_{H^{-1/2}(\partial\Omega_i)} \le C\|\mu_{\mathcal{F}}\|_{H^{-1/2}(\partial\Omega_j)}.$$

*Proof.* For simplicity, we assume that the two substructures are elements of the coarse mesh, obtained by mapping the reference cube $\hat{\Omega}$ with two affine mappings $F_i$ and $F_j$. We have

$$\begin{aligned}
\|\mu_{\mathcal{F}}\|_{H^{-1/2}(\partial\Omega_i)} &= \sup_{\phi\in H^{1/2}(\partial\Omega_i)} \frac{\langle\mu_{\mathcal{F}},\phi\rangle}{\|\phi\|_{H^{1/2}(\partial\Omega_i)}} \\
&\le C_1 \sup_{\hat{\phi}\in H^{1/2}(\partial\hat{\Omega})} \frac{H_i^2\langle\hat{\mu},\hat{\phi}\rangle}{H_i^{1/2}\|\hat{\phi}\|_{H^{1/2}(\partial\hat{\Omega})}} \\
&\le C_1\, C_2 (H_i/H_j))^{3/2} \sup_{\phi\in H^{1/2}(\partial\Omega_j)} \frac{\langle\mu_{\mathcal{F}},\phi\rangle}{\|\phi\|_{H^{1/2}(\partial\Omega_j)}},
\end{aligned}$$

where the constants $C_1$ and $C_2$ only depend on the aspect ratios of the substructures and $\hat{\mu}(\hat{x}) := \mu_{\mathcal{F}}(F_i(\hat{x})) = \mu_{\mathcal{F}}(F_j(\hat{x}))$. The product $C_1C_2(H_i/H_j)^{3/2}$ can then be bounded by a constant that only depends on the aspect ratios of the two substructures. □

### Stable Extensions

As for the case of continuous finite element spaces, we need an extension operator of boundary functions in $Q_i$; cf. Lemma 4.6. The stable extension operator, defined in the next lemma, provides a divergence-free extension of boundary functions in the subspace $Q_{i;0}$. This fact ensures that the stability constant will be independent of the diameter of $\Omega_i$. This will not be true for some other extension procedures.

**Lemma 10.19** *Let $\Omega_i$ be a substructure. Then, there exists an extension operator $\widetilde{\mathcal{H}}_i : Q_{0;h} \to RT^h(\Omega_i)$, such that, for $\mu \in Q_{0;h}$, $\widetilde{\mathcal{H}}_i\mu$ has vanishing divergence and*

$$\|\widetilde{\mathcal{H}}_i\mu\|_{L^2(\Omega_i)} \leq C\|\mu\|_{H^{-1/2}(\Omega_i)}, \tag{10.47}$$

*where $C$ is independent of $h$, $H_i$, and $\mu$.*

*Proof.* The proof is similar to that of Lemma 4.6. It is enough to find the solution $u$ of a Laplace problem with Neumann boundary datum $\mu$ and interpolate $\mathbf{grad}\, u$ into the Raviart-Thomas space $RT^h(\Omega_i)$. The application of the regularity result in Lemma A.53 and the error estimates in (B.19) conclude the proof; see [470, Lem. 4.3] for details. □

### Decomposition Results

In our analysis, we need a decomposition result for trace functions on the boundary of a substructure. In contrast to the case of continuous, piecewise linear functions, we here only have contributions associated with the faces. We refer to Sect. 4.6 for decomposition results in $H^1$ into face, wire basket, edge, and vertex components.

Given a substructure $\Omega_i$, we define partition of unity functions associated with its faces; cf. Lemmas 4.25 and 4.24. Here, we work with discontinuous functions and for a face $\mathcal{F}$, we let $\zeta_{\mathcal{F}}$ be the characteristic function of $\mathcal{F}$, i.e., the function that is identically one on $\mathcal{F}$ and zero on $\partial\Omega_i \setminus \mathcal{F}$. We clearly have

$$\sum_{\mathcal{F}} \zeta_{\mathcal{F}}(x) = 1, \quad \text{almost everywhere on } \partial\Omega_i \setminus \partial\Omega.$$

Given a function $\mu \in Q_i$ and a face of $\Omega_i$, we denote

$$\mu_{\mathcal{F}} := \zeta_{\mathcal{F}}\mu \in Q_i.$$

We have the following estimates for the face components of the particular functions in $Q_i$ with vanishing average on the subdomain faces. A proof was originally given in [470].

**Lemma 10.20** *Let $\mu \in Q_i$. Then,*

$$\sum_{\mathcal{F}} \mu_{\mathcal{F}}(x) = \mu(x), \quad \textit{almost everywhere in } \partial\Omega_i \setminus \partial\Omega.$$

*If, in addition,* $\mu \in Q_{i;0}$ *and*

$$\int_{\mathcal{F}} \mu \, ds = \int_{\mathcal{F}} \mu_{\mathcal{F}} \, ds = 0, \quad \mathcal{F} \subset \partial\Omega_i,$$

*then, there exists a constant* $C$*, independent of* $h$ *and* $\mu_H$*, such that, for* $\mu_H \in Q^H$*,*

$$\begin{aligned} &||\mu_{\mathcal{F}}||^2_{H^{-1/2}(\partial\Omega_i)} \\ &\leq C\left(1+\log\left(\tfrac{H}{h}\right)\right)\left(\left(1+\log\left(\tfrac{H}{h}\right)\right)||\mu+\mu_H||^2_{H^{-1/2}(\partial\Omega_i)} + ||\mu||^2_{H^{-1/2}(\partial\Omega_i)}\right). \end{aligned}$$

*Proof.* The first property follows directly from the definition of $\mu_{\mathcal{F}}$. The proof of the second property is carried out in several steps.

**Step 1.** Since $\mu_{\mathcal{F}} \in Q_{i;0}$, Lemma 10.17 ensures that

$$||\mu_{\mathcal{F}}||_{H^{-1/2}(\partial\Omega_i)} \leq C \sup_{\substack{\phi\in\tilde{V}_i \\ \phi\neq\text{const.}}} \frac{\langle \mu_{\mathcal{F}}, \phi\rangle}{|\phi|_{H^{1/2}(\partial\Omega_i)}}.$$

By using Lemma 10.16, we find that any function $\phi \in \tilde{V}_i$ can be decomposed uniquely into $\phi_V + \phi_B$, with $\phi_V \in V_i$ and $\phi_B \in B_i$ and thus

$$\begin{aligned} ||\mu_{\mathcal{F}}||_{H^{-1/2}(\partial\Omega_i)} \leq\ & C \sup_{\substack{\phi_V\in V_i \\ \phi_V\neq\text{const.}}} \frac{\langle \mu_{\mathcal{F}}, \phi_V\rangle}{|\phi_V|_{H^{1/2}(\partial\Omega_i)}} \\ & + C \sup_{\substack{\phi_B\in B_i \\ \phi_B\neq 0}} \frac{\langle \mu_{\mathcal{F}}, \phi_B\rangle}{|\phi_B|_{H^{1/2}(\partial\Omega_i)}}. \end{aligned} \tag{10.48}$$

**Step 2.** We first consider the first term on the right hand side of (10.48). We consider the continuous, piecewise linear partition of unity functions $\{\theta_{\mathcal{F}}\}$, associated with the faces $\{\mathcal{F}\}$ of $\Omega_i$, constructed in Lemma 4.25. For $\phi_V \in V_i$, we now define a weighted average $c_{\phi_V}$ by

$$c_{\phi_V} \int_{\mathcal{F}} \theta_{\mathcal{F}} \, ds = \int_{\mathcal{F}} I^h(\theta_{\mathcal{F}}\phi_V) \, ds,$$

with $I^h$ is the nodal interpolation operator onto $V_i$. Then, the supremum in the first term on the right in (10.48) can be replaced by

$$\begin{aligned} \sup_{\substack{\phi_V\in V_i \\ \phi_V\neq\text{const.}}} \frac{\langle \mu_{\mathcal{F}}, \phi_V\rangle}{|\phi_V|_{H^{1/2}(\partial\Omega_i)}} &= \sup_{\substack{\phi_V\in V_i \\ \phi_V\neq\text{const.}}} \frac{\langle \mu_{\mathcal{F}}, \phi_V - c_{\phi_V}\rangle}{|\phi_V - c_{\phi_V}|_{H^{1/2}(\partial\Omega_i)}} \\ &= \sup_{\substack{\phi_V\in V_i \\ \phi_V\neq 0,\, c_{\phi_V}=0}} \frac{\langle \mu_{\mathcal{F}}, \phi_V\rangle}{|\phi_V|_{H^{1/2}(\partial\Omega_i)}}, \end{aligned} \tag{10.49}$$

i.e., we need only consider functions $\phi_V$ which have a zero weighted average. The following norm equivalence can be proven using Lemma A.17:

$$\begin{aligned} c(|\phi_V|^2_{H^{1/2}(\partial\Omega_i)} + Hc^2_{\phi_V}) &\le \|\phi_V\|^2_{H^{1/2}(\partial\Omega_i)} \\ &\le C(|\phi_V|^2_{H^{1/2}(\partial\Omega_i)} + Hc^2_{\phi_V}). \end{aligned} \tag{10.50}$$

We remark that, because of (10.50), the $H^{1/2}$-seminorm in the last term of (10.49) can be replaced by the full norm. We now decompose $\phi_V$ into components associated with the faces and the wire basket of $\Omega_i$:

$$\phi_V = \sum_{\mathcal{F}\subset\partial\Omega_i} \phi_{V,\mathcal{F}} + \phi_{V,\mathcal{W}} := \sum_{\mathcal{F}\subset\partial\Omega_i} I^h(\theta_{\mathcal{F}}\phi_V) + \phi_{V,\mathcal{W}}. \tag{10.51}$$

Lemmas 4.19 and A.6 give

$$|\phi_{V,\mathcal{W}}|^2_{H^{1/2}(\partial\Omega_i)} \le C\|\phi_{V,\mathcal{W}}\|^2_{L^2(\mathcal{W}^i)}.$$

Using Lemmas 4.17 and 4.10 then gives

$$|\phi_{V,\mathcal{W}}|^2_{H^{1/2}(\partial\Omega_i)} \le C(1+\log(H/h))\ \|\phi_V\|^2_{H^{1/2}(\partial\Omega_i)}. \tag{10.52}$$

For the face components, Lemma 4.24 and the same trace and extension results employed for $\phi_{V,\mathcal{W}}$ give

$$\|\phi_{V,\mathcal{F}}\|^2_{H^{1/2}(\partial\Omega_i)} \le C(1+\log(H/h))^2\|\phi_V\|^2_{H^{1/2}(\partial\Omega_i)}. \tag{10.53}$$

We find, by using the splitting (10.51), that

$$\begin{aligned} \langle\mu_{\mathcal{F}}, \phi_V\rangle &= \sum_{\hat{\mathcal{F}}\subset\partial T} \langle\mu_{\mathcal{F}}, \phi_{V,\hat{\mathcal{F}}}\rangle + \langle\mu_{\mathcal{F}}, \phi_{V,\mathcal{W}}\rangle \\ &= \langle\mu, \phi_{V,\mathcal{F}}\rangle + \langle\mu_{\mathcal{F}}, \phi_{V,\mathcal{W}}\rangle. \end{aligned} \tag{10.54}$$

Since $I^h(\theta_{\mathcal{F}}\phi_V) = \phi_{V,\mathcal{F}} = I^h(\theta_{\mathcal{F}}\phi_{V,\mathcal{F}})$ and since we can always assume that $c_{\phi_V} = 0$, we obtain

$$\langle\mu_H, \phi_{V,\mathcal{F}}\rangle = 0, \quad \mu_H \in Q^H,$$

and $c_{\phi_{V,\mathcal{F}}} = 0$. The first term on the right side of (10.54) can be bounded by means of (10.53). For $\mu_H \in Q^H$, we have

$$\begin{aligned} |\langle\mu, \phi_{V,\mathcal{F}}\rangle| &= |\langle\mu+\mu_H, \phi_{V,\mathcal{F}}\rangle| \\ &\le C(1+\log(H/h))\|\phi_V\|_{H^{1/2}(\partial\Omega_i)}\ \|\mu+\mu_H\|_{H^{-1/2}(\partial\Omega_i)}. \end{aligned} \tag{10.55}$$

In order to bound the second term on the right side of (10.54), we note that, for each $\phi_{V,\mathcal{W}}$, there is a unique $\tilde{\phi}_{B,\mathcal{F}} \in B_i$, such that

$$\int_f \phi_{V,\mathcal{W}}\ ds = \int_f \tilde{\phi}_{B,\mathcal{F}}\ ds, \quad f \in \mathcal{F}_h,$$

with $\tilde{\phi}_{B,\mathcal{F}} = 0$ on $\partial\Omega_i \setminus \mathcal{F}$. This mapping is clearly continuous in $L^2(\partial\Omega_i)$. Since in addition $\phi_{V,\mathcal{W}}$ vanishes outside a strip of width $h$ around $\partial\mathcal{F}$, an

inverse inequality (cf. Lemma B.27) and a Friedrichs' inequality (cf. Lemma A.17) give

$$\begin{aligned}||\tilde{\phi}_{B,\mathcal{F}}||^2_{H^{1/2}(\partial\Omega_i)} &\leq C\frac{1}{h}||\tilde{\phi}_{B,\mathcal{F}}||^2_{L^2(\partial\Omega_i)} \\ &\leq C\frac{1}{h}||\phi_{V,\mathcal{W}}||^2_{L^2(\partial\Omega_i)} \leq C||\phi_{V,\mathcal{W}}||^2_{H^{1/2}(\partial\Omega_i)}.\end{aligned}$$

By means of this bound and (10.52), we finally obtain

$$\begin{aligned}&|\langle\mu_{\mathcal{F}},\phi_{V,\mathcal{W}}\rangle| = \left|\langle\mu_{\mathcal{F}},\tilde{\phi}_{B,\mathcal{F}}\rangle\right| = \left|\langle\mu,\tilde{\phi}_{B,\mathcal{F}}\rangle\right| \\ &\leq C||\mu||_{H^{-1/2}(\partial\Omega_i)}||\phi_{V,\mathcal{W}}||_{H^{1/2}(\partial\Omega_i)} \\ &\leq C(1+\log(H/h))^{1/2}||\mu||_{H^{-1/2}(\partial\Omega_i)}||\phi_V||_{H^{1/2}(\partial\Omega_i)}.\end{aligned} \tag{10.56}$$

Combining (10.54), (10.55), and (10.56) yields

$$\begin{aligned}&|\langle\mu_{\mathcal{F}},\phi_V\rangle| \leq C(1+\log(H/h))^{1/2}||\phi_V||_{H^{1/2}(\partial\Omega_i)} \\ &\times(||\mu+\mu_H||_{H^{-1/2}(\partial\Omega_i)} + (1+\log(H/h))^{1/2}||\mu||_{H^{-1/2}(\partial\Omega_i)}).\end{aligned} \tag{10.57}$$

**Step 3.** We now consider the second term on the right hand side of (10.48). We decompose $\phi_B$ into the sum of terms $\phi_{B,\mathcal{F}}$ supported on individual faces $\mathcal{F}\subset\partial\Omega_i$:

$$\phi_B = \sum_{\mathcal{F}\subset\partial\Omega_i}\phi_{B,\mathcal{F}}. \tag{10.58}$$

Bounds for the $L^2$-norm can easily be found for $\phi_{B,\mathcal{F}}$. For the $H^1$-seminorm, we employ an inverse inequality and Corollary A.15 on each fine face $f$:

$$|\phi_{B,\mathcal{F}}|^2_{H^1(\partial\Omega_i)} \leq C|\phi_B|^2_{H^1(\partial\Omega_i)}.$$

An interpolation argument using Lemma A.11 then gives a bound for the $H^{1/2}$-norms:

$$||\phi_{B,\mathcal{F}}||^2_{H^{1/2}(\partial\Omega_i)} \leq C||\phi_B||^2_{H^{1/2}(\partial\Omega_i)}. \tag{10.59}$$

Using (10.59), we find:

$$\begin{aligned}|\langle\mu_{\mathcal{F}},\phi_B\rangle| = |\langle\mu,\phi_{B,\mathcal{F}}\rangle| &\leq ||\mu||_{H^{-1/2}(\partial\Omega_i)}\ ||\phi_{B,\mathcal{F}}||_{H^{1/2}(\partial\Omega_i)} \\ &\leq C||\mu||_{H^{-1/2}(\partial\Omega_i)}||\phi_{B,\mathcal{F}}||_{H^{1/2}(\partial\Omega_i)} \\ &\leq C||\mu||_{H^{-1/2}(\partial\Omega_i)}||\phi_B||_{H^{1/2}(\partial\Omega_i)}.\end{aligned} \tag{10.60}$$

**Step 4.** The proof is completed by combining (10.48), (10.57), and (10.60). □

### 10.2.2 A Face-Based Method

We now construct a two-level preconditioner where local components are associated with single faces of the partition and the coarse solve relies on the solution of a problem on a standard coarse space. These algorithms were originally presented and studied in [449] in two dimensions and in [470] in three dimensions. Following the ideas in Sect. 5.3, the local component of the preconditioner is obtained by replacing $S^{-1} = S_{\mathcal{F}\mathcal{F}}^{-1}$ by a block diagonal matrix $\bar{S}_{\mathcal{F}\mathcal{F}}^{-1}$. Each block in $\bar{S}_{\mathcal{F}\mathcal{F}}$ is given by $S_{\mathcal{F}^i\mathcal{F}^i}$, the submatrix of $S$ associated with the face $\mathcal{F}^i$. We define $Q_{\mathcal{F}^i}$ as the space of piecewise constant functions on $\mathcal{F}^i$ and $R_{\mathcal{F}^i} : Q \to Q_{\mathcal{F}^i}$, as the rectangular restriction matrix which returns only the components of a global vector associated with $\mathcal{F}^i$.

The coarse component is the solution of a coarse problem for the space $Q_0 := Q^H$ on the coarse mesh $\mathcal{T}_H$. Here, we have assumed, for simplicity, that the single substructures are mapped cubes of $\mathcal{T}_H$. If $R_0^T : Q_0 \to Q$, is the natural interpolation operator from the coarse to the fine mesh, our additive preconditioned operator $P = \sum_j P_j$ can be written as

$$P = B^{-1}S = R_0^T S_0^{-1} R_0 S + \sum_{\mathcal{F}^i} R_{\mathcal{F}^i}^T S_{\mathcal{F}^i\mathcal{F}^i}^{-1} R_{\mathcal{F}^i} S,$$

with $S_0 = R_0 S R_0^T$ the restriction of $S$ to $Q_0$. This corresponds to the decomposition

$$Q = R_0^T Q_0 + \sum_{\mathcal{F}^i} R_{\mathcal{F}^i}^T Q_{\mathcal{F}^i}.$$

We note that our preconditioner employs exact solvers on all the subspaces and that the action of $S_{\mathcal{F}^i\mathcal{F}^i}^{-1}$ on a vector can be calculated by solving a Dirichlet problem on the union of the two substructures that share $\mathcal{F}^i$; cf. equation (5.3). Numerical results given in [470] show that $S_0$ can be replaced by the stiffness matrix $A_H$ relative to the coarse mesh $\mathcal{T}_H$.

We have the following result:

**Theorem 10.21** *The condition number of the additive Schwarz operator $P$ satisfies*

$$\kappa(P) \leq C\xi(1 + \log(H/h))^2,$$

*with*

$$\xi := \max_{1 \leq i \leq N} \max \left\{ \frac{\gamma_i}{\beta_i}, \frac{H_i^2 \gamma_i}{a_i} \right\}$$

*and $a_i$, $\beta_i$, and $\gamma_i$ defined in (10.39) and (10.38).*

*Proof.* We employ the abstract Schwarz theory in Sect. 2.3; see Theorem 5.3 for a similar proof. A bound on $\rho(\mathcal{E})$ (and thus on the largest eigenvalue of $P$) can be found by using a standard coloring argument; cf. Sect. 2.5.1. Indeed, the supports of functions associated with the local subspaces (unions of pairs of substructures that have a common face) form an overlapping cover

of the domain and every point in the domain is covered by a finite, uniformly bounded number of such subregions.

For the smallest eigenvalue, we need to find a stable decompositions, as required by Assumption 2.2. We will work with discrete harmonic functions. Let $u \in Q$ and $\mathbf{u}$ be the corresponding discrete harmonic function. For every face $\mathcal{F}^i$, we fix a unit normal $\mathbf{n}_i$ to $\mathcal{F}^i$. We define:

$$\mathbf{u}_0 := \tilde{\Pi}^H_{RT}\mathbf{u},$$
$$\mathbf{u}_i := \mathcal{H}(\zeta_{\mathcal{F}^i}(u - u_0)),$$

where $u_0 \in Q_0$ is equal to $\mathbf{u}_0 \cdot \mathbf{n}_i$ on every face $\mathcal{F}^i$ and $\tilde{\Pi}^H_{RT}$ was defined in Sect. 10.2.1 as the discrete harmonic function that has the same normal component on the interface as the coarse interpolant. This splitting clearly gives a decomposition of $\mathbf{u}$, and thus of the corresponding traces a decomposition of $u$.

Bounds for $\mathbf{u}_0$ follow from Lemma 10.14 and the minimizing properties of the harmonic extension. We find

$$a(\mathbf{u}_0, \mathbf{u}_0) \leq C\xi(1 + \log(H/h))a(\mathbf{u}, \mathbf{u}). \tag{10.61}$$

Let now $\mathcal{F}^i$ be a face shared by two substructures $\Omega_j$ and $\Omega_l$. We assume that, e.g., $\mathbf{n}_i = \mathbf{n}_j = -\mathbf{n}_l$. We define the normal component by

$$u_i := \zeta_{\mathcal{F}^i}(u - u_0).$$

We note that, from the definition of the degrees of freedom on the coarse space $RT_0^H(\Omega)$, $u_i$ has vanishing mean value on $\mathcal{F}^i$, and thus on $\partial\Omega_j$ and $\partial\Omega_l$. Using the extension result in Lemma 10.19 and the decomposition result in Lemma 10.20, we find, for any $\mu_H \in Q_0$,

$$\begin{aligned} &a_j(\mathbf{u}_i, \mathbf{u}_i) \leq a_j(\widetilde{\mathcal{H}}_j u_i, \widetilde{\mathcal{H}}_j u_i) = \|B_j^{1/2}(\widetilde{\mathcal{H}}_j u_i)\|^2_{L^2(\Omega_j)} \\ &\leq C\gamma_j \|\zeta_{\mathcal{F}^i}(u - u_0)\|^2_{H^{-1/2}(\partial\Omega_j)} \leq C\gamma_j(1 + \log(H/h)) \\ &\times \Big((1 + \log(H/h))\|(u - u_0) + \mu_H\|^2_{H^{-1/2}(\partial\Omega_j)} + \|u - u_0\|^2_{H^{-1/2}(\partial\Omega_j)}\Big). \end{aligned}$$

For the second term on the right, we apply the trace estimate in (10.40) to the vector $\mathbf{u} - \Pi^H_{RT}\mathbf{u}$ and use Lemma 10.14. For the first term, it is enough to choose $\mu_H = -u_0$. We obtain:

$$a_j(\mathbf{u}_i, \mathbf{u}_i) \leq C\xi(1 + \log(H/h))^2 a(\mathbf{u}, \mathbf{u}). \tag{10.62}$$

A similar bound can be found for $a_l(\mathbf{u}_i, \mathbf{u}_i)$. Combining (10.61) and (10.62) and using Lemma 4.9 yields

$$s(R_0^T u_0, R_0^T u_0) + \sum_{\mathcal{F}^i} s(R^T_{\mathcal{F}^i} u_i, R^T_{\mathcal{F}^i} u_i) \leq C\xi(1 + \log(H/h))^2 s(u, u),$$

which concludes the proof. □

We end this section with some comments on the theoretical result of the previous lemma. As opposed to the case of conforming approximations in $H^1(\Omega)$ in three dimensions, a standard coarse space on a coarse triangulation ensures a logarithmic bound in $(H/h)$; cf. Lemmas 10.14 and 4.12. The bound obtained is independent of possibly large jumps of *both* coefficients $a$ and $B$ of the original problem. It only depends on their ratio and on the largest and smallest eigenvalue of the tensor $B$. Indeed the numerical results reported in [470] show that in practice no deterioration is observed when the ratio $B/a$ becomes small.

### 10.2.3 A Neumann-Neumann Method

We now consider a balancing Neumann-Neumann preconditioner similar to that in Sect. 6.2. The coarse space is now the standard coarse space $Q^H$ already employed in the previous section, while the local solves consist of Neumann problems on single substructures. Our hybrid preconditioned operator is defined in exactly the same way as for scalar and vector problems in $H^1(\Omega)$:

$$P_{hy1} = P_0 + (I - P_0)(\sum_{i=1}^{N} P_i)(I - P_0);$$

see (6.10). We refer to Sect. 2.5.2 for a general treatment of these hybrid operators and to section 6.2 for the precise definition of each component. Here, we use the same notations as in Sect. 6.2 to be able to employ the same definitions and results. The local spaces $W_i := Q_i$ are associated with single substructures and the coarse space $W_0 := Q_H$, built on the coarse mesh $\mathcal{T}_H$, is employed. The conforming space $Q$ is now denoted by $\widehat{W}$, while the product space $\prod_{i=1}^{N} W_i$ is denoted by $W$. In addition, we redefine $S : W \to W$, as the block diagonal matrix where each block corresponds to a substructure $\Omega_i$ and is equal to $S^{(i)}$. We note that $S$ is not singular in this case since the local bilinear forms are not. In $W$, we will employ the norm $|\cdot|_S$ and the scalar product $s(\cdot,\cdot)$ induced by $S$, while for the local spaces $W_i$, we will use $|\cdot|_{S^{(i)}}$ and the scalar product $s_i(\cdot,\cdot)$ induced by the local Schur complement $S^{(i)}$.

An exact solver is employed in $W_0$ and the projection $P_0$ is defined as in the previous section. Once a set of scaling functions $\{\delta_i^\dagger\}$ is introduced, the local solvers and the projection-like operators $P_i = R_i^T \tilde{P}_i$ are defined by (6.6) and (6.7), respectively. Here, $R_i^T$ is the extension by zero from $\partial\Omega_{i,h}$ to $\Gamma_h$.

We are left with the definition of the scaling functions $\{\delta_i^\dagger\}$. In this case, they are defined using only the coefficient $B$. In addition, since the degrees of freedom are associated only with the faces of $\mathcal{T}$, a degree of freedom in $\Gamma_h$ is shared by exactly two substructures. For a substructure $\Omega_i$ and for a face $\mathcal{F}^{ij}$ shared by $\Omega_i$ and $\Omega_j$, we define, for $\chi \geq 1/2$,

$$\delta_i^\dagger = \delta_{ij}^\dagger := \frac{\gamma_i^\chi}{\gamma_i^\chi + \gamma_j^\chi}, \quad \text{on } \mathcal{F}^{ij}. \tag{10.63}$$

These functions provide a partition of unity; see (6.3) and (6.11). This allows us to prove a lower bound for the operator $P_{hy1}$; see Assumption 2.2 and Lemma 2.12.

A bound for the condition number of $P_{hy1}$ was proven in [438] using the Schwarz theory in Sect. 2.5.2. Here, we follow Sect. 6.2 and give a different proof. For the upper bound, we employ the averaging operator $E_D : W \to \widehat{W}$ defined as in (6.4). Indeed, it is enough to estimate the norm of $E_D$; see Lemma 6.17. We need to modify our analysis slightly.

**Lemma 10.22** *There is a constant $C$, independent of the coefficients $a$ and $B$, $h$, and $H$, such that*

$$|(I - P_0)E_D w|_S^2 \leq C\xi(1 + \log(H/h))^2 |w|_S^2, \quad w \in W,$$

*with $\xi$ defined as in Theorem 10.21.*

*Proof.* We first note that

$$\gamma_i \delta_j^{\dagger 2} \leq \min(\gamma_i, \gamma_j); \tag{10.64}$$

see (6.19).

We set

$$\mathbf{w}_i := \mathcal{H} w_i, \qquad \mathbf{w} := \mathcal{H} w$$

and, for $i = 1, \ldots, N$,

$$v_i := (E_D w)_i, \qquad \mathbf{v} := \mathcal{H} v.$$

The representation formula (6.18) remains valid and is indeed simpler, since a fine face in $\Gamma_h$ belongs to exactly two substructures. For a face $\mathcal{F}^{ij} \subset \partial\Omega_i$ that is also shared by a second substructure $\Omega_j$, we have

$$v_i = \delta_{ij}^{\dagger} w_i + \delta_{ji}^{\dagger} w_j = v_j, \quad \text{on } \mathcal{F}^{ij}. \tag{10.65}$$

We note that $v \in \widehat{W}$.

We next introduce two coarse functions. Let $\mathbf{w}_0$ be the discrete harmonic vector obtained by first interpolating $\mathbf{w}$ on the coarse mesh on each subdomain:

$$(\mathbf{w}_0)_i := \tilde{\Pi}_{RT}^{H_i} \mathbf{w}_i, \qquad w_0 := \mathbf{w}_0 \cdot \mathbf{n},$$

and then

$$v_0 := E_D w_0, \quad \mathbf{v}_0 := \mathcal{H} v_0.$$

The local interpolants were defined in Sect. 10.2.1. The trace $w_0$ is not in general continuous across the subdomain boundaries, but $v_0$ is continuous and belongs to the coarse space $W_0$. It is easy to check that $\mathbf{v}_0$ coincides with the interpolant $\tilde{\Pi}_{RT}^{H} \mathbf{v}$ and thus the function

$$u := v - v_0 = E_D(w - w_0) \tag{10.66}$$

has vanishing mean value on each face of the decomposition:

$$\int_{\mathcal{F}^{ij}} u\,ds = \int_{\mathcal{F}^{ij}} u_{\mathcal{F}^{ij}}\,ds = 0. \tag{10.67}$$

We finally set $\mathbf{u} := \mathcal{H}u$.

Since

$$(I - P_0)v = (I - P_0)(v - v_0) = (I - P_0)u,$$

it is enough to find a bound for $u$ or, equivalently, for the corresponding discrete harmonic extension. The function $u$ has vanishing mean value on each face $\mathcal{F}^{ij} \subset \partial\Omega_i$ and thus on $\partial\Omega_i$. We can use Lemma 10.19 and write

$$|u_i|^2_{S^{(i)}} = a_i(\mathbf{u},\mathbf{u}) \leq \gamma_i \|u\|^2_{H^{-1/2}(\partial\Omega_i)} \leq C\gamma_i \sum_{\mathcal{F}^{ij}\subset\partial\Omega_i} \|u_{\mathcal{F}^{ij}}\|^2_{H^{-1/2}(\partial\Omega_i)}.$$

Using (10.65), (10.64), and (10.66), and Lemma 10.18, we can write

$$\begin{aligned}
\gamma_i \|u_{\mathcal{F}^{ij}}\|^2_{H^{-1/2}(\partial\Omega_i)} &\leq C\gamma_i \|\delta^{\dagger}_{ij}\,\zeta_{\mathcal{F}^{ij}}(w_i - w_0)\|^2_{H^{-1/2}(\partial\Omega_i)} \\
&\quad + C\gamma_i \|\delta^{\dagger}_{ji}\,\zeta_{\mathcal{F}^{ij}}(w_j - w_0)\|^2_{H^{-1/2}(\partial\Omega_i)} \\
&\leq C\gamma_i \|\zeta_{\mathcal{F}^{ij}}(w_i - w_0)\|^2_{H^{-1/2}(\partial\Omega_i)} \\
&\quad + C\gamma_j \|\zeta_{\mathcal{F}^{ij}}(w_j - w_0)\|^2_{H^{-1/2}(\partial\Omega_j)}.
\end{aligned}$$

Each of the two terms on the right hand side can be bounded using Lemma 10.20; the choice $\mu_H = w_0$ and a triangle inequality yield

$$\begin{aligned}
&\|\zeta_{\mathcal{F}^{ij}}(w_i - w_0)\|^2_{H^{-1/2}(\partial\Omega_i)} \\
&\leq C(1 + \log(H/h))\Big((1 + \log(H/h))\|w_i\|^2_{H^{-1/2}(\partial\Omega_i)} + \|w_0\|^2_{H^{-1/2}(\partial\Omega_i)}\Big),
\end{aligned}$$

and a similar estimate holds for the contribution on $\Omega_j$. Using the trace estimate in (10.40) and Lemma 10.14, we find

$$\begin{aligned}
|u_i|^2_{S^{(i)}} &\leq C\xi(1 + \log(H/h))^2(a_i(\mathbf{w},\mathbf{w}) + a_j(\mathbf{w},\mathbf{w})) \\
&\leq C\xi(1 + \log(H/h))^2(|w_i|^2_{S^{(i)}} + |w_j|^2_{S^{(j)}}).
\end{aligned}$$

Summing over the faces $\mathcal{F}^{ij} \subset \partial\Omega_i$ and then over the substructures finally yields

$$|(I - P_0)E_D w|^2_S = |(I - P_0)u|^2_S \leq |u|^2_S \leq C\xi(1 + \log(H/h))^2 |w|^2_S.$$

□

We are now able to prove the main result of this section.

**Theorem 10.23** *The condition number of the hybrid Schwarz operator $P_{hy1}$ satisfies*

$$\kappa(P_{hy1}) \leq C\xi(1 + \log(H/h))^2.$$

*Proof.* Let $u \in range(I - P_0)$. We have shown that it is enough to find a bound on the largest eigenvalue. As for scalar, piecewise linear finite element functions, the representation formula (6.15) remains valid if we define $w_i := {D^{(i)}}^{-1} R_i P_i u \in W_i$ with $D^{(i)}$ the diagonal matrix corresponding to the multiplication by $\delta_i^\dagger$. We can then write

$$|(I - P_0) \sum_{i=1}^{N} P_i u|_S^2 = |(I - P_0) E_D w|_S^2 \leq C\xi(1 + \log(H/h))^2 |w|_S^2.$$

Using (6.23), and the fact that $P_0$ is an orthogonal projection, we can then write

$$\begin{aligned} |(I - P_0) \textstyle\sum_{i=1}^{N} P_i u|_S^2 &\leq C\xi(1 + \log(H/h))^2 \langle \textstyle\sum_i P_i u, u\rangle_S \\ &= C\xi(1 + \log(H/h))^2 \langle (I - P_0) \textstyle\sum_i P_i u, u\rangle_S, \end{aligned}$$

and thus

$$|(I - P_0) \sum_{i=1}^{N} P_i u|_S^2 \leq C\xi(1 + \log(H/h))^2 |u|_S^2, \quad u \in range(I - P_0).$$

An upper bound for the largest eigenvalue of $P_{hy1}$ can then be found by using that $P_0$ is an orthogonal projection. □

A balancing Neumann-Neumann method with a minimal coarse space can also be devised in this case. We recall that for scalar diffusive problems and linear elasticity problems, we can associate with each subdomain a set of basis functions, by multiplying the kernel of the local problems by the partition of unity functions $\delta_i^\dagger$. For scalar diffusive problems, one can choose one constant function for each subdomain (see Sect. 6.2 and equation (6.5)), while for linear elasticity problems, the rigid body modes (three in two and six in three dimension) are employed (see Sect. 8.5).

Here, it is enough to choose one function for each substructure corresponding to the outward unit normal vector $\mathbf{n}_i$ to the boundary:

$$\widetilde{W}_0 = \{R_i^T n_i\}. \tag{10.68}$$

The new coarse space is strictly contained in the standard one previously defined. Numerical results in [441] show that the same bound holds as in Theorem 10.23. However, the analysis carried out here and in [438] does not seem to apply to this case in a straightforward way; see also the numerical results in the next subsection.

In addition, one-level FETI methods (see Sect. 6.3) can be defined and analyzed for Raviart-Thomas approximations. This was carried out in [443] and the same bound as in Theorem 10.23 was proven for the corresponding preconditioned operator. In particular, we note that the local Neumann problems are not singular and modifications in the algorithms and in the analysis are

needed; see [191, 443]. We refer to the next subsection for numerical results. Similar FETI preconditioners have also been applied to mortar approximations; see [393]. We finally mention that FETI-DP methods have also been proposed and studied for two-dimensional problems in [447, 448]; the primal constraints are averages over the edges and a coarse space problem of the same dimension and sparsity pattern as for the balancing Neumann-Neumann method in this section needs to be solved.

We note that, as for the face based method in the previous section, the bound obtained for the Neumann-Neumann preconditioner is also independent of possibly large jumps of *both* the coefficients $a$ and $B$ of the original problem, even though the scaling functions $\delta_i^\dagger$ are only constructed using values of $B$. Numerical results in [438] show that in practice no deterioration is observed when the ratio $B/a$ becomes small, at least for two-dimensional problems; cf. the numerical results in the next subsection.

### 10.2.4 Remarks on Two-Dimensional Problems and Numerical Results

The algorithms presented in Sect. 10.2.2 and 10.2.3 can be applied to two-dimensional Raviart-Thomas approximations in a straightforward way. In addition, since, in two dimensions, Nédélec finite element vectors can be obtained from those of Raviart-Thomas spaces by a simple rotation by ninety degrees, these algorithms can also be defined for problems in $H(\mathbf{curl}\,;\Omega)$ in two dimensions. Indeed, the two dimensional version of the face-based algorithm in Sect. 10.2.2 was originally studied in [449] using simpler techniques.

Here, we present some numerical tests for a balancing Neumann-Neumann and a one-level FETI methods for $H(\mathbf{curl}\,;\Omega)$ in two dimensions; cf. (10.4). The first corresponds to the coarse space in (10.68), while the second was studied already in [443]. The results presented here were originally given in [441].

The normal vectors $\mathbf{n}_i$ employed to define a balancing Neumann-Neumann method in $H(\mathrm{div}\,;\Omega)$ with the coarse space in (10.68), are now replaced by unit tangential vectors $\mathbf{t}_i$ on the subdomain boundaries; see [441]. Similarly, the same vectors (suitably scaled) are employed for the one-level FETI method in order to construct the coarse matrix $R$ in (6.28). Since the local Neumann problems are nonsingular, the modifications for FETI methods mentioned in Remark 6.7 need to be used. Finally, we note that the two algorithms have coarse problems of the same size; we have one coarse function for each substructure

We consider the domain $\Omega = (0,1)^2$ and two uniform triangulations $\mathcal{T}_h$ and $\mathcal{T}_H = \{\Omega_i\}$. The fine triangulation is made of triangles for the FETI method, and squares for the balancing method. It consists of $2 * n^2$ triangles and $n^2$ squares, respectively, with $h = 1/n$. Lowest order edge elements are considered. We note that, in contrast to the case of nodal elements, for a fixed value of $n$, triangular and rectangular meshes do not give rise to edge element

spaces of the same dimension. Nevertheless, the mesh size $h$ and the order of accuracy is the same and the comparison of the two methods is reasonably fair. The coarse triangulation consists of $nc^2$ squares which are unions of fine elements, with $H = 1/nc$. The substructures $\Omega_i$ are the elements of the coarse triangulation. Throughout, we use the value $\chi = 1/2$ (cf. (10.63)) and

$$B = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix}.$$

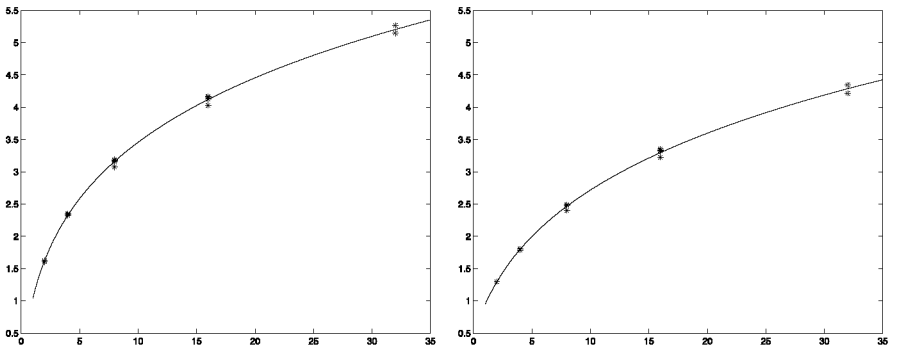

**Fig. 10.2.** Case with $a = 1$, $b = 1$, $n = 32$, 64, 128, 192, 256. Estimated condition number (asterisk) and least-square second order logarithmic polynomial (solid line), versus $\rho = H/h$ for the balancing NN method (*left*) and the one-level FETI method (*right*).

We first consider a case with constant coefficients and meshes with $n = 32$, 64, 128, 192, 256. Fig. 10.2 shows the estimated condition number (asterisks), for $a = b = 1$, as a function of $\rho = H/h$, for different values of $n$. For a fixed value of $\rho$, the condition number is quite insensitive to the dimension of the fine mesh. We have also plotted the best second order logarithmic polynomial least-square fits; our numerical results for both methods are consistent with the condition number bound in Theorem 10.23 and suggest that this bound is sharp. This bound was proved in [443] for the FETI method.

We then consider some cases where the coefficients have jumps. In Table 10.2, we show some results when the coefficient $b$ has jumps across the substructures. We consider the checkerboard distribution shown in Fig. 10.3, where $b$ is equal to $b_1$ in the shaded area and to $b_2$ elsewhere. For a fixed value of $n = 128$, $b_1 = 100$, and $a = 1$, the estimated condition number and the number of conjugate gradient iterations in order to obtain a reduction of the norm of the preconditioned residual by a factor $10^{-6}$, are shown as a function of $\rho = H/h$ and $b_2$. For $b_2 = 100$, the coefficient $b$ has a uniform distribution, and this corresponds to a minimum for the condition number and the number of iterations for both methods. When $b_2$ decreases or increases, the

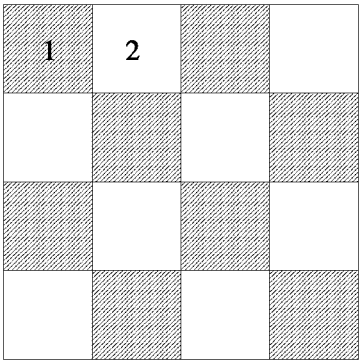


**Fig. 10.3.** Checkerboard distribution of the coefficients in the unit square.

condition number and the number of iterations also increase, but they can still be bounded independently of $b_2$. We observe that the two methods give comparable iteration counts.

**Table 10.2.** Checkerboard distribution for $b$: $(b_1, b_2)$. Estimated condition number and number of CG iterations to obtain a relative preconditioned residual less than $10^{-6}$ (in parentheses), versus $\rho = H/h$ (columns) and $b_2$ (rows), for the Balancing (*left*) and the FETI (*right*) methods. Case of $n = 128$, $a = 1$, and $b_1 = 100$.

| $b_2, \rho$ | 4 | 8 | 16 | 4 | 8 | 16 |
|---|---|---|---|---|---|---|
| $10^{-4}$ | 15.6 (22) | 13.4 (22) | 12.1 (22) | 4.12 (17) | 5.99 (22) | 8.42 (26) |
| $10^{-3}$ | 15.1 (21) | 13.2 (21) | 12.1 (23) | 4.09 (16) | 5.96 (20) | 8.37 (25) |
| $10^{-2}$ | 13.8 (20) | 12.5 (21) | 11.9 (23) | 4.04 (15) | 5.88 (19) | 8.25 (23) |
| $10^{-1}$ | 10.8 (19) | 10.8 (21) | 11.5 (22) | 3.88 (13) | 5.65 (17) | 7.91 (21) |
| 1 | 6.31 (17) | 7.55 (19) | 10.2 (21) | 3.44 (12) | 5.02 (15) | 6.99 (18) |
| 10 | 3.87 (13) | 5.41 (15) | 7.36 (18) | 2.56 (9) | 3.73 (12) | 5.16 (14) |
| $10^2$ | 2.33 (8) | 3.12 (10) | 3.87 (11) | 1.76 (7) | 2.41 (8) | 3.10 (10) |
| $10^3$ | 3.70 (12) | 4.77 (14) | 5.56 (16) | 2.51 (9) | 3.37 (11) | 3.99 (12) |
| $10^4$ | 3.96 (14) | 4.33 (14) | 4.64 (15) | 2.74 (10) | 3.09 (11) | 3.51 (11) |
| $10^5$ | 3.27 (12) | 3.55 (13) | 4.34 (14) | 2.20 (9) | 2.73 (10) | 3.35 (11) |
| $10^6$ | 2.99 (12) | 3.44 (13) | 4.28 (14) | 2.09 (9) | 2.65 (10) | 3.34 (12) |

In Table 10.3, we show some results when the coefficient $a$ has jumps. We consider the checkerboard distribution shown in Fig. 10.3, where $a$ is equal to $a_1$ in the shaded area and to $a_2$ elsewhere. For a fixed value of $n = 128$, $a_1 = 0.01$, and $b = 1$, the estimated condition number and the number of iterations are shown as a function of $\rho = H/h$ and $a_2$. We remark that for $a_2 = 0.01$, the coefficient $a$ has a uniform distribution. For both methods,

a slight increase in the number of iterations and the condition number is observed, when $a_2$ is decreased or increased and when $H/h$ is large.

**Table 10.3.** Checkerboard distribution for $a$: $(a_1, a_2)$. Estimated condition number and number of CG iterations to obtain a relative preconditioned residual less than $10^{-6}$ (in parentheses), versus $\rho = H/h$ (columns) and $a_2$ (rows), for the Balancing (*left*) and the FETI (*right*) methods. Case of $n = 128$, $b = 1$, and $a_1 = 0.01$.

| $a_2, \rho$ | 4 | 8 | 16 | 4 | 8 | 16 |
|---|---|---|---|---|---|---|
| $10^{-7}$ | 2.56 (10) | 4.33 (13) | 8.02 (17) | 2.80 (8) | 4.49 (12) | 7.29 (15) |
| $10^{-6}$ | 2.56 (10) | 4.32 (13) | 8.01 (17) | 2.41 (8) | 3.81 (11) | 6.21 (14) |
| $10^{-5}$ | 2.56 (10) | 4.30 (13) | 7.96 (17) | 1.82 (7) | 2.65 (9) | 4.05 (11) |
| $10^{-4}$ | 2.52 (9) | 4.13 (13) | 7.49 (16) | 1.79 (7) | 2.45 (8) | 3.23 (10) |
| $10^{-3}$ | 2.39 (9) | 3.51 (12) | 5.59 (14) | 1.78 (7) | 2.42 (8) | 3.07 (9) |
| $10^{-2}$ | 2.34 (9) | 3.16 (11) | 4.14 (13) | 1.76 (7) | 2.40 (8) | 3.25 (10) |
| $10^{-1}$ | 2.32 (8) | 3.12 (10) | 3.87 (12) | 1.77 (7) | 2.41 (8) | 3.10 (10) |
| 1 | 2.33 (8) | 3.12 (10) | 3.87 (11) | 1.77 (7) | 2.46 (8) | 3.26 (10) |
| 10 | 2.34 (8) | 3.16 (10) | 4.11 (11) | 1.77 (7) | 2.46 (8) | 3.26 (10) |
| $10^2$ | 2.34 (8) | 3.16 (10) | 4.14 (12) | 1.77 (7) | 2.46 (8) | 3.26 (10) |
| $10^3$ | 2.34 (8) | 3.17 (10) | 4.14 (12) | 1.77 (7) | 2.46 (8) | 3.26 (10) |

### 10.2.5 Iterative Substructuring for Nédélec Approximations in Three Dimensions

The iterative substructuring methods presented in the previous sections cannot be defined for Nédélec approximations in three dimensions.

The interpolant $\Pi^H_{ND}$ onto a Nédélec space on a coarse triangulation $\mathcal{T}_H$ does not exhibit the favorable bounds of $\Pi^H_{RT}$ in Lemma 10.14. Bounds at least as bad as those for the nodal interpolant onto finite element spaces of continuous functions are expected; see Lemma 4.12. Indeed, from the commuting-diagram property (B.23) and Lemma 4.12, we find

$$\begin{aligned}\|\Pi^H_{ND}(\mathbf{grad}\, u)\|^2_{L^2(K)} &= \|\mathbf{grad}\,(I^H u)\|^2_{L^2(K)} \\ &= |I^H u|^2_{H^1(K)} \leq C\frac{H}{h}\|\mathbf{grad}\, u\|^2_{L^2(K)},\end{aligned}$$

for $K \in \mathcal{T}_H$. Here, we have the same problem as for approximations on continuous nodal spaces in Sect. 5.4 and more exotic coarse spaces need to be employed. A first step in this direction was the work in [265] where a wire basket algorithm was proposed and studied. See also [266] for a generalization to a saddle-point problem. Local components are associated with the faces of the partition as in the methods in Sect. 5.4.2 and 10.2.2. Two wire basket spaces are then considered. The corresponding preconditioner has the form

$$B^{-1} = R^T_{\mathcal{W},1}\tilde{S}^{-1}_{\mathcal{W},1}R_{\mathcal{W},1} + R^T_{\mathcal{W},2}\tilde{S}^{-1}_{\mathcal{W},2}R_{\mathcal{W},2} + \sum_i R^T_{\mathcal{F}^i}S^{-1}_{\mathcal{F}^i\mathcal{F}^i}R_{\mathcal{F}^i},$$

As before, the local spaces consist of interface functions with vanishing degrees of freedom outside the faces $\mathcal{F}^i$. The first coarse space consists of the gradients of the wire basket functions in the space $\tilde{V}_{\mathcal{W}}^h$ defined in Sect. 5.4.2. The second coarse space consists of tangential vectors on $\Gamma$ that are nonvanishing only in a suitable neighborhood of the wire basket $\mathcal{W}$. Approximate solvers that are defined in terms of averages as in (5.7) are employed; we refer to [265] for details. Here we only note that the underlying idea is to employ two coarse spaces that should reduce two error components associated with the Helmholtz decomposition in (10.8). The corresponding preconditioned operator is shown to be scalable and to have a condition number that grows at most cubically in $\log(H/h)$. Independence of coefficient jumps is not guaranteed.

More recently, scalable and quasi-optimal FETI-DP algorithms, which are also robust with respect to coefficient jumps, have been found and analyzed for three-dimensional $h$ edge element approximations; see [442]. These methods ensure a logarithmic bound in $H/h$ on the condition number and employ a relatively small number of primal constraints. As noticed in [442], the difficulty of devising efficient and robust iterative substructuring methods for three-dimensional edge element approximations lies in the strong coupling between the degrees of freedom associated with the subdomain edges and faces. The situation is similar to that for $p$ and $hp$ approximations that rely on more general bases (see Sect. 7.5) and a local change of basis appears to be necessary. This observation may shed light to the reason why efficient and robust iterative substructuring algorithms for three dimensional problems have been so hard to devise for a long time. We note that, this difficulty was overcome in [265] by choosing a sufficiently large coarse space. We refer to [442] for additional details and for some discrete Sobolev type inequalities for the appropriate trace spaces.

We conclude this section by mentioning the work of Alonso and Valli who studied certain domain decomposition iterative methods for various types of three-dimensional electromagnetic problems; see, e.g., [16, 17] and the references therein.

# 11

# Indefinite and Nonsymmetric Problems

## 11.1 Introduction

In this chapter, we will present generalizations of some Schwarz preconditioners to nonsymmetric and indefinite problems. Such extensions were originally carried out in [99, 116, 117]. We also refer to [424, Sect. 5.4] for a fine presentation.

The Schwarz theory for nonsymmetric problems is relatively less satisfactory than that for positive definite symmetric problems, presented in Chap. 2 and 3. There are two main reasons. The first is related to the bounds for the convergence rates of GMRES which are far from sharp; see appendix C.6. In particular, we need to assume that the symmetric part of the preconditioned operator is positive definite. The second relies on the assumption that the indefiniteness and nonsymmetry of the differential operator is small with respect to the symmetric, positive definite part, typically a low-order (relatively compact) perturbation. The need of this assumption, for instance, prevents this theory from explaining the behavior of certain domain decomposition preconditioners for convection-dominated problems.

The result that we present here requires that the subdomains and the coarse elements are sufficiently small. This restriction appears natural for indefinite problems for which the coarse problem is well-posed only if the coarse mesh is fine enough. Indeed experience with domain decomposition preconditioners shows that such restriction is not required for nonsymmetric problems that are positive definite; see. e.g., [116] and Sect. 11.4. In addition, the performance of multiplicative methods are far better than that of additive preconditioners. This has no analog in the symmetric, positive definite case and a theory justifying such behavior is still lacking. In the last Sect. 11.5, we will give a brief overview of some important methods for nonsymmetric or indefinite problems, for which a comprehensive theory is mostly missing but which are often used in some large scale computations for the solution of large problems.

Given a bounded polygonal or polyhedral domain $\Omega \subset \mathbb{R}^n$, we will consider the homogeneous Dirichlet boundary value problem: find $u \in H_0^1(\Omega)$, such that

$$\begin{aligned} Lu &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{11.1}$$

The elliptic operator $L$ is defined as

$$Lu = -\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}(a_{ij}\frac{\partial u}{\partial x_j}) + 2\sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} + c\,u; \tag{11.2}$$

see (A.51) in Appendix A.7.

We assume that conditions (A.52) on the coefficients hold. In addition, the matrix $\{a_{ij}(x)\}$ is symmetric and uniformly positive definite for $x \in \Omega$. The right hand side $f \in L^2(\Omega)$; see Appendix A.7. We also assume that problem (11.1) has a unique solution; see Lemma A.46.

The weak form of equation (11.1) is: find $u \in H_0^1(\Omega)$ such that

$$b(u,v) = (f,v)_{L^2(\Omega)}, \quad v \in H_0^1(\Omega), \tag{11.3}$$

where the bilinear form $b(u,v)$ is defined in (A.54):

$$b(u,v) \;=\; \sum_{i,j=1}^{n} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx + \sum_{i=1}^{n} 2\int_\Omega b_i \frac{\partial u}{\partial x_i} v dx + \int_\Omega c\,uv dx.$$

We consider a shape-regular, not necessarily quasi uniform, triangulation $\mathcal{T} = \mathcal{T}_h$ of $\Omega$ and the space $V = V^h$ of piecewise linear functions on $\mathcal{T}$ that vanish on $\partial\Omega$; see appendix B.1. The Galerkin approximation of equation (11.3) is defined by: find $u \in V$ such that

$$b(u,v) \;=\; (f,v)_{L^2(\Omega)}, \quad v \in V. \tag{11.4}$$

If the mesh size $h$ is small enough, it follows from Theorem B.11 that this problem has a unique solution. By using nodal basis functions to span the finite element space, equation (11.4) is transformed into a linear system of algebraic equations

$$Bu = f, \tag{11.5}$$

which is large, sparse, nonsymmetric, possibly indefinite, and relatively ill-conditioned.

We also use two other bilinear forms

$$a(u,v) \;=\; \sum_{i,j=1}^{n} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx$$

and

$$s(u,v) \;=\; \sum_{i=1}^{n} \int_{\Omega} \left( b_i \frac{\partial u}{\partial x_i} v + \frac{\partial (b_i u)}{\partial x_i} v \right) dx \;=\; \sum_{i=1}^{n} \int_{\Omega} b_i \left( \frac{\partial u}{\partial x_i} v - \frac{\partial v}{\partial x_i} u \right) dx,$$

which correspond to the second order terms and the skew-symmetric part of $L$, respectively. The bilinear form $a(\cdot,\cdot)$ defines a norm, which we denote by $\|\cdot\|_a$. Under our assumptions on the coefficients $a_{ij}$, this norm is equivalent to the $H_0^1$-norm by Friedrichs inequality; see Appendix A.4. Let $A$ be the matrix representation of $a(\cdot,\cdot)$.

It is immediate that

$$s(u,v) \;=\; -s(v,u), \quad u,v \in H_0^1(\Omega).$$

We note that, by integration by parts, we can write

$$b(u,v) = a(u,v) + s(u,v) + \int_{\Omega} \tilde{c}\, uv \, dx,$$

with

$$\tilde{c} = c - \sum_{i=1}^{n} \frac{\partial b_i}{\partial x_i}.$$

We note that we make no assumptions on the sign of $\tilde{c}$ but only assume that the continuous and discrete problems are well-posed.

Using elementary, standard tools, it is easy to establish the following inequalities; see also Appendix A.7.

(i) Continuity: there exists a constant $C$, such that

$$| \, b(u,v) \, | \leq C \|u\|_a \|v\|_a, \quad u,v \in H_0^1(\Omega). \tag{11.6}$$

(ii) A Gårding inequality: there exists a constant $C$, such that

$$\|u\|_a^2 - C\|u\|_{L^2(\Omega)}^2 \leq b(u,u), \quad u \in H_0^1(\Omega) \; . \tag{11.7}$$

(iii) There exists a constant $C$, such that

$$\begin{aligned} &|s(u,v)| \leq C\|u\|_a \|v\|_{L^2(\Omega)}, \quad u,v \in H_0^1(\Omega), \\ &|s(u,v)| \leq C\|v\|_a \|u\|_{L^2(\Omega)}, \quad u,v \in H_0^1(\Omega), \\ &\left| \int_{\Omega} \tilde{c}\, uv \, dx \right| \leq C\|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}, \quad u,v \in L^2(\Omega). \end{aligned} \tag{11.8}$$

We note that the bounds for $b(\cdot,\cdot)$ and $s(\cdot,\cdot)$ are different, since each of the terms in $s(\cdot,\cdot)$ contains a factor which is of zero order. This enables us to control the skew-symmetric term and makes the analysis possible.

We also use the following regularity result which is valid for polygonal or polyhedral domains; see Lemma A.49.

(iv) The solution $w$ of the adjoint equation

$$b(\phi, w) \;=\; (g, \phi)_{L^2(\Omega)}, \quad \phi \in H_0^1(\Omega)$$

satisfies

$$\|w\|_{H^{1+\gamma}(\Omega)} \leq C\|g\|_{L^2(\Omega)}\ , \tag{11.9}$$

where $\gamma$ depends on the interior angles of $\partial\Omega$, is independent of $g$, and is greater than $1/2$.

To define the additive Schwarz algorithms, we need a partition into subdomains and a coarse triangulation; see Sect. 3.2 and 3.3. We first partition $\Omega$ into substructures $\{\Omega_i,\ i = 1, \cdots, N\}$, which are unions of fine elements of $\mathcal{T}$. Each $\Omega_i$ is assumed to be shape regular and has a diameter $H_i$; let $H$ be the maximum diameter of the substructures.

Following Sect. 3.3, we next introduce a shape-regular coarse mesh $\mathcal{T}_H$ of $\Omega$ and the finite element space $V_0 = V^H$ of continuous, piecewise linear functions on $\mathcal{T}_H$, vanishing on $\partial\Omega$. For $K \in \mathcal{T}_H$, let $H_K$ denote the diameter of $K$. The fine mesh $\mathcal{T}$ need not be a refinement of $\mathcal{T}_H$. Bilinear or trilinear finite element spaces could also be considered on quadrilateral or hexagonal coarse meshes. We define an interpolation operator

$$R_0^T : V_0 \longrightarrow V,$$

by interpolating the coarse functions onto the fine mesh. For $u \in V_0$, we define

$$R_0^T u = I^h u.$$

## 11.2 Algorithms on Overlapping Subregions

In this section, we introduce two variants of an additive Schwarz algorithm and provide bounds on their convergence rates; see Theorem 11.1 below. The theory that we present here was originally given in [99, 116]. Multiplicative methods can also be defined and analyzed and we refer to [117] for an abstract theorem resembling Theorem 2.9.

We consider the same overlapping partition $\{\Omega_i'\}$, introduced in Sect. 3.2, and obtained by extending the substructures $\{\Omega_i\}$. We assume that the overlap is small, i.e., that Assumption 3.1 holds. We also assume that a finite covering property holds (see Assumption 3.2) and that the subregions and the coarse elements are locally of comparable size (see Assumption 3.5). We will also use the notation $\Omega_0' = \Omega$ and refer to Chap. 3 for additional details.

As before, the local spaces are defined as finite element spaces consisting of functions that are piecewise linear on the local meshes:

$$V_i = \{u \in H_0^1(\Omega_i'), \quad u_{|_k} \in \mathbb{P}^1,\ k \in \mathcal{T}_i\}, \quad 1 \leq i \leq N.$$

We recall that these subspaces are contained in the original space $V$, in the sense that local functions extended by zero to all of $\Omega$ belong to $V$. Let

$$R_i^T : V_i \longrightarrow V, \quad 1 \leq i \leq N,$$

represent these zero extensions.

The space $V$ admits the following decomposition

$$V = R_0^T V_0 + \sum_{i=1}^{N} R_i^T V_i, \tag{11.10}$$

which is the same as that considered in Chap. 3.

As in the positive definite case (see Sect. 2.2), we can define projection-like operators, which will be the main building blocks of our algorithms. These operators map the finite element space $V$ into the subspaces $V_i$ and are defined in terms of local bilinear forms. For simplicity, we only consider the case of exact local bilinear forms defined, for $u_i,\ v_i \in V_i$, by

$$b_i(u_i, v_i) = b(R_i^T u_i, R_i^T v_i),$$
$$a_i(u_i, v_i) = a(R_i^T u_i, R_i^T v_i).$$

Let

$$B_i = R_i B R_i^T, \quad A_i = R_i A R_i^T$$

be the matrix representations of these local bilinear forms.

For $i = 0, \ldots, N$, we define $\tilde{P}_i : V \to V_i$, by

$$b_i(\tilde{P}_i u, v_i) = b(u, R_i^T v_i), \quad v_i \in V_i,$$

and $\tilde{Q}_i : V \to V_i$, by

$$a_i(\tilde{Q}_i u, v_i) = b(u, R_i^T v_i), \quad v_i \in V_i.$$

If $H$ is small enough, the operators $\tilde{P}_i$ are well-defined since $B_i$ is invertible; see Lemmas 11.3 and 11.4. For $\tilde{Q}_i$, no restriction on $H$ is necessary since $A_i$ is symmetric and positive definite. We now set

$$P_i = R_i^T \tilde{P}_i = R_i^T B_i^{-1} R_i B,$$
$$Q_i = R_i^T \tilde{Q}_i = R_i^T A_i^{-1} R_i B.$$

and introduce the additive operators

$$P^{(1)} = \textstyle\sum_{i=0}^{N} P_i = \left(\sum_{i=0}^{N} R_i^T B_i^{-1} R_i\right) B = (B^{(1)})^{-1} B,$$
$$P^{(2)} = P_0 + \textstyle\sum_{i=1}^{N} Q_i = \left(R_0^T B_0^{-1} R_0 + \sum_{i=1}^{N} R_i^T A_i^{-1} R_i\right) B = (B^{(2)})^{-1} B.$$

The only difference between $P^{(1)}$ and $P^{(2)}$ is that, for $i > 0$, we replace the projections $P_i$, corresponding to the $\Omega_i'$, by the $Q_i$. The coarse mesh projection is not changed.

We finally apply GMRES to the preconditioned systems

$$P^{(1)}u = g^{(1)}, \quad g^{(1)} = (B^{(1)})^{-1}f, \tag{11.11}$$

$$P^{(2)}u = g^{(2)}, \quad g^{(2)} = (B^{(2)})^{-1}f. \tag{11.12}$$

In order for equations (11.11) and (11.12) to have unique solutions, the operators $P^{(1)}$ and $P^{(2)}$ must be invertible. This follows from Theorem 11.1 given below.

The main result of this chapter is Theorem 11.1. Together with the error estimates for GMRES in Appendix C.6, it allows us to establish that the two algorithms converge at a rate which is independent of the mesh parameters $h$ and $H$, if the coarse mesh is fine enough.

**Theorem 11.1** *There exist constants $H_0 > 0$, $c(H_0) > 0$, and $C(H_0) > 0$, such that if $H \leq H_0$, then, for $i = 1, 2$ and $u \in V$,*

$$c(H_0)\, C_0^{-2} a(u,u) \leq a(u, P^{(i)}u)$$

*and*

$$a(P^{(i)}u, P^{(i)}u) \leq C(H_0)\, a(u,u),$$

*where the constant $C_0$ will be introduced in Lemma 11.2.*

The operator $P_0$ is very important, since it provides global transportation of information. All the other projections are local mappings. Without using $P_0$, information would travel only from one subregion to its neighbors in each iteration and it would take $O(1/H)$ iterations for the information to propagate across the region. Without such a global mechanism, it would also be impossible to confine the spectrum to the right half plane. To see this, we consider a symmetric, indefinite case. If the subregions are small enough, all the local elliptic problems are positive definite, symmetric and, in the absence of a global part, the preconditioner defined by the Schwarz algorithm is symmetric, positive definite. Therefore, by Sylvester's inertia theorem, cf. [428, Th. 6F], the additive operators $P^{(1)}$ and $P^{(2)}$ have as many negative eigenvalues as the original discrete elliptic problem.

The constant $H_0$ determines the minimal size of the coarse mesh problem and it depends on the operator $L$. In practice, this restriction appears necessary only for indefinite problems, but not for nonsymmetric, positive definite ones, and in general $H_0$ decreases as $\tilde{c}$ becomes more negative, while it increases if we increase the overlap. $H_0$ also depends on the shape of the domain $\Omega$. We refer to Sect. 11.4 for some numerical results. If the domain is not convex, the estimate of $H_0$, implicit in our proof of Lemma 11.6, depends on the parameter $\gamma$ in (11.9). We do not know of any explicit formula for $H_0$ but it is known from experience that it can be determined by numerical experiments. We refer to [116, 117] and [424] for further comments and details.

The proof of Theorem 11.1 is based on the following results. The first is the existence of a stable decomposition; see Lemma 3.12.

**Lemma 11.2** *There exists a constant $C_0$, which is independent of $h$ and $H$, such that, for all $u \in V$, there exist $u_i \in V_i$ with*

$$u = \sum_{i=0}^{N} R_i^T u_i$$

*and*

$$\sum_{i=0}^{N} a(R_i^T u_i, R_i^T u_i) = \sum_{i=0}^{N} a_i(u_i, u_i) \leq C_0^2 a(u,u), \qquad C_0^2 \leq \left(1 + \frac{H}{\delta}\right).$$

We note that this lemma is independent of the skew-symmetric and zero order terms of the elliptic operator since only $a(\cdot,\cdot)$ is involved.

**Lemma 11.3** *There exist constants $H_0 > 0$ and $c(H_0) > 0$, such that if $H \leq H_0$, then, for $u \in V$,*

$$\|P_0 u\|_a \leq c(H_0)\|u\|_a$$

*and*

$$\|P_0 u - u\|_{L^2(\Omega)} \leq C(H_0) H^\gamma \|P_0 u - u\|_a.$$

*Proof.* The first bound follows directly from Schatz's work, cf. Theorem B.11 and [414], by replacing the approximate solution by the coarse mesh solution and the exact solution of the continuous problem by the finite element solution in $V$.

In order to obtain a bound for the error $u - P_0 u$, we consider the auxiliary problem

$$L^* w = P_0 u - u \quad \text{in } \Omega, \quad w = 0 \quad \text{on } \partial\Omega,$$

where $L^*$ is the adjoint of $L$. We have for any $w_0 \in V_0$,

$$\begin{aligned} \|P_0 u - u\|^2_{L^2(\Omega)} &= (P_0 u - u, L^* w)_{L^2(\Omega)} = b(P_0 u - u, w) \\ &= b(P_0 u - u, w - w_0) \leq C \, \|P_0 u - u\|_a \|w - w_0\|_a. \end{aligned}$$

Since $P_0 u - u \in L^2(\Omega)$, then $w \in H^{1+\gamma}(\Omega)$ for an $\gamma > 1/2$ follows from the regularity result (11.9). The Sobolev embedding theorem in Lemma A.5 then implies that $w - w_0$ is continuous. The approximation estimates in Lemma B.6 yield the existence of $w_0 \in V_0$ such that

$$\|w - w_0\|_a \leq C \, H^\gamma \, \|w\|_{1+\gamma,\Omega}; \tag{11.13}$$

see, e.g., [414]. Therefore,

$$\|P_0 u - u\|^2_{L^2(\Omega)} \leq C \, H^\gamma \, \|P_0 u - u\|_a \|P_0 u - u\|_{L^2(\Omega)},$$

which gives the $L^2$-bound. □

We note that in order for the error estimate (11.13) to hold, it is necessary that the coarse space $V_0$ contains not only the constant functions but also the linear ones. This appears to be necessary only for the theory.

**Lemma 11.4** *The restriction of the quadratic form $b(\cdot,\cdot)$ to the subspaces $V_i$, $i > 0$, is strictly positive definite for $H$ sufficiently small, i.e., there exists a constant $c > 0$ such that*

$$c\, a(R_i^T u_i, R_i^T u_i) = c\, a_i(u_i, u_i) \leq b_i(u_i, u_i) = b(R_i^T u_i, R_i^T u_i), \quad u_i \in V_i.$$

*Proof.* We have to prove that the second order terms dominate the other symmetric term, since the contribution from the skew-symmetric term vanishes. This follows from the Friedrichs inequality for the region $\Omega_i'$ given in Sect. A.4:

$$b_i(u_i, u_i) = a_i(u_i, u_i) + \int_{\Omega_i'} \tilde{c}\, u_i^2\, dx \geq (1 - CH_i^2)\, a_i(u_i, u_i).$$

□

We note that it is not necessary to require the positive definiteness of the local bilinear forms, but that a suitable inf-sup condition, that implies invertibility, is sufficient.

**Lemma 11.5** *Let $v = \sum_{i=1}^N R_i^T v_i$ , where $v_i \in V_i$. Then there exists a constant $C > 0$, such that*

$$\|v\|_a^2 = \|\textstyle\sum_{i=1}^N R_i^T v_i\|_a^2 \leq C \sum_{i=1}^N \|R_i^T v_i\|_a^2$$
$$\|v\|_{L^2(\Omega)}^2 = \|\textstyle\sum_{i=1}^N R_i^T v_i\|_{L^2(\Omega)}^2 \leq C \sum_{i=1}^N \|R_i^T v_i\|_{L^2(\Omega)}^2.$$

*Proof.* The proof follows from Assumption 3.2, since for each $x \in \Omega$, the number of terms in the sum, which differ from zero, is uniformly bounded. □

**Lemma 11.6** *There exist constants $H_0 > 0$, $c(H_0) > 0$, and $C(H_0) > 0$, such that if $H \leq H_0$, then, for $u \in V$,*

$$c(H_0)\, C_0^{-2} a(u,u) \leq \sum_{i=0}^N a(P_i u, P_i u) \leq C(H_0)\, a(u,u)$$

*and*

$$c(H_0)\, C_0^{-2} a(u,u) \leq a(P_0 u, P_0 u) + \sum_{i=1}^N a(Q_i u, Q_i u) \leq C(H_0)\, a(u,u).$$

*Proof.* An upper bound for $a(P_0 u, P_0 u)$ is given in Lemma 11.3. To obtain an upper bound for the sum of the other terms, we use Lemma 11.4 and the definition of the $P_i$ to show that

$$c \sum_{i=1}^N a(P_i u, P_i u) \leq \sum_{i=1}^N b(P_i u, P_i u) = b(u, \sum_{i=1}^N P_i u).$$

The right hand side can be estimated by using inequality (11.6) and Lemma 11.5. The other upper bound is established in a similar way.

To prove the lower bounds, we begin by using Lemma 11.3 and obtain

$$\|u\|^2_{L^2(\Omega)} \leq 2\|u - P_0 u\|^2_{L^2(\Omega)} + 2\|P_0 u\|^2_{L^2(\Omega)} \leq C(H^{2\gamma} a(u,u) + \|P_0 u\|^2_{L^2(\Omega)}).$$

Since $P_0 u$ vanishes on $\partial\Omega$, we can use a Friedrichs inequality and Lemma 11.3 and replace the last term by $C\|P_0 u\|_a \|u\|_a$. By using Gårding's inequality, (11.7), it follows that

$$(1 - CH^{2\gamma}) a(u,u) \leq b(u,u) + C\|P_0 u\|_a \|u\|_a \ .$$

By the definition of the operators $P_i$ and Lemma 11.2, we find that

$$b(u,u) = \sum_{i=0}^{N} b(u, R_i^T u_i) = \sum_{i=0}^{N} b_i(\tilde{P}_i u, u_i) = \sum_{i=0}^{N} b(P_i u, R_i^T u_i).$$

The boundedness of $b(\cdot,\cdot)$, (11.6), can now be used to obtain

$$\sum_{i=0}^{N} b(P_i u, R_i^T u_i) \leq C \sum_{i=0}^{N} \|P_i u\|_a \|R_i^T u_i\|_a,$$

which by Lemma 11.2 and the Cauchy-Schwarz inequality can be bounded from above by

$$CC_0 \left( \sum_{i=0}^{N} \|P_i u\|_a^2 \right)^{1/2} \|u\|_a \ .$$

We finally obtain

$$a(u,u) \leq CC_0^2 \sum_{i=0}^{N} a(P_i u, P_i u),$$

for a sufficiently small $H$.

The proof of the other lower bound is quite similar. □

*Proof of Theorem 11.1:* The upper bounds on the norms of the operators follow immediately from Lemmas 11.5 and 11.6.

In order to obtain the lower bounds, we first consider

$$a(u, P^{(1)} u) = \sum_{i=0}^{N} a(u, P_i u) = \sum_{i=0}^{N} a(P_i u, P_i u) + \sum_{i=0}^{N} (a(u, P_i u) - a(P_i u, P_i u)).$$

Using Lemma 11.6, we see that it suffices to show that

$$|\sum_{i=0}^{N} (a(u - P_i u, P_i u))| \leq CHa(u,u). \tag{11.14}$$

The definition of the quadratic forms yields

$$a(u - P_i u, P_i u) = b(u - P_i u, P_i u) - s(u - P_i u, P_i u) - (\tilde{c}(u - P_i u), P_i u)_{L^2(\Omega)}. \tag{11.15}$$

By using the definition of $P_i$, the first term of the right hand side is seen to vanish. For $i = 0$, the second term can be bounded using inequality (11.8) and Lemma 11.3:

$$|s(u - P_0 u, P_0 u)| \leq C H^{2\gamma} a(u, u).$$

Since $s(P_i u, P_i u) = 0$, there remains to consider $s(u, \sum_1^N P_i u)$. Using inequality (11.8) and Lemma 11.5, we find

$$|\sum_{i=1}^{N} s(u - P_i u, P_i u)| \leq C\|u\|_a \, \|\sum_{i=1}^{N} P_i u\|_{L^2(\Omega)} \leq C\|u\|_a \, (\sum_{i=1}^{N} \|P_i u\|^2_{L^2(\Omega)})^{1/2}.$$

Using a Friedrichs inequality on each subdomain $\Omega_i'$ and Lemma 11.6, the desired inequality follows.

The third term in (11.15) is written as the difference of two expressions, which can be handled with exactly the same tools. This proves (11.14) and thus the lower bound for $P^{(1)}$. The estimate for the operator $P^{(2)}$ is obtained similarly. □

## 11.3 An Iterative Substructuring Method

We conclude our discussion by outlining how some other results, previously analyzed for the positive definite, symmetric case, can be extended to the class of elliptic problems 11.1. We confine our discussion to problems in the plane and only consider a simple iterative substructuring method as an example; see Remark 5.4.

Primal iterative substructuring methods for symmetric, positive definite problems are treated in Chap. 5. They are iterative methods for the solution of Schur complement equations, once the degrees of freedom internal to the substructures $\{\Omega_i\}$ have been eliminated. They can also be written as iterative methods for the original system (11.5).

We assume that the initial nonoverlapping partition coincides with the coarse mesh $\mathcal{T}_H$, i.e., the substructures $\{\Omega_i\}$ are the coarse elements. Given two subregions $\Omega_i$ and $\Omega_j$ that have a common edge, let $\mathcal{E}^{ij}$ be this edge. We use subspaces corresponding to the subregions $\Omega_{ij} = \Omega_i \bigcup \mathcal{E}^{ij} \bigcup \Omega_j$. These subregions play the same role as the $\Omega_i'$ in the overlapping method previously defined. We note that an interior substructure is covered by three such regions. The local subspaces are $V_{ij} = H_0^1(\Omega_{ij}) \cap V$. We employ the same coarse space $V_0 = V^H$ as before.

Compared with the case considered previously, we use less overlap in the sense that only the elements of $V^H$ can differ from zero at the vertices of the

substructures. This is reflected in a poorer bound for the constant of Lemma 11.2:

$$C_0^2 \leq C(1 + \log(H/h))^2;$$

see Remark 5.4 in Chap. 5. Lemma 11.2 is modified accordingly. The rest of the proof carries over without change. In Theorem 11.7, we use the notation $\tilde{P}_{ad} = \tilde{P}_0 + \sum \tilde{P}_{ij}$.

**Theorem 11.7** *For the iterative substructuring method, introduced as an additive Schwarz method with the subspaces $V_0$ and $V_{ij}$, there exist constants $H_0 > 0$, $c(H_0) > 0$, and $C(H_0) > 0$, such that if $H \leq H_0$,*

$$c(H_0)\,(1 + \log(H/h))^{-2} a(u,u) \leq a(u, \tilde{P}_{ad} u)$$

*and*

$$a(\tilde{P}_{ad} u, \tilde{P}_{ad} u) \leq C(H_0)\, a(u,u).$$

## 11.4 Numerical Results

In this section, we present some numerical results for some two-level overlapping and one iterative substructuring preconditioners, in order to show that in this case the rate of convergence often depends on the same geometric quantities as in the symmetric, positive definite case (i.e. the relative overlap $\delta/H$ or $log(H/h)$). The results given here are from Cai, Gropp, and Keyes [108]; courtesy of the authors; see Sect. 4.2-4.4 of their paper for nonsymmetric and/or indefinite problems. Results for symmetric, positive definite problems have already been given in Sect. 3.8 of this monograph.

We use the same two-dimensional discretization as in Sect. 3.8, together with the same subdomain partitions. We consider homogeneous Dirichlet problems as in (11.2) with the second order part of the operator discretized by a five-point central or upwind finite difference approximation. The right hand side $f$ is chosen such that $u = \exp(xy)\sin(\pi x)\sin(\pi y)$ is the solution.

We recall that the methods employed are MSR (Richardson's method without relaxation with a nonsymmetric, multiplicative, two-level overlapping preconditioner), MSM (GMRES with the same multiplicative preconditioner), and ASM (GMRES with the symmetric, additive, two-level overlapping preconditioner; cf. $P_{ad}^{(1)}$ in Sect. 11.2). In addition, we consider a hybrid variant of the iterative substructuring preconditioner in Sect. 11.3, denoted here by CGK; cf. Cai, Gropp, and Keyes [107, 108] for details. The subdomain partition coincides with the coarse mesh. A logarithmic growth in $H/h$ is expected for the residual reduction factor of GMRES. Results for an incomplete LU (ILU) factorization with different levels of fill in are also presented as a comparison.

### 11.4.1 A Nonsymmetric Problem

We first consider the nonsymmetric, constant coefficient problem corresponding to a uniform convection making a 45° angle with the coordinate axes:

$$Lu = -\Delta u + b\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y}.$$

We note that this problem is nonsymmetric but positive definite; cf. Lemma A.44 in appendix A.7. The first-order terms of the elliptic operator are discretized by two schemes, namely, the central difference method or the upwind-difference method.

Iteration counts are shown in Table 11.1, as functions of the overlap, for different partitions and different values of the convection strength.

The behavior of the preconditioners is different for central-difference and upwind-difference approximations. When using the former method, for a fixed fine mesh size $h^{-1} = 128$ and $b$ is increased beyond a certain size (near 10), all methods, except MSM with sufficient overlap, show a sharp upturn in the iteration count. The MSR fails to converge if $b$ is larger than this transitional $b$ for essentially all overlapping sizes. All other GMRES-based methods continue to converge but at a slower rate, especially the nonoverlapping method, which appears to have difficulty handling large convection terms. A better performance is observed for a fixed value of $b$ when the overlap is increased.

The situation changes when we switch to the upwind-difference method. For a fixed value of $b$, the iteration counts do decrease when the overlap is increased. However, they appear to be less sensitive to the overlap for larger values of $b$ than for central finite difference approximations. We note that even with modest overlap, the iteration counts are independent of $b$ for MSR, MSM, and ASM. The iteration counts of the iterative substructuring method increase considerably with $b$.

There seems to be a connection between the stability of the discretization and the convergence rate of overlapping domain decomposition preconditioners. The theory presented in this chapter very well predicts the behavior of algorithms with central difference discretizations. However, it offers little insight into the case of upwind finite differencing.

It is important to note that multiplicative preconditioners presented here show a much bigger improvement over additive methods than for symmetric, positive definite problems, when the same Krylov space method is employed; cf. Table 3.1 in Sect. 3.8. The poor performance of MSR may also partially derive from the fact that relaxation must be employed for the Richardson's method; it is however difficult to find an optimal value for the relaxation parameter in practice. Finally, the nonoverlapping method does not behave well if $b$ is large with either discretization scheme.

**Table 11.1.** Iteration count for solving the nonsymmetric model equation. The fine mesh size is uniform with $1/h = 128$. The number, such as $(2h)$, which appears next to the name of a method, indicates the overlap $\delta$ for overlapping algorithms and the level of fill in for ILU preconditioners.

| Methods | $H = 1/4$ | | | | | | $H = 1/8$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Central-difference Method** | | | | | | | | | | | | |
| $b =$ | 1 | 5 | 10 | 50 | 100 | 150 | 1 | 5 | 10 | 50 | 100 | 150 |
| MSR($h$) | 19 | 18 | 15 | $\infty$ | $\infty$ | $\infty$ | 10 | 10 | 10 | 13 | $\infty$ | $\infty$ |
| MSR($2h$) | 12 | 11 | 9 | 21 | $\infty$ | $\infty$ | 7 | 7 | 7 | 14 | $\infty$ | $\infty$ |
| MSR($4h$) | 7 | 8 | 8 | 22 | $\infty$ | $\infty$ | 6 | 6 | 6 | 10 | 35 | $\infty$ |
| MSR($8h$) | 6 | 7 | 8 | 24 | $\infty$ | $\infty$ | 5 | 5 | 5 | 7 | 10 | 21 |
| MSM($h$) | 7 | 7 | 7 | 10 | 10 | 9 | 5 | 5 | 5 | 8 | 10 | 12 |
| MSM($2h$) | 6 | 6 | 6 | 8 | 8 | 8 | 4 | 4 | 4 | 7 | 8 | 11 |
| MSM($4h$) | 5 | 5 | 6 | 7 | 7 | 7 | 4 | 4 | 4 | 5 | 7 | 9 |
| MSM($8h$) | 5 | 5 | 5 | 6 | 6 | 6 | 4 | 4 | 4 | 4 | 5 | 7 |
| ASM($h$) | 15 | 17 | 18 | 22 | 22 | 21 | 11 | 12 | 12 | 20 | 26 | 32 |
| ASM($2h$) | 13 | 15 | 15 | 20 | 20 | 21 | 10 | 10 | 11 | 18 | 23 | 27 |
| ASM($4h$) | 12 | 13 | 13 | 18 | 19 | 20 | 10 | 11 | 11 | 15 | 20 | 23 |
| ASM($8h$) | 11 | 12 | 12 | 16 | 17 | 17 | 10 | 11 | 12 | 14 | 16 | 19 |
| CGK | 13 | 14 | 16 | 28 | 35 | 47 | 11 | 12 | 13 | 26 | 36 | 50 |
| ILU(0) | 60 | 84 | 81 | 59 | 41 | 27 | | | | | | |
| ILU(1) | 38 | 53 | 51 | 34 | 22 | 15 | | | | | | |
| ILU(2) | 31 | 46 | 42 | 28 | 19 | 13 | | | | | | |
| **Upwind-difference Method** | | | | | | | | | | | | |
| $b =$ | 10 | 50 | 100 | 500 | $10^3$ | $10^4$ | 10 | 50 | 100 | 500 | $10^2$ | $10^4$ |
| MSR($h$) | 18 | 14 | 13 | 18 | 18 | 18 | 10 | 13 | 14 | 21 | 23 | 23 |
| MSR($2h$) | 14 | 14 | 15 | 16 | 16 | 16 | 10 | 13 | 16 | 17 | 16 | 15 |
| MSR($4h$) | 12 | 13 | 14 | 13 | 12 | 12 | 9 | 12 | 12 | 12 | 12 | 12 |
| MSR($8h$) | 10 | 10 | 10 | 11 | 11 | 11 | 8 | 9 | 9 | 8 | 8 | 8 |
| MSM($h$) | 9 | 9 | 8 | 7 | 7 | 7 | 7 | 9 | 9 | 10 | 11 | 11 |
| MSM($2h$) | 8 | 8 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 9 | 9 | 9 |
| MSM($4h$) | 7 | 7 | 6 | 6 | 6 | 6 | 7 | 7 | 6 | 6 | 6 | 6 |
| MSM($8h$) | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 |
| ASM($h$) | 19 | 20 | 19 | 18 | 17 | 17 | 14 | 19 | 21 | 22 | 22 | 23 |
| ASM($2h$) | 17 | 18 | 16 | 16 | 17 | 17 | 14 | 17 | 19 | 19 | 20 | 19 |
| ASM($4h$) | 15 | 16 | 16 | 16 | 16 | 16 | 14 | 15 | 16 | 17 | 17 | 18 |
| ASM($8h$) | 13 | 14 | 14 | 14 | 14 | 14 | 13 | 14 | 15 | 15 | 16 | 16 |
| CGK | 17 | 22 | 25 | 41 | 47 | 49 | 16 | 23 | 38 | 41 | 50 | 60 |
| ILU(0) | 82 | 61 | 50 | 23 | 16 | 6 | | | | | | |
| ILU(1) | 51 | 36 | 28 | 12 | 9 | 4 | | | | | | |
| ILU(2) | 42 | 30 | 24 | 11 | 8 | 4 | | | | | | |

### 11.4.2 The Helmholtz Equation

We consider then the Helmholtz equation with constant coefficients and homogeneous Dirichlet conditions:

$$Lu = -\triangle u - \omega^2 u.$$

This problem is self-adjoint, but indefinite. The eigenvalues of the continuous equation are

$$\frac{(i^2 + j^2)\pi^2}{H_\Omega^2} - \omega^2, \tag{11.16}$$

where $i$ and $j$ are positive integers and $H_\Omega$ is the side of $\Omega$. The number $\omega^2$ is chosen so as to avoid having eigenvalues in a small neighborhood of zero (and thus making the problem nearly singular), but there may be several negative eigenvalues.

Iteration counts are shown in Table 11.2 as functions of the overlap, for different partitions and different values of $\omega^2$.

For slightly indefinite (small $\omega^2$) problems, all methods exhibit the same behavior as when $\omega$ is zero (for which the problem is positive definite); convergence improves when the overlap is increased and iteration counts appear to be independent of the number of subdomains and the diameter of the fine mesh. However, when $\omega$ increases, iteration counts grow rapidly, *unless* a sufficiently fine coarse mesh is employed. This shows that a restriction on $H$ may be required in practice, as predicted by Theorem 11.1. A finer coarse mesh (more coarse-mesh points per wavelength) is needed for this problem in order to counteract the higher wavenumber. However, to the best of our knowledge, no quantitative theoretical results are available, to predict this threshold value of $H$.

The requirement for the coarse mesh to be sufficiently fine can be clearly seen for instance for MSM in the two entries with $\omega^2 = 300$ with an overlap of $4h$; with $H = 1/8$ more than 100 iterations are required for convergence, while 9 suffice with $H = 1/16$. We note that the first choice ensures about 3 coarse points per wavelength, while for the second, we have about 6.

As for the other test cases, MSM provides a performance superior to MSR and ASM. We finally note that increasing overlap seems to degrade convergence in the strongly indefinite case, whereas it always improves the convergence of definite operators. The iterative substructuring method appears to be sensitive to $\omega^2$ for $H$ large, but it performs fairly well when the coarse mesh is sufficiently fine.

### 11.4.3 A Variable-Coefficient, Nonsymmetric Indefinite Problem

Our last test problem has variable (oscillatory) coefficients and is nonsymmetric and indefinite:

**Table 11.2.** Iteration count for solving the Helmholtz equation. The fine mesh size is uniform with $1/h = 128$.

| $\omega^2 =$ | 0 | 30 | 70 | 110 | 150 | 300 | 0 | 30 | 70 | 110 | 150 | 300 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Methods | | | $H = 1/8$ | | | | | | $H = 1/16$ | | | |
| MSR($h$) | 10 | 11 | 20 | $\infty$ | $\infty$ | $\infty$ | 6 | 6 | 6 | 7 | 21 | $\infty$ |
| MSR($2h$) | 7 | 7 | 16 | $\infty$ | $\infty$ | $\infty$ | 5 | 5 | 5 | 5 | 14 | $\infty$ |
| MSR($4h$) | 6 | 6 | 14 | $\infty$ | $\infty$ | $\infty$ | 4 | 4 | 5 | 5 | 12 | $\infty$ |
| MSM($h$) | 5 | 5 | 7 | 9 | 13 | 35 | 3 | 4 | 4 | 4 | 6 | 8 |
| MSM($2h$) | 4 | 4 | 6 | 8 | 12 | 37 | 3 | 3 | 4 | 4 | 6 | 9 |
| MSM($4h$) | 4 | 4 | 5 | 8 | 13 | $> 100$ | 3 | 3 | 4 | 4 | 6 | 9 |
| ASM($h$) | 11 | 12 | 14 | 19 | 23 | 62 | 8 | 9 | 9 | 10 | 11 | 16 |
| ASM($2h$) | 10 | 10 | 14 | 18 | 23 | 61 | 8 | 8 | 9 | 10 | 10 | 16 |
| ASM($4h$) | 10 | 10 | 13 | 15 | 22 | 78 | 8 | 9 | 10 | 10 | 12 | 17 |
| CGK | 11 | 13 | 18 | 25 | 31 | 80 | 10 | 10 | 12 | 14 | 16 | 23 |

$$Lu = -\frac{\partial}{\partial x}\left(1 + \frac{1}{2}\sin(50\pi x)\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(1 + \frac{1}{2}\sin(50\pi x)\sin(50\pi y)\frac{\partial u}{\partial y}\right)$$
$$+20\sin(10\pi x)\cos(10\pi y)\frac{\partial u}{\partial x} - 20\cos(10\pi x)\sin(10\pi y)\frac{\partial u}{\partial y} - 70u.$$

The coefficients of the second-order terms oscillate but remain negative. The coefficients of the first-order terms physically represent a ten-by-ten array of closed convection cells, with no convective transport between cells. However, the subdomain boundaries do not in general align with the convection cell boundaries, so this zero-convective-flux property is not exploited. The operator $L$ is discretized by the five-point central difference method. A fixed overlapping factor of $\delta/H = 0.25$ in both $x$ and $y$ directions is employed in all overlapping methods.

Iteration counts are shown in Table 11.3 for different partitions and different fine meshes. This problem is difficult for all of the methods, but the iteration count for MSM is smaller than those of the others by almost a factor of 2, or more. MSR diverges in all cases. For a fixed coarse-mesh size $H$, some methods tend to require fewer iterations when the fine mesh is refined; others require more. This behavior is believed to be related to the oscillatory coefficients in the second-order terms of $L$. The discretization becomes more stable when $h$ gets smaller relative to the coefficient oscillation wavelength.

The nonoverlapping method CGK based solely on the diffusive terms of $L$, behaves reasonably well, probably because the magnitude of the convection is not large and averages to zero over the domain.

For this variable-coefficient problem, the domain decomposition preconditioners overwhelmingly outperformed the global ILU preconditioners for fine mesh sizes.

**Table 11.3.** Iteration counts for solving the variable-coefficient, nonsymmetric indefinite problem.

| | $H = 1/4$ | | | $H = 1/8$ | | | $H = 1/16$ | |
|---|---|---|---|---|---|---|---|---|
| $h =$ | 1/32 | 1/64 | 1/128 | 1/32 | 1/64 | 1/128 | 1/64 | 1/128 |
| MSR | $\infty$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ |
| MSM | 15 | 15 | 14 | 16 | 15 | 15 | 10 | 10 |
| ASM | 33 | 35 | 35 | 29 | 26 | 25 | 19 | 18 |
| CGK | 38 | 37 | 37 | 33 | 30 | 33 | 22 | 24 |
| | $H$=1 | | | | | | | |
| ILU(0) | 44 | 78 | 312 | | | | | |
| ILU(1) | 28 | 44 | 99 | | | | | |
| ILU(2) | 22 | 36 | 76 | | | | | |

## 11.5 Additional Topics

In this section, we provide some more details on certain domain decomposition algorithms for some important nonsymmetric or indefinite problems. We recall that the numerical results in the previous section showed that the abstract Schwarz theory is able to predict the performance of algorithms only in some situations, namely, normally when the indefinite part of the operator is a low-order, small perturbation of the second order part and when the indefiniteness is mild. There are however important practical cases where these assumptions do not hold. Already the numerical results highlighted this fact. Thus,

1. for convection-dominated problems (at least with convective fields with closed streamlines), the rate of convergence of some overlapping preconditioners appears to be independent of possibly strong convection strengths and might be quite insensitive to the overlap, if a stable discretization method is employed (cf. Table 11.1);
2. for highly indefinite problems, convergence may deteriorate when the overlap is increased (cf. Table 11.2);
3. multiplicative preconditioners can be dramatically better than additive ones for certain highly nonsymmetric or indefinite problems.

In the following subsections, we will briefly review some of the most popular domain decomposition preconditioners for some nonsymmetric or indefinite problems. A mathematical theory is mostly missing for these algorithms. We will not go into details but we will refer to appropriate references.

### 11.5.1 Convection-Diffusion Problems

Here, we consider the scalar convection-diffusion operator

$$Lu = -\nu\Delta u + 2\mathbf{b} \cdot \nabla u + cu,$$

with $\mathbf{b}$ a given vector field; cf. (11.2). We assume that the magnitude of $\mathbf{b}$ is much larger than the diffusion coefficient $\nu$, which, only for simplicity, is assumed to be constant. The ratio between $|\mathbf{b}|$ and $\nu$ is proportional to the Reynolds number.

We first recall that for moderate or high Reynolds numbers standard finite difference or finite element approximations need to be suitably modified; here, we only mention, among many others, the stabilized methods in [267, 207] and the discontinuous Galerkin methods in [276, 264]. We refer to, e.g., [391, ch. 8] and [140] for an introduction and additional references. In this review section, we will not go into details and will only consider standard finite element formulations, with the implicit understanding that in practice they must be replaced by a more appropriate approximation when the Reynolds number is moderate or large. While this does not present particular problems for stabilized methods that rely on continuous finite element spaces, the definition of the algorithms for discontinuous Galerkin approximations, which rely on discontinuous functions, is often far from trivial. In particular, we do not know that any iterative substructuring algorithm have been proposed, so far, for such discretizations.

We have already shown in Sect. 11.4 that overlapping preconditioners may perform quite well even for large Reynolds numbers, at least for some types of flows. Dirichlet problems for the overlapping subdomains appear to be appropriate in many cases. We refer to, e.g., Cai [99], Cai and Widlund [116, 117], Cai, Gropp, and Keyes [108], and Hebeker and Kuznetsov [254]. Different boundary conditions can also be employed for local problems in order to accelerate convergence; see, e.g., Garbey [216], Garbey and Kaper [217], for additional details and numerical tests. Overlapping methods for discontinuous Galerkin approximations of convection-diffusion problems have been considered by Lasser and Toselli [307, 306] and analyzed by generalizing the tools developed in this chapter for conforming approximations. We note that overlapping preconditioners for discontinuous Galerkin approximations of symmetric, positive definite problems were previously developed and analyzed by Feng and Karakashian [202]. Finally, we recall that restricted Schwarz algorithms have been employed for the solution of very large nonsymmetric problems; see Sect. 3.9.

Standard iterative substructuring methods do not appear to perform well if convection is strong; cf. Table 11.1. Dirichlet problems do not appear to be entirely appropriate for iterative substructuring algorithms and certain types of flows. We note that Neumann problems are not appropriate either; indeed, a Neumann problem on a substructure $\Omega_i$

$$\begin{aligned} Lu &= f && \text{in } \Omega_i, \\ -\nu \frac{\partial u}{\partial n} &= g && \text{on } \partial\Omega_i, \end{aligned} \tag{11.17}$$

corresponds to a variational formulation involving the bilinear form

$$
\begin{aligned}
b_i(u,v) &= \int_{\Omega_i} (\nu \nabla u \cdot \nabla v + \mathbf{b} \cdot \nabla u v + c\, uv)\, dx \\
&= \int_{\Omega_i} (\nu \nabla u \cdot \nabla v + (c - \nabla \cdot \mathbf{b}) uv)\, dx + \int_{\partial\Omega_i} \mathbf{b} \cdot \mathbf{n}\, uv\, dS \qquad (11.18) \\
&+ \int_{\Omega_i} (\mathbf{b} \cdot \nabla u v - \mathbf{b} \cdot \nabla v u)\, dx.
\end{aligned}
$$

We note that the symmetric contributions, corresponding to the second line, are all positive (semi)definite under the usual assumption that $c - \nabla \cdot \mathbf{b}$ is positive, except for the boundary term. Indeed, the sign of it depends on the angle between the outward normal and the flow field $\mathbf{b}$ and it can be negative unless Neumann conditions are imposed only on the outflow part of the boundary, where $\mathbf{b} \cdot \mathbf{n} > 0$. This shows that the local Neumann problems, required by some algorithms, might not be appropriate for convection-dominated problems; cf., e.g., the Dirichlet-Neumann algorithm in Sect. 1.3.3, or the Neumann-Neumann and FETI algorithms in Sect. 1.3.4 and and 1.3.5 and Chap. 6.

In view of these remarks, more general boundary conditions need to be considered in order to make local problems more stable or to accelerate convergence. We note that this is an idea that was already proposed in some very early works; see Hagstrom, Tewarson, and Jazcilevich [253], for a convection-diffusion problem, and Lions [321] for the Laplace equation. One of the motivations of this early work was the fact that the original Schwarz alternating method, which employs local Dirichlet problems, cannot be used for nonoverlapping partitions into subdomains.

In this section, we review some of the methods specifically designed for convection-diffusion problems. Others, which rely on optimized interface conditions, are mentioned in Sect. 11.5.3 below. A first class of methods rely on adaptively choosing the boundary conditions on the subdomain interfaces according to the sign of $\mathbf{b} \cdot \mathbf{n}$, in such a way that stable local problems are obtained. In particular, this gives rise to modifications of the Dirichlet-Neumann algorithm in Sect. 1.3.3, such as the adaptive Dirichlet-Neumann, Robin-Neumann, and $\beta$ Robin-Neumann methods. We refer to Carlenzoli and Quarteroni [125], Ciccoli [138], Trotta [451], Gastaldi, Gastaldi, and Quarteroni [218] as well as to Nataf and Rogier [358] and Auge, Lube, and Otto [27].

A second strategy is to ensure that the local problems are coercive independently of the sign of $\mathbf{b} \cdot \mathbf{n}$ on the subdomain interface. This gives rise to different modifications of the original Dirichlet-Neumann algorithm for symmetric, positive definite problems, such as the $\gamma$ Dirichlet-Robin and the $\gamma$ Robin-Robin algorithms; see Alonso, Trotta, and Valli [14]. We also refer to [392, Chap. 6] for a fine presentation of these two classes of methods.

A different approach, which has much in common with some of the previous works, is to generalize the Neumann-Neumann methods of Sect. 1.3.5 and Chap. 6, by suitably modifying the local Neumann problems in order to make

them positive definite. This, to the best of our knowledge, was originally proposed by Achdou and Nataf [7], and then extended by Achdou, Le Tallec, Nataf, and Vidrascu [2, 5]. Looking at the bilinear form in (11.18), we see that the possible indefiniteness and instability arises from the boundary term. We then form a new coercive bilinear form by subtracting the offending term:

$$\tilde{b}_i(u,v) = \int_{\Omega_i} (\nu \nabla u \cdot \nabla v + \mathbf{b} \cdot \nabla uv + c\,uv)\,dx - \int_{\partial\Omega_i} \mathbf{b} \cdot \mathbf{n}\,uv\,dS. \quad (11.19)$$

This new bilinear form corresponds to the *Robin* problem

$$\begin{aligned} Lu &= f && \text{in } \Omega_i, \\ -\nu \frac{\partial u}{\partial n} + \mathbf{b} \cdot \mathbf{n}\,u &= g && \text{on } \partial\Omega_i. \end{aligned} \quad (11.20)$$

We note that similar Robin conditions had already been employed; see [392, Chap. 6] and the references therein.

With this new bilinear form, we can now define a variant of the balancing Neumann-Neumann algorithm in exactly the same way as in Sect. 6.2; see in particular the hybrid operator in (6.10) and Sect. 6.2.2. The only difference is in the definition of the local components of the preconditioner, for which the pseudoinverses $S^{(i)\dagger}$, the application of which amounts to solving local Neumann problems of the form (11.17), are replaced by solving Robin problems as in (11.20). We refer to [2, 5] for details and some numerical results. Since this method generalises the Neumann-Neumann algorithm for symmetric problems, it is called the *Robin-Robin* algorithm. However, we note that it is a different algorithm than the $\gamma$ Robin-Robin method of Alonso, Trotta, and Valli [14]. For the case of two nonoverlapping subdomains, the latter involves the solution of two local problems with two *different* types of Robin conditions in the spirit of the Dirichlet-Neumann algorithm in Sect. 1.3.3, where different local problems are solved in each step, while the former involves the solution of four local problems; two Dirichlet problems followed by two others with the *same* type of Robin conditions in the spirit of the Neumann-Neumann algorithm in Sect. 1.3.4.

The same type of coarse space as in (6.5) can be employed. We recall that it is constructed by taking one constant function for a floating subdomain and multiplying it by the scaling function $\delta_i^T$. A more effective coarse space is proposed in [2, 5]:

$$W_0 = \text{span}\,\{R_i^T(\delta_i^\dagger v_i)\},$$

with $v_i$ the solution of the adjoint of the Robin problem (11.20) with $g = 0$ and $f = 1$.

In the same spirit as the Robin-Robin method, a one-level FETI algorithm has been developed by one of the authors in [440]. The point here is that the discrete convection-diffusion problem can be written as a FETI saddle-point system by working in a space of functions that are discontinuous across the subdomain boundaries; cf. (6.27) in Sect. 6.3.1. Here $S$ is a block diagonal

matrix where each block consists of the local Schur complement $S^{(i)}$ for the local *Robin* problem (11.20). When subassembling the local contributions, the Robin terms involving $\mathbf{b} \cdot \mathbf{n} u$ cancel at the subdomain interfaces. A FETI preconditioner can then be defined exactly as in Sect. 6.3.2; cf. (6.37). The coarse matrix $R$ in (6.28) employs the local functions $v_i$ of the Robin-Robin method: $R^{(i)} = v_i$. Different choices for the scaling matrix $Q$, needed for the construction of the projection $P$, have also been proposed and tested. We refer to [440], for details, numerical tests on various types of flows $\mathbf{b}$, and comparisons with the Robin-Robin method. Since the local bilinear forms employed for the construction of the preconditioners are positive definite, modifications mentioned in Remark 6.7 need to be employed.

### 11.5.2 The Helmholtz Equation

We now consider the scalar model problem

$$\begin{aligned} Lu := -\triangle u - \omega^2 u &= f \qquad \text{in } \Omega, \\ \frac{\partial u}{\partial n} + i\omega u &= 0 \qquad \text{on } \partial\Omega \end{aligned} \tag{11.21}$$

in more detail We note that this problem is well-posed for every $\omega$ because of the Sommerfeld boundary condition, which involves an imaginary coefficient $i\omega$; see, e.g., [268, 269]. The problem remains well posed if we replace $i\omega$ by $-i\omega$. These boundary conditions are first order radiation conditions and are often replaced by higher order or nonlocal ones. Additional Dirichlet or Neumann conditions are usually considered on the boundary of scatterers in practical problems.

Helmholtz type problems are particularly difficult to precondition effectively for two main reasons:

1. in order to be effective, coarse solvers usually require a rather fine coarse mesh thus leading to potentially large coarse problems;
2. in case local Dirichlet or Neumann problems are employed on the subdomains, they may be close to singular and are not always appropriate for domain decomposition preconditioners.

The first restriction can already be predicted by the analysis in this chapter and was seen in the numerical results in Sect. 11.4.2; cf. Table 11.2. Indeed, it is reasonable to assume that the coarse problem should be well-posed and this is ensured only if the coarse mesh is sufficiently fine; cf., e.g., [268, 269]. Concerning the second restriction, we note that $\omega^2$ may be close to one of the eigenvalues of the Laplace operator with Dirichlet or Neumann boundary conditions; see (11.16) for the case of a square subdomain and a Dirichlet condition. We note that these eigenvalues scale like $1/H_i^2$, with $H_i$ the typical size of the subdomain and, even though the local problems can be made positive definite by making the subdomain small enough (cf. Lemma 11.4), this is often impractical. For more general subdomains a formula for the eigenvalues

is not available and it is also impractical to determine them numerically. We finally note that in practice it is very unlikely that a local problem is singular; however, local problems that are close to singular may give rise to instabilities of domain decomposition algorithms and therefore slow the convergence.

For these reasons, much of the effort has gone into making local problems more stable, in a spirit similar as for the convection-diffusion problems of the previous section. One of the first important contributions was given by Després in his Ph.D. thesis [157]. We consider for simplicity the case of two nonoverlapping subdomains, $\Omega_1$ and $\Omega_2$, with an interface $\Gamma$, and modify the iterative method of Sect. 1.3.3; cf. Fig. 1.1. In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$
(H_1)\begin{cases} -\Delta u_1^{n+1/2} - \omega^2 u_1^{n+1/2} = f & \text{in } \Omega_1, \\ \dfrac{\partial u_1^{n+1/2}}{\partial n_1} + i\omega u_1^{n+1/2} = 0 & \text{on } \partial\Omega_1 \setminus \Gamma, \\ \dfrac{\partial u_1^{n+1/2}}{\partial n_1} + i\omega u_1^{n+1/2} = -\dfrac{\partial u_2^n}{\partial n_2} + i\omega u_2^n & \text{on } \Gamma, \end{cases}
$$

$$
(H_2)\begin{cases} -\Delta u_2^{n+1} - \omega^2 u_2^{n+1} = f & \text{in } \Omega_2, \\ \dfrac{\partial u_2^{n+1}}{\partial n_2} + i\omega u_2^{n+1} = 0 & \text{on } \partial\Omega_2 \setminus \Gamma, \\ \dfrac{\partial u_2^{n+1}}{\partial n_2} + i\omega u_2^{n+1} = -\dfrac{\partial u_1^{n+1/2}}{\partial n_1} + i\omega u_1^{n+1/2} & \text{on } \Gamma, \end{cases} \tag{11.22}
$$

Some remarks are necessary. First we note that on each subdomain, Helmholtz problems with boundary conditions of the form

$$
\frac{\partial u}{\partial n} + i\omega u = g
$$

are solved and they are therefore well-posed. In addition, in each fractional step different fluxes are matched on the interface $\Gamma$. The algorithms can therefore be regarded as a generalization of the Dirichlet-Neumann algorithm of Sect. 1.3.3. An additive version can also be considered. We refer to [157, 159, 158, 40, 41, 42] and [141, 142] for additional comments, extensions, and some theoretical results. Here we only mention that in the case of one dimensional problems, for which the transmission condition in (11.22) is exact, and $N$ subdomains, the algorithm for the continuous problem converges in a number of iterations that is at most $N$.

Overlapping methods that employ transmission conditions on the subdomain boundaries have also been considered for Helmholtz problems. Here, we mention [103] and [350]. We also note that perfectly matched layers have been considered for the construction of stable local problems in [437]. We also

mention the two level method in [455], which employs an algebraic multigrid technique for the construction of a coarse problem.

We end this section with the important FETI-H method, a generalization of the FETI method for Helmholtz type problems. This method was, to the best of our knowledge, originally presented in [196, 197]. The starting point is always an equivalent formulation of the discrete Helmholtz problem as a saddle point problem with continuity constraints across the interfaces of nonoverlapping subdomains; cf. (6.27) in Sect. 6.3.1. Here we do not eliminate degrees of freedom in the interior of the substructures (which would amount to solving potentially almost singular Dirichlet problems) but work with all degrees of freedom:

$$\left.\begin{aligned} Au + B^T\lambda &= f, \\ Bu &= 0. \end{aligned}\right\}. \tag{11.23}$$

where

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{pmatrix}, \text{ and } A = \begin{pmatrix} A^{(1)} & O & \cdots & O \\ O & A^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & A^{(N)} \end{pmatrix}.$$

The local problems $A^{(i)}$ are chosen in such a way that they are invertible by employing suitable radiation or transparency conditions for the construction of the local problems $A^{(i)}$. We must ensure, in particular:

1. that the local problems are invertible (cf. the boundary conditions in (11.21));
2. that after subassembling, the original discrete Helmholtz problem be recovered (and consequently any additional terms in the local bilinear forms on the internal subdomain boundaries must cancel).

We also recall that a Sommerfeld condition as in (11.21) leads to a well-posed problem with either a plus or a minus sign employed on the subdomain boundary, but not necessarily with different signs on different parts of the boundary of the same domain.

The discretization of the following local problems are therefore employed for the local matrices $A^{(i)}$:

$$\begin{aligned} -\Delta u - \omega^2 u &= f && \text{in } \Omega_i, \\ \frac{\partial u}{\partial n} + i\,\varepsilon_i\,\omega\,u &= g && \text{on } \partial\Omega_i. \end{aligned}$$

We assume that on $\partial\Omega_i \cap \partial\Omega$ the same boundary condition as for the original problem is employed. The coefficients $\varepsilon_i$ is either $-1$ or 1 and must be chosen in such a way that, if two subdomains $\Omega_i$ and $\Omega_j$ have a common interface, $\varepsilon_i = -\varepsilon_j$ on that interface. With this restriction however, it is not possible to choose $\varepsilon_i$ strictly positive (or strictly negative) on the whole of $\partial\Omega_i$ for

a general subdomain partition. If a sign conflict cannot be resolved on the interface between two subdomains, then $\varepsilon_i = \varepsilon_j = 0$ must be employed. In this way, it is possible to choose $\varepsilon_i \geq 0$ (or $\varepsilon_i \leq 0$) in such a way that it does not vanish on at least a subset of $\partial\Omega_i$ with nonvanishing measure. We refer to [196, Fig. 2] for an illustration.

Once we have ensured that the local matrices are invertible, we can eliminate the primal variable $u$ and obtain an equation for the Lagrange multipliers:

$$F\lambda = d. \tag{11.24}$$

The surprising fact is that when a Krylov type method is employed for the system (11.24) without any preconditioner, the rate of convergence is independent of the diameter of the fine mesh $h$; see [196, Sect. 3.2]. However, it depends on the number of subdomains and the frequency $\omega$. A suitable projection $P$ onto a coarse space is therefore employed:

$$P^T F P\lambda = P^T d, \qquad \lambda = \lambda_0 + P\lambda, \quad \lambda_0 = (I - P)\lambda.$$

The coarse space is spanned by the columns of a matrix $Q$. We have

$$P = I - Q(Q^T F Q)^{-1} Q^T F, \qquad \lambda_0 = Q(Q^T F Q)^{-1} Q^T d,$$

and the application of $P$ and $P^T$ requires the solution of a coarse problem. Instead of using simple coarse functions built from constant functions on the subdomains, as in the original one-level FETI method, plane waves are employed, which help reduce the size of the coarse space; if $N_\theta$ is a number of angular directions in the plane, the column $Q^{(j)}$ of $Q$ corresponds to the function

$$\exp(i\omega(x \cos\theta_j + y \sin\theta_j)), \qquad \theta_j = 2\pi\,(j-1)/N_\theta, \quad 1 \leq \theta \leq N_\theta.$$

The generalization to the three dimensional case is straightforward.

When the projected operator $P^T F P$ is employed without any additional preconditioner, the rate of convergence is found to be independent of $h$, the number of subdomains, and the wave number $\omega$, in many practical problems. We also refer to [196] for details, additional comments, and numerical tests. We also note that the choice of the Krylov type method appears to play an important role; see [196, 327]. FETI-H algorithms have been employed for the solution of very large problems; see, e.g., [196, 197, 228, 256]. We also refer to [152] for an earlier FETI method which employs two sets of Lagrange multipliers.

### 11.5.3 Optimized Interface Conditions

The basic idea of optimized interface conditions is to solve an optimization problem for the interface condition in order to maximize the rate of convergence of a domain decomposition iteration. To the best of our knowledge, this

was first proposed by Tan and Borsboom [434]. The connection that was later made between interface conditions in domain decomposition and absorbing boundary conditions for propagation problems made the optimization problem tractable and solvable. A typical approach consists in choosing higher order boundary conditions involving a certain number of coefficients, which are then determined by solving an optimization problem. The parameters need to be computed at the interface nodes. The optimization procedure can be considered either at the continuous or at the discrete level.

One of the first applications in this direction was to convection-diffusion problems and is due to Japhet in her Ph.D. thesis [271] and to Japhet, Nataf, and Rogier [273]; see also [272, 274].

Other applications of optimized interface conditions have been made for the Laplace equation by Engquist and Zhao [187], for Helmholtz problems by Chevalier and Nataf [135], Gander [212], Gander, Magoulès, and Nataf [215], for flows in porous media by Willien, Faille, Nataf, and Schneider [468], for Navier-Stokes equations by Lube, Müller, and Müller [322], for sedimentary basin modeling by Faille, Flauraud, Nataf, Schneider, Willien [190], for evolution problems by Gander and Halpern [214], and for systems of equations by Dolean, Lanteri, and Nataf [166]. See also Japhet, Nataf, and Roux [274], Genseberger [219], and Gander and Golub [213], for additional references.

### 11.5.4 Nonlinear and Eigenvalue Problems

We end this section by briefly mentioning some applications of domain decomposition preconditioners to the solution of some nonlinear problems. We stress that the references given here are only a small part of the work that is available; the interested reader should look at the proceedings of the annual domain decomposition conferences for a broad overview of applications; see, e.g, [56, 338, 304, 130, 155, 255, 297]. One of the first applications to nonlinear problems was carried out in [253].

The solution of an algebraic nonlinear problem arising from the discretization of a nonlinear partial differential equation using a Newton method requires the solution of a linear system with a Jacobian matrix in each step. In order to do so, a Krylov type iterative method can be employed together with a domain decomposition (Schwarz) preconditioner. The overall procedures are commonly referred to as *Newton-Krylov-Schwarz* algorithms and have now been used quite extensively for the solution of large scale problems. Here, we mention applications to compressible flows by Cai, Gropp, Keyes, Melvin, and Young [109], Kaushik, Keyes, and Smith [277], Anderson, Gropp, Kaushik, Keyes, and Smith [19], Gropp, Kaushik, Keyes, and Smith [239], Sala [408], to time-dependent problems by Cai, Keyes, and Venkatakrishnan [113], and to some reacting flow problems by McHugh, Knoll, and Keyes [349]; see also Lee and Yu [311].

The general framework of Newton-Krylov-Schwarz algorithms was first introduced by Cai, Gropp, Keyes, and Tidriri [110]. Some theory was developed by Cai and Dryja [105], and Dryja and Hackbusch [175].

Nonlinear preconditioning strategies have also been proposed more recently in the domain decomposition context by Cai and Keyes [111], followed by Cai, Keyes, and Marcinkowski [112], and Cai, Keyes, and Young [114]. We note that the idea of nonlinear preconditioning had been already employed in the multigrid context; see, e.g., [82, 251].

Contact problems have also been solved efficiently using domain decomposition algorithms. In particular, the inequality constraints that appear in these problems can be dealt with using modifications of FETI methods. Here, we only mention Dostál, Friedlander, and Santos [167], Dostál, Gomes Neto, and Santos [169, 168], and Dostál and Horák [170]. Different domain decomposition preconditioners for contact problems have been considered by Barboteu, Alart, and Vidrascu [36], and Alart, Barboteu, Le Tallec, and Vidrascu [13].

We also mention Zou and Huang [482], Xu [473], Tai [431], Tai and Espedal [432], Barth, Chan, and Tang [37], Abdoulaev, Achdou, Hontand, Kuznetsov, Pironneau, and Prud'homme [1], Palansuriya, Lai, Ierotheou, and Pericleous [370], Tai and Xu [433], and Lui [323, 325, 324] for some other results or applications for nonlinear problems.

Many problems in structural mechanics require the solution of large eigenvalue problems. Domain decomposition algorithms, for an underlying positive definite problem, are often combined with a Lanczos or block Lanczos method. There are other alternatives such as the method of component mode synthesis, which has been analyzed systematically by Bourquin; see, e.g., [66, 67]. An even more accurate method is due to Bennighof and Lehoucq [43].

# A

# Elliptic Problems and Sobolev Spaces

In this appendix, we present, mostly without proof, some standard results about elliptic problems and Sobolev spaces. We only discuss results that are employed in this monograph and we refer to [361, 317, 8] for more rigorous and thorough treatments of these topics. We also refer to [391, Ch. 1] and [48, Ch. 1] for concise introductions. For simplicity, we will only present definitions for real-valued functions; the definitions can easily be extended to the complex-valued ones.

## A.1 Sobolev Spaces

Here and in the following, we assume that $\Omega \subset \mathbb{R}^n$ is a Lipschitz domain, i.e., a bounded open set with Lipschitz continuous boundary; see, e.g., [317, 8].

**Definition A.1.** *The boundary $\partial\Omega$ is Lipschitz continuous if there exists a finite number of open sets $\mathcal{O}_i$, $i = 1, \ldots, m$, that cover $\partial\Omega$, such that, for every $i$, the intersection $\partial\Omega \cap \mathcal{O}_i$ is the graph of a Lipschitz continuous function and $\Omega \cap \mathcal{O}_i$ lies on one side of this graph.*

We recall that for a Lipschitz domain it is possible to define a unit outward normal $\mathbf{n}$ almost everywhere on $\partial\Omega$.

The space of square-summable functions on $\Omega$ is defined as

$$L^2(\Omega) = \left\{ u : \Omega \mapsto \mathbb{R} \,|\quad \int_\Omega |u|^2 \, dx < \infty \right\}.$$

It is a Hilbert space with the scalar product

$$(u, v)_{L^2(\Omega)} = \int_\Omega u \, v \, dx$$

and an induced norm given by

$$\|u\|^2_{L^2(\Omega)} = (u,u)_{L^2(\Omega)} = \int_\Omega |u|^2\,dx.$$

We introduce the space $\mathcal{D}(\Omega) = C_0^\infty(\Omega)$ consisting of functions in $C^\infty(\Omega)$ with compact support in $\Omega$. Given a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $|\alpha| = \alpha_1 + \ldots + \alpha_n$, we define the derivative

$$D^\alpha u = \partial^{\alpha_1}_{x_1} \ldots \partial^{\alpha_n}_{x_n} u := \frac{\partial^{|\alpha|} u}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n}.$$

If $\alpha = (0, \ldots, 0)$, we mean $D^\alpha u = u$.

We define a pseudo-topology in $\mathcal{D}(\Omega)$ and say that a sequence $\{u_n\} \subset \mathcal{D}(\Omega)$ converges if there exists a function $u \in \mathcal{D}(\Omega)$ such that the supports of the $\{u_n\}$ are all contained in a compact subset of $\Omega$ and their derivatives $\{D^\alpha u_n\}$ of any order converge uniformly to $D^\alpha u$. The space of *distributions* $\mathcal{D}'(\Omega)$ is the dual of $\mathcal{D}(\Omega)$, i.e., the space of the linear functionals on $\mathcal{D}(\Omega)$ that are continuous with respect to the pseudo-topology just defined. We note that we can associate a distribution with every $u \in L^2(\Omega)$, still denoted by $u$ and defined by

$$\langle u, \phi \rangle = \int_\Omega u\phi\,dx, \quad \phi \in \mathcal{D}(\Omega),$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $\mathcal{D}'(\Omega)$ and $\mathcal{D}(\Omega)$.

Given a multi-index $\alpha$ and a distribution $u$, we can define its derivative $D^\alpha u \in \mathcal{D}'(\Omega)$ by

$$\langle D^\alpha u, \phi \rangle = (-1)^{|\alpha|} \langle u, D^\alpha \phi \rangle, \quad \phi \in \mathcal{D}(\Omega).$$

We are now ready to define the Sobolev space $H^k(\Omega)$ for any integer $k \geq 1$. A function $u$ belongs to $H^k(\Omega)$ if, for every multi-index $\alpha$, with $|\alpha| \leq k$, there exists $u_\alpha \in L^2(\Omega)$, such that

$$\langle D^\alpha u, \phi \rangle = \int_\Omega u_\alpha \phi\,dx, \quad \phi \in \mathcal{D}(\Omega).$$

We can then identify $D^\alpha u$ with $u_\alpha$. The space $H^k(\Omega)$ is a Hilbert space with the scalar product

$$(u,v)_{H^k(\Omega)} = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}.$$

and an induced norm $\|\cdot\|_{H^k(\Omega)}$ given by

$$\|u\|^2_{H^k(\Omega)} = (u,u)_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u|^2\,dx.$$

A seminorm $|\cdot|_{H^k(\Omega)}$ is given by

$$|u|^2_{H^k(\Omega)} = \sum_{|\alpha|=k} \int_\Omega |D^\alpha u|^2 \, dx.$$

In case $k = 1$, we can write

$$\|u\|^2_{H^1(\Omega)} = \|u\|^2_{L^2(\Omega)} + |u|^2_{H^1(\Omega)} = \|u\|^2_{L^2(\Omega)} + \int_\Omega |\nabla u|^2 \, dx,$$

with

$$\nabla = \mathbf{grad} = \left( \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right).$$

The space $H^s(\Omega)$ can also be defined for any real number $s \geq 0$. We do not enter into details but refer to, e.g., [317, 8] for a complete treatment. Among the many possible (equivalent) definitions, we choose one that employs the K-interpolation method. Let $X \subset Y$ be two Hilbert spaces with continuous embedding. For $u \in Y$ and $t > 0$, we define

$$K(t, u; X, Y)^2 = \inf_{u_0 + u_1 = u} \{ \|u_0\|^2_X + t^2 \|u_1\|^2_Y \}, \quad u_0 \in X, \ \ u_1 \in Y.$$

An interpolation space can be defined for $\theta \in [0, 1]$ as

$$[X, Y]_\theta = \{ u \in Y \mid \ \ t^{-(\theta+1/2)} \, K(t, u; X, Y) \in L^2(0, \infty) \}$$

with a norm given by

$$\|u\|^2_{[X,Y]_\theta} = \|u\|^2_Y + \int_0^\infty t^{-(2\theta+1)} \, K(t, u; X, Y)^2 \, dt.$$

$[X, Y]_\theta$ is a Hilbert space. We have

$$X \subset [X, Y]_\theta \subset Y,$$

with $[X, Y]_\theta$ coinciding with $X$ or $Y$ if $\theta = 0$ or $\theta = 1$, respectively.

Let now $k_1 > k_2$ be two non-negative integers and let $s = (1 - \theta)k_1 + \theta k_2$, $\theta \in [0, 1]$. We then define

$$H^s(\Omega) = [H^{k_1}(\Omega), H^{k_2}(\Omega)]_\theta,$$

with the norm

$$\|u\|_{H^s(\Omega)} = \|u\|_{[H^{k_1}(\Omega), H^{k_2}(\Omega)]_\theta};$$

see, e.g., Sect. 1.9 and 1.15 in [317]. This definition is independent of $k_1$ and $k_2$ and depends only on $s$. In addition, for the case of integer $s$ it gives the Sobolev spaces of integer order previously defined and an equivalent norm.

An equivalent intrinsic norm defined in terms of integrals over $\Omega$ can also be found for Sobolev spaces of fractional order as stated in the following lemma.

**Lemma A.2** *Let $\sigma \in (0,1)$. Then, the norm*

$$\left(\|u\|^2_{L^2(\Omega)} + |u|^2_{H^\sigma(\Omega)}\right)^{1/2}, \tag{A.1}$$

*with the seminorm*

$$|u|^2_{H^\sigma(\Omega)} = \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^2}{|x-y|^{2\sigma+n}}\, dxdy,$$

*provides an equivalent norm in $H^\sigma(\Omega)$. Let $s > 0$, with $[s]$ the integer part of $s$ and $\sigma = s - [s]$. Then, an equivalent norm in $H^s(\Omega)$ is given by*

$$\left(\|u\|^2_{H^{[s]}(\Omega)} + |u|^2_{H^s(\Omega)}\right)^{1/2}, \tag{A.2}$$

*with the seminorm*

$$|u|^2_{H^s(\Omega)} = \sum_{|\alpha|=[s]} \int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{2\sigma+n}}.$$

For simplicity, we employ the same notation for the equivalent interpolation norms and the intrinsic norms defined in (A.1) and (A.2).

We define $H^s_0(\Omega)$ as the closure of $\mathcal{D}(\Omega)$ in $H^s(\Omega)$. One can prove that $H^s_0(\Omega)$ is a proper subspace of $H^s(\Omega)$ if and only if $s > 1/2$:

$$\begin{cases} H^s_0(\Omega) = H^s(\Omega), & s \le 1/2, \\ H^s_0(\Omega) \neq H^s(\Omega), & s > 1/2. \end{cases}$$

For a discussion of the case $s = 1/2$, see, e.g., [351, Page 106].

*Remark A.3.* If $\tilde{\Omega}$ is a proper subset of $\Omega$, relatively open with respect to $\Omega$, and $s$ is a non-negative number different from an integer plus $1/2$ (in particular, from $1/2$), then the space $H^s_0(\tilde{\Omega})$ coincides with the space of functions on $\tilde{\Omega}$ such that their extension by zero to $\Omega$ belongs to $H^s(\Omega)$.

Let now $s$ be a non-negative real number. We define the space $H^{-s}(\Omega)$ as the dual space of $H^s_0(\Omega)$, i.e., as the space of continuous linear functionals on $H^s_0(\Omega)$. Given a functional $u \in H^{-s}(\Omega)$ and a function $v \in H^s_0(\Omega)$, we denote the value of $u$ at $v$ by $\langle u, v\rangle$. The space $H^{-s}(\Omega)$ is then equipped with the dual norm

$$\|u\|_{H^{-s}(\Omega)} = \sup_{v \in H^s_0(\Omega)} \frac{\langle u, v\rangle}{\|v\|_{H^s(\Omega)}}.$$

The following properties hold.

**Lemma A.4** *The space $C^\infty(\overline{\Omega})$ is dense in $H^s(\Omega)$ for every $s \in \mathbb{R}$. In addition, if $s_1 < s_2$ are two real numbers, then $H^{s_2}(\Omega)$ is continuously embedded and dense in $H^{s_1}(\Omega)$. In particular, there exists a constant depending only on $\Omega$, $s_1$, and $s_2$, such that*

$$\|u\|_{H^{s_1}(\Omega)} \leq C \|u\|_{H^{s_2}(\Omega)}, \quad u \in H^{s_2}(\Omega).$$

*For every non-negative integer $k$, the embedding $H^{k+1}(\Omega) \subset H^k(\Omega)$ is compact.*

If $C^0(\overline{\Omega})$ is the space of continuous functions in $\overline{\Omega}$ endowed with the maximum norm, we have the following regularity result.

**Lemma A.5** *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. If $s > n/2$, then*

$$H^s(\Omega) \subset C^0(\overline{\Omega})$$

*and the embedding is compact.*

The space $L^p(\Omega)$, for $p \geq 1$, is defined as

$$L^p(\Omega) = \left\{ u : \Omega \mapsto \mathbb{R} \,|\, \int_\Omega |u|^p \, dx < \infty \right\}.$$

It is a Banach space with norm

$$\|u\|_{L^p(\Omega)} = \left( \int_\Omega |u|^p \, dx \right)^{1/p}.$$

As before, for $p = 2$, $H^s(\Omega) = W^{s,2}(\Omega)$, $s \in \mathbb{R}$, we can define the spaces $W^{s,p}(\Omega)$. For a non-negative integer $k$ and $p = \infty$, $W^{k,\infty}(\Omega)$ is the space of functions that are bounded together with their derivatives up to order $k$.

We can also define spaces of vector functions: $H^s(\Omega)^n$ is the space of vector functions in $\mathbb{R}^n$ such that each component belongs to $H^s(\Omega)$. Scalar products and norms can be defined in a straightforward way.

## A.2 Trace Spaces

In this section, we define some Sobolev spaces on a set $\Gamma \subseteq \partial\Omega$. If $\partial\Omega$ is Lipschitz continuous, integrals over $\partial\Omega$ can be uniquely defined using a partition of unity relative to the covering $\{\mathcal{O}_i\}$ and systems of local coordinates on $\{\partial\Omega \cap \mathcal{O}_i\}$; see Definition A.1. We refer to [361, Sect. 3.1] for details. We can then define the space $H^s(\partial\Omega)$, $s \geq 0$, consisting of functions on $\partial\Omega$, such that

$$\|u\|^2_{H^s(\partial\Omega)} = \|u\|^2_{H^{[s]}(\partial\Omega)} + |u|^2_{H^s(\partial\Omega)} < \infty,$$

with the seminorm

$$|u|^2_{H^s(\partial\Omega)} = \sum_{|\alpha|=[s]} \int_{\partial\Omega} \int_{\partial\Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{2\sigma+n-1}} \, dS_x dS_y$$

and $\sigma = s - [s]$.

The following two lemmas can be found in [237].

**Lemma A.6** *Let $\Omega$ be a Lipschitz region and let $s > 1/2$. Then, the operator $\gamma_0 : C^\infty(\overline{\Omega}) \to C^\infty(\partial\Omega)$, mapping a function into its restriction on the boundary, can be extended continuously to an operator $\gamma_0 : H^s(\Omega) \to H^{s-1/2}(\partial\Omega)$.*

**Lemma A.7** *With the same assumptions as in Lemma A.6, there exists a continuous lifting operator $\mathcal{R}_0 : H^{s-1/2}(\partial\Omega) \to H^s(\Omega)$, such that $\gamma_0(\mathcal{R}_0 u) = u$, $u \in H^{s-1/2}(\partial\Omega)$.*

The previous definitions and properties can easily be generalized to a proper subset $\Gamma \subset \partial\Omega$ with non-vanishing $(n-1)$-dimensional measure and which is relatively open with respect to $\partial\Omega$.

The space $H_0^s(\Omega)$ defined in the previous section coincides with the kernel of $\gamma_0$ for $1/2 < s < 3/2$. For larger $s$, it consists of functions that vanish on $\partial\Omega$ together with some additional derivatives. Similarly as in Lemmas A.6 and A.7, one can define a trace on $\Gamma \subset \partial\Omega$, of positive measure, as a continuous operator from $H^s(\Omega)$ onto $H^{s-1/2}(\Gamma)$. For $1/2 < s < 3/2$, we denote the subspace of $H^s(\Omega)$ of functions that vanish on $\Gamma$ by $H_0^s(\Omega, \Gamma)$.

Since $\partial\Omega$ has no boundary, $H_0^s(\partial\Omega)$ coincides with $H^s(\partial\Omega)$, for an arbitrary $s \geq 0$, where the former is defined as the closure of $\mathcal{D}(\partial\Omega)$, the space of $C^\infty$ functions on $\partial\Omega$ with compact support. The Sobolev space $H^{-s}(\partial\Omega)$, $s > 0$, can be defined as in the previous section, as the dual space of $H_0^s(\partial\Omega) = H^s(\partial\Omega)$.

Even if $\Gamma$ is a proper subset of $\partial\Omega$, $H_0^s(\Gamma)$ coincides with $H^s(\Gamma)$, for $s \leq 1/2$. However the extensions by zero of functions in $H_0^{1/2}(\Gamma) = H^{1/2}(\Gamma)$ do not in general belong to $H^{1/2}(\partial\Omega)$. We therefore define the space

$$H_{00}^{1/2}(\Gamma) = \{u \in H^{1/2}(\Gamma) \mid \mathcal{E}u \in H^{1/2}(\partial\Omega)\},$$

where $\mathcal{E}u$ is the extension by zero of $u$ to $\partial\Omega$. We have the following lemma.

**Lemma A.8** *Let $\Gamma \subset \partial\Omega$. Then, $H_{00}^{1/2}(\Gamma)$ coincides with the interpolation space $[H_0^1(\Gamma), L^2(\Gamma)]_{1/2}$. With*

$$\|u\|_{H_{00}^{1/2}(\Gamma)} := \|u\|_{[H_0^1(\Gamma), L^2(\Gamma)]_{1/2}},$$

*there then exist constants, such that, for $u \in H_{00}^{1/2}(\Gamma)$,*

$$c\|\mathcal{E}u\|^2_{H^{1/2}(\partial\Omega)} \leq \|u\|^2_{H_{00}^{1/2}(\Gamma)} \leq C\|\mathcal{E}u\|^2_{H^{1/2}(\partial\Omega)}.$$

*Remark A.9.* If $u \in H^{1/2}(\partial\Omega)$ vanishes almost everywhere on $\partial\Omega \setminus \Gamma$, then Lemma A.8 ensures that

$$\|u\|_{H^{1/2}(\partial\Omega)} \quad \text{and} \quad \|u\|_{H_{00}^{1/2}(\Gamma)}$$

are equivalent norms. In addition, the immersion $H_{00}^{1/2}(\Gamma) \subset H^{1/2}(\Gamma)$ is continuous:

$$\|u\|_{H^{1/2}(\Gamma)} \leq \|u\|_{H^{1/2}(\partial\Omega)} \leq C\|u\|_{H_{00}^{1/2}(\Gamma)},$$

as it can easily be seen by using the definition of the $H^{1/2}$-norm. It is also straightforward to show that, even though $u$ vanishes on $\partial\Omega \setminus \Gamma$, the seminorms $|u|_{H^{1/2}(\partial\Omega)}$ and $|u|_{H^{1/2}(\Gamma)}$ are not equivalent and the former cannot be bounded from above by a multiple of the latter.

The dual space of $H_{00}^{1/2}(\Gamma)$ is denoted by $H_{00}^{-1/2}(\Gamma)$. We note that since $H_0^{1/2}(\Gamma)$ coincides with $H^{1/2}(\Gamma)$, $H^{-1/2}(\Gamma)$ is a proper subset of $H_{00}^{-1/2}(\Gamma)$.

The Green's formula can be generalized to functions in $H^1(\Omega)^n$.

**Lemma A.10** *For $\mathbf{u} \in H^1(\Omega)^n$ and $v \in H^1(\Omega)$, we have*

$$\int_\Omega (\nabla \cdot \mathbf{u})\, v\, dx + \int_\Omega \mathbf{u} \cdot \nabla v\, dx = \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) v\, dS.$$

## A.3 Linear Operators

Let $X$ and $Y$ be two Banach spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. We define $\mathcal{L}(X,Y)$ as the space of linear continuous operators from $X$ into $Y$. The space $\mathcal{L}(X,Y)$ is a Banach space endowed with the operator norm

$$\|A\| = \|A\|_{\mathcal{L}(X,Y)} = \sup_{u \in X\setminus\{0\}} \frac{\|Au\|_Y}{\|u\|_X}, \quad A \in \mathcal{L}(X,Y).$$

We recall that a linear operator is continuous if and only if it is bounded. We also note that linear functionals are particular linear operators defined from a Banach space $X$ into $\mathbb{C}$. We use the notation $X' = \mathcal{L}(X,\mathbb{C})$.

The following lemma provides a useful tool for the proof of certain bounds in intermediate spaces.

**Lemma A.11** *Let $X$ and $Y$ be two separable Hilbert spaces such that $X$ is continuously embedded and dense in $Y$. Let the same properties hold for a second pair of Hilbert spaces $\tilde{X}$ and $\tilde{Y}$. If*

$$A \in \mathcal{L}(X,\tilde{X}) \cap \mathcal{L}(Y,\tilde{Y}), \qquad a_X = \|A\|_{\mathcal{L}(X,\tilde{X})}, \quad a_Y = \|A\|_{\mathcal{L}(Y,\tilde{Y})},$$

*then, for $\theta \in [0,1]$, $A$ belongs to $\mathcal{L}([X,Y]_\theta, [\tilde{X},\tilde{Y}]_\theta)$, with*

$$\|A\|_{\mathcal{L}([X,Y]_\theta,[\tilde{X},\tilde{Y}]_\theta)} \leq a_X^{1-\theta}\, a_Y^{\theta}.$$

## A.4 Poincaré and Friedrichs Type Inequalities

Poincaré and Friedrichs type inequalities provide powerful tools for the analysis of variational problems, finite element approximations, and domain decomposition methods. These inequalities derive from the following general result.

**Theorem A.12** *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and let $f_i$, $i = 1, \ldots, L$, $L \geq 1$, be functionals (not necessarily linear) in $H^1(\Omega)$, such that, if $v$ is constant in $\Omega$,*

$$\sum_{i=1}^{L} |f_i(v)|^2 = 0 \quad \Leftrightarrow \quad v = 0.$$

*Then, there exist constants, depending only on $\Omega$ and the functionals $f_i$, such that, for $u \in H^1(\Omega)$,*

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \sum_{i=1}^{L} |f_i(u)|^2.$$

Despite the generality of this theorem, its proof is a simple application of Rellich's theorem and can be found in [361, Th. 7.1]. We also remark that it is valid for certain more general domains and functionals; see [361, Sec. 7.1].

**Lemma A.13 (Poincaré inequality)** *Let $u \in H^1(\Omega)$. Then, there exist constants, depending only on $\Omega$, such that*

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \left( \int_{\Omega} u \, dx \right)^2.$$

*Proof.* We only need apply Theorem A.12 with $L = 1$ and the linear functional

$$f_1(u) = \int_{\Omega} u \, dx.$$

□

**Lemma A.14 (Friedrichs inequality)** *Let $\Gamma \subseteq \partial\Omega$ have nonvanishing $(n-1)$-dimensional measure. Then, there exist constants, depending only on $\Omega$ and $\Gamma$, such that, for $u \in H^1(\Omega)$,*

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \|u\|^2_{L^2(\Gamma)}.$$

*In particular, if $u$ vanishes on $\Gamma$,*

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)}$$

*and thus*

$$|u|^2_{H^1(\Omega)} \leq \|u\|^2_{H^1(\Omega)} \leq (C_1 + 1) \, |u|^2_{H^1(\Omega)}.$$

*Proof.* We need only apply Theorem A.12 with $L = 1$ and $f_1(u) = \|u\|_{L^2(\Gamma)}$. □

The following corollary can be obtained by simple *scaling arguments*: More precisely, it can be obtained by mapping the domain $\Omega$, of diameter $H$,

into $\hat{\Omega}$, of unit diameter, by an isotropic dilation. One can then employ the previous results for $\hat{\Omega}$: mapping the functions back into $\Omega$ gives an explicit dependence on $H$. The constants in the estimates depend only on $\hat{\Omega}$, i.e., the shape of $\Omega$.

**Corollary A.15** *Let $\Omega$ be Lipschitz continuous with diameter $H$. Then, there exists a constant $\hat{C}_1$, that depends only on the shape of $\Omega$ but not on its size, such that*

$$\|u\|^2_{L^2(\Omega)} \leq \hat{C}_1 H^2 |u|^2_{H^1(\Omega)},$$

*for $u \in H^1(\Omega)$ with vanishing mean value on $\Omega$. Similarly, if $\Gamma \subset \partial\Omega$ is defined as in Lemma A.14 and has a diameter of order $H$, then*

$$\|u\|^2_{L^2(\Omega)} \leq \hat{C}_2 H^2 |u|^2_{H^1(\Omega)} + \hat{C}_3 H \|u\|^2_{L^2(\Gamma)}.$$

*for $u \in H^1(\Omega)$.*

*Remark A.16.* We note that Lemmas A.13 and A.14 are not valid in general for non Lipschitz domains, as can easily be seen by considering a domain which is the union of two triangles with only a vertex in common.

In the analysis of some iterative substructuring methods, we need some inequalities involving functions on the boundary. The following result can easily be proven using Theorem A.12 and the operators $\gamma_0$ and $\mathcal{R}_0$ of Lemmas A.6 and A.7. A similar result holds in two dimensions.

**Lemma A.17** *Let $\Omega \subset \mathbb{R}^3$ be a Lipschitz continuous polyhedron and let $f_i$, $i = 1, \ldots, L$, $L \geq 1$, be functionals in $H^{1/2}(\partial\Omega)$, such that, if $v$ is constant on $\partial\Omega$, then*

$$\sum_{i=1}^{L} |f_i(v)|^2 = 0 \quad \Leftrightarrow \quad v = 0.$$

*Then, there exist constants, depending only on $\Omega$ and the functionals $f_i$, such that, for $u \in H^{1/2}(\partial\Omega)$,*

$$\|u\|^2_{L^2(\partial\Omega)} \leq C_1 |u|^2_{H^{1/2}(\partial\Omega)} + C_2 \sum_{i=1}^{L} |f_i(u)|^2.$$

*A similar result holds if $\partial\Omega$ is replaced by one of its faces $\mathcal{F}$. Consequently, there exists a constant $\hat{C}_4$, that depends only on the shape of $\Omega$ but not on its size, such that*

$$\|u\|^2_{L^2(\partial\Omega)} \leq \hat{C}_4 H |u|^2_{H^{1/2}(\partial\Omega)},$$

*if $u \in H^{1/2}(\partial\Omega)$ either has a vanishing mean value on $\partial\Omega$ or belongs to the closure of the space of $C^\infty(\partial\Omega)$ functions that vanish on a face of $\Omega$. Similarly, if $\mathcal{F} \subset \partial\Omega$ is one of the faces of $\Omega$ of diameter $H$, then there exists a constant $\hat{C}_5$, that depends only on the shape of $\mathcal{F}$ but not on its size, such that*

$$\|u\|^2_{L^2(\mathcal{F})} \leq \hat{C}_5\, H\, |u|^2_{H^{1/2}(\mathcal{F})},$$

*if* $u \in H^{1/2}(\mathcal{F})$ *either has a vanishing mean value on* $\mathcal{F}$ *or belongs to* $H^{1/2}_{00}(\mathcal{F})$.

We end this section with the following theorem.

**Theorem A.18 (Poincaré inequality)** *Let* $\Omega \subset \mathbb{R}^n$ *be a bounded Lipschitz domain and* $V$ *be a closed subspace of* $H^1(\Omega)$ *that contains* $\mathbb{P}_0(\Omega)$, *the space of constant functions on* $\Omega$. *Let* $W$ *be a Hilbert space with a norm* $\|\cdot\|_W$ *and let*

$$B : V \longrightarrow W,$$

*be a bounded linear operator, such that*

$$Bv = 0, \qquad v \in \mathbb{P}_0(\Omega).$$

*If*

$$\|Bu\|_W \leq \|B\| \|u\|_{H^1(\Omega)}, \qquad u \in W, \tag{A.3}$$

*then*

$$\|Bu\|_W \leq \|B\| C_\Omega\, |u|_{H^1(\Omega)}, \qquad u \in W,$$

*where* $C_\Omega$ *depends only on the domain* $\Omega$, *but is otherwise independent of* $u$, $B$, *and of the spaces* $V$ *and* $W$. *The theorem remains valid if* $H^1(\Omega)$ *is replaced by* $H^{1/2}(\partial\Omega)$ *or* $H^{1/2}(\mathcal{F})$, *with* $\mathcal{F}$ *a face of* $\Omega$.

*Proof.* Let $u \in V$ and let $\bar{u}_\Omega$ be its mean value on $\Omega$. Lemma A.13 ensures

$$\|Bu\|^2_W = \|B(u - \bar{u}_\Omega)\|^2_W = \|B\|^2 (C_1 + 1) |u|^2_{H^1(\Omega)}.$$

For the spaces of traces, Lemma A.17 can be employed. □

The previous result is employed extensively in this monograph. We also refer to it as a *Poincaré inequality*, of which it is a direct consequence. It basically states that if the left hand side of the inequality (A.3) does not change if we add a constant to $u$, then the norm on the right hand side can be replaced by the seminorm. Since the norms we employ are mainly defined by integrals, an additional scaling argument usually provides an explicit dependence of the constants on the diameter of $\Omega$.

## A.5 Spaces of Vector-Valued Functions

In this section we introduce some Sobolev spaces of vector-valued functions. We mention, e.g., [223, 151, 95, 354] as general references for this material.

### A.5.1 The Space $H(\mathrm{div}\,;\Omega)$

Given a vector function $\mathbf{u} \in \mathbb{R}^n$, $n = 2, 3$, the divergence operator is defined as

$$\mathrm{div}\,\mathbf{u} = \nabla \cdot \mathbf{u} = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i}, \tag{A.4}$$

where $u_i$ is the $i$-th component of $\mathbf{u}$. The space $H(\mathrm{div}\,;\Omega)$ consists of square-integrable vectors, with square-integrable divergence. It is a Hilbert space with scalar product and graph norm defined by

$$(\mathbf{u}, \mathbf{v})_{\mathrm{div}\,;\Omega} = \int_\Omega \mathbf{u} \cdot \mathbf{v}\, dx + \int_\Omega \mathrm{div}\,\mathbf{u}\, \mathrm{div}\,\mathbf{v} dx, \quad \|\mathbf{u}\|^2_{\mathrm{div}\,;\Omega} = (\mathbf{u}, \mathbf{u})_{\mathrm{div}\,;\Omega}.$$

Given a vector $\mathbf{u} \in H(\mathrm{div}\,;\Omega)$, it is possible to define its normal component $\mathbf{u} \cdot \mathbf{n}$ on the boundary $\partial\Omega$; see [223, Ch. I, Th. 2.5 and Cor. 2.8].

**Lemma A.19** *Let $\Omega$ be Lipschitz continuous. Then, the operator $\gamma_n : C^\infty(\overline{\Omega}) \to C^\infty(\partial\Omega)$, mapping a vector into its normal component on the boundary, can be extended continuously to an operator $\gamma_n : H(\mathrm{div}\,;\Omega) \to H^{-1/2}(\partial\Omega)$. In addition, there exists a continuous lifting operator $\mathcal{R}_n : H^{-1/2}(\partial\Omega) \to H(\mathrm{div}\,;\Omega)$, such that $\gamma_n(\mathcal{R}_n\phi) = \phi$, $\phi \in H^{-1/2}(\partial\Omega)$. The following Green's formula holds: for $\mathbf{u} \in H(\mathrm{div}\,;\Omega)$ and $q \in H^1(\Omega)$,*

$$\int_\Omega \mathbf{u} \cdot \mathbf{grad}\, q\, dx + \int_\Omega \mathrm{div}\,\mathbf{u}\, q\, dx = \int_{\partial\Omega} \mathbf{u} \cdot \mathbf{n}\, q\, dS,$$

*where the integral on the right hand side is understood as the duality pairing between $H^{-\frac{1}{2}}(\partial\Omega)$ and $H^{\frac{1}{2}}(\partial\Omega)$.*

The subspace of vectors in $H(\mathrm{div}\,;\Omega)$ with vanishing normal component on $\partial\Omega$ is denoted by $H_0(\mathrm{div}\,;\Omega)$, the subspace of vectors in $H(\mathrm{div}\,;\Omega)$ with vanishing divergence by $H(\mathrm{div}_0;\Omega)$:

$$H(\mathrm{div}_0;\Omega) = \{\mathbf{u} \in H(\mathrm{div}\,;\Omega),\ \mathrm{div}\,\mathbf{u} = 0\},$$

and the subspace of vectors in $H_0(\mathrm{div}\,;\Omega)$ with vanishing divergence by $H_0(\mathrm{div}_0;\Omega)$

$$H_0(\mathrm{div}_0;\Omega) = \{\mathbf{u} \in H_0(\mathrm{div}\,;\Omega),\ \mathrm{div}\,\mathbf{u} = 0\}.$$

The space $L^2(\Omega)^n$ has the following orthogonal decompositions, see [151, vol. 3, p. 215, Proposition 1],

$$L^2(\Omega)^n = H(\mathrm{div}_0;\Omega) \oplus \mathbf{grad}\, H^1_0(\Omega), \tag{A.5}$$

$$L^2(\Omega)^n = H_0(\mathrm{div}_0;\Omega) \oplus \mathbf{grad}\, H^1(\Omega). \tag{A.6}$$

These decompositions are the generalization of the Helmholtz decomposition for a smooth vector into a divergence-free and a curl-free part.

Since $H(\mathrm{div}\,;\Omega) \subset L^2(\Omega)^n$, the decompositions (A.5) and (A.6) give rise to an orthogonal decompositions of $H(\mathrm{div}\,;\Omega)$ and $H_0(\mathrm{div}\,;\Omega)$, into the kernel of the divergence operator and its orthogonal complement:

$$H(\mathrm{div}\,;\Omega) = H(\mathrm{div}_0;\Omega) \oplus H^{\perp}(\mathrm{div}\,;\Omega), \tag{A.7}$$

$$H_0(\mathrm{div}\,;\Omega) = H_0(\mathrm{div}_0;\Omega) \oplus H_0^{\perp}(\mathrm{div}\,;\Omega), \tag{A.8}$$

where

$$H^{\perp}(\mathrm{div}\,;\Omega) = H(\mathrm{div}\,;\Omega) \cap \mathbf{grad}\, H_0^1(\Omega),$$
$$H_0^{\perp}(\mathrm{div}\,;\Omega) = H_0(\mathrm{div}\,;\Omega) \cap \mathbf{grad}\, H^1(\Omega).$$

We stress the fact that (A.7) and (A.8) are orthogonal both with respect to the $(\cdot,\cdot)_{L^2(\Omega)}$ and $(\cdot,\cdot)_{\mathrm{div};\Omega}$ inner products. The decompositions (A.7) and (A.8) ensure that

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \le C_\Omega \,\|\mathrm{div}\,\mathbf{u}\|_{L^2(\Omega)}, \quad \mathbf{u} \in H^{\perp}(\mathrm{div}\,;\Omega) \cup H_0^{\perp}(\mathrm{div}\,;\Omega). \tag{A.9}$$

### A.5.2 The Space $H(\mathbf{curl}\,;\Omega)$ in Two Dimensions

We now consider the case $\Omega \subset \mathbb{R}^2$. Given a scalar function $q$ and a vector $\mathbf{u}$, the vector and scalar curl operators are defined, respectively, by

$$\mathbf{curl}\, q = \left(\frac{\partial q}{\partial x_2}, -\frac{\partial q}{\partial x_1}\right),$$

and

$$\mathrm{curl}\,\mathbf{u} = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}. \tag{A.10}$$

The space $H(\mathbf{curl}\,;\Omega)$ consists of square-integrable vectors, with square-integrable curl. This is a Hilbert space with scalar product and graph norm defined by

$$(\mathbf{u},\mathbf{v})_{\mathbf{curl}\,;\Omega} = \int_\Omega \mathbf{u}\cdot\mathbf{v}\,dx + \int_\Omega \mathrm{curl}\,\mathbf{u}\,\mathrm{curl}\,\mathbf{v}, dx, \quad \|\mathbf{u}\|^2_{\mathbf{curl}\,;\Omega} = (\mathbf{u},\mathbf{u})_{\mathbf{curl}\,;\Omega}.$$

We define the unit tangent vector $\mathbf{t}$ on the boundary $\partial\Omega$ by

$$\mathbf{t} = (-n_2, n_1).$$

For a generic vector $\mathbf{u}$, its tangential component on the boundary is

$$\mathbf{u}\cdot\mathbf{t} = \mathbf{n}\times\mathbf{u}.$$

By using the definitions (A.4) and (A.10), we have the following result.

**Lemma A.20** *A vector $\mathbf{u} = (u_1, u_2)$ belongs to $H(\mathbf{curl}\,;\Omega)$ if and only if the vector $\mathbf{v} = (-u_2, u_1)$ belongs to $H(\mathrm{div}\,;\Omega)$. In addition,*

$$\mathbf{v}\cdot\mathbf{n} = -\mathbf{u}\cdot\mathbf{t}.$$

It is then clear that, using the results for $H(\mathrm{div}\,;\Omega)$, for $\mathbf{u} \in H(\mathbf{curl}\,;\Omega)$, it is possible to define its tangential component on the boundary; see also [223, Ch. I, Th. 2.11].

**Lemma A.21** *Let $\Omega \subset \mathbb{R}^2$ be Lipschitz continuous. Then, the operator $\gamma_t : C^\infty(\overline{\Omega}) \to C^\infty(\partial\Omega)$, mapping a vector into its tangential component on the boundary, can be extended continuously to an operator $\gamma_t : H(\mathbf{curl}\,;\Omega) \to H^{-1/2}(\partial\Omega)$. In addition, there exists a continuous lifting operator $\mathcal{R}_t : H^{-1/2}(\partial\Omega) \to H(\mathbf{curl}\,;\Omega)$ such that $\gamma_t(\mathcal{R}_t\phi) = \phi$, $\phi \in H^{-1/2}(\partial\Omega)$. The following Green's formula holds, for $\mathbf{u} \in H(\mathbf{curl}\,;\Omega)$ and $q \in H^1(\Omega)$,*

$$\int_\Omega \mathrm{curl}\,\mathbf{u}\, q\, dx - \int_\Omega \mathbf{u}\cdot\mathbf{curl}\, q\, dx = \int_{\partial\Omega} \mathbf{u}\cdot\mathbf{t}\, q\, dS,$$

The subspace of vectors in $H(\mathbf{curl}\,;\Omega)$ with vanishing tangential component on $\partial\Omega$ is denoted by $H_0(\mathbf{curl}\,;\Omega)$, the subspace of vectors in $H(\mathbf{curl}\,;\Omega)$ with vanishing curl by $H(\mathbf{curl}_0;\Omega)$, and the subspace of vectors in $H_0(\mathbf{curl}\,;\Omega)$ with vanishing curl by $H_0(\mathbf{curl}_0;\Omega)$.

Since $H(\mathbf{curl}\,;\Omega) \subset L^2(\Omega)^n$ and $\mathbf{grad}\, H^1(\Omega) \subset H(\mathbf{curl}\,;\Omega)$, the decompositions (A.5) and (A.6) give rise to the following orthogonal decompositions of $H_0(\mathbf{curl}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$,

$$H_0(\mathbf{curl}\,;\Omega) = \mathbf{grad}\, H_0^1(\Omega) \oplus H_0^\perp(\mathbf{curl}\,;\Omega), \tag{A.11}$$
$$H(\mathbf{curl}\,;\Omega) = \mathbf{grad}\, H^1(\Omega) \oplus H^\perp(\mathbf{curl}\,;\Omega), \tag{A.12}$$

where

$$H_0^\perp(\mathbf{curl}\,;\Omega) = H(\mathrm{div}_0;\Omega) \cap H_0(\mathbf{curl}\,;\Omega), \tag{A.13}$$
$$H^\perp(\mathbf{curl}\,;\Omega) = H_0(\mathrm{div}_0;\Omega) \cap H(\mathbf{curl}\,;\Omega).$$

We note that (A.11) and (A.12) are orthogonal both with respect to the $(\cdot,\cdot)_{L^2(\Omega)}$ and $(\cdot,\cdot)_{\mathbf{curl}\,;\Omega}$ inner products, and that they are valid also for $n = 3$; see Sect. A.5.3.

### A.5.3 The Space $H(\mathbf{curl}\,;\Omega)$ in Three Dimensions

We will now consider the case $\Omega \subset \mathbb{R}^3$. Given a vector $\mathbf{u}$, the vector curl operator is defined by

$$\mathbf{curl}\,\mathbf{u} = \nabla \times \mathbf{u} = \begin{bmatrix} \dfrac{\partial u_3}{\partial x_2} - \dfrac{\partial u_2}{\partial x_3} \\ \dfrac{\partial u_1}{\partial x_3} - \dfrac{\partial u_3}{\partial x_1} \\ \dfrac{\partial u_2}{\partial x_1} - \dfrac{\partial u_1}{\partial x_2} \end{bmatrix}.$$

The space $H(\mathbf{curl};\Omega)$ consists of square-integrable vectors, with a square-integrable curl. This is a Hilbert space with scalar product and graph norm defined by

$$(\mathbf{u},\mathbf{v})_{\mathbf{curl};\Omega} = \int_\Omega \mathbf{u}\cdot\mathbf{v}\,dx + \int_\Omega \mathbf{curl}\,\mathbf{u}\cdot\mathbf{curl}\,\mathbf{v},dx, \quad \|\mathbf{u}\|^2_{\mathbf{curl};\Omega} = (\mathbf{u},\mathbf{u})_{\mathbf{curl};\Omega}.$$

We remark that we will use the same notation $H(\mathbf{curl};\Omega)$ for the space in two and three dimensions.

The tangential component of a vector $\mathbf{u}$ on the boundary $\partial\Omega$ is defined by

$$\mathbf{u}_t = \mathbf{u} - (\mathbf{u}\cdot\mathbf{n})\mathbf{n} = (\mathbf{n}\times\mathbf{u})\times\mathbf{n}.$$

Since $|\mathbf{u}_t| = |\mathbf{n}\times\mathbf{u}|$, the vector $\mathbf{u}$ has vanishing tangential component if and only if $\mathbf{n}\times\mathbf{u} = 0$. With an abuse of terminology, we will also refer to $\mathbf{n}\times\mathbf{u}$ as the *tangential component* in this monograph.

Given a vector $\mathbf{u}\in H(\mathbf{curl};\Omega)$, it is possible to define its tangential component $\mathbf{n}\times\mathbf{u}$ on the boundary; see, e.g., [223, Ch. I, Th. 2.11].

**Lemma A.22** *Let $\Omega\subset\mathbb{R}^3$ be Lipschitz continuous. Then, the operator $\gamma_t : C^\infty(\overline{\Omega})^3 \to C^\infty(\partial\Omega)^3$, mapping a vector into its tangential component on the boundary, can be extended continuously to an operator $\gamma_t : H(\mathbf{curl};\Omega)\to H^{-1/2}(\partial\Omega)^3$. The following Green's formula holds, for $\mathbf{u}\in H(\mathbf{curl};\Omega)$ and $\mathbf{v}\in H^1(\Omega)^3$,*

$$\int_\Omega \mathbf{curl}\,\mathbf{u}\cdot\mathbf{v}\,dx - \int_\Omega \mathbf{u}\cdot\mathbf{curl}\,\mathbf{v}\,dx = \int_{\partial\Omega}(\mathbf{n}\times\mathbf{u})\cdot\mathbf{v}\,dS.$$

Thus the operator $\gamma_t$ is continuous, but it is not surjective. The space of tangential traces of $H(\mathbf{curl};\Omega)$ is a proper subspace of $H^{-1/2}(\partial\Omega)^3$ and it can be fully characterized; see, e.g., [15, 97, 98] and the references therein for a more detailed analysis.

As in the two-dimensional case, the subspace of vectors in $H(\mathbf{curl};\Omega)$ with vanishing tangential component on $\partial\Omega$ is denoted by $H_0(\mathbf{curl};\Omega)$, the subspace of vectors in $H(\mathbf{curl};\Omega)$ with vanishing curl by $H(\mathbf{curl}_0;\Omega)$, and the subspace of vectors in $H_0(\mathbf{curl};\Omega)$ with vanishing curl by $H_0(\mathbf{curl}_0;\Omega)$.

In the three-dimensional case, Equations (A.5) and (A.6) give rise to the same orthogonal decompositions, as given in (A.11) and (A.12).

### A.5.4 The Kernel and Range of the Curl and Divergence Operators

In this section, we will characterize the kernel and the range of the curl and divergence operators. This characterization depends on the domain $\Omega$ where the Sobolev spaces, previously introduced, are defined. We will review some results valid for a particular set of Lipschitz domains and refer to [151, vol. 3,

Sect. IX.1.3] for the case of more general smooth domains, and to [18] for a generalization to a larger class of Lipschitz domains.

We assume that $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, is an open bounded connected set, with Lipschitz continuous boundary. We only consider the case where $\Omega$ is simply connected and its boundary consists of one connected component and note that these two conditions are equivalent for $n = 2$. The results presented in this section can be found in [151, vol. 3, Sect. IX.1.3], to which we refer for the proofs.

The kernel of the gradient operator consists of the constants, when defined in $H^1(\Omega)$, and of the zero function, when defined in $H_0^1(\Omega)$.

The following lemma characterizes the kernel of the curl operator as the range of the gradient operator for simply-connected domains; see [151, vol. 3, pp. 217-221].

**Lemma A.23** *If $\Omega$ is simply connected, then, for $n = 2, 3$,*

$$H(\mathbf{curl}_0; \Omega) = \mathbf{grad}\, H^1(\Omega). \tag{A.14}$$

*Remark A.24.* In the case where $\Omega$ is not simply connected, $\mathbf{grad}\, H^1(\Omega)$ is a proper subspace of $H(\mathbf{curl}_0; \Omega)$, and its orthogonal complement in $H(\mathbf{curl}_0; \Omega)$ can be fully characterized and is of finite dimension which equals the number of cuts necessary to make $\Omega$ simply connected; see [151, vol. 3, p. 219, Prop. 2]. We also note that (A.12) is a decomposition of the space $H(\mathbf{curl}; \Omega)$ into the kernel of the curl operator and its orthogonal complement, only with the hypothesis of Lemma A.23. In this case, we have, for $n = 2$,

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C_\Omega \, \|\mathrm{curl}\, \mathbf{u}\|_{L^2(\Omega)}, \quad \mathbf{u} \in H^\perp(\mathbf{curl}; \Omega), \tag{A.15}$$

and, for $n = 3$,

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C_\Omega \, \|\mathbf{curl}\, \mathbf{u}\|_{L^2(\Omega)^n}, \quad \mathbf{u} \in H^\perp(\mathbf{curl}; \Omega). \tag{A.16}$$

The next proposition characterizes the kernel of the curl operator in $H_0(\mathbf{curl}; \Omega)$ and, for the case $n = 2$, that of the divergence in $H_0(\mathrm{div}; \Omega)$; see Lemma A.20 for the relation between $H(\mathbf{curl}; \Omega)$ and $H(\mathrm{div}; \Omega)$ in two dimensions. This result can be found in [151, vol 3, p. 222, Prop. 3] and [151, vol 3, p. 224, Cor. 5].

**Lemma A.25** *If the boundary $\partial\Omega$ is connected, then*

$$H_0(\mathbf{curl}_0; \Omega) = \mathbf{grad}\, H_0^1(\Omega), \quad n = 2, 3, \tag{A.17}$$

$$H_0(\mathrm{div}_0; \Omega) = \mathbf{curl}\, H_0^1(\Omega), \quad n = 2. \tag{A.18}$$

*Remark A.26.* In the case where $\partial\Omega$ consists of more than one connected component, the subspace of $H^1(\Omega)$ of functions that are constant on each connected component of $\partial\Omega$ replaces $H_0^1(\Omega)$ in (A.17) and (A.18). When there is only one connected component, the gradients (and curls, for $n = 2$) of these

two spaces coincide and (A.11) is a decomposition of the space $H_0(\mathbf{curl}\,;\Omega)$ into the kernel of the curl operator and its orthogonal complement. In this case, we have, for $n = 2$,

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C_\Omega \, \|\mathrm{curl}\,\mathbf{u}\|_{L^2(\Omega)}, \quad \mathbf{u} \in H_0^\perp(\mathbf{curl}\,;\Omega), \tag{A.19}$$

and, for $n = 3$,

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C_\Omega \, \|\mathbf{curl}\,\mathbf{u}\|_{L^2(\Omega)^n}, \quad \mathbf{u} \in H_0^\perp(\mathbf{curl}\,;\Omega). \tag{A.20}$$

The following two propositions complete the characterization of the kernel of the divergence operator. For $n = 3$, we will need the space of vectors in $H^1(\Omega)^3$, with vanishing tangential component

$$H^1_{0t}(\Omega)^3 = \{\mathbf{u} \in H^1(\Omega)^3, \ \mathbf{n} \times \mathbf{u}|_{\partial\Omega} = 0\}.$$

**Lemma A.27** *If $\Omega$ is simply connected and $n = 3$, we have*

$$H_0(\mathrm{div}_0;\Omega) = \mathbf{curl}\, H^1_{0t}(\Omega)^3 = \mathbf{curl}\, H^1_0(\Omega)^3. \tag{A.21}$$

A proof can be found in [151, vol. 3, p. 224, Prop. 4 and Rem. 5]. We note that in the case where $\Omega$ is not simply connected, $\mathbf{curl}\, H^1_{0t}(\Omega)^3$ is a proper subspace of $H_0(\mathrm{div}_0;\Omega)$ and its orthogonal complement is the same finite dimensional space as mentioned in Remark A.24.

**Lemma A.28** *If the boundary $\partial\Omega$ is connected, we have*

$$H(\mathrm{div}_0;\Omega) = \mathbf{curl}\, H^1(\Omega), \quad n = 2,$$
$$H(\mathrm{div}_0;\Omega) = \mathbf{curl}\, H^1(\Omega)^3, \quad n = 3.$$

A proof can be found in [151, vol. 3, p. 222, Prop. 3]. We note that in the case where $\partial\Omega$ consists of $m > 1$ connected components, $\mathbf{curl}\, H^1(\Omega)$ and $\mathbf{curl}\, H^1(\Omega)^3$ are proper subspaces of $H(\mathrm{div}_0;\Omega)$ and their orthogonal complements have dimension $m - 1$.

The following corollary is a direct consequence of Lemma A.28 and gives a necessary and sufficient condition for the existence of a vector potential for a divergence-free vector.

**Corollary A.29** *Let $\partial\Omega$ be connected. A necessary and sufficient condition for a vector $\mathbf{u} \in L^2(\Omega)^n$ to be of the form $\mathbf{u} = \mathbf{curl}\, q$, with $q \in H^1(\Omega)$ if $n = 2$, or $\mathbf{u} = \mathbf{curl}\,\mathbf{v}$, with $\mathbf{v} \in H^1(\Omega)^3$ if $n = 3$, is that*

$$\mathrm{div}\,\mathbf{u} = 0.$$

*Remark A.30.* It follows from Lemmas A.28 and A.27, for $n = 3$, that the range of the curl operator on $H(\mathbf{curl}\,;\Omega)$ and $H_0(\mathbf{curl}\,;\Omega)$ coincides with $\mathbf{curl}\, H^1(\Omega)^3$ and $\mathbf{curl}\, H^1_{0t}(\Omega)^3$, respectively.

The following proposition generalizes Corollary A.29 to the case of more regular vectors, and ensures that a divergence-free vector potential can be found in three dimensions; see [223, Ch. I, Th. 3.4, Cor. 3.3 and Rem. 3.12].

**Lemma A.31** *Let $n = 3$ and let $\partial\Omega$ be connected. For $s \in [0,1]$, a necessary and sufficient condition for a vector $\mathbf{u} \in H^s(\Omega)^3$ to be of the form $\mathbf{u} = \mathbf{curl}\,\mathbf{v}$, with $\mathbf{v} \in H^{1+s}(\Omega)^3$ and $\operatorname{div}\mathbf{v} = 0$, is*

$$\operatorname{div}\mathbf{u} = 0.$$

We end this section with a characterization of the range of the divergence operator. We set

$$L^2_0(\Omega) = \left\{ q \in L^2(\Omega) \mid \int_\Omega q\,dx = 0 \right\}. \tag{A.22}$$

**Lemma A.32** *The divergence operator is continuous and surjective from $H(\operatorname{div};\Omega)$ into $L^2(\Omega)$, and from $H_0(\operatorname{div};\Omega)$ into $L^2_0(\Omega)$.*

*Proof.* The proof is quite simple and only the surjectivity needs to be checked. Let $q \in L^2(\Omega)$. If $p \in H^1(\Omega)$ is the solution of the following Dirichlet problem

$$\begin{aligned} \Delta p &= q, \quad \text{in } \Omega, \\ p &= 0, \quad \text{on } \partial\Omega, \end{aligned}$$

then $\mathbf{u} = \mathbf{grad}\,p \in H(\operatorname{div};\Omega)$ and $\operatorname{div}\mathbf{u} = q$. If, on the other hand, $q \in L^2_0(\Omega)$ and $p \in H^1(\Omega)$ is the solution of the following Neumann problem

$$\begin{aligned} \Delta p &= q, \quad \text{in } \Omega, \\ \frac{\partial p}{\partial n} &= 0, \quad \text{on } \partial\Omega, \end{aligned}$$

then $\mathbf{u} = \mathbf{grad}\,p \in H_0(\operatorname{div};\Omega)$ and $\operatorname{div}\mathbf{u} = q$ . □

## A.6 Positive Definite Problems

Let $H$ be a Hilbert with scalar product $(\cdot,\cdot)_H$ and corresponding induced norm $\|\cdot\|_H$. We introduce a real bilinear form $a(\cdot,\cdot) : H \times H \to \mathbb{R}$. We assume that $a(\cdot,\cdot)$ is *symmetric*,

$$a(u,v) = a(v,u), \quad u,v \in H,$$

*continuous*,

$$|a(u,v)| \le \alpha \|u\|_H \|v\|_H, \quad u,v \in H,$$

and *coercive* (or, equivalently, *elliptic*),

$$a(u,u) \geq \beta \|u\|_H^2, \quad u \in H, \quad \beta > 0.$$

We note that, since $a(\cdot,\cdot)$ is continuous, we can associate a continuous, linear operator $A: H \to H'$, with it, such that

$$\langle Au, v\rangle = a(u,v), \quad u,v \in H, \tag{A.23}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $H'$ and $H$.

*Remark A.33.* We can also define a different operator $\tilde{A}: H \to H$, such that

$$(\tilde{A}u, v)_H = a(u,v), \quad u,v \in H.$$

If $R: H' \to H$ is the linear map in the Riesz representation theorem (cf. [477, P. 90]), then $RA = \tilde{A}$.

The bilinear form $a(\cdot,\cdot)$ provides a second scalar product in $H$ since it is symmetric, continuous, and coercive; we denote the corresponding induced norm by $\|\cdot\|_a$. We have the equivalence

$$\beta \|u\|_H^2 \leq \|u\|_a^2 \leq \alpha \|u\|_H^2, \quad u \in H.$$

Given a linear functional $F \in H'$, we consider the following problem: find $u \in H$, such that

$$a(u,v) = \langle F, v\rangle, \quad v \in H. \tag{A.24}$$

We note that problem (A.24) can equivalently be written as

$$Au = F, \tag{A.25}$$

or

$$\tilde{A}u = RF$$

We have the following result, see, e.g., [477, Sect. III.7], [391, Sect. 5.1.1] or [188, Sect. 6.2.1].

**Lemma A.34 (Lax-Milgram lemma)** *Let $a(\cdot,\cdot)$ be a symmetric, continuous, and coercive bilinear form defined on a Hilbert space $H$ and let $F$ be a linear functional in $H'$. Then, problem (A.24) has a unique solution, satisfying*

$$\|u\|_H \leq (1/\beta)\|F\|_{H'},$$
$$\|u\|_a \leq (1/\sqrt{\beta})\|F\|_{H'},$$

*where $\beta$ is the coercivity constant. In addition, $u$ is the unique function in $H$, such that*

$$J(u) \leq J(v), \quad v \in H,$$

*with*

$$J(u) = \frac{1}{2}a(u,u) - \langle F, u\rangle.$$

*The operator $A$ defines an isomorphism from $H$ onto $H'$.*

### A.6.1 Scalar Problems

We now consider a general, second order, elliptic, scalar partial differential equation involving the operator

$$Lu = -\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}(a_{ij}\frac{\partial u}{\partial x_j}). \tag{A.26}$$

Let $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, be a Lipschitz region of unit diameter. We first impose homogeneous Dirichlet conditions on $\partial\Omega$. We look for a function $u \in H_0^1(\Omega)$, such that

$$\begin{aligned} Lu &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{A.27}$$

By multiplying the first equation of (A.27) by a test function $v$ and integrating by parts, we find the weak form of problem (A.27): find $u \in H_0^1(\Omega)$ such that

$$a(u, v) =< F, v >, \quad v \in H_0^1(\Omega), \tag{A.28}$$

where

$$a(u, v) = \sum_{i,j=1}^{n} \int_{\Omega} a_{ij}(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i}\, dx,$$

and

$$< F, v >= \int_{\Omega} f v\, dx.$$

We assume that the matrix $\mathcal{A} = [a_{ij}]$ is symmetric and uniformly bounded from above and below:

$$A_1\, \xi^T \xi \leq \xi^T \mathcal{A}(x)\, \xi \leq A_2\, \xi^T \xi, \quad \xi \in \mathbb{R}^n, \quad \text{almost all } x \in \Omega, \tag{A.29}$$

with $A_1 > 0$, and that the right hand side $f$ belongs to $H^{-1}(\Omega)$.

Our assumptions on the matrix $\mathcal{A}$ and the Friedrichs inequality in A.14 ensure that the bilinear form $a(\cdot, \cdot)$ is coercive and that the Lax-Milgram lemma A.34 therefore can be applied.

**Theorem A.35 (Dirichlet problem)** *There exists a unique $u \in H_0^1(\Omega)$ satisfying (A.28) and constants, such that*

$$\|u\|_{H^1(\Omega)} \leq C_1 \|F\|_{H^{-1}(\Omega)}, \quad \|u\|_a \leq C_2 \|F\|_{H^{-1}(\Omega)},$$

*with*

$$\|F\|_{H^{-1}(\Omega)} = \|f\|_{H^{-1}(\Omega)}.$$

*Remark A.36.* We can also consider a non-homogeneous Dirichlet condition

$$u = g_D \qquad \text{on } \partial\Omega,$$

with $g_D \in H^{1/2}(\partial\Omega)$. In this case, we look for a function $u \in H^1(\Omega)$ that is equal to $g_D$ on $\partial\Omega$ and that satisfies (A.28). The well-posedness of this problem can be proven by using the continuous lifting $\mathcal{R}_0 g_D$ of Lemma A.7 and then solving for $w = u - \mathcal{R}_0 g_D \in H_0^1(\Omega)$.

*Remark A.37.* We can also consider the more general case of mixed boundary conditions. Let $\partial\Omega_D \subset \partial\Omega$ be a closed set of positive measure, and let $\partial\Omega_N = \partial\Omega \setminus \partial\Omega_D$ be its complement. We impose homogeneous Dirichlet conditions on $\partial\Omega_D$ and a general Neumann condition

$$\sum_{i,j=1}^{n} a_{ij}\frac{\partial u}{\partial x_j}n_i = g_N \qquad \text{on } \partial\Omega_{\mathrm{N}},$$

where $n_i$ is the $i$-th component of the unit vector $\mathbf{n}$. In this case the solution $u$ belongs to $H_0^1(\Omega, \partial\Omega_D)$, the subspace of functions in $H^1(\Omega)$ that vanish on $\partial\Omega_D$, the right hand side $f$ is assumed to belong to the dual of $H_0^1(\Omega, \partial\Omega_D)$, which is a proper subspace of $H^{-1}(\Omega)$, and the Neumann data $g_N$ to $H_{00}^{-1/2}(\partial\Omega_N)$. The bilinear form $a(\cdot,\cdot)$ is the same as before and the right hand side is

$$< F, v >= \int_\Omega fv\,dx + \int_{\partial\Omega_N} g_N v\,ds,$$

where the integrals are understood as duality pairings. We have

$$\|F\|_{H'} \leq C\left(\|f\|_{H_0^1(\Omega,\partial\Omega_D)'} + \|g_N\|_{H_{00}^{-1/2}(\partial\Omega_N)}\right).$$

We next consider the Neumann problem

$$\begin{aligned} Lu &= f \qquad \text{in } \Omega, \\ \textstyle\sum_{i,j=1}^{n} a_{ij}\frac{\partial u}{\partial x_j}n_i &= g_N \qquad \text{on } \partial\Omega. \end{aligned} \tag{A.30}$$

It is clear that if there is a solution $u$ then it is defined up to an additive constant. Moreover, by integration by parts, we find that the following compatibility condition must hold:

$$\int_\Omega f\,dx + \int_{\partial\Omega} g_N\,ds = 0. \tag{A.31}$$

We then consider the quotient space $H = H^1(\Omega)/\mathbb{R}$, defined as a space of equivalence classes, where two functions in $H^1(\Omega)$ are equivalent if they differ by a constant. We note that in an equivalence class, we can always choose the function that has vanishing mean value on $\Omega$ as the representative of the class. The bilinear form $(u, v)_H = (\nabla u, \nabla v)$ uniquely defines a scalar product in $H$ and, by using the Poincaré inequality in Lemma A.13, the corresponding induced norm $\|u\|_H$ is equivalent to

$$\inf_{c\in\mathbb{R}} \|u - c\|_{H^1(\Omega)}.$$

We assume that $f$ belongs to the dual space of $H^1(\Omega)$ (a proper subspace of $H^{-1}(\Omega)$) and $g_N$ to $H^{-1/2}(\partial\Omega)$. With these assumptions and condition (A.31), the expression

$$\langle F, u\rangle = \int_\Omega f u\, dx + \int_{\partial\Omega} g_N u\, ds, \quad u \in H,$$

uniquely defines a linear functional of $H'$.

Using the same definition of $a(\cdot,\cdot)$, we consider the following variational problem: find $u \in H$, such that

$$a(u, v) =< F, v >, \quad v \in H. \tag{A.32}$$

The Lax-Milgram lemma A.34 ensures the well-posedness of problem (A.32):

**Theorem A.38 (Neumann Problem)** *Assume that $f$ and $g_N$ satisfy the regularity assumptions given above and that (A.31) holds. Then, there exists a unique $u \in H$ satisfying (A.32) and constants, such that*

$$\|u\|_H \le C_1 \|F\|_{H'}, \quad \|u\|_a \le C_2 \|F\|_{H'},$$

*with*

$$\|F\|_{H'} \le C \left( \|f\|_{H^1(\Omega)'} + \|g_N\|_{H^{-1/2}(\partial\Omega)} \right).$$

### A.6.2 Linear Elasticity

We will consider the equilibrium equations for a linear isotropic elastic material. We refer to [362] for a detailed analysis of general elastic problems.

Linear elastic materials are described by two generalized Lamé coefficients $\lambda(x)$ and $\mu(x)$. The pure displacement model is a good model for compressible materials, for which the Poisson ratio $\nu = \frac{\lambda}{2(\lambda+\mu)}$ is bounded away from $1/2$; we will consider almost incompressible elasticity in appendix A.7.2. We note that since $\lambda = \frac{2\nu}{1-2\nu}\mu$, then $\lambda$ is uniformly bounded in terms of $\mu$ if the Poisson ratio is bounded away from $1/2$.

Given a displacement vector field $\mathbf{u} \in H^1(\Omega)^n$, we denote its $i$-th component by $u_i$ and define the corresponding strain tensor by

$$\epsilon = \epsilon(\mathbf{u}) = \left[ \epsilon_{ij} = \frac{1}{2}\left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right].$$

We note that

$$\operatorname{div} \mathbf{u} = \sum_{i=1}^{n} \epsilon_{ii}(\mathbf{u})$$

and that $\epsilon$ is symmetric: $\epsilon_{ij} = \epsilon_{ji}$. In addition, $\epsilon(\mathbf{u}) = 0$ if and only if $\mathbf{u} \in \mathcal{RB}$, where

$$\begin{cases} \mathcal{RB} = \{\mathbf{a} + b[x_2, -x_1]^T\}, & n = 2, \\ \mathcal{RB} = \{\mathbf{a} + \mathbf{b} \times [x_1\ x_2\ x_3]^T\}, & n = 3 \end{cases}$$

is the space of rigid body modes. The dimension of $\mathcal{RB}$ is 3 for $n = 2$ and 6 for $n = 3$. Translations correspond to $b = 0$ or $\mathbf{b} = 0$, while rotations correspond to $\mathbf{a} = 0$.

We next define the stress tensor

$$\tau = \tau(\mathbf{u}) = 2\mu\,\epsilon(\mathbf{u}) + \lambda\,\mathrm{div}\,\mathbf{u}\,I,$$

where $I$ is the identity matrix in $\mathbb{R}^n$. The stress tensor is also symmetric. We introduce the second-order operator

$$L\mathbf{u} = -\mathrm{div}\,\tau(\mathbf{u}) = \left[-\sum_{j=1}^{n}\frac{\partial \tau_{ij}}{\partial x_j} = -\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(2\mu\,\epsilon_{ij}(\mathbf{u}) + \lambda\,\mathrm{div}\,\mathbf{u}\,\delta_{ij}\right)\right].$$

Given a load vector $\mathbf{f}$, we first consider homogeneous Dirichlet conditions on $\partial\Omega$: find $\mathbf{u}$, such that

$$\begin{aligned} L\mathbf{u} &= \mathbf{f} && \text{in } \Omega, \\ \mathbf{u} &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{A.33}$$

The weak form of problem (A.33) is: find $\mathbf{u} \in H_0^1(\Omega)^n$ such that

$$a(\mathbf{u},\mathbf{v}) = <\mathbf{F},\mathbf{v}>, \quad \mathbf{v} \in H_0^1(\Omega)^n, \tag{A.34}$$

where

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega (2\mu\epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\,dx + \lambda\,\mathrm{div}\,\mathbf{u}\,\mathrm{div}\,\mathbf{v})\,dx$$

and

$$<\mathbf{F},\mathbf{v}> = \int_\Omega \mathbf{f}\cdot\mathbf{v}\,dx,$$

with

$$\epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) := \sum_{i,j=1}^{n} \epsilon_{ij}(\mathbf{u})\,\epsilon_{ij}(\mathbf{v}).$$

We also introduce the quotient space $H = H^1(\Omega)^n/\mathcal{RB}$, defined as a space of equivalence classes, where two vectors in $H^1(\Omega)^n$ are equivalent if they differ by a rigid body mode. In analogy with the Friedrichs and Poincaré inequalities for scalar functions, we have two Korn inequalities for the strain tensor; see [362, Sect. 6.3].

**Lemma A.39 (Korn inequalities)** *Let $\Omega$ be a bounded Lipschitz domain. Then,*

$$|\mathbf{u}|^2_{H^1(\Omega)^n} \le 2\int_\Omega |\epsilon(\mathbf{u})|^2\,dx = 2\sum_{i,j=1}^{n}\int_\Omega \epsilon_{ij}(\mathbf{u})^2\,dx, \quad \mathbf{u} \in H_0^1(\Omega)^n. \tag{A.35}$$

*There exists a constant, depending only on $\Omega$, such that,*

$$\|\mathbf{u}\|^2_{H^1(\Omega)^n} \le C\int_\Omega |\epsilon(\mathbf{u})|^2\,dx, \quad \mathbf{u} \in H. \tag{A.36}$$

The first Korn inequality (A.35) and the Friedrichs inequality in Lemma A.14 allow us to apply the Lax-Milgram lemma to the Dirichlet problem (A.33).

**Theorem A.40 (Dirichlet Problem)** *Let* $\mathbf{f} \in H^{-1}(\Omega)^n$. *Then, there exists a unique* $\mathbf{u} \in H_0^1(\Omega)^n$ *satisfying (A.34) and constants, such that*

$$\|\mathbf{u}\|_{H^1(\Omega)^n} \leq C_1 \|\mathbf{F}\|_{H^{-1}(\Omega)^n}, \quad \|\mathbf{u}\|_a \leq C_2 \|\mathbf{F}\|_{H^{-1}(\Omega)^n},$$

*with*

$$\|\mathbf{F}\|_{H^{-1}(\Omega)^n} = \|\mathbf{f}\|_{H^{-1}(\Omega)^n}.$$

Neumann problems can also be considered

$$\begin{array}{ll} L\mathbf{u} = \mathbf{f} & \text{in } \Omega, \\ \tau(\mathbf{u})\,\mathbf{n} = \mathbf{g}_N & \text{on } \partial\Omega. \end{array} \tag{A.37}$$

It is clear that, if a solution $\mathbf{u}$ of problem (A.37) exists, then it is defined only up to a rigid body mode in $\mathcal{RB}$. Moreover, the following compatibility condition must hold

$$\int_\Omega \mathbf{f} \cdot \mathbf{r}\, dx + \int_{\partial\Omega} \mathbf{g}_N \cdot \mathbf{r}\, ds = 0, \quad \mathbf{r} \in \mathcal{RB}. \tag{A.38}$$

The bilinear form

$$(\mathbf{u}, \mathbf{v})_H = \int_\Omega \sum_{i,j=1}^{n} \epsilon_{ij}(\mathbf{u})\, \epsilon_{ij}(\mathbf{v})\, dx$$

defines a scalar product in $H$ and, by using the second Korn inequality (A.36), the corresponding induced norm $\|\mathbf{u}\|_H$ is equivalent to

$$\inf_{\mathbf{r} \in \mathcal{RB}} \|\mathbf{u} - \mathbf{r}\|_{H^1(\Omega)^n}.$$

We assume that $\mathbf{f}$ belongs to the dual space of $H^1(\Omega)^n$ and $\mathbf{g}_N$ to $H^{-1/2}(\partial\Omega)^n$. From these assumptions and condition (A.38), the expression

$$\langle \mathbf{F}, \mathbf{u} \rangle = \int_\Omega \mathbf{f} \cdot \mathbf{u}\, dx + \int_{\partial\Omega} \mathbf{g}_N \cdot \mathbf{u}\, ds, \quad \mathbf{u} \in H,$$

defines a linear functional of $H'$.

Using the same definition of $a(\cdot,\cdot)$, we consider the following variational problem: find $\mathbf{u} \in H$, such that

$$a(\mathbf{u}, \mathbf{v}) = < \mathbf{F}, \mathbf{v} >, \quad \mathbf{v} \in H. \tag{A.39}$$

The Lax-Milgram lemma and the second Korn inequality (A.36) ensure the well-posedness of problem (A.39):

**Theorem A.41 (Neumann Problem)** *Assume that* $\mathbf{f}$ *and* $\mathbf{g}_N$ *satisfy the regularity assumptions given above and that (A.38) holds. Then, there exists a unique* $\mathbf{u} \in H$ *satisfying (A.39) and constants, such that*

$$\|\mathbf{u}\|_H \leq C_1 \|\mathbf{F}\|_{H'}, \quad \|\mathbf{u}\|_a \leq C_2 \|\mathbf{F}\|_{H'},$$

*with*

$$\|\mathbf{F}\|_{H'} \leq C \left( \|\mathbf{f}\|_{(H^1(\Omega)^n)'} + \|\mathbf{g}_N\|_{H^{-1/2}(\partial\Omega)^n} \right).$$

*Remark A.42.* As for the case of the Laplace equation, non-homogeneous Dirichlet or mixed boundary conditions can also be considered.

### A.6.3 Problems in $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$

Given a bounded polyhedral domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, we introduce the boundary value problems

$$\begin{aligned} Lu = -\mathbf{grad}\,(a \,\mathrm{div}\,\mathbf{u}) + B\,\mathbf{u} &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{A.40}$$

for $n = 2, 3$,

$$\begin{aligned} Lu = \mathbf{curl}\,(a \,\mathrm{curl}\,\mathbf{u}) + B\,\mathbf{u} &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} \times \mathbf{n} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{A.41}$$

for $n = 2$, and

$$\begin{aligned} Lu = \mathbf{curl}\,(A\,\mathbf{curl}\,\mathbf{u}) + B\,\mathbf{u} &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} \times \mathbf{n} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{A.42}$$

for $n = 3$. The coefficient matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are symmetric uniformly positive definite with $a_{i,j}$ and $b_{i,j} \in L^\infty(\Omega)$, $1 \leq i, j \leq n$, and $a \in L^\infty(\Omega)$ is a positive function bounded away from zero.

For the weak formulation of problems (A.40), (A.41), and (A.42), we employ the Hilbert spaces $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$, defined in Appendix A.5 together with their scalar products and graph norms. We first consider problem (A.40). If we define the bilinear form

$$a_{\mathrm{div}}(\mathbf{u}, \mathbf{v}) = \int_\Omega (a \,\mathrm{div}\,\mathbf{u} \,\mathrm{div}\,\mathbf{v} + B\,\mathbf{u} \cdot \mathbf{v})\, dx, \quad \mathbf{u}, \mathbf{v} \in H(\mathrm{div}\,;\Omega), \tag{A.43}$$

for $\mathbf{f} \in L^2(\Omega)^n$, the variational formulation of Equation (A.40) is:
find $\mathbf{u} \in H_0(\mathrm{div}\,;\Omega)$ such that

$$a_{\mathrm{div}}(\mathbf{u}, \mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v}\, dx, \quad \mathbf{v} \in H_0(\mathrm{div}\,;\Omega). \tag{A.44}$$

For problem (A.42), we define the bilinear form

$$a_{\mathbf{curl}}(\mathbf{u},\mathbf{v}) = \int_\Omega (A\,\mathbf{curl}\,\mathbf{u}\cdot\mathbf{curl}\,\mathbf{v} + B\,\mathbf{u}\cdot\mathbf{v})\,dx, \quad \mathbf{u},\,\mathbf{v} \in H(\mathbf{curl};\Omega), \tag{A.45}$$

and, for $\mathbf{f} \in L^2(\Omega)^n$, the variational formulation of Equation (A.42) is:
find $\mathbf{u} \in H_0(\mathbf{curl};\Omega)$ such that

$$a_{\mathbf{curl}}(\mathbf{u},\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v}\,dx, \quad \mathbf{v} \in H_0(\mathbf{curl};\Omega). \tag{A.46}$$

For problem (A.41), an analogous definition in terms of the scalar curl holds. With an abuse of notation, we denote the corresponding bilinear form by $a_{\mathbf{curl}}(\cdot,\cdot)$.

We associate an energy norm, defined by $\|\cdot\|_a^2 = a_{\mathrm{div}}(\cdot,\cdot)$ or $\|\cdot\|_a^2 = a_{\mathbf{curl}}(\cdot,\cdot)$, with these bilinear forms; our assumptions on the coefficients guarantee that these norms is equivalent to the graph norms in $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$, respectively.

The Lax-Milgram lemma A.34 ensures that problems (A.44) and (A.46) are well posed:

**Theorem A.43** *Let the above assumptions on the coefficients A, B, and a be satisfied and* $\mathbf{f} \in L^2(\Omega)^n$. *Then, problems (A.44) and (A.46) have unique solutions.*

Considering problem (A.40), we note that if there is a $g$ such that $\mathbf{f} = -\nabla(ag)$, then (A.44) is equivalent to a mixed variational formulation of the following elliptic equation

$$\begin{aligned} -\mathrm{div}\,(B^{-1}\nabla w) + a^{-1}w &= g, \quad \text{in } \Omega,\\ B^{-1}\nabla w\cdot\mathbf{n} &= 0, \quad \text{on } \partial\Omega. \end{aligned}$$

To see this, we introduce a flux $\mathbf{q} = -B^{-1}\nabla w$ as an additional unknown. The corresponding mixed variational problem can be written as:
find $(\mathbf{q}, w) \in H_0(\mathrm{div};\Omega) \times L^2(\Omega)$ such that

$$\begin{aligned} &\int_\Omega B\mathbf{q}\cdot\mathbf{p}\,dx - \int_\Omega w\,\mathrm{div}\,\mathbf{p}\,dx = 0, \quad \mathbf{p} \in H_0(\mathrm{div};\Omega),\\ &\int_\Omega \mathrm{div}\,\mathbf{q}\,v\,dx + \int_\Omega a^{-1}wv\,dx = \int_\Omega gv\,dx, \quad v \in L^2(\Omega). \end{aligned} \tag{A.47}$$

The second equation gives

$$w = ag - a\mathrm{div}\,\mathbf{q}$$

and thus, using the first equation, $\mathbf{q} = \mathbf{u}$. We refer to Sect. A.7.2 for the analysis of the well-posedness of Problem (A.47).

Another application is provided by stabilized mixed formulations of the Stokes problem; see [95, Ch. IV] and the references therein. Still other applications of the space $H(\mathrm{div}\,;\Omega)$ are given in [21].

As for problems (A.41) and (A.42), when time-dependent Maxwell's equations are considered, the electric field $\mathbf{u}$ satisfies the following equation

$$\mathbf{curl}\,(\mu^{-1}\mathrm{curl}\,\mathbf{u}) + \varepsilon\frac{\partial^2\mathbf{u}}{\partial t^2} + \sigma\frac{\partial\mathbf{u}}{\partial t} = -\frac{\partial\mathbf{J}}{\partial t}, \quad \text{in } \Omega. \tag{A.48}$$

Here $\mathbf{J}(\mathbf{x},t)$ is the current density and $\varepsilon$, $\mu$, $\sigma$ are non-negative functions that describe the electromagnetic properties of the medium. For their meaning and for a general discussion of Maxwell's equations, see [151, 354]. A similar equation holds for the magnetic field. For a perfect conducting boundary, the electric field satisfies the essential boundary condition

$$\mathbf{u}\times\mathbf{n} = 0, \quad \text{on } \partial\Omega. \tag{A.49}$$

Natural boundary conditions

$$\mathbf{curl}\,\mathbf{u}\times\mathbf{n} = 0, \quad \text{on } \partial\Omega,$$

can also be considered; see [151, 354]. For low-frequency fields in conducting media, the term involving the second derivative in time in (A.48) can be neglected, and a parabolic equation is obtained, known as an eddy current problem. Variational problems as (A.46), involving the bilinear form $a_{\mathbf{curl}}(\cdot,\cdot)$, arise, for instance, when Equation (A.48) is discretized with a finite difference scheme which is implicit in time.

## A.7 Non-Symmetric and Indefinite Problems

### A.7.1 Generalizations of the Lax-Milgram Lemma

The Lax-Milgram lemma can be generalized to more general problems. Let $H$ be a Hilbert space with scalar product $(\cdot,\cdot)_H$ and corresponding induced norm $\|\cdot\|_H$. We introduce a complex sesquilinear form $b(\cdot,\cdot) : H\times H\to\mathbb{C}$. We assume that $b(\cdot,\cdot)$ is *continuous*,

$$|b(u,v)| \le \alpha\|u\|_H\|v\|_H, \quad u,v\in H,$$

and *coercive*,

$$|b(u,u)| \ge \beta\|u\|_H^2, \quad u\in H, \quad \beta>0.$$

We note that $b(\cdot,\cdot)$ is not required to be symmetric and that we cannot in general associate a scalar product with it. We can still define operators $B$ and $\tilde{B}$; cf. Equation (A.23) and Remark A.33.

We have the following generalization, see, e.g., [391, Sect. 5.1.1]:

**Lemma A.44 (Lax-Milgram lemma)** *Let $b(\cdot,\cdot)$ be a continuous and coercive sesqilinear form defined on a Hilbert space $H$ and let $F$ be a linear functional in $H'$. Then, problem (A.24) has a unique solution, satisfying*

$$\|u\|_H \leq (1/\beta)\,\|F\|_{H'}.$$

*The operator $B$ therefore defines an isomorphism from $H$ onto $H'$.*

There is however a wide class of problems that are not coercive and to which the Lax-Milgram lemma cannot be applied. Given a bounded polygonal or polyhedral domain $\Omega$, we consider the homogeneous Dirichlet boundary value problem: find $u \in H_0^1(\Omega)$, such that

$$\begin{aligned} Lu &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{A.50}$$

where the operator $L$ has the form

$$Lu = -\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}(a_{ij}\frac{\partial u}{\partial x_j}) + 2\sum_{i=1}^{d} b_i \frac{\partial u}{\partial x_i} + c\,u. \tag{A.51}$$

We make the following assumptions on the coefficients

$$a_{ij} \in L^\infty(\Omega), \quad b_i \in W^{1,\infty}(\Omega), \quad c \in L^\infty(\Omega), \qquad 1 \leq i,j \leq n. \tag{A.52}$$

In addition, the matrix $\mathcal{A} = [a_{ij}]$ is symmetric and uniformly positive definite in $\Omega$: cf. (A.29). The right hand side $f \in L^2(\Omega)$.

The weak form of Equation (A.50) is: find $u \in H_0^1(\Omega)$ such that

$$b(u,v) = (f,v)_{L^2(\Omega)}, \quad v \in H_0^1(\Omega). \tag{A.53}$$

The bilinear form $b(u,v)$ is defined by

$$b(u,v) \;=\; \sum_{i,j=1}^{n} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx + \sum_{i=1}^{n} 2\int_\Omega b_i \frac{\partial u}{\partial x_i} v dx + \int_\Omega cuvdx \tag{A.54}$$

or

$$b(u,v) \;=\; \sum_{i,j=1}^{n} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx + \sum_{i=1}^{n} \int_\Omega \left( b_i \frac{\partial u}{\partial x_i} v - b_i \frac{\partial v}{\partial x_i} u \right) dx + \int_\Omega \tilde{c}uvdx.$$

Here, $\tilde{c}(x) = c(x) - \sum_{i=1}^{n} \partial b_i(x)/\partial x_i$.

We also consider the adjoint problem: find $w \in H_0^1(\Omega)$ such that

$$b(v,w) = (f,v)_{L^2(\Omega)}, \quad v \in H_0^1(\Omega). \tag{A.55}$$

This problem corresponds to a homogeneous Dirichlet problem for the adjoint operator $L^*$, defined by

$$(L^*u, v)_{L^2(\Omega)} = b(v,u) = (u, Lv)_{L^2(\Omega)}, \quad u,v \in H_0^1(\Omega).$$

Under particular assumptions on $\tilde{c}$ the bilinear form is coercive and the Lax-Milgram lemma A.44 can be applied:

**Lemma A.45** *If $\tilde{c} \geq 0$, then $b(\cdot,\cdot)$ is coercive and problem (A.53) is well posed.*

In the more general case, a *Gårding inequality* holds:

$$\beta \|u\|^2_{H^1(\Omega)} \leq |b(u,u)| + \gamma \|u\|^2_{L^2(\Omega)}, \quad u \in H^1_0(\Omega), \tag{A.56}$$

for suitable $\beta > 0$ and $\gamma \in \mathbb{R}$.

The following result relies on the Gårding inequality (A.56) and the compactness of the embedding $H^1(\Omega) \subset L^2(\Omega)$: cf. Lemma A.4. We refer to, e.g., [188, Sect. 6.2.3] for a proof.

**Lemma A.46 (Fredholm Alternative)** *Let the Gårding's inequality (A.56) hold. Then,*

1. *For every $\mu \geq \gamma$, the problem of finding $u \in H^1_0(\Omega)$, such that,*

$$b(u,v) + \mu\,(u,v)_{L^2(\Omega)} = (f,v)_{L^2(\Omega)}, \quad v \in H^1_0(\Omega)$$

   *is well-posed.*
2. *Either problem (A.53) has a unique solution for every $f \in L^2(\Omega)$, such that,*

$$\|u\|_{H^1(\Omega)} \leq C \|f\|_{L^2(\Omega)},$$

   *or there exists a non-trivial solution $u \neq 0$ in $H^1_0(\Omega)$ for $f = 0$.*
3. *The solutions of (A.53) for $f = 0$ form a finite dimensional space $N \subset H^1_0(\Omega)$. Its dimension is equal to the dimension of $N^* \subset H^1_0(\Omega)$ of solutions of the adjoint problem (A.55) for $f = 0$.*
4. *Problem (A.53) has a solution if and only if*

$$(f,v)_{L^2(\Omega)} = 0, \quad v \in N^*.$$

### A.7.2 Saddle-Point Problems

Let $V$ and $Q$ two Hilbert spaces, endowed with scalar products $(\cdot,\cdot)_V$ and $(\cdot,\cdot)_Q$, and corresponding induced norms $\|\cdot\|_V$ and $\|\cdot\|_Q$, respectively.

Given bilinear forms $a : V \times V \to \mathbb{R}$ and $b : V \times Q \to \mathbb{R}$, and functionals $F \in V'$ and $G \in Q'$, we consider the problem: find $(u,p) \in V \times Q$, such that,

$$\begin{aligned} a(u,v) + b(v,p) &= F(v), \quad v \in V, \\ b(u,q) \qquad\quad &= G(q), \quad q \in Q. \end{aligned} \tag{A.57}$$

If $b(\cdot,\cdot)$ is continuous, we can associate a continuous, linear operator $B : V \to Q'$, with it, such that

$$\langle Bu, q\rangle = b(u,q), \qquad u \in V, \quad q \in Q.$$

We make some assumptions on the bilinear forms.

(i) $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are continuous

$$\begin{aligned} |a(u,v)| &\leq \alpha_1 \|u\|_V \, \|v\|_V, \quad u,v \in V \\ |b(u,p)| &\leq \alpha_2 \|u\|_V \, \|p\|_Q, \quad u \in V, \quad p \in Q. \end{aligned} \tag{A.58}$$

(ii) $a(\cdot,\cdot)$ is coercive on $Z$:

$$a(u,u) \geq \beta \|u\|_V^2, \quad u \in Z, \tag{A.59}$$

where $Z \subset V$ is defined as

$$Z = ker(B) = \{v \in V \mid \quad b(v,q) = 0, \; q \in Q\}.$$

(iii) $b(\cdot,\cdot)$ satisfies the inf-sup condition

$$\inf_{0 \neq q \in Q} \sup_{0 \neq v \in V} \frac{b(v,q)}{\|v\|_V \, \|q\|_Q} \geq \gamma > 0. \tag{A.60}$$

We have the following result; see [95, Sect. II.1].

**Lemma A.47** *Let $F \in V'$ and $G \in Range(B)$. If conditions (i), (ii), and (iii) hold, then there exists a unique solution to the mixed problem (A.57) and constants, depending only on $\alpha_1$, $\alpha_2$, $\beta$, and $\gamma$, such that*

$$\|u\|_V \leq C_1(\|F\|_{V'} + \|G\|_{Q'}), \quad \|p\|_Q \leq C_2(\|F\|_{V'} + \|G\|_{Q'}).$$

*If $G = 0$, the solution $u$ is the unique function in $Z \subset V$ that satisfies*

$$a(u,v) = F(v), \quad v \in Z,$$

*and, if $a(\cdot,\cdot)$ is symmetric, it is the unique function in $Z$, such that*

$$J(u) \leq J(v), \quad v \in Z,$$

*with*

$$J(u) = \frac{1}{2} a(u,u) - \langle F, u \rangle.$$

Given a continuous, symmetric, positive semi-definite bilinear form

$$c : Q \times Q \to \mathbb{R},$$

we can also consider the problem

$$\begin{aligned} a(u,v) + b(v,p) &= F(v), \quad v \in V, \\ b(u,q) \; - c(p,q) &= G(q), \quad q \in Q. \end{aligned} \tag{A.61}$$

Under the the same assumptions (i), (ii), and (iii), one can show that problem (A.61) is well posed; see [95, Sect. II.1.2]. We note in particular that if

$$c(\cdot,\cdot) = \epsilon \, (\cdot,\cdot)_Q, \quad \epsilon \geq 0,$$

the constants appearing in the stability estimates are independent of $\epsilon$.

**The Stokes System and Nearly Incompressible Elasticity**

Given a viscosity $\nu > 0$, $\mathbf{f} \in H^{-1}(\Omega)^n$, and $\mathbf{g} \in H^{1/2}(\partial\Omega)^n$, such that

$$\int_{\partial\Omega} \mathbf{g} \cdot \mathbf{n}\, ds = 0,$$

we consider the Stokes problem: find $(\mathbf{u}, p) \in H^1(\Omega)^n \times L_0^2(\Omega)$, such that $\mathbf{u} = \mathbf{g}$ on $\partial\Omega$ and

$$\begin{aligned} -\nu\Delta\mathbf{u} + \nabla p &= \mathbf{f}, \\ \nabla \cdot \mathbf{u} \qquad\quad &= 0. \end{aligned} \tag{A.62}$$

Here the space $L_0^2(\Omega)$ was defined in (A.22). This problem can be rewritten in the mixed variational form

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) &= F(\mathbf{v}), \quad \mathbf{v} \in V, \\ b(\mathbf{u}, q) \qquad\quad &= G(q), \quad q \in Q, \end{aligned} \tag{A.63}$$

by defining

$$a(\mathbf{u}, \mathbf{v}) = \nu \int_\Omega \nabla\mathbf{u} : \nabla\mathbf{v}\, dx, \quad b(\mathbf{u}, p) = -\int_\Omega \nabla \cdot \mathbf{u}\, p\, dx,$$
$$F(\mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v}\, dx, \quad G(q) = 0,$$

and

$$V = H_0^1(\Omega)^n, \quad Q = L_0^2(\Omega).$$

Here, given two vectors $\mathbf{u}$ and $\mathbf{v}$, we have used the notation

$$\nabla\mathbf{u} : \nabla\mathbf{v} = \sum_{i,j=1}^{n} \frac{\partial u_i}{\partial x_j} \frac{\partial u_i}{\partial x_j}.$$

Lemma A.47 and the theory of the previous section can be applied to the Stokes system (A.63). In particular, we note that by the Friedrichs inequality in Lemma A.14, the bilinear form $a(\cdot, \cdot)$ is coercive on the whole space $V$. In addition, an inf-sup condition holds for the bilinear form $b(\cdot, \cdot)$, since, for every $q \in L_0^2(\Omega)$, it is possible to find $\mathbf{u} \in H_0^1(\Omega)^n$, such that

$$-\nabla \cdot \mathbf{u} = p, \qquad |\mathbf{u}|_{H^1(\Omega)^n} \leq \gamma^{-1} \|p\|_{L^2(\Omega)}; \tag{A.64}$$

see [95].

Nearly incompressible elasticity problems give rise to a similar mixed system:

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) \quad &= F(\mathbf{v}), \quad \mathbf{v} \in V, \\ b(\mathbf{u}, q) \; - t^2 c(p, q) &= G(q), \quad q \in Q, \end{aligned} \tag{A.65}$$

by defining, for $\mu > 0$,

$$a(\mathbf{u},\mathbf{v}) = 2\mu \int_\Omega \epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\, dx, \qquad c(p,q) = \int_\Omega pq\, dx, \qquad G(q) = 0$$

and with $b(\cdot,\cdot)$, $F(\cdot)$, $V$, and $Q$ as before. The strain tensor $\epsilon(\cdot)$ is defined in Sect. A.6.2 and the incompressible limit corresponds to the value $t = 0$. We note that this problem can be obtained from the linear elasticity problem (A.33), by setting $p = \lambda \operatorname{div} \mathbf{u}$. We have $t^2 = (1/\lambda)$ and the limit case $t \to 0$ corresponds to the incompressible limit $\lambda \to \infty$, when the Poisson ratio $\nu$ approaches $1/2$.

The coercivity of $a(\cdot,\cdot)$ on the whole $V$ is ensured by the first Korn inequality (A.35) and an inf-sup condition for $b(\cdot,\cdot)$ follows from the previous analysis of the Stokes problem. The framework of the previous section can then be applied in order to prove the well-posedness of problem (A.65) and the constants involved in the estimates are independent of $t \geq 0$, or, equivalently, of $\lambda > 0$.

### Mixed Formulation of the Laplace Equation: Flows in Porous Media

In certain application involving the diffusive equations of section A.6.1, it is important to have a reliable approximation of the flux itself. If $\mathcal{A} = [a_{ij}]$ is the diffusive matrix of the operator $L$ defined in (A.26), we can reformulate the scalar elliptic problem as

$$\begin{aligned} \mathbf{u} &= -\mathcal{A}\nabla p, \\ \nabla \cdot \mathbf{u} &= f, \\ (-\mathcal{A}\nabla p)\cdot \mathbf{n} &= -\textstyle\sum_{i,j=1}^n a_{ij} \dfrac{\partial p}{\partial x_j} n_i = \mathbf{u}\cdot\mathbf{n}, \end{aligned} \tag{A.66}$$

with $f \in L^2(\Omega)$. The first equation of (A.66) is called *Darcy's law* and relates the pressure $p$ to the velocity $\mathbf{u}$. Since $\mathcal{A}$ is uniformly positive definite because of (A.29), we can multiply the first equation by $\mathcal{A}^{-1}$, and then both equations by test functions, and integrate over $\Omega$. We obtain

$$\begin{aligned} \int_\Omega (\mathcal{A}^{-1}\mathbf{u})\cdot\mathbf{v}\, dx - \int_\Omega \nabla\cdot\mathbf{v}\, p\, dx &= -\int_{\partial\Omega} \mathbf{v}\cdot\mathbf{n}\, p\, ds \\ -\int_\Omega \nabla\cdot\mathbf{u}\, q\, dx \qquad\qquad &= \qquad -\int_\Omega f q\, dx. \end{aligned} \tag{A.67}$$

We then define the forms

$$\begin{aligned} a(\mathbf{u},\mathbf{v}) &= \int_\Omega (\mathcal{A}^{-1}\mathbf{u})\cdot\mathbf{v}\, dx, \quad b(\mathbf{u},p) = -\int_\Omega \nabla\cdot\mathbf{u}\, p\, dx, \\ G(q) &= -\int_\Omega f q\, dx, \end{aligned} \tag{A.68}$$

and the spaces

$$V = H(\mathrm{div}\,;\Omega), \quad Q = L^2(\Omega). \tag{A.69}$$

We note that the bilinear form $b(\cdot,\cdot)$ is the same as that of the Stokes problem in the previous section but it acts on the larger space $V$.

We start by considering general *Neumann conditions* for the function $p$:

$$p = g_N \qquad \text{on } \partial\Omega,$$

with $g_N \in H^{1/2}(\partial\Omega)$. Substituting this expression into the first equation of (A.67), gives rise to a problem of the form (A.63), with

$$F(\mathbf{v}) := -\int_{\partial\Omega} \mathbf{v}\cdot\mathbf{n}\, g_N\, ds. \tag{A.70}$$

We note that $F(\cdot)$ is well-defined in $V$ because of Lemma A.19.

The well-posedness of problem (A.63) is ensured by Lemma A.47. More precisely, the continuity conditions in (A.58) clearly hold. We next note that the operator $B : V \to Q$, defined by the bilinear form $b(\cdot,\cdot)$ coincides with the divergence operator and therefore

$$Z = ker(B) = H(\mathrm{div}_0;\Omega).$$

The bilinear form $a(\cdot,\cdot)$ is thus coercive on $Z$ with $\beta \geq A_2^{-1}$; see (A.59) and (A.29). The inf-sup condition (A.60) follows from the surjectivity of the divergence operator; see Lemma A.32 and its proof. The same lemma ensures that $G \in Range(B)$ for $f \in L^2(\Omega)$.

We next consider the *Dirichlet condition*

$$-\sum_{i,j=1}^{n} a_{ij}\frac{\partial p}{\partial x_j} n_i = \mathbf{u}\cdot\mathbf{n} = g_D,$$

with $g_D \in H^{-1/2}(\partial\Omega)$. We assume that $g_D$ and $f$ satisfy the compatibility condition (A.31). In this case, the solution $p$ is defined up to an additive constant and we therefore require that it has mean value zero on $\Omega$. The appropriate spaces that ensure the well-posedness of this problems are thus

$$V = H_0(\mathrm{div}\,;\Omega), \quad Q = L^2_0(\Omega). \tag{A.71}$$

We then look for functions $\mathbf{u} \in H(\mathrm{div}\,;\Omega)$ and $p \in Q$, such that $\mathbf{u}\cdot\mathbf{n} = g_D$ on $\partial\Omega$ and (A.63) is satisfied with

$$F(\mathbf{v}) = 0. \tag{A.72}$$

The well-posedness of this problem can be proven as before.

We remark that essential (Dirichlet) boundary condition on the scalar function $p$ in the original scalar elliptic problem give rise to a mixed problem with natural boundary conditions (Neumann) on the flux $\mathbf{u}$ and viceversa. We

finally remark that a more general diffusion-reaction problem can be considered:

$$-\nabla \cdot (\mathcal{A}\nabla p) + cp = f.$$

With the same expression for the flux $\mathbf{u}$, the corresponding mixed formulation has the form (A.61); see also Sect. A.6.3 and Problem (A.47).

*Remark A.48.* We can also consider the more general case of mixed boundary conditions. Let $\partial\Omega_D \subset \partial\Omega$ be a closed set of positive measure, and let $\partial\Omega_N = \partial\Omega \backslash \partial\Omega_D$ be its complement. We impose, for simplicity, homogeneous Dirichlet conditions on $\partial\Omega_D$ and a general Neumann condition $p = g_N$ on $\partial\Omega_N$. In this case the solution $\mathbf{u}$ and the test functions $\mathbf{v}$ belong to $H_0(\mathrm{div}; \Omega, \partial\Omega_D)$, the subspace of vectors in $H(\mathrm{div}; \Omega)$ with vanishing normal component on $\partial\Omega_D$ and the Neumann data $g_N$ is supposed to belong to $H^{1/2}(\partial\Omega_N)$. In this case,

$$F(\mathbf{v}) = -\int_{\partial\Omega_N} \mathbf{v} \cdot \mathbf{n}\, g_N \, dx,$$

and no compatibility condition on $f$ is required, and the pressure is uniquely determined.

## A.8 Regularity Results

In this section, we will give some regularity results for scalar and vector-valued functions. We only report on results that we need in this monograph.

The following regularity result for the Laplace operator can be found by combining [238, Corollary 2.6.7] with the existence of the extension operator given in Lemma A.7. It provides a regularity result for the Poisson problem with nonhomogeneous Dirichlet conditions.

**Lemma A.49** *Let $\Omega \subset \mathbb{R}^3$ be a bounded polyhedron with Lipschitz continuous boundary. Let, in addition, $f \in L^2(\Omega)$ and $\phi \in H^{1+s}(\partial\Omega)$. Then, there exists $s_\Omega > 0$, such that, if $s < s_\Omega$, the solution $u$ of*

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u &= \phi \qquad \text{on } \partial\Omega, \end{aligned}$$

*belongs to $H^{3/2+s}(\Omega)$.*

The result can be improved for convex domains; see [150, Corollary 18.18].

**Lemma A.50** *Let $\Omega \subset \mathbb{R}^3$ be a bounded, open, convex polyhedron and let $s \neq -\frac{1}{2}$ be a real number, such that*

$$s < \min\left\{\frac{3}{2}, \frac{\pi}{\omega} - 1\right\}, \tag{A.73}$$

*where $\omega$ is the largest angle between the faces of $\Omega$. Then, the Laplace operator $\Delta$ defines an isomorphism:*

$$\Delta : H^{2+s}(\Omega) \cap H_0^1(\Omega) \longleftrightarrow H^s(\Omega). \tag{A.74}$$

*Remark A.51.* Since, for every fixed bounded, convex polyhedron the maximum angle $\omega$ is strictly smaller than $\pi$, Lemma A.50 implies that there exists a real number $s_\Omega > 0$, such that the mapping (A.74) is an isomorphism, for any $s \in [0, s_\Omega)$.

The following corollary is a direct consequence of Lemma A.50 and the existence of the extension operator given in Lemma A.7.

**Corollary A.52** *Let $\Omega \subset \mathbb{R}^3$ be a polyhedron satisfying the same assumptions as in Lemma A.50 and let $s < s_\Omega$, $s \neq -1/2$, with $s_\Omega$ given in Remark A.51. Let, in addition, $f \in H^s(\Omega)$ and $\phi \in H^{3/2+s}(\partial\Omega)$. Then, the solution $u$ of*

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= \phi && \text{on } \partial\Omega, \end{aligned}$$

*belongs to $H^{2+s}(\Omega)$.*

The following regularity result for the Laplace operator can be found in [150, Corollary 23.5]. It provides a regularity result for the Laplace problem (A.30) with a Neumann boundary condition. See also [238, Corollary 2.6.7].

**Lemma A.53** *Let $\Omega \subset \mathbb{R}^3$ be a bounded polyhedron with a Lipschitz continuous boundary, not necessarily convex. Let, in addition, $f \in H^{-1/2+s}(\Omega)$ and $\psi \in H^s(\partial\Omega)$, satisfy the compatibility condition (A.31). Then, there exists $s'_\Omega > 0$, such that, if $s \neq 0$ and $s < s'_\Omega$, the solution $u$ of*

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ \frac{\partial u}{\partial n} &= \psi && \text{on } \partial\Omega, \end{aligned}$$

*belongs to $H^{3/2+s}(\Omega)$.*

We finally provide some results on $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$. Given the spaces

$$\begin{aligned} H_T(\Omega) &= H_0(\mathrm{div}\,;\Omega) \cap H(\mathbf{curl}\,;\Omega), \\ H_N(\Omega) &= H(\mathrm{div}\,;\Omega) \cap H_0(\mathbf{curl}\,;\Omega), \end{aligned}$$

the following proposition is a classical result and a proof can be found in [18, Th. 2.17].

**Lemma A.54** *If the domain $\Omega$ is convex, then the spaces $H_T(\Omega)$ and $H_N(\Omega)$ are continuously embedded in $H^1(\Omega)^n$.*

We remark that the conclusion of Lemma A.54 is in general false for a nonconvex polyhedron.

# B

# Galerkin Approximations

In this appendix, we recall some well-known results on finite and spectral element approximations. We mention [136, 391, 90, 50] as general references, [121, 48] for spectral element approximations, [223, 95] for approximations of saddle-point problems, and [223, 95, 359, 391, 354] for conforming approximations in $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$.

## B.1 Finite Element Approximations

### B.1.1 Triangulations

Let $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, be a bounded polygonal or polyhedral domain with Lipschitz continuous boundary.

A *triangulation* (or, equivalently, mesh) is a non-overlapping partition of $\Omega$ into *elements*. We consider meshes consisting of triangles or affinely mapped rectangles in two dimensions, and of tetrahedra or affinely mapped parallelepipeds in three dimensions. More precisely, let the *reference* triangle (tetrahedron) have vertices $(0,0)$, $(0,1)$, $(1,0)$ (or $(0,0,0)$, $(0,0,1)$, $(0,1,0)$, $(1,0,0)$, respectively). The *reference* square and cube are $(-1,1)^n$. Throughout this monograph, a reference element $\hat{K}$ is one of the four regions defined above and elements are always open sets. An *affine* mapping from $\hat{K}$ onto an element $K$ is defined by

$$F_K : \hat{K} \to K, \quad F_K(x) = B_K x + \mathbf{b}_K,$$

with $B_K$ a linear mapping and $\mathbf{b}_K$ a constant vector. We define a family of triangulations $\mathcal{T}_h$, $h > 0$:

**Definition B.1.** *Let $h > 0$. A family of triangulations of $\Omega$ is a partition of $\Omega$*

$$\mathcal{T}_h = \{K = F_K(\hat{K})\},$$

*such that,*

$$\bigcup_{K\in\mathcal{T}_h} \overline{K} = \overline{\Omega}; \qquad K \cap K' = \emptyset \quad \text{if } K \neq K';$$

$\hat{K}$ is a reference element and $F_K$ is an affine mapping;

$$h = \max_{K\in\mathcal{T}_h} h_K, \quad h_K = \operatorname{diam}(K).$$

*$h$ is called the diameter of $\mathcal{T}_h$. The family $\mathcal{T}_h$ is called* geometrically conforming *(briefly,* conforming*), if the intersection between the closure of two different elements is either empty, a vertex, an edge, or a face that is common to both elements.*

We consider particular triangulations.

**Definition B.2.** *A family of triangulations $\mathcal{T}_h$ is called* shape-regular *if there exists a constant independent of $h$, such that*

$$h_K \leq C\rho_K, \quad K \in \mathcal{T}_h,$$

*where $\rho_K$ is the radius of the largest circle or sphere contained in $K$. The ratio $h_K/\rho_K$ is called the* aspect ratio *of $K$.*

**Definition B.3.** *A family of triangulations $\mathcal{T}_h$ is called* quasi-uniform *if it is shape-regular and if there exists a constant independent of $h$, such that*

$$h_K \geq Ch, \quad K \in \mathcal{T}_h.$$

### B.1.2 Finite Element Spaces

Given an open set $\mathcal{D}$ in $\mathbb{R}^n$, $n = 1, 2, 3$, we now define some polynomial spaces. Let $\mathbb{P}_k(\mathcal{D})$, $k \geq 0$, be the set of polynomials of total degree at most $k$ defined on $\mathcal{D}$, and let $\mathbb{P}_k(\mathcal{D})^n$, for $n = 2, 3$, be the set of vectors of $\mathbb{R}^n$, the components of which belong to $\mathbb{P}_k(\mathcal{D})$. In addition, let $\mathbb{Q}_k(\mathcal{D})$ be the set of polynomials of degree at most $k$ in each variable.

Let $\mathcal{T}_h$ be a conforming triangulation. We have the following result; cf., e.g., [391, Pr. 3.2.1].

**Lemma B.4** *A function $u : \Omega \to \mathbb{R}$ belongs to $H^1(\Omega)$ if and only if the restriction of $u$ to every $K \in \mathcal{T}_h$ belongs to $H^1(K)$, and, for each common face (or edge in two dimensions) $f = \overline{K_1} \cap \overline{K_2}$, we have*

$$u_{|K_1} = u_{|K_2}, \quad \text{on } f.$$

Finite element spaces of continuous, piecewise polynomial functions are therefore contained in $H^1(\Omega)$. For $k \geq 1$, we define (see [391, Sect. 3.2])

$$V^h = V_k^h(\Omega) := \left\{u \in C^0(\Omega) \mid u_{|K} \in \mathbb{P}_k(K),\ K \in \mathcal{T}_h\right\},$$
$$V_0^h = V_{k;0}^h(\Omega) := V_k^h(\Omega) \cap H_0^1(\Omega),$$

if $\mathcal{T}_h$ consists of triangles or tetrahedra, and

$$V^h = V_k^h(\Omega) := \left\{u \in C^0(\Omega) \mid u_{|_K} \in \mathbb{Q}_k(K),\ K \in \mathcal{T}_h\right\},$$
$$V_0^h = V_{k;0}^h(\Omega) := V_k^h(\Omega) \cap H_0^1(\Omega),$$

if $\mathcal{T}_h$ is made of affinely mapped rectangles or parallelepipeds.

For a fixed polynomial degree $k$, the set of Lagrangian basis functions $\{\phi_i^h\}$ associated to a set of nodes $\{P_i\}$ of the triangulation can be introduced. The degrees of freedom are then the values of a function at these nodes. We have

$$u(x) = \sum_i u(P_i)\,\phi_i^h(x), \quad u \in V^h,$$

and the basis functions are uniquely defined by

$$\phi_i^h(P_j) = \delta_{ij}.$$

There is of course a one-to-one correspondence between functions in $V^h$ and vectors of degrees of freedom. Throughout this monograph, we use the same notation for finite element functions $u$ and vectors of degrees of freedom, and for finite element spaces and spaces of vectors of degrees of freedom.

The support of the nodal basis function $\phi_i^h$ is contained in the union of the elements that share the node $P_i$. A scaling argument allows us to prove the following property; see [391, Prop. 3.4.1].

**Lemma B.5** *Let $\phi_i^h$ be a basis function associated to a node of $K \in \mathcal{T}_h$. Then there exist constants independent of $h_K$, and $h$, such that*

$$c_1 h_K^n \le \|\phi_i^h\|_{L^2(K)}^2 \le C_1 h_K^n,$$
$$c_2 h_K^{n-2} \le |\phi_i^h|_{H^1(K)}^2 \le C_2 h_K^{n-2},$$
$$c_3 h_K^{n-1} \le |\phi_i^h|_{H^{1/2}(K)}^2 \le C_3 h_K^{n-1},$$

*where $C_1$ is also independent of the aspect ratio of $K$.*

A nodal interpolation operator $I^h = I_k^h$ can be defined for functions that are continuous in $\overline{\Omega}$ by

$$I^h u = \sum_i u(P_i)\,\phi_i^h, \quad u \in C^0(\overline{\Omega});$$

see Lemma A.5. Error estimates can also be found; see [391, Sect. 3.4.1].

**Lemma B.6** *Given a mesh $\mathcal{T}_h$, for $u \in H^s(\Omega)$ and*

$$K \in \mathcal{T}_h, \quad \frac{n}{2} < s \le k+1, \quad 0 \le m \le s,$$

*there exists a constant, depending only on $m$, $s$, and the aspect ratio of $K$, such that,*

$$|u - I_k^h u|_{H^m(K)} \leq C\, h_K^{s-m} |u|_{H^s(K)}.$$

*If $\mathcal{T}_h$ is conforming and shape-regular, we have*

$$|u - I_k^h u|_{H^m(\Omega)} \leq C\, h^{s-m} |u|_{H^s(\Omega)}.$$

We also need some finite element spaces that consist of discontinuous functions and are conforming in $L^2(\Omega)$. For $k \geq 0$, see [391, Sect. 3.2],

$$Q^h = Q_k^h(\Omega) := \left\{ u \in L^2(\Omega) \mid u_{|_K} \in \mathbb{P}_k(K),\ K \in \mathcal{T}_h \right\},$$
$$Q_0^h = Q_{k;0}^h(\Omega) := Q_k^h(\Omega) \cap L_0^2(\Omega),$$

if $\mathcal{T}_h$ is made of triangles or tetrahedra, and an analogous definition holds if $\mathcal{T}_h$ is made of rectangles or parallelepipeds.

In the remainder of this appendix, we will always assume that triangulations $\mathcal{T}_h$ are conforming, unless it is otherwise stated, in such a way that the finite element spaces introduced in this section are well defined.

### B.1.3 Symmetric, Positive Definite Problems

We first consider the finite element spaces $V^h = V_k^h(\Omega)$ and $V_0^h = V_{k;0}^h(\Omega)$, defined in Sect. B.1.2 and the scalar elliptic problem (A.27). Given its variational formulation (A.28), we consider a *conforming* approximation in the subspace $V_0^h$: find $u_h \in V_0^h$, such that,

$$a(u_h, v_h) =< F, v_h >, \quad v_h \in V_0^h. \tag{B.1}$$

The well-posedness of problem (B.1) is ensured by the Lax-Milgram Lemma A.34. Error estimates can be found using Lemma B.6; see, e.g., [391, Sect. 6.2.1].

**Theorem B.7** *Let $\mathcal{T}_h$ be a conforming triangulation of $\Omega$ and $k \geq 1$. Then, problem (B.1) is well-posed: there exists a unique solution, such that*

$$\|u_h\|_{H^1(\Omega)} \leq C_1 \|F\|_{H^{-1}(\Omega)}, \quad \|u_h\|_a \leq C_2 \|F\|_{H^{-1}(\Omega)},$$

*where the constants are the same as in Theorem A.35. The finite element solution satisfies*

$$a(u - u_h, v_h) = 0, \quad v_h \in V_0^h, \tag{B.2}$$

*or, equivalently*

$$\|u - u_h\|_a = \inf_{v_h \in V_0^h} \|u - v_h\|_a.$$

*If the mesh $\mathcal{T}_h$ is shape-regular, we have*

$$|u - u_h|_{H^1(\Omega)} \leq C\, h^{s-1} |u|_{H^s(\Omega)}, \quad \frac{n}{2} < s \leq k+1$$

*and if, in addition, $\Omega$ is convex,*

$$\|u - u_h\|_{L^2(\Omega)} \leq C\, h^s |u|_{H^s(\Omega)}.$$

Property (B.2) is referred to as *Galerkin orthogonality* of the finite element solution.

*Remark B.8.* We note that the error estimates in the previous lemma ensure that the discrete solution converges to the exact $u$ when $h$ tends to zero, if $u$ is sufficiently regular. Convergence can be proven even in the case when the solution $u$ only belongs to $H^1(\Omega)$; see [391, Th. 6.2.1]. For the Neumann problem (A.30), we choose the finite element space $H_h = V^h/\mathbb{R}$ and a result analogous to Theorem B.7 can be proved.

*Remark B.9.* For the linear elasticity problems of Sect. A.6.2, we employ spaces of vector functions in $V_k^h(\Omega)$, the components of which belong to $V_k^h(\Omega)$. Results analogous to Theorem B.7 can be proved in this case as well.

### B.1.4 Non-Symmetric and Indefinite Problems

We now consider the more general Dirichlet problem (A.50). We consider the following discrete problem: find $u_h \in V_0^h$, such that,

$$b(u_h, v_h) = (f, v_h), \quad v_h \in V_0^h; \tag{B.3}$$

see the variational problem (A.53).

Under the assumption that the bilinear form $b(\cdot,\cdot)$ is coercive, the Lax-Milgram Lemma A.44 can be applied.

**Theorem B.10** *Let $\mathcal{T}_h$ be a conforming triangulation of $\Omega$ and let the conditions stated in Lemma A.45 be satisfied. For $k \geq 1$, problem (B.3) is well-posed: there exists a unique solution, such that*

$$|u_h|_{H^1(\Omega)} \leq C_1 \|f\|_{L^2(\Omega)}.$$

*The finite element solution satisfies*

$$|u - u_h|_{H^1(\Omega)} \leq C \inf_{v_h \in V_0^h} |u - v_h|_{H^1(\Omega)}.$$

*If the mesh $\mathcal{T}_h$ is shape-regular, we have*

$$|u - u_h|_{H^1(\Omega)} \leq C\, h_K^{s-1} |u|_{H^s(K)}, \quad \frac{n}{2} < s \leq k+1.$$

We remark that under the same assumptions on the coefficients of Lemma A.45, the discrete problem is well-posed without any restriction on the mesh size $h$. However, in the more general case where only a Gårding inequality (A.56) holds, well-posedness can only be ensured if the finite element space is sufficiently large and thus $h$ is sufficiently small. We refer to [414] for a proof of the following result.

**Theorem B.11** *Let the continuous problem (A.53) have a unique solution $u$ and let $k \geq 1$. Then, there exists a constant $h_0 > 0$, such that problem (B.3) is well-posed for $h < h_0$: there exists a unique solution, such that*

$$|u_h|_{H^1(\Omega)} \leq C_1 \|f\|_{L^2(\Omega)}$$

*and*

$$|u - u_h|_{H^1(\Omega)} \leq C \inf_{v_h \in V_0^h} |u - v_h|_{H^1(\Omega)}.$$

*Finally, if the mesh $\mathcal{T}_h$ is shape-regular, we have*

$$|u - u_h|_{H^1(\Omega)} \leq C\, h^{s-1} |u|_{H^s(\Omega)}, \quad \frac{n}{2} < s \leq k+1.$$

## B.2 Spectral Element Approximations

Spectral element methods are particular $p$-version finite element approximations where higher accuracy is achieved by increasing the polynomial degree of the approximation spaces inside each element of a given, fixed mesh, as opposed to the $h$-version where a more accurate solution is obtained by refining the mesh while keeping the degree fixed. Nodal basis functions are associated to a special set of nodes which allow us to use quadrature formulas of high precision for the evaluation of the stiffness matrix and the load vector. We refer to [46, 48, 50] as general references for this section.

We consider $\mathcal{T}$, a conforming, shape-regular triangulation of $\Omega$ consisting of (mapped) rectangles or cubes as in Definitions B.1 and B.2. This mesh is considered fixed. We will denote the elements of $\mathcal{T}$, by $\{\Omega_i,\ i = 1, \ldots, N\}$ and the reference element by $\hat{\Omega} = (-1, 1)^n$.

Given a polynomial degree $k \geq 1$, the discrete space $V^k \subset H^1(\Omega)$ is the one defined in Sect. B.1.2, i.e., the space of continuous, piecewise $\mathbb{Q}_k$ elements, constructed from a tensor product of degree $k$ polynomials of one variable. Since, so far, we have only considered elements that are affinely mapped, our spectral element space can also be written as

$$\begin{aligned} V^k &= V^k(\Omega) = \left\{ u_k \in C^0(\Omega) \mid u_k(F_i(\hat{x})) \in \mathbb{Q}_k(\hat{\Omega}),\ i = 1. \ldots, N \right\}, \\ V_0^k &= V^k \cap H_0^1(\Omega), \end{aligned} \tag{B.4}$$

where $F_i : \hat{\Omega} \to \Omega_i$ is an affine mapping; cf. Definition B.1. We note that not every quadrilateral can be obtained from the reference square through an affine mapping. For general quadrilaterals the definition (B.4) still holds, but $F_i$ must be replaced by a *bilinear* mapping. Similar considerations hold in three dimensions.

We now introduce a particular nodal basis of $V^k$. Basis functions will be defined on the reference element and then mapped onto the current element.

We denote by $GLL(k)$ the set of Gauss-Lobatto points $\{\xi_j;\ 0 \le j \le k\}$ on $\bar{\Lambda} = [-1, 1]$ in increasing order and by $\{w_j > 0\}$ the corresponding quadrature weights; see [48, Sect. 4]. The $GLL(k)$ nodes are the (distinct and real) zeros of $(1 - x^2)L_k'(x)$, with $L_k$ the Legendre polynomial of degree $k$, cf. [48, Sect. 3]. They also satisfy

$$\xi_j = -\cos(\theta_j), \qquad j\pi/k < \theta_j < (j+1)\pi/k, \quad 1 \le j \le k-1. \tag{B.5}$$

This follows from observing that the zeros of $L_k'(x)$ are bracketed by pairs of zeros of $L_k(x)$ and by using bounds for these zeros given in [429, Thm. 6.21.3]. There is also a constant $C > 0$ such that

$$\frac{1}{C}\frac{(1-\xi_j^2)^{1/2}}{k} \le w_j \le C\frac{(1-\xi_j^2)^{1/2}}{k}; \tag{B.6}$$

see Szego [429, (15.3.14)]. In particular, we have (see, e.g., [46, Page 76])

$$\frac{2}{k(k+1)} \le w_i \le \frac{C}{k}, \qquad 0 \le i \le k. \tag{B.7}$$

In the theory for domain decomposition methods for spectral element methods, the following important result is used:

**Lemma B.12** *Let $\Lambda = (-1, 1)$ and let $I^k : C(\Lambda) \to \mathbb{P}_k(\Lambda)$, be the polynomial interpolation operator which interpolates at the GLL(k) points. Then, there exists a constant $C$, which is independent of $k$, such that*

$$\|I^k\phi\|_{H^1(\Lambda)} \le C\|\phi\|_{H^1(\Lambda)}.$$

This result is due to Bernardi and Maday; cf. [47, Corollary 4.6]. The proof is quite technical.

For the cube $\hat{\Omega} = (-1, 1)^3$, we set $GLL(k)^3 = \{\xi_{ijl} = (\xi_i, \xi_j, \xi_l);\ 0 \le i, j, l \le k\}$ and denote the corresponding weight by $\{w_{ijl} = w_i w_j w_l > 0\}$. Similar definitions can be given in two dimensions with the obvious modifications. The following result, which is relatively easy to prove, cf. [48, Rem. 13.3], allow us to prove the ellipticity of the spectral element approximation obtained by using GLL quadrature:

$$\|u_k\|^2_{L^2(\hat{\Omega})} \le \sum_{i=0}^{k} u_k(\xi_i)^2\, w_i \le 3^n \|u_k\|^2_{L^2(\hat{\Omega})}, \quad u_k \in \mathbb{Q}_k(\Lambda); \tag{B.8}$$

see [48, Rem. 13.3]. In the following, we use the same notation for the mapped Gauss-Lobatto nodes and corresponding weights for an element $\Omega_i = F_i(\hat{\Omega}) \in \mathcal{T}$.

Given the nodes $GLL(k)^n$, our basis functions on $V^k(\hat{\Omega})$ are the tensor product of $k$-th order Lagrange interpolating polynomials on $GLL(k)$, defined by

$$\hat{l}_i(\xi_j) = \delta_{ij}. \tag{B.9}$$

On the reference element in three dimensions, we can write

$$u_k(x,y,z) = \sum_{i=0}^{k}\sum_{j=0}^{k}\sum_{l=0}^{k} u_k(\xi_{ijl})\, \hat{l}_i(x)\hat{l}_j(y)\hat{l}_l(z), \tag{B.10}$$

and an analogous representation holds in two dimensions. For a general element in $\mathcal{T}$, basis functions are obtained by mapping those on the reference element. Three dimensional basis functions can be divided into:

1. **interior**, if the indices $i$, $j$, $k$ are all different from 0 and $k$,
2. **face**, if exactly one of the indices is 0 or $k$,
3. **edge**, if exactly two of the indices are 0 or $k$,
4. **vertex**, if all indices are 0 and/or $k$.

In two dimensions, we distinguish between interior, vertex, and edge basis functions.

Equation (B.10) defines an interpolation operator $I^k$ on the reference element. By using the mappings $\{F_i\}$, a global interpolant, still denoted by $I^k$, can be obtained on $\Omega$. It is clear that $I^k$ can, in general, only be defined for continuous functions. On $\hat{\Omega}$, we have

$$I^k u(x,y,z) := \sum_{i=0}^{k}\sum_{j=0}^{k}\sum_{l=0}^{k} u(\xi_{ijl})\, \hat{l}_i(x)\hat{l}_j(y)\hat{l}_l(z),$$

and an analogous definition holds in two dimensions. We have the following error estimates, cf. [48, Th. 14.2].

**Lemma B.13** *For any real numbers $r$ and $s$ satisfying $s > (n+r)/2$ and $0 \le r \le 1$, there exists a positive constant depending only on $s$ such that*

$$|u - I^k u|_{H^r(\hat{\Omega})} \le \hat{C} k^{r-s} |u|_{H^s(\hat{\Omega})}.$$

*In addition,*

$$|u - I^k u|_{H^r(\Omega)} \le C k^{r-s} |u|_{H^s(\Omega)}.$$

Given $f \in L^2(\Omega)$, we now consider the model problem: find $u \in H_0^1(\Omega)$, such that

$$a(u,v) := \int_\Omega \rho \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx, \quad v \in H_0^1(\Omega). \tag{B.11}$$

The coefficient $\rho(x) \ge \rho_{\min} > 0$ can be discontinuous, with very different values for different elements, but we allow it to vary only moderately within each subregion. Here, we restrict ourselves to a piecewise constant coefficient $\rho(x) = \rho_i$ for $x \in \Omega_i$, a particular case of problem (A.28).

A discrete problem is obtained by restricting problem (B.11) to $V_0^k$ and by using Gauss-Lobatto-Legendre numerical quadrature on each element. We obtain the discrete bilinear form

$$a_Q(u_k, v_k) := (\rho \nabla u_k, \nabla v_k)_Q = \sum_{i=1}^{N} \rho_i (\nabla u_k, \nabla v_k)_{Q,\Omega_i}, \quad u_k, v_k \in V^k(\Omega),$$

and the discrete right hand side

$$(f, v_k)_Q = \sum_{i=1}^{N} (f, v_k)_{Q,\Omega_i}, \quad v_k \in V^k(\Omega).$$

The local contributions of the new inner product are defined in, e.g., three dimensions by

$$(u, v)_{Q,\Omega_i} = \sum_{m=0}^{k} \sum_{j=0}^{k} \sum_{l=0}^{k} u(\xi_{mjl}) v(\xi_{mjl}) w_{mjl}, \tag{B.12}$$

As before, we have used the same notation for the mapped nodes $GLL(k)^n$ and corresponding weights on an element $\Omega_i$. We note that, in order for the discrete right hand side $(f, v_k)_Q$ to be defined, $f$ needs to be continuous on $\overline{\Omega}$. This is certainly so if $f \in H^t(\Omega)$, $t > (n/2)$; cf. Lemma A.5.

The inequalities in (B.8) ensure that the discrete inner product is uniformly equivalent to the standard $L^2$-inner product on $V^k(\Omega)$, i.e.,

$$\|u_k\|^2_{L^2(\Omega)} \leq (u_k, u_k)_Q \leq C \|u_k\|^2_{L^2(\Omega)}, \quad u_k \in V^k(\Omega), \tag{B.13}$$

with a constant $C$ that only depends on the shape of the elements in $\mathcal{T}$.

The spectral element approximation of (B.11) consists in finding $u_k \in V_0^k$, such that,

$$a_Q(u_k, v_k) = (f, v_k)_Q, \quad v_k \in V_0^k. \tag{B.14}$$

Thanks to (B.8), we have

$$a(u_k, u_k) \leq a_Q(u_k, u_k) \leq C a(u_k, u_k), \quad u_k \in V^k(\Omega), \tag{B.15}$$

and the new bilinear form is therefore continuous and coercive. A detailed analysis of this method, including a discussion of existence, uniqueness, and error estimates for an individual element, is given in Bernardi and Maday [46, Sect. 15]. Using the coercivity of the discrete bilinear form, we can prove the following result.

**Theorem B.14** *Let $\mathcal{T}$ consist of one element $\Omega = \hat{\Omega}$. If $f \in H^t(\Omega)$, $t > (n/2)$, then problem (B.14) is well posed: there exists a unique solution, such that*

$$\|u_k\|_{H^1(\Omega)} \leq C_{1,\rho} \|I^k f\|_{L^2(\Omega)},$$

*where $C_{1,\rho}$ depends on the coefficient $\rho$. The spectral element solution satisfies*

$$|u - u_k|_{H^1(\Omega)} \leq C\,(k^{1-s}\|u\|_{H^s(\Omega)} + k^{-t}\|u\|_{H^t(\Omega)} + C_{2,\rho}k^{-\tau_\rho}), \quad s \geq 1,$$

*where $C_{2,\rho}$ and $\tau_\rho$ depend on $\rho$, and $C_{2,\rho} = 0$ in case $\rho(x) \equiv 1$. In addition,*

$$\|u - u_k\|_{L^2(\Omega)} \leq C\,(k^{-s}\|u\|_{H^s(\Omega)} + k^{-t}\|u\|_{H^t(\Omega)} + C_{2,\rho}k^{-\tau_\rho-1}).$$

*Remark B.15.* The case of more than one spectral element is treated in [330]. In addition, more general boundary conditions can also be considered, as well as more general second-order, positive definite problems; see section A.6.

## B.3 Divergence and Curl Conforming Finite Elements

### B.3.1 Raviart-Thomas Elements

The Raviart-Thomas spaces are conforming in $H(\mathrm{div}\,;\Omega)$ and were originally introduced in [394] in two dimensions and then extended in [359] to the three dimensional case. We mention [223, 95, 354] as good references for this section.

We only consider the case of a triangulation $\mathcal{T}_h$ made of triangles, for $n = 2$, and of tetrahedra, for $n = 3$, as introduced in section B.1.1. Given a triangle or a tetrahedron $K$, we consider the polynomial space

$$\mathbb{D}_k(K) := \mathbb{P}_{k-1}(K)^n \oplus \mathbf{x}\tilde{\mathbb{P}}_{k-1}(K), \quad k \geq 1,$$

where $\mathbf{x}$ is the position vector in $\mathbb{R}^n$ and $\tilde{\mathbb{P}}_{k-1}(K)$ is the space of homogeneous polynomials of degree $k-1$ on $K$. A function $\mathbf{u}$ in $\mathbb{D}_k(K)$ is uniquely defined by the following degrees of freedom

$$\int_f \mathbf{u}\cdot\mathbf{n}\,p\,ds, \quad p \in \mathbb{P}_{k-1}(f), \tag{B.16}$$

for each edge ($n = 2$) or face ($n = 3$) $f$ of $K$. For $k > 1$, we add

$$\int_K \mathbf{u}\cdot\mathbf{p}\,dx, \quad \mathbf{p} \in \mathbb{P}_{k-2}(K)^n.$$

We have the following result; cf., e.g., [359, 95].

**Lemma B.16** *A vector function $\mathbf{u} : \Omega \to \mathbb{R}^n$ belongs to $H(\mathrm{div}\,;\Omega)$ if and only if the restriction of $\mathbf{u}$ to every $K \in \mathcal{T}_h$ belongs to $H(\mathrm{div}\,;K)$, and, for each common face (or edge in two dimensions) $f = \overline{K_1} \cap \overline{K_2}$, we have*

$$\mathbf{u}\cdot\mathbf{n}_{|K_1} = \mathbf{u}\cdot\mathbf{n}_{|K_2}, \quad \text{on } f,$$

*with, e.g., $\mathbf{n} = \mathbf{n}_1 = -\mathbf{n}_2$.*

It can be proven that the following spaces are well-defined; see [359, 95]:

$$RT_k^h(\Omega) := \left\{\mathbf{u} \in H(\mathrm{div}\,;\Omega) \mid \mathbf{u}_{|K} \in \mathbb{D}_k(K),\ K \in \mathcal{T}_h\right\},$$
$$RT_{k;0}^h(\Omega) := \left\{\mathbf{u} \in H_0(\mathrm{div}\,;\Omega) \mid \mathbf{u}_{|K} \in \mathbb{D}_k(K),\ K \in \mathcal{T}_h\right\}.$$

The corresponding interpolation operator is denoted by $\Pi_{RT_k}^h$. When there is no ambiguity, we simply use the notations $RT^h(\Omega)$, $RT_0^h(\Omega)$, and $\Pi_{RT}^h$.

For the case $k = 1$, the elements of the local space have the simple form

$$\mathbb{D}_1(K) = \{\mathbf{u} = \mathbf{a} + b\,\mathbf{x} \mid \mathbf{a} \in \mathbb{P}_0(K)^n,\ b \in \mathbb{P}_0(K)\}.$$

It is immediate to check that the normal components of a vector in $\mathbb{D}_1(K)$ are constant on each edge ($n = 2$) or face ($n = 3$) $f$. These values

$$\lambda_f(\mathbf{u}) = \frac{1}{|f|}\int_f \mathbf{u}\cdot\mathbf{n}\,ds, \quad f \subset \partial K, \tag{B.17}$$

can be taken as the degrees of freedom. As in the case of nodal elements, the $L^2$-norm of a vector $\mathbf{u} \in \mathbb{D}_1(K)$ can be bounded from above and below by means of its degrees of freedom. The proof of Lemma B.5 can easily be adapted to the following result, and similar estimates can also be obtained for the case $k > 1$.

**Lemma B.17** *Let $K \in \mathcal{T}_h$. Then, there exist constants only depending on the aspect ratio of $K$, such that*

$$c\sum_{f\subset\partial K} h_f^n\,\lambda_f(\mathbf{u})^2 \le \|\mathbf{u}\|_{L^2(K)}^2 \le C\sum_{f\subset\partial K} h_f^n\,\lambda_f(\mathbf{u})^2, \tag{B.18}$$

*where $h_f$ is the diameter of $f$.*

Finite elements built on triangulations made of rectangles ($n = 2$) or parallelepipeds ($n = 3$) can also be considered. We refer to [394, 359, 95, 391] for details.

For the spaces defined in this section, the corresponding nodal interpolation operator $\Pi_{RT_k}^h$ is not defined in the whole space $H(\mathrm{div}\,;\Omega)$; some additional regularity is required. In particular, the normal component on $\partial K$ of a vector $\mathbf{u} \in H(\mathrm{div}\,;\Omega)$ generally only belongs to $H^{-1/2}(\partial K)$ and not to $H^{-1/2}(f)$, and the degrees of freedom (B.16) are not defined in general. They are certainly well-defined if the trace of a vector $\mathbf{u}$ on the boundary of a generic element is sufficiently regular, that is, if $\mathbf{u}$ belongs to $H^r(\Omega)^n$ for $r > 1/2$. The following error estimate can be proven using standard arguments as in [391, Sect. 3.4.2]

$$\left\|\mathbf{u} - \Pi_{RT_k}^h\mathbf{u}\right\|_{L^2(\Omega)} \le Ch^r|\mathbf{u}|_{H^r(\Omega)}, \quad \frac{1}{2} < r \le k. \tag{B.19}$$

The constant $C$ depends only on the aspect ratios of the elements of $\mathcal{T}_h$ and the exponent $r$. See also [359] for other error estimates and [95, Sect. III.3.3] for additional comments.

### B.3.2 Nédélec Elements in Two Dimensions

Nédélec elements (also called *edge* elements) are finite elements which are conforming in $H(\mathbf{curl}\,;\Omega)$. Just as, in two dimensions, vectors in $H(\mathbf{curl}\,;\Omega)$ are obtained from those in $H(\mathrm{div}\,;\Omega)$ by a rotation of 90 degrees, Nédélec finite element vectors are obtained from those in the Raviart-Thomas spaces in the same way. However, we will see that the two families of elements are completely different in three dimensions.

In two dimensions ($n = 2$), it is enough to rotate the vectors in the Raviart-Thomas spaces introduced in the previous section. In the case of triangles the local spaces are thus

$$\mathcal{R}_k(K) := \left\{ \mathbf{u} + \mathbf{v} \mid \mathbf{u} \in \mathbb{P}_{k-1}(K)^2,\ \mathbf{v} \in \tilde{\mathbb{P}}_k(K)^2,\ \mathbf{v} \cdot \mathbf{x} = 0 \right\}, \quad k \geq 1,$$

and a function $\mathbf{u}$ in $\mathcal{R}_k(K)$ is uniquely defined by the following degrees of freedom:

$$\int_e \mathbf{u} \cdot \mathbf{t}_e\, p\, ds, \quad p \in \mathbb{P}_{k-1}(e),$$

for each edge $e$, and, in addition, for $k > 1$,

$$\int_K \mathbf{u} \cdot \mathbf{p}\, dx, \quad \mathbf{p} \in \mathbb{P}_{k-2}(K)^n.$$

Here $\mathbf{t}_e$ is a unit vector that is tangent to $e$.

In the case $k = 1$, the local space $\mathcal{R}_k(K)$ has the form

$$\mathcal{R}_1(K) = \left\{ \mathbf{u} = \begin{bmatrix} a_1 + bx_2 \\ a_2 - bx_1 \end{bmatrix},\ \mathbf{a} \in \mathbb{P}_0(K)^2,\ b \in \mathbb{P}_0(K) \right\}.$$

It is easy to check that a vector in $\mathcal{R}_1(K)$ has a constant tangential component on each edge of $K$, and the corresponding degrees of freedom $\{\lambda_e(\mathbf{u}),\ e \subset \partial K\}$ can be taken as the values of the tangential component on the three edges $e$ of $K$. They can be written as

$$\lambda_e(\mathbf{u}) = \frac{1}{h_e} \int_e \mathbf{n} \times \mathbf{u}\, ds, \quad e \subset \partial K,$$

and the degrees of freedom in the global space are associated to the edges of the triangulation. By direct computation, the curl of a vector in $\mathcal{R}_1(K)$ is constant.

We have the following result; cf., e.g., [359].

**Lemma B.18** *A vector function $\mathbf{u} : \Omega \to \mathbb{R}^n$ belongs to $H(\mathbf{curl}\,;\Omega)$ if and only if the restriction of $\mathbf{u}$ to every $K \in \mathcal{T}_h$ belongs to $H(\mathbf{curl}\,;K)$, and, for each common face (or edge in two dimensions) $f = \overline{K_1} \cap \overline{K_2}$, we have*

$$\mathbf{u} \times \mathbf{n}_{|K_1} = \mathbf{u} \times \mathbf{n}_{|K_2}, \quad \text{on } f,$$

*with, e.g., $\mathbf{n} = \mathbf{n}_1 = -\mathbf{n}_2$*

Given the local spaces $\mathcal{R}_k(K)$, the global ones, conforming in $H(\mathbf{curl};\Omega)$ and $H_0(\mathbf{curl};\Omega)$, are defined in the obvious way and are denoted by $ND_k^h(\Omega)$ and $ND_{k;0}^h(\Omega)$, respectively, or by $ND^h(\Omega)$ and $ND_0^h(\Omega)$, when there is no ambiguity. Clearly, Lemma B.17 and inequality (B.19) still hold for the Nédélec spaces in two dimensions.

### B.3.3 Nédélec Elements in Three Dimensions

We now consider the three dimensional case $n = 3$. We give [359, 223, 354] as general references for this section.

For triangulations made of tetrahedra, the local spaces on a generic tetrahedron $K$ are defined as

$$\mathcal{R}_k(K) := \left\{ \mathbf{u} + \mathbf{v} \mid \mathbf{u} \in \mathbb{P}_{k-1}(K)^3,\ \mathbf{v} \in \tilde{\mathbb{P}}_k(K)^3,\ \mathbf{v}\cdot\mathbf{x} = 0 \right\}, \quad k \geq 1;$$

see [223, Sect. III.5.3]. A vector in $\mathbf{u} \in \mathcal{R}_k(K)$ is uniquely defined by the following degrees of freedom, see [359],

$$\int_e \mathbf{u}\cdot\mathbf{t}_e\, p\, ds, \quad p \in \mathbb{P}_{k-1}(e), \tag{B.20}$$

for the six edges $e$ of $K$, and, for $k > 1$,

$$\int_f (\mathbf{u}\times\mathbf{n})\cdot\mathbf{p}\, dS, \quad \mathbf{p} \in \mathbb{P}_{k-2}(f)^2,$$

for the four faces $f$ of $K$, and, additionally, for $k > 2$,

$$\int_K \mathbf{u}\cdot\mathbf{p}\, dx, \quad \mathbf{p} \in \mathbb{P}_{k-3}(K)^3.$$

Here $\mathbf{t}_e$ denotes a unit vector in the direction of the edge $e$. We recall that Lemma B.18 also holds in three dimensions. It can be proven that the following finite element spaces are well defined (see [359]):

$$ND_k^h(\Omega) := \left\{ \mathbf{u} \in H(\mathbf{curl};\Omega) \mid \mathbf{u}_{|K} \in \mathcal{R}_k(K),\ K \in \mathcal{T}_h \right\},$$
$$ND_{k;0}^h(\Omega) := \left\{ \mathbf{u} \in H_0(\mathbf{curl};\Omega) \mid \mathbf{u}_{|K} \mid \in \mathcal{R}_k(K),\ K \in \mathcal{T}_h \right\}.$$

The corresponding interpolation operator is denoted by $\Pi_{ND_k}^h$, and, as usual, we will drop the superscript $k$, when there is no ambiguity.

In the case $k = 1$, the elements of the local space $\mathcal{R}_k(K)$ have the simple form

$$\mathcal{R}_1(K) = \left\{ \mathbf{u} = \mathbf{a} + \mathbf{b}\times\mathbf{x} \mid \mathbf{a},\mathbf{b} \in \mathbb{P}_0(K)^3 \right\}.$$

It is immediate to see that the tangential components of a vector in $\mathcal{R}_1(K)$ are constant on the six edges $e$ of $K$. These values

$$\lambda_e(\mathbf{u}) = \frac{1}{|e|} \int_e \mathbf{u} \cdot \mathbf{t}_e \, ds, \quad e \subset \partial K, \tag{B.21}$$

can be taken as the degrees of freedom. Similarly to Lemma B.17, we have

$$c \sum_{e \subset \partial K} h_e^3 \, \lambda_e(\mathbf{u})^2 \leq \|\mathbf{u}\|^2_{L^2(K)} \leq C \sum_{e \subset \partial K} h_e^3 \, \lambda_e(\mathbf{u})^2, \tag{B.22}$$

where $h_e$ is the length of $e$ and the constants $c$ and $C$ only depend on the aspect ratio of the element $K$.

Nédélec finite elements can also be built on triangulations made of parallelepipeds. We refer to [359] for details.

The interpolation operator on the space $ND_k^h(\Omega)$ is not defined in the whole space $H(\mathbf{curl}\,;\Omega)$; additional regularity is required. Different choices, and consequently different error estimates, are possible; see [359, 223, 353, 15, 18, 23, 354]. We remark that the interpolant $\Pi^h_{ND_k}\mathbf{u}$ is not defined for a generic vector $\mathbf{u}$ in $H^r(\Omega)^3$ for $1/2 < r \leq 1$, as was the case for the Raviart-Thomas spaces, since the degrees of freedom (B.20) involve the tangential component of $\mathbf{u}$ on the edges, which is not necessarily defined. If we require that $\mathbf{u} \in H^r(\Omega)^3$ for $r > 1$, then the trace of $\mathbf{u}$ is defined on the edges.

We only give the following result: the first part is proven in [18, Lem. 4.7], while the second follows from a Sobolev embedding theorem.

**Lemma B.19** *For any $p > 2$ and any element $K \in \mathcal{T}_h$, the interpolant $\Pi^h_{ND_k}$ is well defined and continuous on the space*

$$\{\mathbf{u} \in L^p(K)^3 \mid \mathbf{curl}\,\mathbf{u} \in L^p(K)^3, \ \mathbf{u} \times \mathbf{n} \in L^p(\partial K)^3\},$$

*and thus on*

$$\{\mathbf{u} \in H^1(K)^3 \mid \mathbf{curl}\,\mathbf{u} \in L^p(K)^3\}.$$

### B.3.4 The Kernel and Range of the Curl and Divergence Operators

We now suppose that the domain $\Omega$ is a simply connected polygon or polyhedron, with a connected boundary. In Sect. A.5.4, we have seen that the range of the divergence operator in $H(\mathrm{div}\,;\Omega)$ is $L^2(\Omega)$ and the kernel of the curl operator in $H(\mathbf{curl}\,;\Omega)$ is the space $\mathbf{grad}\,H^1(\Omega)$. In two dimensions, the kernel of the divergence operator is $\mathbf{curl}\,H^1(\Omega)$, and in three dimensions it coincides with $\mathbf{curl}\,H(\mathbf{curl}\,;\Omega)$. Similar properties hold for the spaces of functions satisfying homogeneous boundary conditions.

In this section, we state analogous properties for the finite element spaces previously introduced. These results are well known and can, for instance, be found in [223, 171, 64, 95, 260, 18, 354]. In particular, we refer to [260] for the proofs of the results in this section. The case when $\Omega$ is not simply connected or its boundary is not connected, is treated in [18] and does not present any particular difficulties.

We denote the $L^2$-projection onto the finite element space $Q_k^h(\Omega)$ by $\Pi_{Q_k}^h = \Pi_Q^h$. We recall that the spaces $Q_k^h(\Omega)$ and $Q_{k;0}^h(\Omega)$ have been defined at the end of Sect. B.1.2. The following lemma is well known and is often referred to as commuting-diagram property; a rigorous proof can be found in, e.g., [260, Th. 2.30].

**Lemma B.20** *Let $\mathcal{T}_h$ be shape-regular and the functions $q$, $\mathbf{u}$, $\mathbf{v}$ sufficiently regular. Then, the following identities hold on each element in $\mathcal{T}_h$ and for $k \geq 1$:*

$$\mathbf{grad}\,(I_k^h q) = \Pi_{ND_k}^h(\mathbf{grad}\, q)\,, \tag{B.23}$$

$$\operatorname{curl}\,(\Pi_{ND_k}^h \mathbf{u}) = \Pi_{Q_{k-1}}^h(\operatorname{curl}\mathbf{u})\,, \quad n = 2, \tag{B.24}$$

$$\mathbf{curl}\,(I_k^h q) = \Pi_{RT_k}^h(\mathbf{curl}\, q)\,, \quad n = 2, \tag{B.25}$$

$$\mathbf{curl}\,(\Pi_{ND_k}^h \mathbf{u}) = \Pi_{RT_k}^h(\mathbf{curl}\,\mathbf{u})\,, \quad n = 3, \tag{B.26}$$

$$\operatorname{div}\,(\Pi_{RT_k}^h \mathbf{v}) = \Pi_{Q_{k-1}}^h(\operatorname{div}\mathbf{v})\,. \tag{B.27}$$

We remark that Proposition B.20 is proved by local arguments on each element of $\mathcal{T}_h$, and the result is thus valid for an arbitrary Lipschitz domain. Proposition B.20 implies that the interpolants of the finite element spaces, that we have introduced, preserve the kernel of the relevant operators. An analogous result is valid for the spaces that satisfy homogeneous boundary conditions.

The following proposition characterizes the kernel of the curl operator; see [260, Th. 2.36] for a proof.

**Lemma B.21** *If $\Omega$ is simply connected, with a connected boundary, the kernels of the curl operator defined in $ND_k^h(\Omega)$ and $ND_{k;0}^h(\Omega)$ are $\mathbf{grad}\, V_k^h(\Omega)$ and $\mathbf{grad}\, V_{k;0}^h(\Omega)$, respectively.*

We can now define the following decompositions of the Nédélec spaces into the kernel of the curl operator and its orthogonal complement:

$$ND_k^h(\Omega) = \mathbf{grad}\, V_k^h(\Omega) \oplus ND_k^{h;\perp}(\Omega), \tag{B.28}$$

$$ND_{k;0}^h(\Omega) = \mathbf{grad}\, V_{k;0}^h(\Omega) \oplus ND_{k;0}^{h;\perp}(\Omega). \tag{B.29}$$

These decompositions are the discrete analogs of (A.12) and (A.11). We note that, in general, the spaces $ND_k^{h;\perp}(\Omega)$ and $ND_{k;0}^{h;\perp}(\Omega)$ are not included in $H^{\perp}(\mathbf{curl}\,;\Omega)$ and $H_0^{\perp}(\mathbf{curl}\,;\Omega)$.

The following Proposition states the analog of (A.15), (A.16), (A.19), and (A.20). Its proof can be found in [18, Prop. 4.6].

**Lemma B.22** *Let $\Omega$ be a Lipschitz domain. Then, for $\mathbf{u} \in ND_k^{h;\perp}(\Omega) \cup ND_{k;0}^{h;\perp}(\Omega)$,*

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C\, H_\Omega\, \|\operatorname{curl}\mathbf{u}\|_{L^2(\Omega)^n}\,, \; n = 2, \tag{B.30}$$

$$\|\mathbf{u}\|_{L^2(\Omega)^n} \leq C\, H_\Omega\, \|\mathbf{curl}\,\mathbf{u}\|_{L^2(\Omega)^n}\,, \; n = 3. \tag{B.31}$$

We now consider the characterization of the kernel of the divergence operator. For the two-dimensional case, we can use the results for $H(\mathbf{curl};\Omega)$ and prove results analogous to Proposition B.21. In particular, the kernels of the divergence operator defined in $RT_k^h(\Omega)$ and $RT_{k;0}^h(\Omega)$ are $\mathbf{curl}\, V_k^h(\Omega)$ and $\mathbf{curl}\, V_{k;0}^h(\Omega)$ and the following decompositions hold for $n = 2$:

$$RT_k^h(\Omega) = \mathbf{curl}\, V_k^h(\Omega) \oplus RT_k^{h;\perp}(\Omega), \tag{B.32}$$

$$RT_{k;0}^h(\Omega) = \mathbf{curl}\, V_{k;0}^h(\Omega) \oplus RT_{k;0}^{h;\perp}(\Omega). \tag{B.33}$$

For the three-dimensional case, the following result can be found in [260, Th. 2.36].

**Lemma B.23** *If $\Omega \subset \mathbb{R}^3$ is simply connected, with a connected boundary, the kernels of the divergence operator defined in $RT_k^h(\Omega)$ and $RT_{k;0}^h(\Omega)$ are $\mathbf{curl}\, ND_k^h(\Omega)$ and $\mathbf{curl}\, ND_{k;0}^h(\Omega)$.*

We can now define the following decompositions of the Raviart-Thomas spaces into the kernel of the divergence operator and its orthogonal complement, for $n = 3$,

$$RT_k^h(\Omega) = \mathbf{curl}\, ND_k^h(\Omega) \oplus RT_k^{h;\perp}(\Omega), \tag{B.34}$$

$$RT_{k;0}^h(\Omega) = \mathbf{curl}\, ND_{k;0}^h(\Omega) \oplus RT_{k;0}^{h;\perp}(\Omega). \tag{B.35}$$

These decompositions are the discrete analogs of (A.7) and (A.8).

We end this section with a characterization of the range of the divergence operator; see [260, Th. 2.36].

**Lemma B.24** *The divergence operator is surjective from $RT_k^h(\Omega)$ into $Q_{k-1}^h(\Omega)$, and from $RT_{k;0}^h(\Omega)$ into $Q_{k-1;0}^h(\Omega)$.*

## B.4 Saddle-Point Problems

We now consider the approximation of problem (A.57). As for indefinite problems, it is not enough to consider subspaces $V^h \subset V$ and $Q^h \subset Q$ that have good approximation properties in order to have a well-posed discrete problem. Given $G \in Q'$, we define the space $Z^h(G) \subset V^h$ as

$$Z^h(G) = \{v_h \in V^h \mid b(v_h, q_h) = G(q_h),\ q_h \in Q^h\},$$

with $Z^h = Z^h(0)$. Since in general $Z^h \not\subset Z$, condition (A.59) does not ensure the coercivity of $a(\cdot,\cdot)$ in $Z^h$. Similarly, the continuous inf-sup condition (A.60) does not translate into an analogous condition for the discrete spaces. We then make two assumptions on the subspaces $V^h$ and $Q^h$:

(i) $a(\cdot,\cdot)$ is coercive on $Z^h$:

$$a(u_h, u_h) \geq \beta_h \|u_h\|_V^2, \quad u_h \in Z^h. \tag{B.36}$$

(ii) $b(\cdot,\cdot)$ satisfies the *discrete* inf-sup condition:

$$\inf_{0 \neq q_h \in Q^h} \sup_{0 \neq v_h \in V^h} \frac{b(v_h, q_h)}{\|v_h\|_V \, \|q_h\|_Q} \geq \gamma_h > 0. \tag{B.37}$$

We note that in most applications, $\beta_h$ is independent of the discretization parameters, but $\gamma_h$ may not be.

We next consider the discrete problem: find $(u_h, p_h) \in V^h \times Q^h$, such that,

$$\begin{aligned} a(u_h, v_h) + b(v_h, p_h) &= F(v_h), \quad v_h \in V^h, \\ b(u_h, q_h) &= G(q_h), \quad q_h \in Q^h. \end{aligned} \tag{B.38}$$

We have the following result; see [95, Sect. II.2.2].

**Lemma B.25** *Let the same assumptions as in Lemma A.47 be satisfied. If (B.36) and (B.37) hold and $Z^h(G)$ is not empty, then problem (B.38) has a unique solution. There is a constant, depending only on $\alpha_1$, $\alpha_2$, $\beta_h$, and $\gamma_h$, such that*

$$\|u_h\|_V \leq C(\|F\|_{V'} + \|G\|_{Q'}), \quad \|p_h\|_Q \leq C(\|F\|_{V'} + \|G\|_{Q'}).$$

*In addition,*

$$\|u - u_h\|_V \leq C\gamma_h^{-1} \inf_{v_h \in V^h} \|u - v_h\|_V + C \inf_{q_h \in Q^h} \|p - q_h\|_Q,$$

$$\|p - p_h\|_Q \leq C\gamma_h^{-2} \inf_{v_h \in V^h} \|u - v_h\|_V + C\gamma_h^{-1} \inf_{q_h \in Q^h} \|p - q_h\|_Q,$$

*where the constants only depend on $\alpha_1$, $\alpha_2$, and $\beta_h$. If, in addition, $Z^h \subset Z$, then*

$$\|u - u_h\|_V \leq \left(1 + \frac{\alpha_1}{\beta_h}\right) \inf_{v_h \in Z^h} \|u - v_h\|_V.$$

*Remark B.26.* Under the same assumptions (B.36) and (B.37), an analogous result to Lemma B.25 can be proven for the more general problem (A.61); see [95, Sect. II.2.4].

### B.4.1 Finite Element Approximations for the Stokes Problem

We now consider some choices of spaces for the Stokes problem; see section A.7.2. We first note that the bilinear form $a(\cdot,\cdot)$ is coercive in any finite dimensional subspace $V^h \subset V = H_0^1(\Omega)$. The main focus is then the choice of subspaces that ensure the discrete inf-sup condition (B.37). For nearly incompressible elasticity the bilinear form $a(\cdot,\cdot)$ is coercive in the whole $V_h$ thanks to the Korn inequality (A.35). Since the bilinear form $b(\cdot,\cdot)$ is the

same as for the Stokes problem, velocity and pressure spaces that work for the Stokes problem also work for these elasticity problems.

In the following, we only mention some particular choices of spaces and refer to the references for a detailed analysis.

We consider a conforming, shape-regular triangulation $\mathcal{T}_h$ consisting of affinely mapped squares or cubes; see Sect. B.1.1. We first consider the so-called $\mathbb{Q}_2(h)$-$\mathbb{Q}_0(h)$ approximations. Here, the velocity space consists of continuous vectors the components of which belong to

$$\left\{u \in C^0(\Omega) \mid u_{|_K} \in \mathbb{Q}_2(K),\ K \in \mathcal{T}_h\right\},$$

while the pressure space consists of discontinuous functions in

$$\left\{p \in L^2(\Omega) \mid p_{|_K} \in \mathbb{Q}_0(K),\ K \in \mathcal{T}_h\right\}.$$

These spaces satisfy an inf-sup condition (B.37), with $\gamma_h$ independent of $h$, but they lead to non-optimal error estimates; see [95, Ch. VI, Pg. 221].

For $\mathbb{Q}_2(h)$-$\mathbb{P}_1(h)$ approximations, the velocity space is the same, but the pressure space is chosen as

$$\left\{p \in L^2(\Omega) \mid p_{|_K} \in \mathbb{P}_1(K),\ K \in \mathcal{T}_h\right\}.$$

These elements also satisfy a uniform inf-sup condition and give optimal error estimates for the velocity and the pressure; see [95, Ch. VI, Pg. 216].

Some choices with continuous pressure spaces are also possible; see [95] for more details.

### B.4.2 Spectral Element Approximations for the Stokes Problem

We consider a conforming, shape-regular triangulation $\mathcal{T}$ consisting of affinely mapped squares or cubes, as described in Sect. B.2.

The $\mathbb{Q}_k$-$\mathbb{Q}_{k-2}$ approximation, with $k \geq 2$, consists of continuous velocities, the components of which belong to

$$\left\{u \in C^0(\Omega) \mid u(F_i(\hat{x})) \in \mathbb{Q}_k(\hat{\Omega}),\ i = 1\ldots, N\right\},$$

and discontinuous pressures in

$$\left\{p \in L^2(\Omega) \mid p(F_i(\hat{x})) \in \mathbb{Q}_{k-2}(\hat{\Omega}),\ i = 1\ldots, N\right\}.$$

For velocities, we consider the tensor product Lagrangian nodal basis based on the one dimensional nodes $GLL(k)$ described in Sect. B.2. Numerical quadrature based on $GLL(k)$ is also employed. A very convenient basis for the pressures consists on tensor product nodal basis functions associated to the *internal* $GLL(k)$ nodes, i.e., the end points $\pm 1$ are excluded.

This method satisfies a non-uniform inf-sup condition, with $\gamma_k = ck^{(1-n)/2}$; see [329, 427]. The choice $\mathbb{Q}_k$-$\mathbb{P}_{k-1}$, on the other hand, ensures a uniform inf-sup condition; see [49].

### B.4.3 Finite Element Approximations for Flows in Porous Media

We now consider some choices of finite element spaces for the saddle point problem (A.67) in Sect. A.7.2, which is written as in (A.63) by defining the (bi)linear forms as in (A.68):

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega (\mathcal{A}^{-1}\mathbf{u})\cdot\mathbf{v}\,dx, \quad b(\mathbf{u},p) = -\int_\Omega \nabla\cdot\mathbf{u}\,p\,dx, \quad G(q) = -\int_\Omega f q\,dx.$$

Here, we only consider the case of Neumann boundary conditions; this corresponds to the choice (A.70) for the functional $F(\cdot)$ and (A.69) for the continuous spaces:

$$F(\mathbf{v}) = -\int_{\partial\Omega} \mathbf{v}\cdot\mathbf{n}\,g_N\,ds, \qquad V = H(\mathrm{div};\Omega), \quad Q = L^2(\Omega)$$

The choices (A.72) and (A.71) can be treated in a similar way.

We consider a conforming, shape-regular triangulation $\mathcal{T}_h$ consisting of affinely mapped triangles or tetrahedra as in Sect. B.1.1. The case of mapped squares or cubes can be treated similarly. We then choose finite element spaces that are conforming in $H(\mathrm{div};\Omega)$ and $L^2(\Omega)$:

$$V^h = RT_k^h(\Omega), \quad Q^h = Q_{k-1}^h(\Omega),$$

for $k \geq 1$; see Sect. B.3.1 and B.1.2.

Lemma B.24 ensures that the inf-sup condition (B.37) holds for the bilinear form $b(\cdot,\cdot)$. In addition, it can be proven that $\gamma_h$ is independent of the meshsize $h$; see, e.g., [394, 95]. The same lemma also ensures that

$$Z^h = Z^h(0) = \{\mathbf{u}\in V^h \mid \nabla\cdot\mathbf{u} = 0\}$$

and that therefore, for $\mathbf{u}\in Z^h$,

$$a(\mathbf{u},\mathbf{u}) = \int_\Omega (\mathcal{A}^{-1}\mathbf{u})\cdot\mathbf{u}\,dx \geq A_2^{-1}\|\mathbf{u}\|^2_{L^2(\Omega)} = A_2^{-1}\|\mathbf{u}\|^2_{\mathrm{div};\Omega},$$

since $\mathcal{A}$ is uniformly positive definite thanks to (A.29). Inequality (B.36) then holds.

For the problem corresponding to the choices (A.72) and (A.71), we employ

$$V^h = RT_{k;0}^h(\Omega), \quad Q^h = Q_{k-1;0}^h(\Omega),$$

## B.5 Inverse Inequalities

Inverse inequalities are powerful tools for the analysis of Galerkin approximations and domain decomposition methods. While Poincaré type inequalities

(see Sect. A.4) are valid for functions in the continuous space $H^1(\Omega)$, inverse inequalities are only valid for finite element or polynomial functions and typically give bounds of certain Sobolev norms in terms of a weaker one.

We first consider a conforming, shape-regular triangulation $\mathcal{T}_h$ and the finite element space $V^h = V_k^h(\Omega)$; see Sect. B.1.2. The following result can be obtained by simple scaling arguments; see, e.g., [391, Sect. 6.3.2].

**Lemma B.27 (Finite element spaces)** *Let $s \geq t \geq 0$ be two real numbers. Then, for $K \in \mathcal{T}_h$, there exists a constants, depending only on $s$, $r$, $k$, and the aspect ratio of $K$, such that*

$$|u_h|_{H^s(K)} \leq C h_K^{-(s-t)} |u_h|_{H^t(K)}, \quad u_h \in V^h. \tag{B.39}$$

*If, in addition, $\mathcal{T}_h$ is quasi-uniform, then*

$$|u_h|_{H^s(\Omega)} \leq C h^{-(s-t)} |u_h|_{H^t(\Omega)}, \quad u_h \in V^h. \tag{B.40}$$

Some inverse inequalities involving the $L^\infty$- and the $H^1$-norms were also given in Equation (4.16) and Lemma 4.15.

Inverse inequalities are also available for polynomials. We refer to [46, Sect. 5] for the following lemma.

**Lemma B.28 (Polynomial spaces)** *Let $m$ be an integer and $s$ a real number with $0 \leq m \leq s$. We have*

$$|u_k|_{H^s(\hat{\Omega})} \leq C k^{2(s-m)} |u_k|_{H^m(\hat{\Omega})}, \quad u_k \in \mathbb{Q}_k(\hat{\Omega}).$$

*In addition,*

$$|u_k|_{H^s(\Omega)} \leq \tilde{C} k^{2(s-m)} |u_k|_{H^m(\Omega)}, \quad u_k \in V^k,$$

*where $\tilde{C}$ depends on the triangulation $\mathcal{T}$.*

We note that the two previous lemmas can be combined in order to obtain bounds that are explicit both in $h$ and $k$. We refer to [415, Sect. 4.6] for a comprehensive presentation. We finally remark that the two previous results are independent of the particular choice of the basis for the finite dimensional subspaces.

## B.6 Matrix Representation and Condition Number

In this section, for simplicity, we only consider the model problem: find $u \in H_0^1(\Omega)$, such that

$$a(u, v) = < F, v >, \quad v \in H_0^1(\Omega),$$

with

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

We first consider finite element approximations; see Sect. B.1. Once a basis $\{\phi_i\}$ is chosen for the finite element space $V = V_0^h$, a finite element function $u$ can be uniquely written as

$$u(x) = \sum_i u_i \phi_i(x).$$

We recall that we use the same notation for a function $u(x)$ and the corresponding column vector of degrees of freedom, and for spaces of functions $V$ and the corresponding spaces of degrees of freedom. The discrete problem (B.1) then gives then rise to a linear system

$$A_h u = f,$$

where the matrix $A = A_h$ and the right hand side $f$ are given by

$$A_{ij} = a(\phi_j, \phi_i), \quad f_i =< F, \phi_i > .$$

The following result is a consequence of the symmetry and coercivity of $a(\cdot, \cdot)$. We note, in particular, that it is independent of the choice of the basis functions.

**Lemma B.29** *The matrix $A_h$ is symmetric and positive definite.*

We look for an upper bound for the condition number $\kappa(A_h)$ of $A_h$. We recall that $\kappa(A_h)$ is defined by

$$\kappa(A_h) = \frac{\lambda_{max}(A_h)}{\lambda_{min}(A_h)},$$

where

$$\lambda_{max}(A_h) = \max_{0 \neq u \in V} \frac{u^T A_h u}{u^T u}, \quad \lambda_{min}(A_h) = \min_{0 \neq u \in V} \frac{u^T A_h u}{u^T u}$$

are the maximum and minimum eigenvalues of $A_h$; see appendix C. Before proceeding, we also define the mass matrix $M = M_h$ as

$$M_{ij} = (\phi_j, \phi_i)_{L^2(\Omega)}.$$

We note that $M$ is also symmetric and positive definite.

Combining the Friedrichs inequality in Lemma A.14 and the inverse estimate in Lemma B.27 for $s = 1$ and $t = 0$, we obtain the following result which gives bounds for the generalized eigenvalue problem involving $A_h$ and $M_h$.

**Lemma B.30** *Let the triangulation $\mathcal{T}_h$ be conforming and quasi-uniform. Then, there exist two constants, such that*

$$c_1 \|u\|^2_{L^2(\Omega)} \leq a(u, u) \leq C_2 h^{-2} \|u\|^2_{L^2(\Omega)}, \quad u \in V,$$

*where $c_1$ only depends on $\Omega$ and both constants are independent of the basis. Equivalently, we have*

$$c_1 \, u^T M u \leq u^T A u \leq C_2 \, h^{-2} \, u^T M u, \quad u \in V.$$

In order to find bounds for $\lambda_{max}(A_h)$ and $\lambda_{min}(A_h)$, it is enough to bound the eigenvalues of $M$. Such bounds in general depend on the choice of the basis. The following result can be found in [391, Pr. 6.3.1] and can be proven using a scaling argument. We note that it remains valid for all the vector finite element spaces introduced in Sect. B.3, thanks to Lemma B.17 and inequality (B.22).

**Lemma B.31** *Let the triangulation $\mathcal{T}_h$ be conforming and quasi-uniform. Let in addition $\{\phi_i = \phi_i^h\}$ be the nodal basis functions introduced in Sect. B.1.2. Then, there exist two constants, such that*

$$c_3\, h^n\, u^T u \le u^T M u \le C_4\, h^n\, u^T u, \quad u \in V.$$

Combining Lemmas B.30 and B.31, we obtain the desired result, which we also express in terms of the number of degrees of freedom $L_h = O(h^{-n})$.

**Theorem B.32** *Let the triangulation $\mathcal{T}_h$ be conforming and quasi-uniform. Let in addition $\{\phi_i = \phi_i^h\}$ be the nodal basis functions introduced in Sect. B.1.2. Then, there exist two constants, such that*

$$c_5\, h^n\, u^T u \le u^T A_h u \le C_6\, h^{n-2}\, u^T u, \quad u \in V.$$

*Consequently*

$$\kappa(A_h) \le C h^{-2} \le \tilde{C}\, L_h^{2/n}.$$

We note that the estimate of the condition number is sharp. Estimates for particular meshes that are not quasi-uniform can also be found; see, e.g., [34, 352].

The analysis can also be carried out for the spectral approximations in Sect. B.2. Once a basis $\{\phi_i\}$ for the space $V = V_0^k$ is employed, problem (B.14) gives rise to a linear system

$$A_k u = f,$$

where $A = A_k$ and $f$ are defined as before, using the approximate bilinear form $a_Q(\cdot,\cdot)$ and right hand side $(f,\cdot)_Q$. The properties of $a_Q(\cdot,\cdot)$ ensure that $A$ is symmetric and positive definite. Here we only mention the following two dimensional result. A proof can be found in [46, Lem. 5.5] or [352, Prop. 2.7]. We note that this bound is sharp.

**Theorem B.33** *Let $n = 2$ and $\{\phi_i = \phi_i^k\}$ be the nodal basis functions introduced in Sect. B.2. Then there exist two constants, such that*

$$c_7\, k^{-2}\, u^T u \le u^T A_k u \le C_8\, k\, u^T u, \quad u \in V.$$

*Consequently,*

$$\kappa(A_k) \le C k^3 \le \tilde{C}\, L_k^{3/n},$$

*with $L_k = O(k^n)$ the number of degrees of freedom.*

We end this section with an estimate for the mass matrix $M$:

$$M_{ij} = (\phi_j^k, \phi_i^k)_Q.$$

We note that $M$ is diagonal thanks to the choice of nodal basis functions on the GLL nodes and the quadrature formula. Indeed, its entries are equal to the quadrature weights. The proof of the following lemma is a direct consequence of the bounds in (B.7).

**Lemma B.34** *Let $\{\phi_i = \phi_i^k\}$ be the nodal basis functions introduced in Sect. B.2. Then there exist two constants, such that*

$$c_9\, k^{-2n}\, u^T u \leq u^T M u \leq C_{10}\, k^{-n}\, u^T u, \quad u \in V.$$

# C

# Solution of Algebraic Linear Systems

Throughout this appendix we consider the solution of linear systems

$$Au = b, \tag{C.1}$$

with $u,\ b \in \mathbb{R}^n$, and $A$ an $n \times n$, real, invertible matrix. We use the notation $< u, v >= u^T v$, for $u,\ v \in \mathbb{R}^n$. We mention, e.g., [456, 231, 28, 252, 450, 233, 156, 363] as good references for this appendix.

## C.1 Eigenvalues and Condition Number

We recall that, given matrix $A \in \mathbb{R}^{n \times n}$, its eigenvalues $\lambda \in \mathbb{C}$ and eigenvectors $u \in \mathbb{C}^n \setminus \{0\}$ are solution of

$$Au = \lambda u.$$

The set of eigenvalues of $A$, also called *spectrum*, is denoted by $\sigma(A)$. The spectral radius $\rho(A)$ is defined as

$$\rho(A) := \max_{\lambda \in \sigma(A)} \{|\lambda|\}.$$

Given a matrix norm $\|\cdot\|_*$, we define the condition number of an invertible matrix $A$ by

$$\kappa_*(A) := \|A\|_* \|A^{-1}\|_*.$$

In the same way, given a second matrix $M$, we can consider the generalized eigenproblem

$$Au = \lambda M u.$$

A matrix $A$ is said *positive definite* if all its eigenvalues have positive real part or, equivalently, if $u^T Au$ has positive real part for $u \in \mathbb{C} \setminus \{0\}$. We note that in this case

$$u^T Au = u^T \frac{A + A^T}{2} u > 0, \quad u \in \mathbb{R}^n \setminus \{0\}. \tag{C.2}$$

If in addition $A$ is symmetric, then its eigenvalues are real and strictly positive. Throughout this monograph, we make use of the following property. It can be proven using simple transformations and well-known properties of the eigenvalues; see, e.g., [231, Sect. 8.1].

**Lemma C.1** *Let $A$ and $M$ be two symmetric, positive definite matrices of order $n$. For an arbitrary matrix $B \in \mathbb{R}^{n\times n}$, let*

$$\|B\|_A := \sup_{u\in\mathbb{R}^n} \frac{\|Bu\|_A}{\|u\|_A},$$

*with $\|u\|_A^2 := u^T Au$, and similarly for $\|B\|_M$. Then:*

1. *The following eigenvalue problems have the same $n$ eigenvalues*

$$Au = \lambda Mu, \tag{C.3}$$
$$M^{-1}Au = \lambda u, \tag{C.4}$$
$$(M^{-1/2}AM^{-1/2})u = \lambda u, \tag{C.5}$$
$$(A^{1/2}M^{-1}A^{1/2})u = \lambda u. \tag{C.6}$$

   *They are all real and strictly positive.*
2. *The smallest and largest eigenvalues of the problems above satisfy*

$$\lambda_{min} = \inf_{u\in\mathbb{R}^n} \frac{u^T Au}{u^T Mu}, \qquad \lambda_{max} = \sup_{u\in\mathbb{R}^n} \frac{u^T Au}{u^T Mu}.$$

3. *We have*

$$\|M^{-1}A\|_A = \|M^{-1}A\|_M = \lambda_{max} = \rho(M^{-1}A),$$
$$\|(M^{-1}A)^{-1}\|_A = \|(M^{-1}A)^{-1}\|_M = 1/\lambda_{min},$$

   *and thus*

$$\kappa_A(M^{-1}A) = \kappa_2(M^{-1/2}AM^{-1/2}) = \lambda_{max}/\lambda_{min}.$$

4. *We have*

$$(u^T Au \geq c\, u^T Mu, \quad u \in \mathbb{R}^n) \qquad \Longrightarrow \qquad \lambda_{min} \geq c.$$

   *Analogously,*

$$(u^T Au \leq C\, u^T Mu, \quad u \in \mathbb{R}^n) \qquad \Longrightarrow \qquad \lambda_{max} \leq C,$$

   *and thus*

$$\kappa_A(M^{-1}A) \leq C/c.$$

In this monograph, we use the notation

$$\kappa(M^{-1}A) = \kappa_A(M^{-1}A) = \kappa_M(M^{-1}A).$$

We note that since $M^{-1}A$ is not symmetric, its norm $\|M^{-1}A\|_2$ is not in general equal to its largest eigenvalue. However, $M^{-1}A$ is symmetric with respect to the scalar products induced by $A$ and $M$.

Lemma C.1 allows us to prove the following corollary, which we use extensively in this monograph. It basically ensures that *a good preconditioner of a good preconditioner remains a good preconditioner.*

**Corollary C.2** *Let $A, B$, and $C$ be three positive definite symmetric matrices. Then, $\kappa(C^{-1}A) \leq \kappa(C^{-1}B)\kappa(B^{-1}A)$.*

## C.2 Direct Methods

### C.2.1 Factorizations

Direct methods for the solution of linear systems usually construct a factorization of the matrix $A$. For a general matrix we can consider an $LU$ decomposition (possibly after a reordering of the equations or unknowns)

$$A = LU,$$

or, for a symmetric, positive definite matrix, a Choleski decomposition

$$A = U^T U = LDL^T.$$

Here, $L$ is a lower triangular matrix with ones on the main diagonal, $U$ is upper triangular, and $D$ is diagonal.

Block factorizations can also be employed. If for instance $A$ is symmetric and is partitioned into four blocks, corresponding to a partition of the unknowns into two sets, we can formally write

$$\begin{aligned} A &= \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{12}^T A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & S \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ A_{12}^T A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix}, \end{aligned} \tag{C.7}$$

with the Schur complement $S = A_{22} - A_{12}^T A_{11}^{-1} A_{12}$. The above factorization is not defined for an arbitrary symmetric matrix. However, we have the following result for positive definite matrices.

**Lemma C.3** *Let $A$ be symmetric and positive definite. Given an arbitrary partition into four blocks as in (C.7), corresponding to a partition of the unknowns into two sets, then the diagonal blocks $A_{11}$ and $A_{22}$ are also symmetric and positive definite and the factorization (C.7) is always defined. In addition $S$ is also symmetric and positive definite.*

*Proof.* The first assertion is easily proven by noting that

$$u_1^T A_{11} u_1 = [u_1^T \ 0] A [u_1^T \ 0]^T > 0, \quad u_1 \neq 0,$$

and similarly for $A_{22}$. The second assertion can be proven by noting that, according to (C.7), if $u_2 \neq 0$,

$$u_2^T S u_2 = \begin{bmatrix} 0 \\ u_2 \end{bmatrix}^T \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} 0 \\ u_2 \end{bmatrix} = v^T A v > 0,$$

where

$$v = \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ u_2 \end{bmatrix} \neq 0.$$

□

We refer to, e.g., [231] for a detailed presentation and discussion of these factorizations.

In some of the methods presented in this monograph, the application of the inverse of the Schur complement $S$ itself is required. Indeed, in order to calculate

$$u_2 = S^{-1} v_2,$$

it is not necessary to form $S^{-1}$ explicitly, but it is enough to solve a linear system involving the original matrix $A$ with a suitable right-hand side, and extract the lower block of the solution. Using (C.7), we can easily show that

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = A^{-1} \begin{bmatrix} 0 \\ v_2 \end{bmatrix}.$$

### C.2.2 Fill-in

Matrices arising from finite or spectral element approximations are generally *sparse*, i.e., the number of their non-zero elements is small compared to the total number of entries. However, the factors of the factorizations above are in general much less sparse. We say that two unknowns $i$ and $j$ are *connected* or *coupled* if either $a_{ij}$ or $a_{ji}$ is different from zero. Fill-in takes place at $(i, j)$, with $i$ and $j$ two uncoupled unknowns, if the corresponding entry $(i, j)$ or $(j, i)$ in one of the factors is non-zero. In the following we briefly describe how fill-in takes place and refer to [220] for additional details.

We consider an $LU$ factorization; see [231, Ch. 3]. Given $A^{(1)} = A$, the algorithm constructs a sequence of matrices $\{A^{(k)},\ k = 2, \ldots, n\}$, such that $U = A^{(n)}$. The entries of $A^{(k)}$ below the main diagonal along the columns from 1 to $k-1$ are zero. For $k = 1, \ldots, n-1$, the entries of the new matrix $A^{(k+1)}$ are constructed from $A^{(k)}$ and are given by

$$a_{ij}^{(k+1)} = \begin{cases} a_{ij}^{(k)}, & i = 1, \ldots, k, \quad j = i, \ldots, n, \\ 0, & j = 1, \ldots, k, \quad i = j+1, \ldots, n, \\ a_{ij}^{(k)} - \frac{a_{ik}^{(k)} a_{kj}^{(k)}}{a_{kk}^{(k)}}, & i, j = k+1, \ldots, n. \end{cases} \tag{C.8}$$

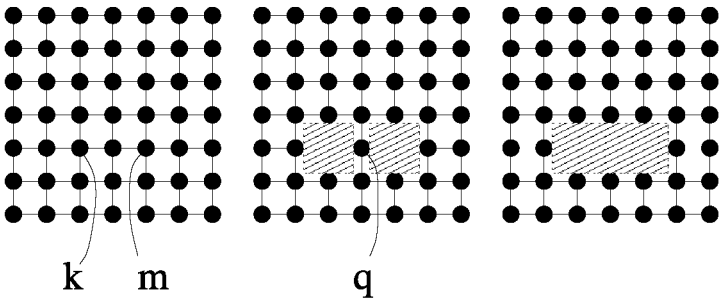


**Fig. C.1.** Fill-in: unknowns shared by the same square/rectangle are coupled.

After the construction of $A^{(k+1)}$, we say that the unknown $k$ has been *eliminated.* The algorithm preserves the entries in the rows from 1 to $k$, and introduces zeros in the $k$-th column below the main diagonal. It is clear that, even if an entry $a_{ij}^{(k)}$, associated with the unknowns $i$ and $j$, is zero, the new entry $a_{ij}^{(k+1)}$ may be different from zero if

$$a_{ik}^{(k)} a_{kj}^{(k)} \neq 0.$$

This is the case if unknowns $i$ and $j$ are not coupled, but they are *both* connected to the newly eliminated unknown $k$. This is described graphically in Fig. C.1, left, where unknowns are represented by points on a lattice. Two unknowns are connected if they belong to the same square or rectangle. Elimination of the unknown $k$ brings in a coupling of the unknowns to which $k$ was connected. Similarly for the unknown $m$; see the two new shaded squares in Fig. C.1, center. The successive elimination of $q$ then couples all the unknowns belonging to the two bigger squares; see Fig. C.1, right. Clearly this process and the resulting factorization depend on the order in which the unknowns are eliminated. However, unknowns $k$ and $m$ are not coupled and their elimination can be made independently and in parallel and the resulting factor is independent of the order in which they are eliminated.

Similar remarks apply for block factorizations . Block elimination of a set of unknowns causes a coupling between all the unknowns that were connected to those in the block; cf., for instance, the block $\{k, m, q\}$ in Fig. C.1.

## C.3 Richardson Method

Given a splitting of the matrix $A$

$$A = M - N, \tag{C.9}$$

with $M$ non-singular, and an initial vector $u^0 \in \mathbb{R}^n$, we can set up an iterative scheme where, for $k \geq 0$,

$$Mu^{k+1} = Nu^k + b. \tag{C.10}$$

Clearly, if the sequence $\{u^k\}$ converges, it must converge to the solution of the linear system

$$Mu = Nu + b, \tag{C.11}$$

and then of (C.1). The iteration scheme (C.10) can be rewritten in a number of different ways: in term of the residual $r^k := b - Au^k$, we have the equation

$$u^{k+1} = u^k + M^{-1}r^k,$$

which shows the increment between two successive iterates. We note that the computation of the new iterate $u^{k+1}$ involves the calculation of the residual $r^k$ (and then the application of the original matrix $A$ to a vector) and the application of the inverse of $M$ to the residual. A more general iteration scheme may be obtained by using a *relaxation* parameter $\alpha \in \mathbb{R}$, which should have the effect of accelerating convergence:

$$u^{k+1} = u^k + \alpha M^{-1}r^k. \tag{C.12}$$

Using (C.1) and (C.12), and the iteration matrix

$$B := M^{-1}((1-\alpha)M + \alpha N) = I - \alpha M^{-1}A, \tag{C.13}$$

we can write

$$e^{k+1} = Be^k, \tag{C.14}$$

which provides an equation for the error $e^k := u - u^k$, and then

$$e^k = B^k e^0. \tag{C.15}$$

Convergence takes place, independently of the choice of $u^0$, if the error tends to zero as $k$ tends to infinity, independently of $e^0$. We note that equation (C.15) only involves the iteration matrix $B$, but not $A$ alone, and is independent of the right-hand side $b$. We have the following conditions for convergence; cf., e.g., [231, Sect. 10.1.2].

**Lemma C.4** *A necessary and sufficient condition for the convergence of (C.10) is that the spectral radius of the iteration matrix $B$ is strictly less than one. A sufficient condition is that*

$$\|B\|_* < 1,$$

*where $\|\cdot\|_*$ is any norm in $\mathbb{R}^{n\times n}$. If $\|\cdot\|_*$ is induced by a vector norm, then we have*

$$\|e^k\|_* \le \eta_*^k \|e^0\|_*, \quad \eta_* := \|B\|_*. \tag{C.16}$$

The quantity $\eta_*$ is called *convergence factor* and gives a bound for the error reduction at each step.

If no relaxation is employed ($\alpha = 1$), it is clear that best convergence is achieved when $M = A$, for which $B = 0$. The matrix $M$ is called *preconditioner*, and for the reasons above, in practice it is chosen in such a way to satisfy the two following, somewhat conflicting, conditions:

1. $M$ should be 'close' to $A$;
2. the inverse of $M$ should be easy to apply.

The case $M = I$ corresponds to no preconditioning.

If we decompose A into its lower diagonal, diagonal, and upper diagonal parts

$$A = L + D + U, \tag{C.17}$$

the Jacobi and Gauss-Seidel preconditioners correspond to the choices $M = D$ and $M = D + L$, respectively. Sufficient conditions for the convergence of the corresponding iterative methods are given in, e.g., [231, Sect. 10.1.2]. We can also consider a block decomposition in (C.17).

In case both $A$ and $M$ are symmetric and positive definite, we can employ Lemma C.1 in order to get more precise estimates in terms of the spectrum of $M^{-1}A$.

**Lemma C.5** *Let $A$ and $M$ be symmetric and positive definite and let $\lambda_{min}$ and $\lambda_{max}$ be the smallest and largest eigenvalues of $M^{-1}A$, respectively. Then the preconditioned Richardson method (C.12) converges for any choice of the initial vector, if and only if*

$$\alpha\mu \in (0, 2), \quad \mu \in \sigma(M^{-1}A),$$

*and thus*

$$\alpha \in (0, 2/\lambda_{max}).$$

*In this case, the convergence factor is*

$$\eta_A(\alpha) = \max\{|1 - \alpha\lambda_{min}|, |\alpha\lambda_{max} - 1|\}.$$

*The minimum value of $\eta_A(\alpha)$ is*

$$\eta_{opt} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} = \frac{\kappa(M^{-1}A) - 1}{\kappa(M^{-1}A) + 1}$$

*and is attained for $\alpha = 2/(\lambda_{max} + \lambda_{min})$.*

*Proof.* Using (C.13) and Lemma C.4, we see that the method converges if and only if

$$\sigma(I - \alpha M^{-1}A) \subset (-1, 1).$$

This is equivalent to

1. Initialize: $r^0 = b - Au^0$

2. Iterate $k = 0, 1, \cdots$ until convergence

$$\begin{aligned} \text{Precondition: } z^k &= M^{-1}r^k \\ u^{k+1} &= u^k + \alpha z^k \\ r^{k+1} &= b - Au^{k+1} \end{aligned}$$

**Fig. C.2.** Preconditioned Richardson iteration.

$$-1 < 1 - \alpha\mu < 1, \quad \mu \in \sigma(M^{-1}A),$$

which proves the first assertion. The bound for $\alpha$ is obtained by noting that the eigenvalues of $M^{-1}A$ are all positive. The convergence factor $\eta_A(\alpha)$ is found by taking the maximum of $|1 - \alpha\mu|$ over $\mu \in \sigma(M^{-1}A)$ and the value of $\eta_{opt}$ by taking the minimum of $\eta_A(\alpha)$ over $\alpha \in (0, 2/\lambda_{max})$. □

We note that in case no relaxation is employed ($\alpha = 1$), the Richardson method converges if and only if $\lambda_{max} < 2$ since all eigenvalues of $M^{-1}A$ are positive.

Given $u^0$ and a choice for $\alpha$, the preconditioned Richardson iteration is given in Fig. C.2.

## C.4 Steepest Descent

We note that if the matrix $A$ is symmetric and positive definite, the solution of the linear system (C.1) is the unique vector that minimizes the quadratic functional

$$\phi(x) = \frac{1}{2}x^T Ax - b^T x.$$

Indeed, the gradient of $\phi(x)$ is equal to $-r(x) = Ax - b$. If we consider the Richardson iteration of the previous section with no preconditioning, we see that the new iterate $u^{k+1}$ is found from $u^k$ by moving along the *search direction* $r^k$. The *step-length* $\alpha$ gives the amount of which one moves along this direction. In this case, the search direction coincides with the direction of $r^k = r(u^k)$, i.e., the direction where, locally, the reduction of $\phi(x)$ is maximum (steepest descent).

Given a direction and a positive definite, quadratic functional, the optimal $\alpha$ that achieves the maximum reduction of $\phi(x)$ at each step can be easily found by solving

$$\min_{\alpha \in \mathbb{R}} \phi(u^k + \alpha r^k\}.$$

The minimum is attained for

1. Initialize: $r^0 = b - Au^0$

2. Iterate $k = 0, 1, \cdots$ until convergence

$$\begin{aligned} \alpha^k &= < r^k, r^k > / < r^k, Ar^k > \\ u^{k+1} &= u^k + \alpha^k r^k \\ r^{k+1} &= r^k - \alpha^k A r^k \end{aligned}$$

**Fig. C.3.** Unpreconditioned Steepest Descent iteration.

$$\alpha = \alpha^k = \frac{< r^k, r^k >}{< r^k, Ar^k >}.$$

In addition, the new residual corresponding to the new iterate $u^{k+1} = u^k + \alpha^k r^k$ is

$$r^{k+1} = r^k - \alpha^k A r^k.$$

We therefore have the algorithm in Fig. C.3.

We note that only one application of the original matrix $A$ is required at each step, since the vector $Ar^k$ can be stored and used to evaluate the new residual. We have the following convergence result; cf. [363, Th. 3.3].

**Lemma C.6** *Let $A$ be symmetric and positive definite. Then, the Steepest Descent method with no preconditioning satisfies the error bound*

$$\|e^k\|_A \leq \eta_A^k \|e^0\|_A,$$

*where the convergence factor is*

$$\eta_A = \frac{\kappa_2(A) - 1}{\kappa_2(A) + 1}.$$

*Consequently, it converges for any choice of $u^0$.*

We note that the Steepest Descent method satisfies the same error bound as the Richardson method with an optimal choice of $\alpha$. However in this case no information about the spectrum of $A$ is required.

## C.5 Conjugate Gradient Method

For this section, we follow the presentation in [363, Sect. 5.1]. We also refer to [231, 450, 233, 156] for somewhat different presentations.

Let $A$ be symmetric and positive definite. Steepest Descent chooses the search directions as those of the residuals. In Conjugate Gradient the search directions $\{p^k\}$ are chosen as *conjugate* with respect to $A$, i.e.,

$$< p^i, Ap^j >=< p^i, p^j >_A = 0, \qquad i \neq j. \tag{C.18}$$

In addition, the algorithm provides a cheap way of generating conjugate directions. We assume now that we are given an initial guess $u^0$ and a set of conjugate directions $\{p^k\}$. We can then generate the sequence

$$u^{k+1} = u^k + \alpha^k p^k, \qquad k = 0, 1, \ldots \tag{C.19}$$

where $\alpha_k$ is the one-dimensional minimizer that solves

$$\min_{\alpha \in \mathbb{R}} \phi(u^k + \alpha p^k\}.$$

We have

$$\alpha = \alpha^k = \frac{< r^k, p^k >}{< p^k, Ap^k >}.$$

Since the vectors $\{p^k\}$ are orthogonal with respect to the scalar product $< \cdot, \cdot >_A$, they are also linear independent and thus provide a basis for $\mathbb{R}^n$. We have the following property; cf. [363, Th. 5.1 and 5.2].

**Lemma C.7** *Let $u^0 \in \mathbb{R}^n$ and $\{p^k\}$ be an arbitrary set of conjugate directions. For $k \geq 1$, we have*

$$< r_k, p_i >= 0, \quad i = 0, 1, \ldots, k-1,$$

*and $u^k$ minimizes $\phi(x)$ over the space*

$$u^0 + \text{span}\,\{p_i,\ i = 0, 1, \ldots, k-1\}.$$

*Consequently, the sequence generated by (C.19) converges to the solution of (C.1) in at most n steps.*

We note that the previous result holds for *any* set of conjugate directions. In addition, since it is not valid for an arbitrary set of linearly independent directions, it highlights the importance of the conjugacy condition (C.18).

The Conjugate Gradient algorithm now provides a particular choice for the search directions: they are of the form

$$p^k = r^k + \beta^k p^{k-1},$$

where the scalar $\beta^k$ is uniquely determined by imposing the condition that the new direction is conjugate to the previous ones; see (C.18). We can then write the algorithm in Fig. C.4.

As for Steepest Descent, only one application of the original matrix $A$ is required at each step. We have the following property; cf.[363, Th. 5.3].

**Lemma C.8** *Suppose that the k-th iterate generated by the Conjugate Gradient method is not the solution of (C.1). Then,*

1. Initialize: $r^0 = b - Au^0$

2. Iterate $k = 1, 2, \cdots$ until convergence

$$\begin{aligned}
\beta^k &= \langle r^{k-1}, r^{k-1}\rangle / \langle r^{k-2}, r^{k-2}\rangle \quad [\beta^1 = 0] \\
p^k &= r^{k-1} + \beta^k p^{k-1} \quad [p^1 = r^0] \\
\alpha^k &= \langle r^{k-1}, r^{k-1}\rangle / \langle p^k, Ap^k\rangle \\
u^k &= u^{k-1} + \alpha^k p^k \\
r^k &= r^{k-1} - \alpha^k Ap^k
\end{aligned}$$

**Fig. C.4.** Unpreconditioned Conjugate Gradient.

$$\begin{aligned}
&< r^k, r^i >= 0, \qquad i = 0, \ldots, k-1, \\
&\text{span}\{r^i, \quad i = 0, 1, \ldots, k\} = \text{span}\{A^i r^0, \quad i = 0, 1, \ldots, k\} \\
&\text{span}\{p^i, \quad i = 0, 1, \ldots, k\} = \text{span}\{A^i r^0, \quad i = 0, 1, \ldots, k\} \\
&< p^k, Ap^i >= 0, \qquad i = 0, \ldots, k-1.
\end{aligned}$$

*Therefore the sequence $\{u^k\}$ converges to the solution of (C.1) in at most n steps.*

We note that the residuals and the conjugate directions both provide bases for the *Krylov spaces* $\mathcal{K}_k = \mathcal{K}_k(r_0, A) = \text{span}\{A^i r^0, \quad i = 0, 1, \ldots, k-1\}$ and that, thanks to Lemma C.7, the iterate $u^k$ minimizes $\phi(x)$ over the space $u^0 + \mathcal{K}_k(r_0, A)$.

Convergence of the unpreconditioned Conjugate Gradient depends on the condition number of $A$: we have the following result. We refer to [450, Th. 38.5] for a proof.

**Lemma C.9** *Let $A$ be symmetric and positive definite. Then, the Conjugate Gradient method satisfies the error bound*

$$\|e^k\|_A \leq 2\eta_A^k \|e^0\|_A,$$

*where the convergence factor is*

$$\eta_A = \frac{\sqrt{\kappa_2(A)} - 1}{\sqrt{\kappa_2(A)} + 1}.$$

We note that the previous result may provide just a crude estimate of the rate of convergence. If the eigenvalues of $A$ are clustered, it is then well-known that convergence is fast; we refer to, e.g., [450, Sect. VI.38] or [156, Sect. 6.6.4] for additional results and comments.

1. Initialize: $r^0 = b - Au^0$

2. Iterate $k = 1, 2, \cdots$ until convergence

Precondition: $z^{k-1} = M^{-1}r^{k-1}$

$$\beta^k = \langle z^{k-1}, r^{k-1}\rangle / \langle z^{k-2}, r^{k-2}\rangle \quad [\beta^1 = 0]$$
$$p^k = z^{k-1} + \beta^k p^{k-1} \quad [p^1 = z^0]$$
$$\alpha^k = \langle z^{k-1}, r^{k-1}\rangle / \langle p^k, Ap^k\rangle$$
$$u^k = u^{k-1} + \alpha^k p^k$$
$$r^k = r^{k-1} - \alpha^k Ap^k$$

**Fig. C.5.** Preconditioned Conjugate Gradient.

The Conjugate Gradient iteration also provides an estimate of the eigenvalues of the matrix $A$ (and thus of $\kappa_2(A)$). Indeed, thanks to Lemma C.8, the columns of

$$R_k = [r_0/\|r_0\|, \ldots, r_{k-1}/\|r_{k-1}\|],$$

provides an orthonormal basis for the Krylov space $\mathcal{K}_k(r_0, A)$. One can prove that the restriction of $A$ to $\mathcal{K}_k(r_0, A)$

$$T_k = R_k^T A R_k$$

is a symmetric, tridiagonal matrix (see [231, Sect. 10.2.5]), the entries of which can be constructed from the coefficients $\alpha^i$ and $\beta^i$ of the Conjugate Gradient iteration. We refer to [430] for the precise formulas. By calculating the eigenvalues of $T_k$ one can easily obtain estimates of the largest and smallest eigenvalues of $A$.

Lemma C.9 provides a better bound than that for Steepest Descent in Lemma C.6, however, in case $\kappa_2(A)$ is large as in the case of finite and spectral element approximations, preconditioning is necessary. Given a symmetric, positive definite matrix $M$, we can consider the modified linear system

$$M^{-1/2}AM^{-1/2}v = M^{-1/2}b, \quad v = M^{1/2}u.$$

We note that $M^{-1/2}AM^{-1/2}$ is symmetric and positive definite and it reduces to the identity in case $M = A$. We can then consider $M$ as a *preconditioner* for $A$ and apply the unpreconditioned Conjugate Gradient method to this modified system. After some manipulations, we find the algorithm in Fig. C.5.

We note that this algorithm does not involve the application of $M^{-1/2}$, but only of $M^{-1}$. It is not however the Conjugate Gradient method applied to a system involving the matrix $M^{-1}A$, which is not symmetric in general.

In addition, it requires one application of the original matrix $A$ and one application of $M^{-1}$ to a vector. Using Lemmas C.9 and C.1, we find the following result.

**Lemma C.10** *Let $A$ and $M$ be symmetric and positive definite. Then, the preconditioned Conjugate Gradient method satisfies the same error bound as in Lemma C.9, with*

$$\eta_A = \frac{\sqrt{\kappa(M^{-1}A)} - 1}{\sqrt{\kappa(M^{-1}A)} + 1}.$$

In Sect. 2.5.2, we give a particular variant of the Conjugate Gradient. Estimates for the eigenvalues of $M^{-1}A$ can also be obtained in this case using the coefficients $\alpha^i$ and $\beta^i$.

## C.6 Methods for Non-Symmetric and Indefinite Systems

### C.6.1 The Generalized Minimal Residual Method

The Generalized Minimal Residual Method (GMRES) is described in [407] and the theory developed in $L^2(\Omega)$ can be found in [185]. Here we briefly describe the GMRES algorithm and state a convergence theorem without proof. We follow the presentation in [407].

Let $A$ be an invertible matrix, not necessarily symmetric or positive definite, and $u^0 \in \mathbb{R}^n$ an initial vector. The GMRES algorithm basically relies on two ideas. The first is to build orthonormal bases $\{v^1, v^2, \ldots, v^k\}$ for the Krylov subspaces

$$\mathcal{K}_k = \mathcal{K}_k(r^0, A) \;=\; \text{span}\,\{r^0, Ar^0, \cdots, A^{k-1}r^0\},$$

where $r^0 = b - Au^0$ is the initial residual. It does so by the so-called *Arnoldi iteration* which also produces the representation of $A$ in $\mathcal{K}_k$. An approximate solution $u^j$ is then found by solving the least-square problem

$$\min_{u \in u^0 + \mathcal{K}_j} \|b - Au\|_2 = \min_{z \in \mathcal{K}_j} \|r^0 - Az\|_2. \tag{C.20}$$

We find the algorithm in Fig. C.6, where, $\beta = \|r^0\|$, $e_1$ is the first column of the $(j+1) \times (j+1)$ identity matrix, $V_j$ is the matrix, the columns of which are the vectors $v^k$, $k = 1, \ldots, j$, and $H_j$ is a $(j+1) \times j$ matrix, the non-zero entries of which are the elements $h_{i,k}$; see [407].

Several comments are required. The second step in the iteration above is a step of the Arnoldi iteration. Indeed, we easily find

$$Av^k = \sum_{i=1}^{k+1} h_{i,k} v^i, k = 1, \ldots, j,$$

1. Initialize: $r^0 = b - Au^0$

$$r^0 = b - Au^0$$
$$v^1 = r^0/\|r^0\|_2$$

2. Iterate $k = 1, 2, \ldots, j$

$$q^k = Av^k$$
$$h_{i,k} = < q^k, v^i >, \quad i = 1, \ldots, k$$
$$\hat{v}^k = q^k - \sum_{i=1}^{k} h_{i,k} v^i$$
$$h_{k+1,k} = \|\hat{v}^k\|_2$$
$$v^{k+1} = \hat{v}^k / h_{k+1,k}$$

3. Form the approximate solution

$$u^j = u^0 + V_j w^j, \quad \text{where } w^j \text{ minimizes } \|\beta e_1 - H_j w\|_2, \; w \in \mathbb{R}^j$$

**Fig. C.6.** GMRES iteration.

and thus

$$AV_j = V_{j+1} H_j \tag{C.21}$$

Since the columns of $V_j$ and $V_{j+1}$ are orthonormal, we find

$$V_j^T A V_j = \tilde{H}_j,$$

where $\tilde{H}_j$ is a $j \times j$ Hessemberg matrix, the non-zero entries of which are the elements $h_{i,k}$. The representation of $A$ in $\mathcal{K}_k$ is thus $\tilde{H}_j$ and can be used to extract information on the spectrum of $A$.

We next remark on the third step of the algorithm. It indeed solves the minimization problem (C.20). In order to see that, we rewrite (C.20) as

$$\min_{w \in \mathbb{R}^j} \|\beta v^1 - AV_j w\|_2.$$

Using the representation (C.21) and the fact that the columns of $V_{j+1}$ are orthonormal, we can finally write

$$\|\beta v^1 - AV_j w\|_2 = \|V_{j+1}(\beta e^1 - H_j w)\|_2 = \|\beta e^1 - H_j w\|_2.$$

Some other important practical aspects are given in [407]. Here we only mention that the least-square problem in the third step of the algorithm can be solved by using a $QR$ factorization of $H_j$. Such factorization can be obtained from that of $H_{j-1}$ at the previous step. In addition, the residual norm of the

approximate solution $u^j$ can be computed from the $QR$ factorization at no extra cost, without finding $u^j$.

Thanks to the minimization property (C.20), the exact solution would be reached in no more than $n$ iterations if we use exact arithmetic. Following [185], the rate of convergence of the GMRES method can be characterized in terms of the minimal eigenvalue of the symmetric part of the operator and the norm of the operator. They are defined by

$$c_p = \min_{u \in \mathbb{R}^n} \frac{< u, Au >}{< u, u >}, \qquad C_p = \max_{u \in \mathbb{R}^n} \frac{\|Au\|_2}{\|u\|_2};$$

cf. (C.2).

By considering the decrease of the norm of the residual in a single step, the following theorem can be established.

**Lemma C.11** *If $c_p > 0$, then, after $k$ steps, the norm of the residual is bounded by*

$$\|r^k\|_2 \leq \left(1 - \frac{c_p^2}{C_p^2}\right)^{k/2} \|r^0\|_2.$$

The rate of convergence can be improved by using a suitable preconditioner. Let $M$ be an invertible matrix. We consider the precondition problem, derived from (C.1),

$$M^{-1}Au = M^{-1}b.$$

The GMRES algorithm can be written in this case as in Fig. C.7.

Here, $\beta$, $V_j$, and $H_j$ are defined as before.

### C.6.2 The Conjugate Residual Method

The preconditioned conjugate residual (PCR) method is a generalization of Conjugate Gradient for symmetric, not necessarily positive definite, systems and symmetric, positive definite preconditioners; see [25, 252]. A fine presentation can be found in [252, sect. 9.5]; see also [280, sect. 2.1.3].

We consider the linear system (C.1) and a preconditioner $M$. One version of PCR is given in Fig. C.8. We note that the algorithm can be implemented in such a way that only one matrix-vector product for $A$ and $M^{-1}$ is needed, by introducing three additional vectors; see, e.g., [280, sect. 2.1.3].

In order to give a convergence result, we define

$$\kappa = \kappa(M^{-1}A) = \frac{\mu_{max}}{\mu_{min}} = \frac{\max\{|\lambda| : \lambda \in \sigma(M^{-1}A)\}}{\min\{|\lambda| : \lambda \in \sigma(M^{-1}A)\}},$$

where $\sigma(M^{-1}A)$ is the spectrum of $M^{-1}A$. We have the following result.

1. Initialize: $r^0 = b - Au^0$

$$\begin{aligned} r^0 &= b - Au^0 \\ \text{Precondition: } z^0 &= M^{-1} r^0 \\ v^1 &= z^0 / \|z^0\|_2 \end{aligned}$$

2. Iterate $k = 1, 2, \ldots, j$

$$\begin{aligned} q^k &= Av^k \\ \text{Precondition: } z^k &= M^{-1} q^k \\ h_{i,k} &= < z^k, v^i >, \quad i = 1, \ldots, k \\ \hat{v}^k &= z^k - \sum_{i=1}^{k} h_{i,k} v^i \\ h_{k+1,k} &= \|\hat{v}^k\|_2 \\ v^{k+1} &= \hat{v}^k / h_{k+1,k} \end{aligned}$$

3. Form the approximate solution

$$u^j = u^0 + V_j w^j, \quad \text{where } w^j \text{ minimizes } \|\beta e_1 - H_j w\|_2, \; w \in \mathbb{R}^j.$$

**Fig. C.7.** Preconditioned GMRES iteration.

1. Initialize: $r^0 = b - Au^0, \quad p^{-1} = 0, \quad p^0 = M^{-1} r^0$

2. Iterate $k = 1, 2, \cdots$ until convergence

$$\begin{aligned} \beta &= \langle r^{k-1}, M^{-1} A p^{k-1} \rangle / \langle A p^{k-1}, M^{-1} A p^{k-1} \rangle \\ u^k &= u^{k-1} + \beta p^{k-1} \\ r^k &= r^{k-1} - \beta A p^{k-1} \\ \alpha_0 &= \langle A M^{-1} A p^{k-1}, M^{-1} A p^{k-1} \rangle / \langle A p^{k-1}, M^{-1} A p^{k-1} \rangle \\ \alpha_1 &= \langle A M^{-1} A p^{k-1}, M^{-1} A p^{k-2} \rangle / \langle A p^{k-2}, M^{-1} A p^{k-2} \rangle \\ p^k &= M^{-1} A p^{k-1} - \alpha_0 p^{k-1} - \alpha_1 p^{k-2} \end{aligned}$$

**Fig. C.8.** PCR iteration.

**Lemma C.12** *Let $A$ be regular and symmetric and $M$ symmetric and positive definite. Then, after $k$ steps, the norm of the residual is bounded by*

$$\|M^{-1/2} r^k\|_2 \leq \frac{2\rho^\mu}{1 + \rho^{2\mu}} \|M^{-1/2} r^0\|_2,$$

*with* $\rho = \frac{\kappa-1}{\kappa+1}$ *and* $\mu \in \mathbb{Z}$, *such that* $\frac{k}{2} - 1 < \mu \leq \frac{k}{2}$.

The proof can be found in, e.g., [252, Th. 9.5.13]. A more general result can also be given in case $A$ is mildly indefinite and bounds for the negative and positive parts of $\sigma(M^{-1}A)$ are available; see [252, Th. 9.5.14].

# References


1. Gassan Abdoulaev, Yves Achdou, Jean-Claude Hontand, Yuri A. Kuznetsov, Olivier Pironneau, and Christophe Prud'homme. Domain decomposition for Navier-Stokes equations. In *ICIAM 99 (Edinburgh)*, pages 191–204. Oxford Univ. Press, Oxford, 2000.
2. Yves Achdou, Caroline Japhet, Patric Le Tallec, Frédéric Nataf, François Rogier, and Marina Vidrascu. Domain decomposition methods for non-symmetric problems. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 3–17. DDM.org, 1999. Greenwich, England, July 20–24, 1998.
3. Yves Achdou and Yuri A. Kuznetsov. Substructuring preconditioners for finite element methods on nonmatching grids. *East-West J. Numer. Math.*, 3(1):1–28, 1995.
4. Yves Achdou, Yuri A. Kuznetsov, and Olivier Pironneau. Substructuring preconditioners for the $Q_1$ mortar element method. *Numer. Math.*, 71:419–449, 1995.
5. Yves Achdou, Patrick Le Tallec, Frédéric Nataf, and Marina Vidrascu. A domain decoposition preconditioner for an advection-diffusion problem. *Comp. Methods Appl. Mech. Engrg*, 184:145–170, 2000.
6. Yves Achdou, Yvon Maday, and Olof B. Widlund. Iterative substructuring preconditioners for mortar element methods in two dimensions. *SIAM J. Numer. Anal.*, 36(2):551–580, 1999.
7. Yves Achdou and Frédéric Nataf. A Robin-Robin preconditioner for an advection-diffusion problem. *C. R. Acad. Sci. Paris*, 325, Série I:1211–1216, 1997.
8. Robert A. Adams. *Sobolev spaces.* Academic Press, New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
9. Valeri I. Agoshkov. Poincaré-Steklov operators and domain decomposition methods in finite dimensional spaces. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 73–112, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.
10. Valeri I. Agoshkov and Vyacheslav I. Lebedev. Poincaré-Steklov operators and the methods of partition of the domain in variational problems. In

Guri I. Marchuk, editor, *Computational Processes and Systems*, pages 173–227, Moscow, 1985. Nauka. In Russian.

11. Mark Ainsworth. A preconditioner based on domain decomposition for *hp*–FE approximation on quasi–uniform meshes. *SIAM J. Numer. Anal.*, 33:1358–1376, 1996.
12. Mark Ainsworth and Spencer Sherwin. Domain decomposition preconditioners for $p$ and $hp$ finite element approximation of Stokes equations. *Comp. Methods Appl. Mech. Engrg.*, 175:243–266, 1999.
13. Pierre Alart, Mikael Barboteu, Patrick Le Tallec, and Marina Vidrascu. Additive Schwarz method for nonsymmetric problems: application to frictional multicontact problems. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, pages 3–13, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
14. Ana Alonso, R. Loredana Trotta, and Alberto Valli. Coercive domain decomposition algorithms for advection-diffusion equations and systems. *J. Comput. Appl. Math.*, 96(1):51–76, 1998.
15. Ana Alonso and Alberto Valli. Some remarks on the characterization of the space of tangential traces of $H(\mathrm{rot};\omega)$ and the construction of an extension operator. *Manuscripta Math.*, 89(2):159–178, 1996.
16. Ana Alonso and Alberto Valli. An optimal domain decomposition preconditioner for low-frequency time-harmonic Maxwell equations. *Math. Comp.*, 68(226):607–631, 1999.
17. Ana Alonso and Alberto Valli. Domain decomposition methods for time-harmonic Maxwell equations: Numerical results. In Luca Pavarino and Andrea Toselli, editors, *Recent Developments in Domain Decomposition Methods*, Lecture Notes in Computational Science and Engineering, Volume 23, pages 157–171. Springer, 2002.
18. Chérif Amrouche, Christine Bernardi, Monique Dauge, and Vivette Girault. Vector potentials in three-dimensional non-smooth domains. *Math. Methods Appl. Sci.*, 21:823–864, 1998.
19. W. Kyle Anderson, William D. Gropp, Dinesh K. Kaushik, David E. Keyes, and Barry F. Smith. Achieving high sustained performance in an unstructured mesh CFD application. In *Proceedings of SC1999*. IEEE Computer Society, 1999. Bell Prize award paper.
20. Douglas N. Arnold and Franco Brezzi. Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO Modél. Math. Anal. Numér.*, 19(1):7–32, 1985.
21. Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. Preconditioning in $H(\mathrm{div})$ and applications. *Math. Comp.*, 66(219):957–984, 1997.
22. Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. Multigrid preconditioning in $H(\mathrm{div})$ on non-convex polygons. *Comput. Appl. Math.*, 17(3):303–315, 1998.
23. Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. Multigrid in H(div) and H(curl). *Numer. Math.*, 85(2):197–217, 2000.
24. Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa. *Studies in linear and non-linear programming*. Stanford Mathematical Studies in the Social Sciences, vol. II. Stanford University Press, Stanford, Calif., 1958.

25. Steven F. Ashby, Thomas A. Manteuffel, and Paul E. Saylor. A taxonomy for conjugate gradient methods. *SIAM J. Numer. Anal.*, 27(6):1542–1568, 1990.
26. Gennady P. Astrakhantsev. Method of fictitious domains for a second-order elliptic equation with natural boundary conditions. *U.S.S.R. Comput. Math. Math. Phys.*, 18:114–121, 1978.
27. Andreas Auge, Gert Lube, and Frank-Christian Otto. A non-overlapping domain decomposition method with adaptive interface conditions for elliptic problems. In *Numerical treatment of multi-scale problems (Kiel, 1997)*, volume 70 of *Notes Numer. Fluid Mech.*, pages 12–23. Vieweg, Braunschweig, 1999.
28. Owe Axelsson. *Iterative Solution Methods.* Cambridge University Press, New York, 1994.
29. Ivo Babuška. Über Schwarzsche Algorithmen in partiellen Differentialgleichungen der mathematischen Physik. *ZAMM*, 37(7/8):243–245, 1957.
30. Ivo Babuška. The finite element method with Lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973.
31. Ivo Babuška, Alan Craig, Jan Mandel, and Juhani Pitkäranta. Efficient preconditioning for the p-version finite element method in two dimensions. *SIAM J. Numer. Anal.*, 28(3):624–661, 1991.
32. Satish Balay, Kris Buschelman, William D. Gropp, Dinesh Kaushik, Matt Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc home page. http://www.mcs.anl.gov/petsc, 2001.
33. Satish Balay, Kris Buschelman, William D. Gropp, Dinesh Kaushik, Matt Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 2.1.5, Argonne National Laboratory, 2002.
34. Randolph E. Bank and L. Ridgway Scott. On the conditioning of finite element equations with highly refined meshes. *SIAM J. Numer. Anal.*, 26(6):1383–1394, 1989.
35. Randolph E. Bank, Bruno D. Welfert, and Harry Yserentant. A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56:645 – 666, 1990.
36. Mikael Barboteu, Pierre Alart, and Marina Vidrascu. A domain decomposition strategy for nonclassical frictional multi-contact problems. *Comp. Methods Appl. Mech. Engrg.*, 190(37–38):4785–4803, 2001.
37. Timothy J. Barth, Tony F. Chan, and Wei-Pai Tang. A parallel non-overlapping domain decomposition algorithm for compressible fluid flow problems on triangulated domains. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 23–41. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
38. Andrew C. Bauer and Abani K. Patra. Performance of parallel preconditioners for adaptive *hp* FEM discretization of incompressible flows. *Comm. Numer. Methods Engrg.*, 18(5):305–313, 2002.
39. Faker Ben Belgacem. Polynomial extensions of compatible polynomial traces in three dimensions. *Comput. Methods Appl. Mech. Engrg.*, 116(1-4):235–241, 1994. ICOSAHOM'92 (Montpellier, 1992).
40. Jean-David Benamou. A domain decomposition method for the optimal control of system governed by the Helmholtz equation. In Gary C. Cohen, editor,

*Third international conference on mathematical and numerical wave propagation phenomena.* SIAM, 1995.

41. Jean-David Benamou. Domain decomposition methods with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. *SIAM J. Numer. Anal.*, 33(6):2401–2416, 1996.
42. Jean-David Benamou and Bruno Després. A domain decomposition method for the Helmholtz equation and related optimal control problems. *J. Comput. Phys.*, 136:68–82, 1997.
43. Jefferey K. Benninghof and Richard B. Lehoucq. An automated multilevel substructuring method for eigenspace computation in linear elastodynamics. *SIAM J Sci. Comput.*, 25(6):2084–2106, 2004.
44. Michele Benzi, Andreas Frommer, Reinhard Nabben, and Daniel B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numer. Math.*, 89(4):605–639, 2001.
45. Christine Bernardi and Yvon Maday. Relèvement polynomial de traces et applications. *RAIRO Modél. Math. Anal. Numér.*, 24(5):557–611, 1990.
46. Christine Bernardi and Yvon Maday. *Approximations Spectrales de Problèmes aux Limites Elliptiques*, volume 10 of *Mathèmatiques & Applications.* Springer-Verlag France, Paris, 1992.
47. Christine Bernardi and Yvon Maday. Polynomial interpolation results in Sobolev spaces. *J. Comput. Appl. Math.*, 43:53 – 80, 1992.
48. Christine Bernardi and Yvon Maday. Spectral methods. In Philippe G. Ciarlet and Jacques-Louis Lions, editors, *Handbook of Numerical Analysis, Volume V: Techniques of Scientific Computing (Part 2)*, pages 209–485. North-Holland, 1997.
49. Christine Bernardi and Yvon Maday. Uniform inf-sup conditions for the spectral element discretization of the Stokes problem. *Math. Models Methods Appl. Sci.*, 9:395–414, 1999.
50. Christine Bernardi, Yvon Maday, and Francesca Rapetti. *Discrétisations variationnelles de problèmes aux limites elliptiques*, volume 45 of *Mathématiques et Applications.* Springer-Verlag, Berlin, 2004.
51. Manoj Bhardwaj, David Day, Charbel Farhat, Michel Lesoinne, Kendall Pierson, and Daniel Rixen. Application of the FETI method to ASCI problems: Scalability results on one thousand processors and discussion of highly heterogeneous problems. *Internat. J. Numer. Methods Engrg.*, 47:513–535, 2000.
52. Ion Bică. *Iterative Substructuring Algorithms for the p-version Finite Element Method for Elliptic Problems.* PhD thesis, Courant Institute of Mathematical Sciences, September 1997. TR-743, Department of Computer Science.
53. Ion Bică. Nonoverlapping Domain Decomposition Algorithms for the p-version Finite Element Method for Elliptic Problems. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 231–237. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
54. Petter E. Bjørstad. Multiplicative and Additive Schwarz Methods: Convergence in the 2 domain case. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods*, pages 147–159, Philadelphia, PA, 1989. SIAM. Los Angeles, California, January 14–16, 1988.

55. Petter E. Bjørstad, Maksymilian Dryja, and Talal Rahman. Additive Schwarz methods for elliptic mortar finite element problems. *Numer. Math.*, 95(3):427–457, 2003.
56. Petter E. Bjørstad, Magne Espedal, and David E. Keyes, editors. *Domain Decomposition Methods in Sciences and Enginnering. Ninth International Conference on Domain Decomposition Methods.* DDM.org, 1998. Ullensvang at the Hardanger Fjord, Norway, June 4–7, 1996.
57. Petter E. Bjørstad and Anders Hvidsten. Iterative methods for substructured elasticity problems in structural analysis. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 301–312, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.
58. Petter E. Bjørstad and Jan Mandel. On the spectra of sums of orthogonal projections with applications to parallel computing. *BIT*, 31:76–88, 1991.
59. Petter E. Bjørstad, Randi Moe, and Morten Skogen. Parallel domain decomposition and iterative refinement algorithms. In Wolfgang Hackbusch, editor, *Parallel Algorithms for PDEs, Proceedings of the 6th GAMM-Seminar held in Kiel, Germany, January 19–21, 1990*, Braunschweig, Wiesbaden, 1990. Vieweg-Verlag.
60. Petter E. Bjørstad and Morten Skogen. Domain decomposition algorithms of Schwarz type, designed for massively parallel computers. In David E. Keyes, Tony F. Chan, Gérard A. Meurant, Jeffrey S. Scroggs, and Robert G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 362–375, Philadelphia, PA, 1992. SIAM. Norfolk, Virgina, May 6–8, 1991.
61. Petter E. Bjørstad and Olof B. Widlund. Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM J. Numer. Anal.*, 23(6):1093–1120, 1986.
62. Petter E. Bjørstad and Olof B. Widlund. To overlap or not to overlap: A note on a domain decomposition method for elliptic problems. *SIAM J. Sci. Statist. Comput.*, 10(5):1053–1061, 1989.
63. Christoph Börgers. The Neumann–Dirichlet domain decomposition method with inexact solvers on the subdomains. *Numer. Math.*, 55:123–136, 1989.
64. Alain Bossavit. A rationale for edge elements in 3d field computations. *IEEE Trans. Mag.*, 24:74–79, 1988.
65. Jean-François Bourgat, Roland Glowinski, Patrick Le Tallec, and Marina Vidrascu. Variational formulation and algorithm for trace operator in domain decomposition calculations. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods*, pages 3–16, Philadelphia, PA, 1989. SIAM. Los Angeles, California, January 14–16, 1988.
66. Frédéric Bourquin. Analysis and comparison of several component mode synthesis methods on one-dimensional domains. *Numer. Math*, 58(1):11–34, 1990.
67. Frédéric Bourquin. Component mode synthesis and eigenvalues of second order operators: Discretization and algorithm. *Rairo Modél. Math. Anal. Numér.*, 26:385–423, 1992.
68. Dietrich Braess and Rüdiger Verfürth. Multigrid methods for nonconforming finite element methods. *SIAM J. Numer. Anal.*, 27(4):979–986, 1990.

69. James H. Bramble. A second order finite difference analogue of the first biharmonic boundary value problem. *Numer. Math.*, 9:236–249, 1966.
70. James H. Bramble and Joseph E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, 50(181):1–17, 1988.
71. James H. Bramble and Joseph E. Pasciak. A domain decomposition technique for Stokes problems. *Appl. Numer. Math.*, 6:251–261, 1989/90.
72. James H. Bramble, Joseph E. Pasciak, and Alfred H. Schatz. The construction of preconditioners for elliptic problems by substructuring, I. *Math. Comp.*, 47(175):103–134, 1986.
73. James H. Bramble, Joseph E. Pasciak, and Alfred H. Schatz. An iterative method for elliptic problems on regions partitioned into substructures. *Math. Comp.*, 46(173):361–369, 1986.
74. James H. Bramble, Joseph E. Pasciak, and Alfred H. Schatz. The construction of preconditioners for elliptic problems by substructuring, II. *Math. Comp.*, 49(179):1–16, 1987.
75. James H. Bramble, Joseph E. Pasciak, and Alfred H. Schatz. The construction of preconditioners for elliptic problems by substructuring, III. *Math. Comp.*, 51(184):415–430, 1988.
76. James H. Bramble, Joseph E. Pasciak, and Alfred H. Schatz. The construction of preconditioners for elliptic problems by substructuring, IV. *Math. Comp.*, 53(187):1–24, 1989.
77. James H. Bramble, Joseph E. Pasciak, and Apostol Vassilev. Non-overlapping domain decomposition preconditioners with inexact solves. In Petter E. Bjørstad, Magne Espedal, and David E. Keyes, editors, *Domain Decomposition Methods in Sciences and Enginnering. Ninth International Conference on Domain Decomposition Methods*, pages 40–52. DDM.org, 1998. Ullensvang at the Hardanger Fjord, Norway, June 4–7, 1996.
78. James H. Bramble, Joseph E. Pasciak, and Apostol T. Vassilev. Analysis of the inexact Uzawa algorithm for saddle point problems. *SIAM J. Numer. Anal.*, 34(3):1072–1092, 1997.
79. James H. Bramble, Joseph E. Pasciak, Junping Wang, and Jinchao Xu. Convergence estimates for multigrid algorithms without regularity assumptions. *Math. Comp.*, 57(195):23–45, 1991.
80. James H. Bramble, Joseph E. Pasciak, Junping Wang, and Jinchao Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57(195):1–21, 1991.
81. James H. Bramble and Jinchao Xu. Some estimates for a weighted $L^2$ projection. *Math. Comp.*, 56(194):463–476, 1991.
82. Achi Brandt. Multi-level adaptive solutions to boundary value problems. *Numer. Math.*, 31:333–390, 1977.
83. Susanne C. Brenner. An optimal-order multigrid method for P1 nonconforming finite elements. *Math. Comp.*, 52(185):1–15, 1989.
84. Susanne C. Brenner. A multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method. *SIAM J. Numer. Anal.*, 29(3):647–678, 1992.
85. Susanne C. Brenner. The condition number of the Schur complement in domain decomposition. *Numer. Math.*, 83:187–203, 1999.
86. Susanne C. Brenner. Lower bounds of two-level additive Schwarz preconditioners with small overlap. *SIAM J. Sci. Comput.*, 21(5):1657–1669, 2000.

87. Susanne C. Brenner. A new look at FETI. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, pages 41–51, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
88. Susanne C. Brenner. An additive Schwarz preconditioner for the FETI method. *Numer. Math.*, 94:1–31, 2003.
89. Susanne C. Brenner and Qingmi He. Lower bounds for three-dimensional nonoverlapping domain decomposition algorithms. *Numer. Math.*, 93(3):445–470, 2003.
90. Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, Berlin, 2002. Second Edition.
91. Susanne C. Brenner and Li-yeng Sung. Discrete Sobolev and Poincaré inequalities via Fourier series. *East-West J. of Numer. Math*, 8:83–92, 2000.
92. Susanne C. Brenner and Li-yeng Sung. Lower bounds for nonoverlapping domain decomposition preconditioners in two dimensions. *Math. Comp*, 69:1319–1339, 2000.
93. Marian Brezina and Petr Vaněk. A black-box iterative solver based on a two-level Schwarz method. *Computing*, 63(3):233–263, 1999.
94. Franco Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Rev. Française Automat. Recherche Opérationale Sér. Rouge*, 8:129–151, 1974.
95. Franco Brezzi and Michel Fortin. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, New York – Berlin – Heidelberg, 1991.
96. Oliver Bröker, Marcus J. Grote, Carsten Mayer, and Arnold Reusken. Robust parallel smoothing for multigrid via sparse approximate inverses. *SIAM J. Sci. Comput.*, 23(4):1396–1417, 2001.
97. Annalisa Buffa and Patrick Ciarlet, Jr. On traces for functional spaces related to Maxwell's equations. I. An integration by parts formula in Lipschitz polyhedra. *Math. Methods Appl. Sci.*, 24(1):9–30, 2001.
98. Annalisa Buffa and Patrick Ciarlet, Jr. On traces for functional spaces related to Maxwell's equations. II. Hodge decompositions on the boundary of Lipschitz polyhedra and applications. *Math. Methods Appl. Sci.*, 24(1):31–48, 2001.
99. Xiao-Chuan Cai. *Some Domain Decomposition Algorithms for Nonselfadjoint Elliptic and Parabolic Partial Differential Equations.* PhD thesis, Courant Institute of Mathematical Sciences, September 1989. TR-461, Department of Computer Science, Courant Institute.
100. Xiao-Chuan Cai. An additive Schwarz algorithm for nonselfadjoint elliptic equations. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 232–244. SIAM, Philadelphia, PA, 1990. Houston, Texas, March 20–22,1989.
101. Xiao-Chuan Cai. Additive Schwarz algorithms for parabolic convection-diffusion equations. *Numer. Math.*, 60(1):41–61, 1991.
102. Xiao-Chuan Cai. An optimal two-level overlapping domain decomposition method for elliptic problems in two and three dimensions. *SIAM J. Sci. Comput.*, 14:239–247, January 1993.
103. Xiao-Chuan Cai, Mario A. Casarin, Frank W. Jr. Elliott, and Olof B. Widlund. Overlapping Schwarz algorithms for solving Helmholtz's equation. In

Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 391–399. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.

104. Xiao-Chuan Cai, Maksymilan Dryja, and Marcus V. Sarkis. Overlapping nonmatching grid mortar element methods for elliptic problems. *SIAM J. Numer. Anal.*, 36:581–606, 1999.
105. Xiao-Chuan Cai and Maksymilian Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. In David Keyes and Jinchao Xu, editors, *Domain decomposition methods in scientific and engineering computing. Seventh International Conference on Domain Decomposition*, pages 21–27. AMS, Contemporary Mathematics 180, 1994. Pennsylvania State University, October 27–30, 1993.
106. Xiao-Chuan Cai, Charbel Farhat, and Marcus Sarkis. A minimum overlap restricted additive Schwarz preconditioner and applications to 3D flow simulations. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 479–485. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
107. Xiao-Chuan Cai, William D. Gropp, and David E. Keyes. Convergence rate estimate for a domain decomposition method. *Numer. Math.*, 61:153–169, 1992.
108. Xiao-Chuan Cai, William D. Gropp, and David E. Keyes. A comparison of some domain decomposition and ILU preconditioned iterative methods for nonsymmetric elliptic problems. *Numer. Linear Algebra Appl*, 1(5):477–504, 1994.
109. Xiao-Chuan Cai, William D. Gropp, David E. Keyes, Robin G. Melvin, and David P. Young. Parallel Newton-Krylov-Schwarz algorithms for the transonic full potential equation. *SIAM J. Sci. Comput.*, 19(1):246–265, January 1998.
110. Xiao-Chuan Cai, William D. Gropp, David E. Keyes, and Molay D. Tidriri. Newton-Krylov-Schwarz methods in CFD. In Rolf Rannacher, editor, *Proceedings of the International Workshop on the Navier-Stokes Equations, Notes in Numerical Fluid Mechanics*, Notes in Numerical Fluid Mechanics, pages 17–30. Vieweg Verlag, Braunschweig, 1994.
111. Xiao-Chuan Cai and David E. Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SIAM J. Sci. Comput.*, 24(1):183–200, 2002.
112. Xiao-Chuan Cai, David E. Keyes, and Leszek Marcinkowski. Non-linear additive Schwarz preconditioners and application in computational fluid dynamics. *Internat. J. Numer. Methods Fluids*, 40(12):1463–1470, 2002. LMS Workshop on Domain Decomposition Methods in Fluid Mechanics (London, 2001).
113. Xiao-Chuan Cai, David E. Keyes, and Venkatasubramanian Venkatakrishnan. Newton-Krylov-Schwarz: An implicit solver for CFD. In Roland Glowinski, Jacques Periaux, Zhong-Ci Shi, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering. Eighth International Conference*, pages 387–402. John Wiley & Sons, Ltd., 1996. Beijing, P.R. China, May 16–20, 1995.
114. Xiao-Chuan Cai, David E. Keyes, and David P. Young. A nonlinearly additive Schwarz preconditioned inexact Newton method for shocked duct flow. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain De-*

*composition Methods*, pages 345–352, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
115. Xiao-Chuan Cai and Marcus Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM J. Sci. Comp.*, 21:239–247, 1999.
116. Xiao-Chuan Cai and Olof Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM J. Sci. Statist. Comput.*, 13(1):243–258, January 1992.
117. Xiao-Chuan Cai and Olof Widlund. Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM J. Numer. Anal.*, 30(4):936–952, August 1993.
118. Zhiqiang Cai, Rossen Parashkevov, Thomas F. Russell, and Xiu Ye. Overlapping domain decomposition for a mixed finite element method in three dimensions. In Petter E. Bjørstad, Magne Espedal, and David E. Keyes, editors, *Domain Decomposition Methods in Sciences and Enginnering. Ninth International Conference on Domain Decomposition Methods*, pages 188–196. DDM.org, 1998. Ullensvang at the Hardanger Fjord, Norway, June 4–7, 1996.
119. Caterina Calgaro and Jacques Laminie. On the domain decomposition method for the generalized Stokes problem with continuous pressure. *Numer. Methods Partial Differential Equations*, 16(1):84–106, 2000.
120. Claudio Canuto. Stablization of spectral methods by finite element bubble functions. *Comput. Methods Appl. Mech. Engrg*, 116:13–26, 1994. Proceedings of ICOSAHOM 92, a conference held in Montpellier, France, June 1992.
121. Claudio Canuto, M. Yousuff Hussaini, Alfio Quarteroni, and Thomas A. Zang. *Spectral Methods in Fluid Dynamics*. Springer-Verlag, 1988.
122. Claudio Canuto and Alfio Quarteroni. Preconditioned minimal residual methods for Chebyshev spectral calculations. *J. Comput. Phys*, 60:315–337, 1985.
123. Weiming Cao and Benqi Guo. Preconditioners on element interfaces for the $p$ version and spectral element method. *SIAM J. Sci. Comput.*, 21:522–551, 1999.
124. Weiming Cao and Benqi Guo. Preconditioning for the $p$-version boundary element method in three dimensions with triangular elements. *J. Korean Math. Soc.*, 41(2):345–368, 2004.
125. Claudio Carlenzoli and Alfio Quarteroni. Adaptive domain decomposition methods for advection-diffusion problems. *The IMA Volumes in Mathematics and its Applications, Springer Verlag*, 75:165–186, 1995.
126. Mario A. Casarin. *Schwarz Preconditioners for Spectral and Mortar Finite Element Methods with Applications to Incompressible Fluids*. PhD thesis, Courant Institute of Mathematical Sciences, March 1996. TR-717, Department of Computer Science.
127. Mario A. Casarin. Quasi-optimal Schwarz methods for the conforming spectral element discretization. *SIAM J. Numer. Anal.*, 34(6):2482–2502, 1997.
128. Mario A. Casarin. Schwarz preconditioners for the spectral element discretization of the steady Stokes and Navier-Stokes equations. *Numer. Math.*, 89(2):307–339, August 2001.
129. Mario A. Casarin and Olof B. Widlund. A hierarchical preconditioner for the mortar finite element method. *ETNA*, 4:75–88, June 1996.
130. Tony Chan, Takashi Kako, Hideo Kawarada, and Olivier Pironneau, editors. *Domain Decomposition Methods in Sciences and Engineering. Twelfth Interna-*

*tional Conference on Domain Decomposition.* DDM.org, 2001. Chiba, Japan, October 25–29, 1999.

131. Tony F. Chan and Tarek P. Mathew. An application of the probing technique to the vertex space method in domain decomposition. In Roland Glowinski, Yuri A. Kuznetsov, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, editors, *Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 101–111, Philadelphia, PA, 1991. SIAM. Moscow, Russia, May 21–25, 1990.
132. Tony F. Chan and Tarek P. Mathew. Domain decomposition algorithms. In *Acta Numerica 1994*, pages 61–143. Cambridge University Press, 1994.
133. Tony F. Chan, Tarek P. Mathew, and Jian-Ping Shao. Efficient variants of the vertex space domain decomposition algorithm. *SIAM J. Sci. Comput.*, 15(6), 1994.
134. Tony F. Chan, Barry F. Smith, and Jun Zou. Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids. *Numer. Math.*, 73(2):149–167, 1996.
135. Philippe Chevalier and Frédéric Nataf. Symmetrized method with optimized second-order conditions for the Helmholtz equation. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 400–407. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
136. Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam, 1978.
137. Philippe G. Ciarlet. *Mathematical Elasticity.* North-Holland, Amsterdam, 1988.
138. Marie-Claude Ciccoli. Adaptive domain decomposition algorithms and finite volume/finite element approximation for advection-diffusion equations. *J. Sci. Comput.*, 11(4):299–341, 1996.
139. Philippe Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.*, 9(R-2):77–84, 1975.
140. Bernardo Cockburn, George E. Karniadakis, and Chi-Wang Shu, editors. *Discontinuous Galerkin Methods*, volume 11 of *Lecture Notes in Computational Science and Engineering.* Springer Verlag, 2000.
141. Francis Collino, Gianluca Delbue, Patrick Joly, and Andrea Piacentini. A new interface condition in the non-overlapping domain decomposition. *Comput. Methods Appl. Mech. Engrg.*, 148:195–207, 1997.
142. Francis Collino, Souad Ghanemi, and Patrick Joly. Domain decomposition method for harmonic wave propagation: a general presentation. *Comput. Methods Appl. Mech. Engrg.*, 184:171–211, 2000.
143. Martin Costabel and Ernst Stephan. Boundary integral equations for mixed boundary value problems in polygonal domains and Galerkin approximation. In *Mathematical models and methods in mechanics*, volume 15 of *Banach Center Publ.*, pages 175–251. PWN, Warsaw, 1985.
144. Wouter Couzy and Michel O. Deville. Spectral-element preconditioners for the Uzawa pressure operator applied to incompressible flows. *J. Sci. Comput.*, 9:107–112, 1994.

145. Wouter Couzy and Michel O. Deville. A fast Schur complement method for the spectral element discretization of the incompressible Navier-Stokes equations. *J. Comput. Phys.*, 116:135–142, 1995.
146. Lawrence C. Cowsar. Domain decomposition methods for nonconforming finite elements spaces of Lagrange-type. Technical Report TR93-11, Department of Mathematical Sciences, Rice University, March 1993.
147. Lawrence C. Cowsar. Dual variable Schwarz methods for mixed finite elements. Technical Report TR93-09, Department of Mathematical Sciences, Rice University, March 1993.
148. Lawrence C. Cowsar, Jan Mandel, and Mary F. Wheeler. Balancing domain decomposition for mixed finite elements. *Math. Comp.*, 64(211):989–1015, July 1995.
149. Michel Crouzeix and Pierre-Arnaud Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.
150. Monique Dauge. *Elliptic boundary value problems on corner domains.* Springer-Verlag, New York, 1988.
151. Robert Dautray and Jacques-Louis Lions. *Mathematical analysis and numerical methods for science and technology.* Springer-Verlag, 1990.
152. Armel de La Bourdonnaye, Charbel Farhat, Antonini Macedo, Frédéric Magoulès, and François-Xavier Roux. A non-overlapping domain decomposition method for exterior Helmholtz problems. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 42–66. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
153. Yann-Hervé De Roeck. *Résolution sur Ordinateurs Multi-Processeurs de Problème d'Elasticité par Décomposition de Domaines.* PhD thesis, Université Paris IX Daupine, 1991.
154. Yann-Hervé De Roeck and Patrick Le Tallec. Analysis and test of a local domain decomposition preconditioner. In Roland Glowinski, Yuri A. Kuznetsov, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, editors, *Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 112–128, Philadelphia, PA, 1991. SIAM. Moscow, Russia, May 21–25, 1990.
155. Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors. *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
156. James W. Demmel. *Applied numerical linear algebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
157. Bruno Després. *Méthodes de décomposition de domaine pour les problèms de propagation d'ondes en régime harmonique.* PhD thesis, Université Paris IX Dauphine, 1991.
158. Bruno Després. Domain decomposition method and the Helmholtz problem.II. In *Second International Conference on Mathematical and Numerical Aspects of Wave Propagation (Newark, DE, 1993)*, pages 197–206, Philadelphia, PA, 1993. SIAM.

159. Bruno Després, Patrick Joly, and Jean E. Roberts. A domain decomposition method for the harmonic Maxwell equations. In *Iterative methods in linear algebra (Brussels, 1991)*, pages 475–484, Amsterdam, 1992. North-Holland.
160. Michel O. Deville, Paul F. Fischer, and Ernest H. Mund. *High-Order Methods for Incompressible Fluid Flow.* Cambridge University Press, Cambridge, England, 2002.
161. Michel O. Deville and Ernest H. Mund. Chebyshev pseudospectral solution of second order elliptic equations with finite element preconditioning. *J. Comput. Phys*, 60:517–533, 1985.
162. Michel O. Deville and Ernest H. Mund. Finite-element preconditioning for pseudospectral solutions of elliptic problems. *SIAM J. Sci. Statist. Comput.*, 11(2):311 – 342, March 1990.
163. Quang Vinh Dihn, Roland Glowinski, and Jacques Périaux. Solving elliptic problems by domain decomposition methods with applications. In Garrett Birkhoff and Arthur Schoenstadt, editors, *Elliptic Problem Solvers II*, pages 395–426, New York, 1984. Academic Press.
164. Clark R. Dohrmann. A preconditioner for substructuring based on constrained energy minimization. *SIAM J. Sci Comput.*, 25(1):246–258, 2003.
165. Clark R. Dohrmann. A study of domain decomposition preconditioners. Technical Report SAND2003-4391, Sandia National Laboratories, Albuquerque, New Mexico, 2003.
166. Victorita Dolean, Stéphane Lanteri, and Frédéric Nataf. Optimized interface conditions for domain decomposition methods in fluid dynamics. *Internat. J. Numer. Methods Fluids*, 40:1539–1550, 2002.
167. Zdeněk Dostál, Ana Friedlander, and Sandra A. Santos. Solution of coercive and semicoercive contact problems by FETI domain decomposition. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 82–93. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
168. Zdeněk Dostál, Francisco A. M. Gomes, and Sandra A. Santos. FETI domain decomposition for contact 3D problems. In *Finite element methods (Jyväskylä, 2000)*, volume 15 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 54–65. Gakkōtosho, Tokyo, 2001.
169. Zdeněk Dostál, Francisco A. M. Gomes Neto, and Sandra A. Santos. Duality-based domain decomposition with natural coarse-space for variational inequalities. *J. Comput. Appl. Math.*, 126(1-2):397–415, 2000.
170. Zdeněk Dostál and David Horák. Scalability and FETI based algorithm for large discretized variational inequalities. *Math. Comput. Simulation*, 61(3-6):347–357, 2003. MODELLING 2001 (Pilsen).
171. Jim Douglas, Jr. and Jean E. Roberts. Global estimates for mixed methods for 2nd order elliptic equations. *Math. Comp.*, 44:39–52, 1985.
172. Maksymilian Dryja. An algorithm with a capacitance matrix for a variational-difference scheme. In Guri I. Marchuk, editor, *Variational-Difference Methods in Mathematical Physics*, pages 63–73, Novosibirsk, 1981. USSR Academy of Sciences.
173. Maksymilian Dryja. A method of domain decomposition for 3-D finite element problems. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and

Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 43–61, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.

174. Maksymilian Dryja. An additive Schwarz algorithm for two- and three-dimensional finite element elliptic problems. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods*, pages 168–172, Philadelphia, PA, 1989. SIAM. Los Angeles, California, January 14–16, 1988.
175. Maksymilian Dryja and Wolfgang Hackbusch. On the nonlinear domain decomposition method. *BIT*, pages 296–311, 1997.
176. Maksymilian Dryja, Wlodek Proskurowski, and Olof Widlund. A method of domain decomposition with crosspoints for elliptic finite element problems. In Blagovest Sendov, editor, *Optimal Algorithms*, pages 97–111, Sofia, Bulgaria, 1986. Bulgarian Academy of Sciences.
177. Maksymilian Dryja, Marcus V. Sarkis, and Olof B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.*, 72(3):313–348, 1996.
178. Maksymilian Dryja, Barry F. Smith, and Olof B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31(6):1662–1694, December 1994.
179. Maksymilian Dryja and Olof B. Widlund. An additive variant of the Schwarz alternating method for the case of many subregions. Technical Report TR-339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, 1987.
180. Maksymilian Dryja and Olof B. Widlund. Some domain decomposition algorithms for elliptic problems. In Linda Hayes and David Kincaid, editors, *Iterative Methods for Large Linear Systems*, pages 273–291, San Diego, California, 1989. Academic Press. Proceeding of the Conference on Iterative Methods for Large Linear Systems held in Austin, Texas, October 19 - 21, 1988, to celebrate the sixty-fifth birthday of David M. Young, Jr.
181. Maksymilian Dryja and Olof B. Widlund. Towards a unified theory of domain decomposition algorithms for elliptic problems. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 3–21. SIAM, Philadelphia, PA, 1990. Houston, Texas, March 20–22,1989.
182. Maksymilian Dryja and Olof B. Widlund. Domain decomposition algorithms with small overlap. *SIAM J. Sci. Comput.*, 15(3):604–620, May 1994.
183. Maksymilian Dryja and Olof B. Widlund. Some recent results on Schwarz type domain decomposition algorithms. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods in Science and Engineering. Sixth International Conference of Domain Decomposition*, pages 53–62. AMS, Contemporary Mathematics 157, 1994. Como, Italy, June 15–19,1992.
184. Maksymilian Dryja and Olof B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. *Comm. Pure Appl. Math.*, 48(2):121–155, February 1995.
185. Stanley C. Eisenstat, Howard C. Elman, and Martin H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.*, 20 (2):345–357, 1983.

186. Howard C. Elman and Gene H. Golub. Inexact and preconditioned Uzawa algorithms for saddle point problems. *SIAM J. Numer. Anal.*, 31(6):1645–1661, 1994.
187. Björn Engquist and Hong-Kai Zhao. Absorbing boundary conditions for domain decomposition. *Appl. Numer. Math.*, 27(4):341–365, 1998.
188. Lawrence C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence, 1998.
189. Richard E. Ewing and Junping Wang. Analysis of the Schwarz algorithm for mixed finite element methods. *RAIRO Mathematical Modelling and Numerical Analysis*, 26(6):739–756, 1992.
190. Isabelle Faille, Eric Flauraud, Frédéric Nataf, Frédéric Schneider, and Françoise Willien. Optimized interface conditions for sedimentary basin modeling. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, pages 463–470, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
191. Charbel Farhat, Po-Shu Chen, and Jan Mandel. A scalable Lagrange multiplier based domain decomposition method for time-dependent problems. *Internat. J. Numer. Methods Engrg*, 38:3831–3853, 1995.
192. Charbel Farhat, Po-Shu Chen, Jan Mandel, and François-Xavier Roux. The two–level FETI method part II: Extension to shell problems, parallel implementation and performance results. *Comput. Methods Appl. Mech. Engrg.*, 155:153–179, 1998.
193. Charbel Farhat, Po-Shu Chen, Franck Risler, and François-Xavier Roux. A unified framework for accelerating the convergence of iterative substructuring methods with Lagrange multipliers. *Internat. J. Numer. Methods Engrg*, 42:257–288, 1998.
194. Charbel Farhat, Michel Lesoinne, Patrick Le Tallec, Kendall Pierson, and Daniel Rixen. FETI-DP: A dual-primal unified FETI method – part I: A faster alternative to the two-level FETI method. *Internat. J. Numer. Methods Engrg.*, 50:1523–1544, 2001.
195. Charbel Farhat, Michel Lesoinne, and Kendall Pierson. A scalable dual-primal domain decomposition method. *Numer. Linear Algebra Appl.*, 7(7–8):687–714, 2000.
196. Charbel Farhat, Antonini Macedo, and Michel Lesoinne. A two-level domain decomposition method for the iterative solution of high-frequency exterior Helmholtz problems. *Numer. Math.*, 85(2):283–303, 2000.
197. Charbel Farhat, Antonini Macedo, and Radek Tezaur. FETI-H: a scalable domain decomposition method for high frequency exterior Helmholtz problems. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 231–241 (electronic). DDM.org, 1999. Greenwich, England, July 20–24, 1998.
198. Charbel Farhat and Jan Mandel. The two–level FETI method for static and dynamic plate problems part I: An optimal iterative solver for biharmonic systems. *Comput. Methods Appl. Mech. Engrg.*, 155:129–151, 1998.
199. Charbel Farhat, Jan Mandel, and François-Xavier Roux. Optimal convergence properties of the FETI domain decomposition method. *Comput. Methods Appl. Mech. Engrg.*, 115:367–388, 1994.

200. Charbel Farhat and François-Xavier Roux. A Method of Finite Element Tearing and Interconnecting and its Parallel Solution Algorithm. *Internat. J. Numer. Methods Engrg.*, 32:1205–1227, 1991.
201. Charbel Farhat and François-Xavier Roux. Implicit parallel processing in structural mechanics. In J. Tinsley Oden, editor, *Computational Mechanics Advances*, volume 2 (1), pages 1–124. North-Holland, 1994.
202. Xiaobing Feng and Ohannes A. Karakashian. Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, 39(4):1343–1365, 2001.
203. Paul F. Fischer. An overlapping Schwarz method for spectral element solution of the incompressible Navier-Stokes equations. *J. Comput. Phys.*, 133:84–101, April 1997.
204. Paul F. Fischer, Neil I. Miller, and Henry M. Tufo. An overlapping Schwarz method for spectral element simulation of three-dimensional incompressible flows. In Petter Bjørstad and Michell Luskin, editors, *Parallel solution of partial differential equations (Minneapolis, MN, 1997)*, volume 120 of *IMA Vol. Math. Appl.*, pages 159–180. Springer, New York, 2000.
205. Paul F. Fischer and Einar Rønquist. Spectral element methods for large scale parallel Navier-Stokes calculations. *Comput. Methods Appl. Mech. Engrg*, 116:69–76, 1994. Proceedings of ICOSAHOM 92, a conference held in Montpellier, France, June 22-26, 1992.
206. Yannis Fragakis and Manolis Papadrakakis. The mosaic of high performance domain decomposition methods for structural mechanics: Formulation, interrelation and numerical efficiency of primal and dual methods. *Comput. Methods Appl. Mech. Engrg*, 192(35–36):3799–3830, 2003.
207. Leopoldo P. Franca, Sérgio L. Frey, and Thomas J. R. Hughes. Stabilized finite element methods. I. Application to the advective-diffusive model. *Comput. Methods Appl. Mech. Engrg.*, 95(2):253–276, 1992.
208. Andreas Frommer and Hartmut Schwandt. A unified representation and theory of algebraic additive Schwarz and multisplitting methods. *SIAM J. Matrix Anal. Appl.*, 18:893–912, 1997.
209. Andreas Frommer and Daniel B. Szyld. Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numer. Math.*, 83:259–278, 1999.
210. Andreas Frommer and Daniel B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. *SIAM J. Numer. Anal.*, 39(2):463–479, 2001.
211. Daniele Funaro, Alfio Quarteroni, and Paola Zanolli. An iterative procedure with interface relaxation for domain decomposition methods. *SIAM J. Numer. Anal.*, 25(6):1213–1236, 1988.
212. Martin J. Gander. Optimized Schwarz methods for Helmholtz problems. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, pages 247–254, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
213. Martin J. Gander and Gene H. Golub. A non-overlapping optimized Schwarz method which converges with an arbitrarily weak dependence on h. In Ismael Herrera, David E. Keyes, Olof B. Widlund, and Robert Yates, editors, *Domain Decomposition Methods in Science and Engineering. Fourteenth International*

*Conference on Domain Decomposition Methods*, pages 281–288, 2003. Cocoyoc in Morelos, Mexico, January 6–12, 2002.
214. Martin J. Gander and Laurence Halpern. Méthodes de décomposition de domaines pour l'équation des ondes en dimension 1. *C. R. Acad. Sci. Paris Sér. I Math.*, 333(6):589–592, 2001.
215. Martin J. Gander, Frédéric Magoulès, and Frédéric Nataf. Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM J. Sci. Comput.*, 24(1):38–60, 2002.
216. Marc Garbey. A Schwarz alternating procedure for singular perturbation problems. *SIAM J. Sci. Comput.*, 17:1175–1201, 1996.
217. Marc Garbey and Hans G. Kaper. Heterogeneous domain decomposition for singularly perturbed elliptic boundary value problems. *SIAM J. Numer. Anal.*, 34(4):1513–1544, 1997.
218. Fabio Gastaldi, Lucia Gastaldi, and Alfio Quarteroni. Adaptive domain decomposition methods for advection dominated equations. *East-West J. Numer. Math.*, 4(3):165–206, 1996.
219. Menno Genseberger. *Domain decomposition in the Jacobi-Davidson method for eigenproblems.* PhD thesis, Utrecht University, September 2001.
220. Alan George. Nested dissection of a regular finite element mesh. *SIAM J. Numer. Anal.*, 10:345–363, 1973.
221. Alan George and Joseph Liu. *Computer Solution of Large Sparse Positive Definite Systems.* Prentice-Hall, Englewood Cliffs, NJ, 1981.
222. Paola Gervasio. *Risoluzione delle equazioni alle derivate parziali con metodi spettrali in regioni partizionate in sottodomini.* PhD thesis, Università di Milano, 1995.
223. Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer-Verlag, New York, 1986.
224. Roland Glowinski and Yuri A. Kuznetsov. On the solution of the Dirichlet problem for linear elliptic operators by a distributed Lagrange multiplier method. *C. R. Acad. Sci. Paris Sér. I Math.*, 327(7):693–698, 1998.
225. Roland Glowinski, Tsorng-Whay Pan, Todd I. Hesla, Daniel D. Joseph, and Jacques Périaux. A fictitious domain method with distributed Lagrange multipliers for the numerical simulation of particulate flow. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 121–137. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
226. Roland Glowinski and Mary F. Wheeler. Domain decomposition and mixed finite element methods for elliptic problems. Technical Report 87-11, Rice University, 1987.
227. Roland Glowinski and Mary F. Wheeler. Domain decomposition and mixed finite element methods for elliptic problems. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 144–172, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.
228. Nabil Gmati, Charbel Farhat, and Ulrich Hetmaniuk. An efficient substructuring method for analyzing acoustics in a concentric hole-cavity resonator. In *Mathematical and numerical aspects of wave propagation (Santiago de Compostela, 2000)*, pages 817–821. SIAM, Philadelphia, PA, 2000.

229. Paulo Goldfeld. *Balancing Neumann-Neumann Preconditioners for Mixed Formulation of Almost-Incompressible Linear Elasticity.* PhD thesis, Courant Institute of Mathematical Sciences, September 2003. TR-847, Department of Computer Science.
230. Paulo Goldfeld, Luca F. Pavarino, and Olof B. Widlund. Balancing Neumann-Neumann preconditioners for mixed approximations of heterogeneous problems in linear elasticity. *Numer. Math.*, 95(2):283–324, 2003.
231. Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* Johns Hopkins Univ. Press, 1989. Second Edition.
232. Jayadeep Gopalakrishnan and Joseph E. Pasciak. Overlapping Schwarz preconditioners for indefinite time harmonic Maxwell equations. *Math. Comp.*, 72(241):1–15, 2003.
233. Anne Greenbaum. *Iterative methods for solving linear systems.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
234. Michael Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystem.* Teubner Skripten zur Numerik. B.G. Teubner, Stuttgart, 1994.
235. Michael Griebel and Peter Oswald. Remarks on the abstract theory of additive and multiplicative Schwarz algorithms. Technical report, Technische Universität München, 1993.
236. Michael Griebel and Peter Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70(2):163–180, 1995.
237. Pierre Grisvard. *Elliptic problems in nonsmooth domains.* Pitman Publishing, Boston, 1985.
238. Pierre Grisvard. *Singularities in boundary value problems*, volume 22 of *RMA.* Masson, Paris, 1992.
239. William D. Gropp, Dinesh K. Kaushik, David E. Keyes, and Barry F. Smith. Performance modeling and tuning of an unstructured mesh CFD application. In *Proceedings of SC2000.* IEEE Computer Society, 2000.
240. William D. Gropp, David E. Keyes, Lois Curfman McInnes, and Moulay D. Tidriri. Parallel implicit PDE computations: Algorithms and software. In A. Ecer et al., editor, *Parallel CFD '97'*, 1997. Held in Manchester, UK, May 19–21, 1997.
241. William D. Gropp and Barry F. Smith. Experiences with domain decomposition in three dimensions: Overlapping Schwarz methods. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods in Science and Engineering. Sixth International Conference of Domain Decomposition*, pages 323–334. AMS, Contemporary Mathematics 157, 1994. Como, Italy, June 15–19,1992.
242. Hervé Guillard and Petr Vaněk. An aggregation multigrid solver for convection-diffusion problems on unstructured meshes. Technical Report 130, Center for Computational Mathematics, University of Colorado, Denver, April 1998.
243. Benqi Guo and Weiming Cao. A preconditioner for the $h-p$ version of the finite element method in two dimensions. *Numer. Math.*, 75(1):59–77, 1996.
244. Benqi Guo and Weiming Cao. Additive Schwarz methods for the $h-p$ version of the finite element method in two dimensions. *SIAM J. Sci. Comput.*, 18(5):1267–1288, 1997.
245. Benqi Guo and Weiming Cao. An iterative and parallel solver based on domain decomposition of the $h-p$ version of the finite element method. *J. Comp. Appl. Math.*, 83:71–85, 1997.

246. Benqi Guo and Weiming Cao. An additive Schwarz method for the $h-p$ version of the finite element method in three dimensions. *SIAM J. Numer. Anal.*, 35(2):632–654, 1998.
247. Benqi Guo and Weiming Cao. Domain decomposition method for the $h-p$ version finite element method. *Comp. Methods Appl. Mech. Engrg.*, 157:425–440, 1998.
248. Benqi Guo and Weiming Cao. A preconditioner with inexact element face solvers for the 3-dimensional $p$-version finite element methods. *J. Comput. Appl. Math.*, 144:131–144, 2002.
249. Gundolf Haase, Michael Kuhn, and Ulrich Langer. Parallel multigrid 3D Maxwell solvers. *Parallel Comput.*, 27(6):761–775, 2001.
250. Gundolf Haase, Ulrich Langer, and Arnd Meyer. A new approach to the Dirichlet domain decomposition method. In S. Hengst, editor, *Fifth Multigrid Seminar, Eberswalde 1990*, pages 1–59, Berlin, 1990. Karl–Weierstrass–Institut. Report R–MATH–09/90.
251. Wolfgang Hackbusch. *Multi-Grid Methods and Applications.* Springer-Verlag, Berlin – Heidelberg – New York, 1985.
252. Wolfgang Hackbusch. *Iterative Solution of Large Sparse Linear Systems of Equations.* Springer-Verlag, Berlin, 1994.
253. Thomas Hagstrom, Reginal P. Tewarson, and Aron Jazcilevich. Numerical experiments on a domain decomposition algorithm for nonlinear elliptic boundary value problems. *Appl. Math. Lett.*, 1(3):299–302, 1988.
254. Friedrich K. Hebeker and Yuri A. Kuznetsov. Unsteady convection and convection-diffusion problems via direct overlapping domain decomposition methods. *Numer. Methods Partial Differential Equations*, 14(3):387–406, 1998.
255. Ismael Herrera, David E. Keyes, Olof B. Widlund, and Robert A. Yates, editors. *Domain Decomposition Methods in Science and Engineering. Fourteenth International Conference on Domain Decomposition Methods*, Mexico City, 2003. UNAM. Cocoyoc in Morelos, Mexico, January 6–12, 2002.
256. Ulrich Hetmaniuk and Charbel Farhat. A fictitious domain decomposition method for the solution of partially axisymmetric acoustic scattering problems. II. Neumann boundary conditions. *Internat. J. Numer. Methods Engrg.*, 58(1):63–81, 2003.
257. Bernhard Hientzsch. *Fast Solvers and Domain Decomposition Preconditioners for Spectral Element Discretizations of Problems in H(curl).* PhD thesis, Courant Institute of Mathematical Sciences, September 2001. TR-823, Department of Computer Science.
258. Bernhard Hientzsch. Schwarz preconditioners for spectral Nédélec elements for a model problem in H(curl). Technical Report TR-834, Courant Institute for Mathematical Sciences, New York University, November 2002. Submitted to Numer. Math.
259. Ralf Hiptmair. *Multilevel preconditioning for Mixed Problems in Three Dimensions.* PhD thesis, Universität Augsburg, Germany, 1996.
260. Ralf Hiptmair. Multigrid methods for Maxwell's equations. *SIAM J. Numer. Anal.*, 36:204–225, 1999.
261. Ralf Hiptmair and Ronald H.W. Hoppe. Multilevel methods for mixed finite elements in three dimensions. *Numer. Math.*, 82(2):253–279, 1999.
262. Ralf Hiptmair and Andrea Toselli. Overlapping and multilevel Schwarz methods for vector valued elliptic problems in three dimensions. In Petter Bjørstad

and Michell Luskin, editors, *Parallel solution of partial differential equations (Minneapolis, MN, 1997)*, volume 120 of *IMA Vol. Math. Appl.*, pages 181–208, New York, 2000. Springer.
263. Ronald H.W. Hoppe and Yuri A. Kuznetsov. Overlapping domain decomposition methods with distributed Lagrange multipliers. *East-West J. Numer. Math.*, 9(4):285–293, 2001.
264. Paul Houston, Christoph Schwab, and Endre Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163, 2002.
265. Qiya Hu and Jun Zou. A non-overlapping domain decomposition method for Maxwell's equations in three dimensions. *SIAM J. Numer. Anal.*, 41:1682–1708, 2003.
266. Qiya Hu and Jun Zou. Substructuring preconditioners for saddle-point problems arising from Maxwell's equations in three dimensions. *Math. Comp.*, 73(245):35–61, 2004.
267. Thomas J.R. Hughes, Leopoldo P. Franca, and Gregory M. Hulbert. A new finite element formulation for computational fluid dynamics. VIII. The Galerkin/least-squares method for advective-diffusive equations. *Comput. Methods Appl. Mech. Engrg.*, 73(2):173–189, 1989.
268. Frank Ihlenburg and Ivo Babuška. Finite element solution of the Helmholtz equation with high wave number, Part I: The h-version of the FEM. *Comput. Math. Appl.*, 30(9):9–37, 1995.
269. Frank Ihlenburg and Ivo Babuška. Finite element solution of the Helmholtz equation with high wave number, Part II: The h-p-version of the FEM. *SIAM J. Numer. Anal.*, 34(1):315–358, 1997.
270. Yuri Iliash, Tuomo Rossi, and Jari Toivanen. Two iterative methods for solving the Stokes problem. Technical Report 2/93, Laboratory of Scientific Computing, University of Jyväskylä, 1993. `http://www.mit.jyu.fi/tene/papers/report93-2.ps.gz`.
271. Caroline Japhet. Conditions aux limites artificielles et décomposition de domaine: Méthode OO2 (optimisé d'ordre 2). Application à la résolution de problèmes en mécanique des fluides. Technical Report 373, CMAP (Ecole Polytechnique), 1997.
272. Caroline Japhet. Optimized Krylov-Ventcell method. Application to convection-diffusion problems. In Petter E. Bjørstad, Magne Espedal, and David E. Keyes, editors, *Domain Decomposition Methods in Sciences and Enginnering. Ninth International Conference on Domain Decomposition Methods*, pages 382–389. DDM.org, 1998. Ullensvang at the Hardanger Fjord, Norway, June 4–7, 1996.
273. Caroline Japhet, Frédéric Nataf, and François Rogier. The optimized order 2 method. Application to convection-diffusion problems. *Future Generation Computer Systems FUTURE*, 18, 2001.
274. Caroline Japhet, Frédéric Nataf, and François-Xavier Roux. Extension of a coarse grid preconditioner to non-symmetric problems. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 279–286,. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
275. Eleanor W. Jenkins, C. Tim Kelley, Cass T. Miller, and Christopher E. Kees. An aggregation-based domain decomposition preconditioner for groundwater

flow. Technical Report TR00–13, Department of Mathematics, North Carolina State University, 2000.
276. Claes Johnson, Uno Nävert, and Juhani Pitkäranta. Finite element methods for linear hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 45(1-3):285–312, 1984.
277. Dinesh K. Kaushik, David E. Keyes, and Barry F. Smith. NKS methods for compressible and incompressible flows on unstructured grids. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 513–520. DDM.org, 1999. Greenwich, England, July 20–24, 1998.
278. David E. Keyes, Dinesh K. Kaushik, and Barry F. Smith. Prospects for CFD on petaflops systems. *CFD Review*, 1997.
279. Axel Klawonn. An optimal preconditioner for a class of saddle point problems with a penalty term, Part II: General theory. Technical Report 14/95, Westfälische Wilhelms-Universität Münster, Germany, April 1995. Also available as TR-683, Department of Computer Science, Courant Institute of Mathematical Sciences.
280. Axel Klawonn. *Preconditioners for Indefinite Problems.* PhD thesis, Westfälische Wilhelms-Universität Münster, February 1996. Also available as TR-716, Department of Computer Science, Courant Institute of Mathematical Sciences.
281. Axel Klawonn. Block-Triangular Preconditioners for Saddle Point Problems with a Penalty Term. *SIAM J. Sci. Comput.*, 19(1):172–184, January 1998.
282. Axel Klawonn. An optimal preconditioner for a class of saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19(2):540–552, March 1998.
283. Axel Klawonn and Luca F. Pavarino. An overlapping additive Schwarz method for a saddle point problem from linear elasticity. In Junping Wang, Myron B. Allen III, Benito M. Chen, and Tarek Mathew, editors, *Iterative Methods in Scientific Computation*, pages 219–224. IMACS Series in Computational and Applied Mathematics, 1998.
284. Axel Klawonn and Luca F. Pavarino. Overlapping Schwarz methods for elasticity and Stokes problems. *Comp. Methods Appl. Mech. Engrg*, 165(1-4):233–245, 1998.
285. Axel Klawonn and Luca F. Pavarino. A comparison of overlapping Schwarz methods and block preconditioners for saddle point problems. *Numer. Linear Algebra Appl.*, 7:1–25, 2000.
286. Axel Klawonn, Oliver Rheinbach, and Olof B. Widlund. Some computational results for Dual-Primal FETI methods for three dimensional elliptic problems. In Ralf Kornhuber, Ronald H.W. Hoppe, Jacques Périaux, Olivier Pironneau, Olof B. Widlund, and Jinchao Xu, editors, *Domain Decomposition Methods in Science and Engineering.* Springer-Verlag, Lecture Notes in Computational Science and Engineering, 2004. Proceedings of the 15th International Conference on Domain Decomposition Methods, Berlin, July 21-25, 2003.
287. Axel Klawonn and Gerhard Starke. Block triangular preconditioners for nonsymmetric saddle point problems: Field-of-values analysis. *Numer. Math.*, 81:577–594, 1999.
288. Axel Klawonn and Olof B. Widlund. A domain decomposition method with Lagrange multipliers and inexact solvers for linear elasticity. *SIAM J. Sci. Comput.*, 22(4):1199–1219, 2000.

289. Axel Klawonn and Olof B. Widlund. FETI and Neumann–Neumann Iterative Substructuring Methods: Connections and New Results. *Comm. Pure Appl. Math.*, 54:57–90, January 2001.
290. Axel Klawonn and Olof B. Widlund. Dual-Primal FETI methods for linear elasticity. In preparation, 2004.
291. Axel Klawonn and Olof B. Widlund. Selecting constraints in Dual-Primal FETI methods for elasticity in three dimensions. In Ralf Kornhuber, Ronald H.W. Hoppe, Jacques Périaux, Olivier Pironneau, Olof B. Widlund, and Jinchao Xu, editors, *Domain Decomposition Methods in Science and Engineering*. Springer-Verlag, Lecture Notes in Computational Science and Engineering, 2004. Proceedings of the 15th International Conference on Domain Decomposition Methods, Berlin, July 21-25, 2003.
292. Axel Klawonn, Olof B. Widlund, and Maksymilian Dryja. Dual-primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients. *SIAM J. Numer. Anal.*, 40(1):159–179, April 2002.
293. Axel Klawonn, Olof B. Widlund, and Maksymilian Dryja. Dual-Primal FETI methods with face constraints. In Luca Pavarino and Andrea Toselli, editors, *Recent Developments in Domain Decomposition Methods*, Lecture Notes in Computational Science and Engineering, Volume 23, pages 27–40. Springer, 2002.
294. Vadim Korneev, Joseph E. Flaherty, J. Tinsley Oden, and Jacob Fish. Additive Schwarz algorithms for solving *hp*-version finite element systems on triangular meshes. *Appl. Numer. Math.*, 43(4):399–421, 2002.
295. Vadim Korneev, Ulrich Langer, and Leonidas Xanthis. On fast domain decomposition solving procedures for hp-discretizations of 3D elliptic problems. *Comput. Methods Appl. Math.*, 3(4):536–559, 2003.
296. Vadim G. Korneev and Søren Jensen. Preconditioning of the $p$-version of the finite element method. *Comput. Methods Appl. Mech. Engrg.*, 150(1-4):215–238, 1997. Symposium on Advances in Computational Mechanics, Vol. 2 (Austin, TX, 1997).
297. Ralf Kornhuber, Ronald H.W. Hoppe, Jacques Périaux, Olivier Pironneau, Olof B. Widlund, and Jinchao Xu, editors. *Domain Decomposition Methods in Sciences and Engineering*, Lecture Notes in Computational Science and Engineering, Berlin, 2004. Springer. Proceedings of the Fifteenth International Conference on Domain Decomposition Methods, Berlin, Germany, July 21 - 25, 2003.
298. Yuri A. Kuznetsov. Efficient iterative solvers for elliptic problems on nonmatching grids. *Russian J. Numer. Anal. Math. Modelling*, 10(3):187–211, 1995.
299. Yuri A. Kuznetsov. Efficient iterative solvers for elliptic problems on nonmatching grids. *Russian J. Numer. Anal. Math. Modelling*, 10(3):187–211, 1995.
300. Yuri A. Kuznetsov. Domain decomposition and fictitious domain methods with distributed Lagrange multipliers. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, pages 67–75, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.

301. Yuri A. Kuznetsov, Petri Manninen, and Yuri Vassilevski. On numerical experiments with Neumann-Neumann and Neumann-Dirichlet domain decomposition preconditioners. Technical report, University of Jyväskylä, 1993.
302. Yuri A. Kuznetsov and Yuri Vassilevski. A Dirichlet-Dirichlet preconditioner for the mortar element method. Technical Report B 12/1999, University of Jyväskylä, 1999.
303. Catherine Lacour. Iterative substructuring preconditioners for the mortar finite element method. In Petter E. Bjørstad, Magne Espedal, and David E. Keyes, editors, *Domain Decomposition Methods in Sciences and Enginnering. Ninth International Conference on Domain Decomposition Methods*, pages 406–412. DDM.org, 1998. Ullensvang at the Hardanger Fjord, Norway, June 4–7, 1996.
304. Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof Widlund, editors. *Eleventh International Conference on Domain Decomposition Methods*, 1998. Proceedings of the 11th International Conference on Domain Decomposition Methods, Greenwich, England, July 20-24, 1998.
305. Caroline Lasser and Andrea Toselli. Convergence of some two-level overlapping domain decomposition preconditioners with smoothed aggregation coarse spaces. In Luca Pavarino and Andrea Toselli, editors, *Recent Developments in Domain Decomposition Methods*, Lecture Notes in Computational Science and Engineering, Volume 23, pages 95–117. Springer, 2002.
306. Caroline Lasser and Andrea Toselli. Overlapping preconditioners for discontinuous Galerkin approximations of second order problems. In Naima Debit, Marc Garbey, Ronald H.W. Hoppe, David E. Keyes, Yuri A. Kuznetsov, and Jacques Périaux, editors, *Domain Decomposition Methods in Sciences and Engineering. Thirteenth International Conference on Domain Decomposition Methods*, pages 77–84, Barcelona, 2002. CIMNE. Lyon, France, October 9–12, 2000.
307. Caroline Lasser and Andrea Toselli. An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comp.*, 72(243):1215–1238, 2003.
308. Patrick Le Tallec. Domain decomposition methods in computational mechanics. In J. Tinsley Oden, editor, *Computational Mechanics Advances*, volume 1 (2), pages 121–220. North-Holland, 1994.
309. Patrick Le Tallec, Yann-Hervé De Roeck, and Marina Vidrascu. Domain decomposition methods for large linearly elliptic three-dimensional problems. *J. Comput. Appl. Math.*, 34(1):93–117, 1991.
310. Patrick Le Tallec and Abani K. Patra. Non-overlapping domain decomposition methods for adaptive *hp* approximations of the Stokes problem with discontinuous pressure fields. *Comput. Methods Appl. Mech. Engrg.*, 145(3-4):361–379, 1997.
311. Daniel Lee and Mulder Yu. A study on parallel flow computation by Newton-Schwarz method. In *Proceeding of the Sixth National Conference in Computational Fluid Dynamics*, Taiwan, R.O.C., August 29 1999.
312. Wassily Leontief. *The structure of the American Economy.* Oxford University Press, New York, 1951.
313. Michel Lesoinne. A FETI-DP corner selection algorithm for three-dimensional problems. In Ismael Herrera, David E. Keyes, Olof B. Widlund, and Robert Yates, editors, *Domain Decomposition Methods in Science and Engineering. Fourteenth International Conference on Domain Decomposition Methods*, pages 217–223, 2003. Cocoyoc in Morelos, Mexico, January 6–12, 2002.

314. Jing Li. A dual-primal FETI method for incompressible Stokes equations. Technical report, Courant Institute of Mathematical Sciences, 2001. TR-816, Department of Computer Science.
315. Jing Li. A dual-primal FETI method for incompressible Stokes and linearized Navier-Stokes equations. Technical report, Courant Institute of Mathematical Sciences, 2002. TR-828, Department of Computer Science.
316. Jing Li. *Dual-Primal FETI Methods for Stationary Stokes and Navier-Stokes Equations*. PhD thesis, Courant Institute of Mathematical Sciences, September 2002. TR-830, Department of Computer Science.
317. Jacques-Louis Lions and Enrico Magenes. *Problèmes aux limites non homogènes et applications*, volume 17-18 of *Travaux et recherches mathematiques*. Dunod, 1968.
318. Pierre-Louis Lions. Interprétation stochastique de la méthode alternée de Schwarz. *C. R. Acad. Sci. Paris*, 268:325–328, 1978.
319. Pierre-Louis Lions. On the Schwarz alternating method. I. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.
320. Pierre-Louis Lions. On the Schwarz alternating method. II. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods*, pages 47–70, Philadelphia, PA, 1989. SIAM. Los Angeles, California, January 14–16, 1988.
321. Pierre-Louis Lions. On the Schwarz alternating method. III: a variant for nonoverlapping subdomains. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 202–223. SIAM, Philadelphia, PA, 1990. Houston, Texas, March 20–22,1989.
322. Gert Lube, Lars Müller, and Hannes Müller. A new non-overlapping domain decomposition method for stabilized finite element methods applied to the non-stationary Navier-Stokes equations. Numerical linear algebra methods for computational fluid flow problems. *Numer. Linear Algebra Appl.*, 7(6):449–472, 2000.
323. Shiu-Hong Lui. On Schwarz alternating methods for nonlinear elliptic PDEs. *SIAM J. Sci. Comput.*, 21(4):1506–1523, 1999/00.
324. Shiu-Hong Lui. On monotone and Schwarz alternating methods for nonlinear elliptic PDEs. *M2AN Math. Model. Numer. Anal.*, 35(1):1–15, 2001.
325. Shiu-Hong Lui. On Schwarz methods for monotone elliptic PDEs. In Tony Chan, Takashi Kako, Hideo Kawarada, and Olivier Pironneau, editors, *Domain Decomposition Methods in Sciences and Engineering. Twelfth International Conference on Domain Decomposition*, pages 55–62 (electronic). DDM.org, 2001. Chiba, Japan, October 25–29, 1999.
326. Robert E. Lynch, John R. Rice, and Donald H. Thomas. Direct solution of partial differential equations by tensor product methods. *Numer. Math.*, 6:185–199, 1964.
327. Antonini Macedo Puppin. On the choice of Krylov methods and preconditioners for a domain decomposed iterative solution of the exterior Helmholtz problem. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B.

Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 531–538. DDM.org, 1999. Greenwich, England, July 20–24, 1998.

328. Yvon Maday. Relèvement de traces polynomiales et interpolations hilbertiennes entre espaces de polynômes. *C. R. Acad. Sci. Paris*, 309, Série I:463–468, 1989.
329. Yvon Maday, Dan Meiron, Anthony T. Patera, and Einar M. Rønquist. Analysis of iterative methods for the steady and unsteady Stokes problem: Application to spectral element discretizations. *SIAM J. Sci. Comput.*, 14(2):310–337, 1993.
330. Yvon Maday and Anthony T. Patera. Spectral element methods for the Navier–Stokes equations. In Ahmed K. Noor and J. Tinsley Oden, editors, *State-of-the Art Surveys in Computational Mechanics*, pages 71–143. ASME, New York, 1989.
331. Jan Mandel. Efficient domain decomposition preconditioning for the p-version finite element method in three dimensions. Technical report, Computational Mathematics Group, University of Colorado at Denver, 1989.
332. Jan Mandel. Iterative solvers by substructuring for the p-version finite element method. *Comp. Methods Appl. Mech. Engrg.*, 80:117–128, 1990.
333. Jan Mandel. Two-level domain decomposition preconditioning for the p-version finite element version in three dimensions. *Int. J. Numer. Meth. Eng.*, 29:1095–1108, 1990.
334. Jan Mandel. Hybrid domain decomposition with unstructured subdomains. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods in Science and Engineering. Sixth International Conference of Domain Decomposition*, pages 103–112. AMS, Contemporary Mathematics 157, 1994. Como, Italy, June 15–19,1992.
335. Jan Mandel and Marian Brezina. Balancing domain decomposition for problems with large jumps in coefficients. *Math. Comp.*, 65:1387–1401, 1996.
336. Jan Mandel and Clark R. Dohrmann. Convergence of a balancing domain decomposition by constraints and energy minimization. *Numer. Linear Algebra Appl.*, 10(7):639–659, 2003.
337. Jan Mandel, Clark R. Dohrmann, and Radek Tezaur. An algebraic theory for primal and dual substructuring methods by constraints. Technical report, Center for Computational Mathematics, University of Colorado at Denver, 2004.
338. Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors. *Domain Decomposition Methods 10.Tenth International Conference on Domain Decomposition Methods*, Providence, RI, 1998. AMS, Contemporary Mathematics 218. Boulder, Colorado, August 10–14, 1997.
339. Jan Mandel and Bohuslav Sekerka. A local convergence proof for the iterative aggregation method. *Linear Algebra Appl.*, 51:163–172, 1983.
340. Jan Mandel and Radek Tezaur. Convergence of a Substructuring Method with Lagrange Multipliers. *Numer. Math.*, 73:473–487, 1996.
341. Jan Mandel and Radek Tezaur. On the convergence of a dual-primal substructuring method. *Numer. Math.*, 88(3):543–558, January 2001.
342. Jan Mandel, Radek Tezaur, and Charbel Farhat. A Scalable Substructuring Method by Lagrange Multipliers for Plate Bending Problems. *SIAM J. Numer. Anal.*, 36(5):1370–1391, 1999.

343. Luisa D. Marini and Alfio Quarteroni. An iterative procedure for domain decomposition methods: a finite element approach. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 129–143, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.
344. Luisa D. Marini and Alfio Quarteroni. A relaxation procedure for domain decomposition methods using finite elements. *Numer. Math*, 55(5):575–598, 1989.
345. Tarek P. Mathew. *Domain Decomposition and Iterative Refinement Methods for Mixed Finite Element Discretizations of Elliptic Problems.* PhD thesis, Courant Institute of Mathematical Sciences, September 1989. TR-463, Department of Computer Science.
346. Tarek P. Mathew. Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part I: Algorithms and Numerical results. *Numer. Math.*, 65(4):445–468, 1993.
347. Tarek P. Mathew. Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, part II: Theory. *Numer. Math.*, 65(4):469–492, 1993.
348. Aleksandr M. Matsokin and Sergey V. Nepomnyaschikh. A Schwarz alternating method in a subspace. *Soviet Math.*, 29(10):78–84, 1985.
349. Paul R. McHugh, Dana A. Knoll, and David E. Keyes. Application of a Schwarz-preconditioned Newton-Krylov algorithm to a low-speed reacting flow problem. *AIAA Journal*, 36:290–292, 1998.
350. Lois Curfman McInnes, Romeo F. Susan-Resigna, David E. Keyes, and Hafiz M. Atassi. Additive Schwarz methods with nonreflecting boundary conditions for the parallel computation of Helmholtz problems. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 325–333. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
351. William McLean. *Strongly elliptic systems and boundary integral equations.* Cambridge University Press, Cambridge, 2000.
352. Jens Markus Melenk. On condition numbers in *hp*-FEM with Gauss-Lobatto-based shape functions. *J. Comput. Appl. Math.*, 139(1):21–48, 2002.
353. Peter Monk. An analysis of Nédélec's method for the spatial discretization of Maxwell's equations. *J. Comput. Appl. Math.*, 47:101–102, 1993.
354. Peter Monk. *Finite Element Methods for Maxwell's Equations.* Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 2003. Oxford Science Publications.
355. Rafael Muñoz-Sola. Polynomial liftings on a tetrahedron and applications to the $h-p$ version of the finite element method in three dimensions. *SIAM J. Numer. Anal.*, 34(1):282–314, February 1997.
356. Reinhard Nabben and Daniel B. Szyld. Convergence theory of restricted multiplicative Schwarz methods. *SIAM J. Numer. Anal.*, 40(6):2318–2336 (electronic) (2003), 2002.
357. Sh.M. Nasibov. On optimal constants in some Sobolev inequalities and their applications to a nonlinear Schrödinger equation. *Soviet Math.*, 40(1):110–115, 1990.

358. Frédéric Nataf and François Rogier. Factorization of the convection-diffusion operator and the Schwarz algorithm. *Math. Models Methods Appl. Sci.*, 5(1):67–93, 1995.
359. Jean-Claude Nédélec. Mixed finite elements in $R^3$. *Numer. Math.*, 35:315–341, 1980.
360. Sergey V. Nepomnyaschikh. *Domain Decomposition and Schwarz Methods in a Subspace for the Approximate Solution of Elliptic Boundary Value Problems.* PhD thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986.
361. Jindřich Nečas. *Les méthodes directes en théorie des équations elliptiques.* Academia, Prague, 1967.
362. Jindřich Nečas and Ivan Hlavácek. *Mathematical theory of elastic and elasto-plastic bodies: an introduction.* Elsevier Scientific Publishing Co., 1980.
363. Jorge Nocedal and Stephen J. Wright. *Numerical optimization.* Springer-Verlag, New York, 1999.
364. J. Tinsley Oden, Abani K. Patra, and Yusheng Feng. Parallel domain decomposition solver for adaptive *hp* finite element methods. *SIAM J. Numer. Anal.*, 34:2090–2118, 1997.
365. Steven A. Orszag. Spectral methods for problems in complex geometry. *J. Comput. Phys*, 37:70–92, 1980.
366. Peter Oswald. On a BPX-preconditioner for P1 elements. Technical Report Math./91/2, Friedrich Schiller Universität, Jena, Germany, 1991. (part 2).
367. Peter Oswald. On a hierarchical basis multilevel method with nonconforming P1 elements. *Numer. Math.*, 62(2):189–212, 1992.
368. Peter Oswald. *Multilevel Finite Element Approximation, Theory and Applications.* Teubner Skripten zur Numerik. B.G. Teubner, Stuttgart, 1994.
369. Shannon S. Pahl. Schwarz type domain decomposition methods for spectral element discretizations. Master's thesis, Department of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa, December 1993.
370. Charaka J. Palansuriya, Choi-Hong Lai, Constantinos S. Ierotheou, and Koulis A. Pericleous. A domain decomposition based algorithm for non-linear 2D inverse heat conduction problems. In Jan Mandel, Charbel Farhat, and Xiao-Chuan Cai, editors, *Domain Decomposition Methods 10. Tenth International Conference on Domain Decomposition Methods*, pages 515–522. AMS, Contemporary Mathematics 218, 1998. Boulder, Colorado, August 10–14, 1997.
371. David Pardo and Leszek Demkowicz. Integration of hp-adabtivity and multigrid. I. A two grid solver for hp finite elements. Technical Report 03-33, TICAM, 2002.
372. Kwang-Chun Park, Manoel R. Justino, and Carlos A. Felippa. An algebraically partitioned FETI method for parallel structural analysis: algorithm description. *Internat. J. Numer. Methods Engrg.*, 40:2717–2737, 1997.
373. Joseph E. Pasciak. Two domain decomposition techniques for Stokes problems. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods*, pages 419–430, Philadelphia, PA, 1989. SIAM. Los Angeles, California, January 14–16, 1988.
374. Anthony T. Patera. A spectral element method for fluid dynamics: Laminar flow in a channel expansion. *J. Comput. Phys*, 54:468–488, 1984.

375. Luca F. Pavarino. *Domain Decomposition Algorithms for the p-version Finite Element Method for Elliptic Problems.* PhD thesis, Courant Institute, New York University, September 1992. TR-616, Department of Computer Science.
376. Luca F. Pavarino. Additive Schwarz methods for the p-version finite element method. *Numer. Math.*, 66(4):493–515, 1994.
377. Luca F. Pavarino. Schwarz methods with local refinement for the p-version finite element method. *Numer. Math.*, 69(2):185–211, 1994.
378. Luca F. Pavarino. Neumann-Neumann algorithms for spectral elements in three dimensions. *RAIRO Modél. Math. Anal. Numér.*, 31:471–493, 1997.
379. Luca F. Pavarino. Indefinite overlapping Schwarz methods for time-dependent Stokes problems. *Comput. Methods Appl. Mech. Engrg.*, 187 (1-2):35–51, 2000.
380. Luca F. Pavarino and Timothy Warburton. Overlapping Schwarz methods for unstructured spectral elements. *J. Comput. Phys.*, 160 (1):298–317, 2000.
381. Luca F. Pavarino and Olof B. Widlund. A polylogarithmic bound for an iterative substructuring method for spectral elements in three dimensions. *SIAM J. Numer. Anal.*, 33(4):1303–1335, August 1996.
382. Luca F. Pavarino and Olof B. Widlund. Iterative substructuring methods for spectral elements: Problems in three dimensions based on numerical quadrature. *Comput. Math. Appl.*, 33(1/2):193–209, 1997.
383. Luca F. Pavarino and Olof B. Widlund. Iterative substructuring methods for spectral element discretizations of elliptic systems. I: Compressible linear elasticity. *SIAM J. Numer. Anal.*, 37(2):353–374, January 2000.
384. Luca F. Pavarino and Olof B. Widlund. Iterative substructuring methods for spectral element discretizations of elliptic systems. II: Mixed methods for linear elasticity and Stokes flow. *SIAM J. Numer. Anal.*, 37(2):375–402, January 2000.
385. Luca F. Pavarino and Olof B. Widlund. Iterative substructuring methods for spectral element discretizations of elliptic systems in three dimensions. In Petter Bjørstad and Michell Luskin, editors, *Parallel solution of partial differential equations (Minneapolis, MN, 1997)*, volume 120 of *IMA Vol. Math. Appl.*, pages 1–30, New York, 2000. Springer.
386. Luca F. Pavarino and Olof B. Widlund. Balancing Neumann-Neumann methods for incompressible Stokes equations. *Comm. Pure Appl. Math.*, 55(3):302–335, March 2002.
387. Gergina Pencheva and Ivan Yotov. Balancing domain decomposition for mortar mixed finite element methods. *Numer. Linear Algebra Appl.*, 10(1-2):159–180, 2003. Dedicated to the 60th birthday of Raytcho Lazarov.
388. Kendall H. Pierson. *A family of domain decomposition methods for the massively parallel solution of computational mechanics problems.* PhD thesis, University of Colorado at Boulder, Aerospace Engineering, 2000.
389. Janusz S. Przemieniecki. Matrix structural analysis of substructures. *Am. Inst. Aero. Astro. J.*, 1:138–147, 1963.
390. Alfio Quarteroni. Domain decomposition algorithms for the Stokes equations. In Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, editors, *Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods*, pages 431–442, Philadelphia, PA, 1989. SIAM. Los Angeles, California, January 14–16, 1988.
391. Alfio Quarteroni and Alberto Valli. *Numerical Approximation of Partial Differential Equations.* Springer-Verlag, Berlin, 1994. Second, Corrected Printing 1997.

392. Alfio Quarteroni and Alberto Valli. *Domain Decomposition Methods for Partial Differential Equations.* Oxford Science Publications, 1999.
393. Francesca Rapetti and Andrea Toselli. A FETI preconditioner for two dimensional edge element approximations of Maxwell's equations on non–matching grids. *SIAM J. Sci. Comput.*, 23(1):92–108, March 2001.
394. Pierre-Arnaud Raviart and Jean-Marie Thomas. A mixed finite element method for 2-nd order elliptic problems. In A. Dold and B. Eckmann, editors, *Mathematical Aspects of Finite Element Methods.* Springer, 1975. Lecture Notes of Mathematics, Volume 606.
395. Stefan Reitzinger and Joachim Schöberl. An algebraic multigrid method for finite element discretizations with edge elements. *Numer. Linear Algebra Appl.*, 9(3):223–238, 2002.
396. Oliver Rheinbach. FETI - a dual iterative substructuring method for elliptic partial differential equations. Master's thesis, Mathematisches Institut, Universität zu Köln, 2002.
397. Theodore J. Rivlin. *The Chebyshev Polynomials.* Wiley Interscience, 1990.
398. Daniel Rixen and Charbel Farhat. Preconditioning the FETI and balancing domain decomposition methods for problems with intra- and inter-subdomain coefficient jumps. In Petter E. Bjørstad, Magne Espedal, and David E. Keyes, editors, *Domain Decomposition Methods in Sciences and Enginnering. Ninth International Conference on Domain Decomposition Methods*, pages 472–479. DDM.org, 1998. Ullensvang at the Hardanger Fjord, Norway, June 4–7, 1996.
399. Daniel Rixen and Charbel Farhat. A simple and efficient extension of a class of substructure based preconditioners to heterogeneous structural mechanics problems. *Internat. J. Numer. Methods Engrg.*, 44:489–516, 1999.
400. Daniel Rixen and Charbel Farhat. A simple and efficient extension of a class of substructure based preconditioners to heterogeneous structural mechanics problems. *Internat. J. Numer. Methods Engrg.*, 44:489–516, 1999.
401. Daniel Rixen, Charbel Farhat, Radek Tezaur, and Jan Mandel. Theoretical Comparison of the FETI and Algebraically Partitioned FETI Methods, and Performance Comparisons with a Direct Sparse Solver. *Internat. J. Numer. Methods Engrg.*, 46:501–534, 1999.
402. Einar Rønquist. A domain decomposition solver for the steady Navier-Stokes equations. In A.V. Ilin and L.R. Scott, editors, *Proc. of ICOSAHOM.95, published by the Houston Journal of Mathematics*, pages 469–485, 1996.
403. Einar M. Rønquist. A domain decomposition solver for the incompressible Navier-Stokes equations. In *Workshop on spectral element methods, North Carolina State University*, 1994.
404. Einar M. Rønquist. Domain decomposition methods for the steady Stokes equations. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 330–340 (electronic). DDM.org, 1999. Greenwich, England, July 20–24, 1998.
405. Torgeir Rusten and Ragnar Winther. Mixed finite element methods and domain decomposition. In *Computational methods in water resources, IX, Vol. 1 (Denver, CO, 1992)*, pages 597–604. Comput. Mech., Southampton, 1992.
406. Torgeir Rusten and Ragnar Winther. A preconditioned iterative method for saddlepoint problems. *SIAM J. Matrix Anal. Appl.*, 13(3):887–904, 1992. Iterative methods in numerical linear algebra (Copper Mountain, CO, 1990).

407. Youcef Saad and Martin H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comp.*, 7:856–869, 1986.
408. Marzio Sala. An algebraic 2-level domain decomposition preconditioner with applications to the compressible Euler equations. *Internat. J. Numer. Methods Fluids*, 40(12):1551–1560, 2002. LMS Workshop on Domain Decomposition Methods in Fluid Mechanics (London, 2001).
409. Marcus Sarkis. Partition of unity coarse spaces and Schwarz methods with harmonic overlap. In Luca Pavarino and Andrea Toselli, editors, *Recent Developments in Domain Decomposition Methods*, Lecture Notes in Computational Science and Engineering, Volume 23, pages 77–94. Springer, 2002.
410. Marcus V. Sarkis. Two-level Schwarz methods for nonconforming finite elements and discontinuous coefficients. In N. Duane Melson, Thomas A. Manteuffel, and Steve F. McCormick, editors, *Proceedings of the Sixth Copper Mountain Conference on Multigrid Methods, Volume 2*, number 3224, pages 543–566, Hampton VA, 1993. NASA.
411. Marcus V. Sarkis. Multilevel methods for $P_1$ nonconforming finite elements and discontinuous coefficients in three dimensions. In David Keyes and Jinchao Xu, editors, *Domain decomposition methods in scientific and engineering computing. Seventh International Conference on Domain Decomposition*, pages 119–124. AMS, Contemporary Mathematics 180, 1994. Pennsylvania State University, October 27–30, 1993.
412. Marcus V. Sarkis. *Schwarz Preconditioners for Elliptic Problems with Discontinuous Coefficients Using Conforming and Non-Conforming Elements.* PhD thesis, Courant Institute of Mathematical Sciences, September 1994. TR-671, Department of Computer Science.
413. Marcus V. Sarkis. Nonstandard coarse spaces and Schwarz methods for elliptic problems with discontinuous coefficients using non-conforming elements. *Numer. Math.*, 77(3):383–406, 1997.
414. Alfred H. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Math. Comp.*, 28(128):959–962, 1974.
415. Christoph Schwab. *p– and hp– finite element methods.* Oxford Science Publications, 1998.
416. Hermann A. Schwarz. *Gesammelte Mathematische Abhandlungen*, volume 2, pages 133–143. Springer, Berlin, 1890. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, volume 15, 1870, pp. 272–286.
417. Spencer J. Sherwin and Mario A. Casarin. Low energy bases preconditioning for elliptic substructured solvers based on spectral/*hp* element discretizations. *J. Comput. Phys.*, 171:1–24, 2001.
418. David J. Silvester and Andrew J. Wathen. Fast iterative solution of stabilised Stokes systems part II: using general block preconditioners. *SIAM J. Numer. Anal.*, 31:1352–1367, 1994.
419. Morten D. Skogen. *Schwarz Methods and Parallelism.* PhD thesis, Department of Informatics, University of Bergen, Norway, February 1992.
420. Barry F. Smith. *Domain Decomposition Algorithms for the Partial Differential Equations of Linear Elasticity.* PhD thesis, Courant Institute of Mathematical Sciences, September 1990. TR-517, Department of Computer Science, Courant Institute.
421. Barry F. Smith. A domain decomposition algorithm for elliptic problems in three dimensions. *Numer. Math.*, 60(2):219–234, 1991.

422. Barry F. Smith. An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems. *SIAM J. Sci. Statist. Comput.*, 13(1):364–378, January 1992.
423. Barry F. Smith. A parallel implementation of an iterative substructuring algorithm for problems in three dimensions. *SIAM J. Sci. Comput.*, 14(2):406–423, March 1993.
424. Barry F. Smith, Petter E. Bjørstad, and William Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, 1996.
425. Sergey L. Sobolev. L'Algorithme de Schwarz dans la Théorie de l'Elasticité. *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS*, IV((XIII) 6):243–246, 1936.
426. Dan Stefanica and Axel Klawonn. The FETI Method for Mortar Finite Elements. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering*, pages 120–127. DDM.org, 1999. Proceedings of the 11th International Conference on Domain Decomposition Methods, Greenwich, UK, July 20–24, 1998, `URL: http://www.ddm.org/DD11/Stefanica.ps.gz`.
427. Rolf Stenberg and Manil Suri. Mixed *hp* finite element methods for problems in elasticity and Stokes flow. *Numer. Math.*, 72(3):367–389, 1996.
428. Gilbert Strang. *Linear algebra and its applications.* Academic-Press, New York, 1976.
429. Gábor Szegő. *Orthogonal Polynomials*, volume 23 of *Colloquium Publications.* American Mathematical Society, Providence, RI, 1975.
430. Daniel B. Szyld and Olof B. Widlund. Variational analysis of some conjugate gradient methods. *East-West J. Numer. Math.*, 1(1):51–74, 1993.
431. Xue-Cheng Tai. Domain decomposition for linear and nonlinear elliptic problems via function or space decomposition. In David Keyes and Jinchao Xu, editors, *Domain decomposition methods in scientific and engineering computing. Seventh International Conference on Domain Decomposition*, pages 335–360. AMS, Contemporary Mathematics 180, 1994. Pennsylvania State University, October 27–30, 1993.
432. Xue-Cheng Tai and Magne S. Espedal. Rate of convergence of some space decomposition method for linear and non-linear elliptic problems. *SIAM J. Numer. Anal.*, 35:1558–1570, 1998.
433. Xue-Cheng Tai and Jinchao Xu. Global convergence of subspace correction methods for some convex optimization problems. *Math. Comput.*, 71(237):105–124, 2002.
434. Kian H. Tan and Mart J. A. Borsboom. On generalized Schwarz coupling applied to advection-dominated problems. In David Keyes and Jinchao Xu, editors, *Domain decomposition methods in scientific and engineering computing. Seventh International Conference on Domain Decomposition*, pages 125–130. AMS, Contemporary Mathematics 180, 1994. Pennsylvania State University, October 27–30, 1993.
435. Radek Tezaur. *Analysis of Lagrange multiplier based domain decomposition.* PhD thesis, University of Colorado at Denver, 1998. `URL:http://www-math.cudenver.edu/graduate/ thesis/rtezaur.ps.gz`.
436. Vidar Thomée. *Galerkin finite element methods for parabolic problems*, volume 1054 of *Lecture notes in mathematics.* Springer-Verlag, Berlin, New York, 1984.

437. Andrea Toselli. Some results on overlapping Schwarz methods for the Helmholtz equation employing perfectly matched layers. In Choi-Hong Lai, Petter E. Bjørstad, Mark Cross, and Olof B. Widlund, editors, *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 539–545. DDM.org, 1999. Greenwich, England, July 20–24, 1998.
438. Andrea Toselli. Neumann–Neumann methods for vector field problems. *ETNA*, 11:1–24, 2000.
439. Andrea Toselli. Overlapping Schwarz methods for Maxwell's equations in three dimensions. *Numer. Math.*, 86(4):733–752, 2000.
440. Andrea Toselli. FETI domain decomposition methods for scalar advection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.*, 190(43-44):5759–5776, 2001.
441. Andrea Toselli. Two iterative substructuring methods for Maxwell's equations with discontinuous coefficients in two dimensions. In Tony Chan, Takashi Kako, Hideo Kawarada, and Olivier Pironneau, editors, *Domain Decomposition Methods in Sciences and Engineering. Twelfth International Conference on Domain Decomposition*, pages 215–222. DDM.org, 2001. Chiba, Japan, October 25–29, 1999.
442. Andrea Toselli. Dual-Primal FETI algorithms for edge element approximations: Three-dimensional $h$ finite elements on shape-regular meshes. In preparation, 2004.
443. Andrea Toselli and Axel Klawonn. A FETI domain decomposition method for edge element approximations in two dimensions with discontinuous coefficients. *SIAM J. Numer. Anal.*, 39:932–956, 2001.
444. Andrea Toselli and Xavier Vasseur. Neumann-Neumann and FETI preconditioners for $hp$-approximations on geometrically refined boundary layer meshes in two dimensions. Technical Report 02–15, Seminar für Angewandte Mathematik, ETH, Zürich, 2002. Submitted to Numer. Math.
445. Andrea Toselli and Xavier Vasseur. A numerical study on Neumann-Neumann and FETI methods for $hp$-approximations on geometrically refined boundary layer meshes in two dimensions. *Comput. Methods Appl. Mech. Engrg.*, 192:4551–4579, 2003.
446. Andrea Toselli and Xavier Vasseur. Domain decomposition preconditioners of Neumann-Neumann type for $hp$-approximations on boundary layer meshes in three dimensions. *IMA J. Numer. Anal.*, 24(1):123–156, 2004.
447. Andrea Toselli and Xavier Vasseur. Dual-Primal FETI algorithms for edge element approximations: Two-dimensional $h$ and $p$ finite elements on shape-regular meshes. Technical Report 04–02, Seminar für Angewandte Mathematik, ETH, Zürich, 2004. To appear in SIAM J. Numer. Anal.
448. Andrea Toselli and Xavier Vasseur. A numerical study on dual-primal FETI algorithms for edge element approximations: two-dimensional $p$ and $hp$ finite elements on shape-regular and anisotropic meshes. In preparation, 2004.
449. Andrea Toselli, Olof B. Widlund, and Barbara I. Wohlmuth. An iterative substructuring method for Maxwell's equations in two dimensions. *Math. Comp.*, 70(235):935–949, 2001.
450. Lloyd N. Trefethen and David Bau, III. *Numerical linear algebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
451. R. Loredana Trotta. Multidomain finite elements for advection-diffusion equations. *Appl. Numer. Math.*, 21(1):91–118, 1996.

452. Petr Vaněk, Marian Brezina, and Jan Mandel. Convergence of algebraic multigrid based on smoothed aggregation. *Numer. Math.*, 88:559–679, 2001.
453. Petr Vaněk, Marian Brezina, and Radek Tezaur. Two-grid method for linear elasticity on unstructured meshes. *SIAM J. Sci. Comput.*, 21:900–923, 1999.
454. Petr Vaněk, Jan Mandel, and Marian Brezina. Algebraic multigrid by smooth aggregation for second and fourth order elliptic problems. *Computing*, 56:179–196, 1996.
455. Petr Vaněk, Jan Mandel, and Marian Brezina. Solving a two-dimensional Helmholtz problem using algebraic multigrid. Technical Report 110, Center for Computational Mathematics, University of Colorado, Denver, October 1997.
456. Richard Varga. *Matrix Iterative Analysis.* Prentice Hall, first edition, 1962.
457. Panayot S. Vassilevski and Junping Wang. An application of the abstract multilevel theory to nonconforming finite element methods. *SIAM J. Numer. Anal.*, 32(1):235–248, 1995.
458. Rüdiger Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques.* Wiley and Teubner, 1996.
459. Timothy Warburton, Luca F. Pavarino, and Jan S. Hesthaven. A pseudo-spectral scheme for the incompressible Navier-Stokes equations using unstructured nodal elements. *J. Comput. Phys.*, 164(1):1–21, 2000.
460. Andrew Wathen and David Silvester. Fast iterative solution of stabilised Stokes systems. I. Using simple diagonal preconditioners. *SIAM J. Numer. Anal.*, 30(3):630–649, 1993.
461. Andrew J. Wathen, Bernhard Fischer, and David J. Silvester. The convergence rate of the minimal residual method for the Stokes problem. *Numer. Math.*, 71:121–134, 1995.
462. Mary Fanett Wheeler and Ruth Gonzalez. Mixed finite element methods for petroleum reservoir engineering problems. In Roland Glowinski and Jacques-Louis Lions, editors, *Computing methods in applied sciences and engineering, VI (Versailles, 1983)*, pages 639–657. North-Holland, Amsterdam, 1984.
463. Olof B. Widlund. An extension theorem for finite element spaces with three applications. In Wolfgang Hackbusch and Kristian Witsch, editors, *Numerical Techniques in Continuum Mechanics*, pages 110–122, Braunschweig/Wiesbaden, 1987. Notes on Numerical Fluid Mechanics, v. 16, Friedr. Vieweg und Sohn. Proceedings of the Second GAMM-Seminar, Kiel, January, 1986.
464. Olof B. Widlund. A comparison of some domain decomposition and iterative refinement algorithms for elliptic finite element problems. Technical Report BSC 88/15, IBM Bergen Scientific Centre, Thormøhlensgaten 55, N-5006 Bergen, Norway, 1988.
465. Olof B. Widlund. Iterative substructuring methods: Algorithms and theory for elliptic problems in the plane. In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 113–128, Philadelphia, PA, 1988. SIAM. Paris, France, January 7–9, 1987.
466. Olof B. Widlund. Some Schwarz methods for symmetric and nonsymmetric elliptic problems. In David E. Keyes, Tony F. Chan, Gérard A. Meurant, Jeffrey S. Scroggs, and Robert G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 19–36, Philadelphia, PA, 1992. SIAM. Norfolk, Virgina, May 6–8, 1991.

467. Olof B. Widlund. Exotic coarse spaces for Schwarz methods for lower order and spectral finite elements. In David Keyes and Jinchao Xu, editors, *Domain decomposition methods in scientific and engineering computing. Seventh International Conference on Domain Decomposition*, pages 131–136. AMS, Contemporary Mathematics 180, 1994. Pennsylvania State University, October 27–30, 1993.
468. Françoise Willien, Isabelle Faille, Frédéric Nataf, and Frédéric Schneider. Domain decomposition methods for fluid flow in porous medium. In *6th European Conference on the Mathematics of Oil Recovery*, September 1998.
469. Barbara I. Wohlmuth. *Discretization Methods and Iterative Solvers Based on Domain Decomposition*, volume 17 of *Lecture Notes in Computational Science and Engineering*. Springer Verlag, 2001.
470. Barbara I. Wohlmuth, Andrea Toselli, and Olof B. Widlund. Iterative substructuring method for Raviart-Thomas vector fields in three dimensions. *SIAM J. Numer. Anal.*, 37(5):1657–1676, 2000.
471. Jinchao Xu. Counter examples concerning a weighted $L^2$ projection. *Math. Comp.*, 57:563–568, 1991.
472. Jinchao Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34(4):581–613, December 1992.
473. Jinchao Xu. Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.*, 33:1759–1777, 1996.
474. Jinchao Xu. The method of subspace corrections. *J. Comput. Appl. Math.*, 128(1–2):335–362, 2001.
475. Jinchao Xu and Ludmil Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Amer. Math. Soc.*, 15(3):573–597, 2002.
476. Jinchao Xu and Jun Zou. Some nonoverlapping domain decomposition methods. *SIAM Review*, 40:857–914, 1998.
477. Kôsaku Yosida. *Functional Analysis*. Springer-Verlag, sixth edition, 1980.
478. Harry Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 49:379–412, 1986.
479. Xuejun Zhang. *Studies in Domain Decomposition: Multilevel Methods and the Biharmonic Dirichlet Problem*. PhD thesis, Courant Institute, New York University, September 1991. TR-584, Department of Computer Science.
480. Xuejun Zhang. Domain decomposition algorithms for the biharmonic Dirichlet problem. In *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 119–126, Philadelphia, PA, 1992. SIAM. Norfolk, Virgina, May 6–8, 1991.
481. Xuejun Zhang. Multilevel Schwarz methods. *Numer. Math.*, 63(4):521–539, 1992.
482. Jun Zou and Hong Ci Huang. Algebraic subproblem decomposition methods and parallel algorithms with monotone convergence. *J. Comput. Math.*, 10:47–59, 1992.

# Index


$H(\mathbf{curl}\,;\Omega)$, 348, 349
  approximation, *see* Nédélec elements
  preconditioning, 271
$H(\mathrm{div}\,;\Omega)$, 347
  approximation, *see* Raviart-Thomas elements
  preconditioning, 271

affine mapping, 371
anisotropic meshes, 202, 206, 215, 273
average operator, 134, 139, 140, 151, 208

basis functions
  Lagrangian, 373
block preconditioners, 24, 235, 241, 258

Cauchy-Schwarz inequalities, strengthened, 40
coarse grid, 59, 90, 219, 274, 289, 314
coloring, 46
condition number, 395
  convergence of conjugate gradient, 407
  of Schur complement, 97, 202
  of stiffness and mass matrix, 390
Conjugate Gradient, 403
Conjugate Residual, 409
convection-diffusion problems, 311, 326
counting functions, 134
  for $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$, 301
  for elasticity, 226
  for scalar elliptic problems, 134
  for spectral elements, 208
curl, *see* $H(\mathbf{curl}\,;\Omega)$

Darcy flows, *see* porous media
Deville-Mund preconditioner, 196
direct methods, 397
Dirichlet-Neumann, 8, 18, 328
discontinuous Galerkin, 82, 327
divergence, *see* $H(\mathrm{div}\,;\Omega)$

edge elements, *see* Nédélec elements
eigenvalue problems, 335
eigenvalues, 395
elasticity, linear
  approximation, 375
  nearly incompressible, 257, 366, 388
  preconditioning, 217
  well-posedness, 357
extensions
  continuous spaces, 342
  discrete harmonic, *see* harmonic extensions, discrete

factorizations, 397
FETI, 131
  DP, *see* dual-primal
  dual-primal, 160
    for elasticity, 227
    acceptable edge path, 172, 229
    analysis, 175
    three dimensions, 167
    two dimensions, 161
  for $H(\mathrm{div}\,;\Omega)$ and $H(\mathbf{curl}\,;\Omega)$, 304
  for convection-diffusion, 329
  for Helmholtz, 332

for porous media, 241
for spectral elements, 208
for Stokes problems, 241
one-level, 143
nonredundant Lagrange multipliers, 150
redundant Lagrange multipliers, 156
one-level for elasticity, 227
two subdomains, 12, 241
fictitious domain methods, 233
fill-in, 398
finite elements
spaces, 372
triangulation, *see* triangulation
floating subdomains, 88
Fredholm alternative, 364
Friedrichs inequalities, 104, 343

Gårding inequality, 313, 364
Galerkin approximations, 371
Gauss-Lobatto
meshes, 194
points, 377
GMRES, 315, 407
Green's formulas, 347, 349, 350

harmonic extensions, discrete, 96
for elasticity, 222
comparison, 264
for $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$, 290, 295
for porous media, 253
for spectral elements, 201
for Stokes problems, 264
Helmholtz equation, 312, 330
FETI-H, 332
overlapping methods, 331

indefinite problems, 362
algebraic, 407
approximation, 375
inexact solvers, 36
interface conditions, *see* transmission conditions
interface, subdomain, 88
interpolant
$L^2$ projection, 101
in $H(\mathbf{curl};\Omega)$, 384
in $H(\mathrm{div};\Omega)$, 381
nodal into and from GLL meshes, 195
quasi-interpolant, 62
standard nodal, 99, 100, 112, 373
standard nodal for nonnested meshes, 64
interpolation
spaces, 339, 343
inverse inequalities, 389

jump operator, 150

Korn inequalities, 358

Lax-Milgram lemma, 354, 363
Lipschitz domain, 337

mass matrix, 391
Maxwell's equations, 362
mesh, *see* triangulation
mixed methods, *see* saddle point problems
mortar methods, 233

Nédélec elements
interpolant, 384
range and kernel of the curl, 384
three dimensions, 383
two dimensions, 382
Neumann-Dirichlet, 15
Neumann-Neumann, 10, 16, 131
balancing, 133, 250, 265, 301
for $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$, 301
for convection-diffusion, 328
for elasticity, 225
for porous media, 250
for spectral elements, 208
for Stokes problems, 265
two subdomains, 10, 241
nonlinear problems, 334
nonnested grids, 64

optimal, method, 9
optimality, 9
optimized interface conditions, 333
overlap, definition/small/generous, 56
overlapping methods, 55
alternative coarse problems, 75
analysis, 67
for $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$, 274

for elasticity, 219
for Helmholtz, 331
for nonsymmetric or indefinite problems, 314
for porous media, 255
for spectral elements, 198
for Stokes problems, 258
local solvers, 56
numerical results, 73
standard coarse spaces, 59

partition of unity
coarse spaces, 84
functions, 57, 105, 204, 295
plane waves, 333
Poincaré inequalities, 104, 343
porous media, flows in
approximation, 389
preconditioners, 239
well-posedness, 367
positive definite problems, 353
algebraic, 395, 403
approximation, 374
preconditioner, 401, 406, 409
projection, *see* interpolant

quasi-optimal, method, 18
quotient space arguments, 60, 346

Raviart-Thomas elements, 380
interpolant, 381
range and kernel of the divergence, 384
reference element, 371
regularity results, 369
Richardson method, 399
Robin-Robin, 328, 329

saddle point problems
approximation, 386
well-posedness, 364
scalability, 17
scalable methods, 17
scaling arguments, 60
Schur complement, 5, 94, 201, 244, 251, 263, 290, 397
application, 94
application of the inverse, 95
condition number, 97, 202
Uzawa algorithm, 233, 237
Schwarz methods
additive, 37
alternating, 21, 27
convergence theory, 39
definition, 35
hybrid, 38, 47, 137
implementation, 52
multiplicative, 37
restricted, 75
smoothed aggregation, 81
Sobolev spaces
definition, 337
immersions, 340
of vector-valued functions, 346
traces, 341
spectral elements, 376
approximations, 378
interpolant, 195, 378
preconditioners, 193
spaces, 376
spectral radius, 395
spectrum, *see* eigenvalues
steepest descent, 402
Steklov-Poincaré operator, 6
stiffness matrix, 390
Stokes problem
finite element approximations, 387
preconditioners, 257
spectral element approximations, 388
well-posedness, 366
subdomain partitions
nonoverlapping, 88
overlapping, 56
substructuring methods
for $H(\mathrm{div};\Omega)$ and $H(\mathbf{curl};\Omega)$, 288
for linear elasticity, 220
for nonsymmetric or indefinite problems, 320, 327, 331, 333
for porous media, 241
for scalar problems, 113, 131
for spectral elements, 200
for Stokes problems, 261
generalities, 87

trace operators
continuous spaces, 341
in $H(\mathbf{curl};\Omega)$, 349, 350
in $H(\mathrm{div};\Omega)$, 347

transmission conditions, 2, 4, 242, 245, 331, 333
triangulation, 371
  quasiuniform, 372
  shape-regular, 372

Uzawa algorithm, 233

weighted average, *see* average operator
wire basket
  an extension from, 222
  definition, 89
  method for $H(\mathbf{curl};\Omega)$, 308
  method for elasticity, 221
  method for spectral elements, 206