On Multilevel Methods Based on Non-Nested Meshes

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Thomas Dickopf

aus

Dernbach

Bonn, August 2010

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Referent: Prof. Dr. Rolf Krause

2. Referent: Prof. Dr. Martin Rumpf

Tag der Promotion: 18. November 2010

Erscheinungsjahr: 2010

Diese Dissertation ist mit Unterstützung durch ein Hausdorff-Stipendium der von der Deutschen Forschungsgemeinschaft getragenen Bonn International Graduate School in Mathematics (BIGS) entstanden. Weitere Mittel der Deutschen Forschungsgemeinschaft wurden durch den Sonderforschungsbereich Singuläre Phänomene und Skalierung in mathematischen Modellen (SFB 611) zur Verfügung gestellt.

Zusammenfassung

Diese Arbeit beschäftigt sich mit Multilevel-Verfahren zur effizienten Lösung von Partiellen Differentialgleichungen im Bereich des Wissenschaftlichen Rechnens. Dabei liegt ein weiterer Schwerpunkt auf der eingehenden Untersuchung des Informationsaustauschs zwischen Finite-Elemente-Räumen zu nicht-geschachtelten Gittern.

Zur Diskretisierung von komplizierten Geometrien mit einer Finite-Elemente-Methode sind unstrukturierte Gitter oft von Vorteil, weil sie der Form des Rechengebiets einfacher angepasst werden können. Solche Gitter, und somit die zugehörigen diskreten Funktionenräume, besitzen im Allgemeinen keine leicht zugängliche Multilevel-Struktur, die sich zur Konstruktion schneller Löser ausnutzen ließe. In der vorliegenden Arbeit stellen wir eine Klasse "semi-geometrischer" Multilevel-Iterationen vor, die auf Hierarchien voneinander unabhängiger, nicht-geschachtelter Gitter beruhen. Dabei bestimmen in einem variationellen Ansatz rekursiv die Bilder geeigneter Prolongationsoperatoren im jeweils folgenden (feineren) Raum die Grobgitterräume. Das semi-geometrische Konzept ist sehr allgemeiner Natur verglichen mit anderen Verfahren, die auf geometrischen Überlegungen beruhen. Dies zeigt sich in der verhältnismäßig losen Beziehung der verwendeten Gitter zueinander. Der konkrete Nutzen des Ansatzes mit nicht-geschachtelten Gittern ist die Flexibilität der Wahl der Grobgitter. Diese können beispielsweise unabhängig mit Standardverfahren generiert werden. Die Auflösung des Randes des tatsächlichen Rechengebiets in den konstruierten Grobgitterräumen ist eine Eigenschaft der entwickelten Verfahrensklasse.

Die flexible Einsetzbarkeit und die Effizienz der vorgestellten Lösungsverfahren zeigt sich in einer Reihe von numerischen Experimenten. Dazu geben wir Hinweise zur praktischen Umsetzung der semi-geometrischen Ideen und konkreter Transfer-Konzepte zwischen nicht-geschachtelten Gittern. Darüber hinaus wird eine Erweiterung zu einem semi-geometrischen monotonen Mehrgitterverfahren zur Lösung von Variationsungleichungen untersucht.

Wir führen die Analysis der Konvergenz- bzw. Vorkonditionierungseigenschaften im Rahmen der Theorie der Teilraumkorrekturmethoden durch. Unsere technische Ausarbeitung liefert ein quasi-optimales Resultat, das wir mithilfe lokaler Argumente für allgemeine, shape-reguläre Gitterfamilien beweisen. Als relevante Eigenschaften der Operatoren zur Prolongation zwischen nicht-geschachtelten Finite-Elemente-Räumen erweisen sich die H^1-Stabilität und eine L^2-Approximationseigenschaft sowie die Lokalität des Transfers.

Diese Arbeit ist ein Beitrag zur Entwicklung schneller Löser für Gleichungen auf komplizierten Gebieten mit Schwerpunkt auf geometrischen Techniken (im Unterschied zu algebraischen). Verbindungen zu anderen Ansätzen werden sorgfältig aufgezeigt. Daneben untersuchen wir den Informationsaustausch zwischen nicht-geschachtelten Finite-Elemente-Räumen als solchen. In einer neuartigen Studie verbinden wir theoretische, praktische und experimentelle Überlegungen. Eine sorgfältige Prüfung der qualitativen Eigenschaften sowie eine quantitative Analyse der Unterschiede verschiedener Transfer-Konzepte zueinander führen zu neuen Ergebnissen bezüglich des Informationsaustauschs selbst. Schließlich erreichen wir durch die Einführung eines verallgemeinerten Projektionsoperators, der Pseudo-L^2-Projektion, eine deutlich bessere Approximation der eigentlichen L^2-orthogonalen Projektion als andere Ansätze aus der Literatur.

Danksagung

Mein erster Dank gilt Herrn Professor Rolf Krause, der mich in das Gebiet des Wissenschaftlichen Rechnens und insbesondere in die Theorie der Teilraumkorrekturmethoden eingeführt hat. Für die vielfältige Förderung und die zahlreichen Hilfestellungen sowie die Möglichkeit der aktiven Teilnahme an internationalen Konferenzen bin ich ihm sehr dankbar. Sein wertvoller Rat hat große Teile dieser Arbeit maßgeblich beeinflusst. Ich danke Herrn Professor Martin Rumpf für die freundliche Fortführung der Betreuung in Bonn verbunden mit der angenehmen Einführung in seine Arbeitsgruppe und der Übernahme des Zweitgutachtens. Herrn Professor Helmut Harbrecht danke ich für sein wohltuendes Interesse an meiner Arbeit.

Der Bonn International Graduate School in Mathematics sei gedankt für die Gewährung eines Hausdorff-Stipendiums sowie großzügiger Reisemittel. Für die Bereitstellung einer hervorragenden Infrastruktur und weiterer finanzieller Mittel danke ich dem Institut für Numerische Simulation der Rheinischen Friedrich-Wilhelms-Universität in Bonn, dem Sonderforschungsbereich Singuläre Phänomene und Skalierung in mathematischen Modellen (SFB 611) und dem Institute of Computational Science der Università della Svizzera italiana in Lugano.

Ein herzlicher Dank gilt allen Kollegen am Institut für Numerische Simulation der Rheinischen Friedrich-Wilhelms-Universität in Bonn und am Institute of Computational Science der Università della Svizzera italiana in Lugano. Ihre außerordentliche Freundlichkeit weiß ich sehr zu schätzen. Vor allem danke ich Frau Mirjam Walloth, die mir bei einer Vielzahl von Details bereitwillig zugehört hat und mir oft weiterhelfen konnte. Darüber hinaus hat sie einen großen Teil dieser Arbeit sehr intensiv korrekturgelesen. Herrn Christian Groß bin ich besonders dankbar für die abwechslungsreiche Zeit in unseren beiden gemeinsamen Büros und auf verschiedenen Konferenzreisen. Die mathematischen, informatischen und weltanschaulichen Diskussionen habe ich sehr genossen. Ich danke Herrn Arne Dirks, mit dem ich in einem Zwischenstadium dieser Arbeit einige Implementierungsfragen diskutieren durfte. Bei Herrn Dorian Krause bedanke ich mich für die stets zuverlässige Administration des Rechenclusters in Lugano und unkomplizierte Hilfe bei dessen Verwendung. Frau Christina Mohr und Herrn Johannes Steiner danke ich für ihr angenehmes Interesse an meiner Arbeit und das sorgfältige Korrekturlesen großer Teile derselben.

I would like to thank Professor David Keyes and his group at the Department of Applied Physics and Applied Mathematics, Columbia University in the City of New York for their kind hospitality during an inspiring research stay in fall 2008, especially Mr. Aron Ahmadia and Mr. Braxton Osting for their continuous help and friendship.

Schließlich habe ich vielen Freunden zu danken, die mich in den letzten Jahren begleitet und unterstützt haben. Ganz besonders danke ich meiner Familie für ihre uneingeschränkte Unterstützung, die mir stets großartigen Rückhalt gibt.

Bonn, August 2010 Thomas Dickopf

Contents

Introduction

This thesis is about multilevel methods for an efficient solution of partial differential equations in complicated domains. We introduce a new class of semi-geometric preconditioners and multigrid methods for problems arising from unstructured finite element discretizations. The multilevel framework is developed from a variational approach based on a hierarchy of non-nested meshes. We present new results on the proposed multilevel iterative methods as well as the actual information transfer between finite element spaces associated with non-nested meshes.

Background

Mathematical models of many phenomena in the natural sciences and engineering are formulated as boundary value problems of partial differential equations. For this class of problems and many others, computer experiments have proved their capability of providing additional insight or even complementing or substituting actual experiments. In the field of scientific computing, beside important modeling aspects, the design of an efficient numerical simulation also comprises an appropriate discrete approximation of the considered quantities. For many problems associated with partial differential equations, finite element methods [25, 39, 56] are popular choices as they have favorable properties from both a theoretical and a practical point of view.

A crucial factor for an efficient numerical treatment of partial differential equations is the solution of the appearing linear systems of equations, where required after a linearization or an implicit time discretization by Rothe's method. Although such a linear system of equations, which is typically large but sparse and in our applications also symmetric positive definite but ill-conditioned, can in principle be solved disregarding the underlying discretization scheme, one may profit from additional insight into the structure of the considered problem. The multilevel methods to be studied in the present thesis do this in a rather sophisticated manner.

Multilevel methods

In scientific computing, the term multilevel appears in many ways. During the last decades, multilevel ideas have influenced the thinking of many researchers in some form or another, ranging from advanced mathematical modeling aspects in order to concurrently describe phenomena on different length or time scales to the design of modern computer architectures. For the development of numerical methods for partial differential equations, the multilevel paradigm is particularly appealing for both analytical and algorithmical reasons.

In this thesis, we focus on the application of multilevel ideas to the development of iterative solvers for elliptic partial differential equations. As a direct solution is usually not feasible for large systems due to enormous time and memory consumption, the discrete systems need to be solved iteratively to achieve a reasonably flexible numerical simulation environment. Multigrid methods [37, 103, 178, 199] turn out to be the fastest solvers in many applications. They show an optimal convergence behavior in the sense that the work

required to reduce the iteration error to a requested accuracy is proportional to the problem size. It is well-known that the performance of the classical linear iterations such as the Jacobi or the Gauß–Seidel method is not satisfactory even for simple equations as it degenerates with increasing problem size. In rather general settings, they have smoothing properties, though. The power of multilevel iterations results from a sophisticated combination of smoothing iterations and coarse level corrections. These ingredients should be complementary in the sense that they reduce different components of the error; at each level a different section of the spectrum should be processed. This paradigm manifests in the multigrid methods. Here, only very few steps of a relaxation method are performed at each level to obtain defect problems where the corresponding errors may be well represented in spaces with less degrees of freedom. An essential element of an efficient algorithm is a methodology of how to realize such a coarse approximation.

Multilevel finite elements

A numerical approximation of the continuous quantities relies on a suitable discrete representation of the computational domain, for instance by a grid or mesh. The finite element method is usually preferred to finite difference schemes in case the resolution of the potentially complicated geometry is of interest. In addition, the variational setting in a Hilbert space allows for a powerful convergence analysis. More precisely, the considered multilevel methods fit into the framework of additive and multiplicative Schwarz methods [177] acting on the residual by parallel and successive subspace correction [194], respectively. The decisive steps for the analysis of the multiplicative case were taken by Bramble, Pasciak, Wang and Xu in [27, 32, 33, 193]. The final breakthrough was made by Oswald [151] and others establishing a fundamental connection of multilevel finite elements to approximation theory. The parallel BPX-preconditioner [34] and the work by Griebel [95, 96] show that the strictly level-oriented view is actually not mandatory provided that suitable multilevel bases are used.

The standard multigrid algorithms for finite element discretizations are based on a hierarchy of finite element spaces associated with a sequence of nested meshes. In this case, the variational approach in a suitable Hilbert space yields a very natural way to realize the coarse space approximation and the information transfer between two successive spaces by the canonical inclusion (coarse-to-fine) and the orthogonal projection (fine-to-coarse). However, many important applications in computational engineering, especially involving complicated geometries in three dimensions, do not allow for straightforward multilevel hierarchies. The treatment of general unstructured finite element meshes is a demanding task for multilevel iterative methods. Such meshes are beneficial for a flexible adaptation of the discrete representation to the computational domain with relatively few degrees of freedom, though. In fact, the shortcomings of standard multigrid methods regarding the handling of complex geometric data may be considered one of the major reasons for multilevel methods not being as prevalent as their powerful convergence or preconditioning properties would certainly justify.

Multilevel methods for complicated geometries

This thesis constitutes a contribution to the development of multilevel methods for complicated domains, which is clearly a topic of current research interest. In the past years, several methodologies have been developed for the application of basic multilevel algorithms to problems with complicated boundaries of the computational domain. The particular strategies to obtain suitable hierarchies or approximation spaces are in part very different.

For instance, the theoretical contribution of Yserentant [200] about so-called boundary fitted elements may be considered a justification of the paradigm to construct a multilevel hierarchy by coarsening away from the boundary. Other methods are based on tailored fine and coarse level discretizations, which are mostly built from structured meshes. The multigrid method based on parametric finite elements [107, 108] we present in Chapter 7 belongs to this class of algorithms. The conceptually simplest method to determine a multilevel hierarchy for complicated domains is a rough approximation by a structured mesh from the inside, which is analyzed by Kornhuber and Yserentant [123] in case of a pure Dirichlet problem. The composite finite element methods by Hackbusch and Sauter [105, 106] also rely on sequences of structured meshes. By sophisticated adaptation procedures of sufficiently fine, structured meshes to the boundaries of the computational domains, logically nested and physically almost nested mesh hierarchies are constructed. A related technique motivated by image based computing in Liehr et al. [137] works with locally cut off basis functions close to the boundary.

All these approaches have in common that the relation of two successive meshes is generally much closer than in the setting that we aim at; the families of meshes exhibit some additional structure.

Further, meshfree and particle methods, which do not rely on a mesh, constitute a different approach to approximate continuous quantities in potentially complicated domains. For the discretization of partial differential equations, for instance, the partition of unity methods have been developed, which glue local approximation spaces by partition of unity functions associated with overlapping decompositions of the computational domains; see Babuška and Melenk [10, 145]. Multilevel methods for partition of unity discretizations of elliptic partial differential equations have been studied by Griebel and Schweitzer [99, 170] and others. We refer to [55, 62] for recent analytical results.

Another branch of research is concerned with the development of algebraic multigrid methods [159, 176], which attempt to construct coarse level hierarchies only by taking the entries of the given stiffness matrix into account, at least in the pure form. Variants for finite element discretizations exist; see [40, 53, 109, 115].

The auxiliary space method by Xu [195] is a general approach to make an existing multilevel preconditioner, which has been designed for a simpler setting, applicable to more complicated problems. One of its main purposes is the reuse of available implementations by transferring the given problem to a more easily treatable auxiliary space. Applied to the issue of finite element spaces associated with unstructured meshes representing complicated domains, this amounts to a non-variational approach using a (multilevel) preconditioner defined with respect to structured meshes. However, both analysis and experiments indicate that the sizes of the original space and of the auxiliary space need to be comparable in a

quite restrictive sense; see [195]. Further applications of this idea to, e. g., higher order elements or other problem classes have been investigated in [195] and then [112, 113, 179].

We turn our special attention to another class of multilevel methods. The idea of accelerating an iterative solution process by adding a correction step from non-nested meshes has probably first been used in practice by aerospace engineers as early as the year 1986; see the reports by Mavriplis and Jameson [143, 144] and also Löhner and Morgan [139]. It has later found its way into domain decomposition methods by Cai [43] and Chan et al. [47, 48, 49, 51, 52].

Purpose of this thesis

This thesis is about multilevel methods based on non-nested meshes, which meet the challenges indicated above. We approach this topic in different ways. Our research has clearly been driven by the desire to construct coarse level approximations to be efficiently used in multilevel preconditioners. However, we also contribute to the research on the actual information transfer between finite element spaces associated with non-nested meshes.

The purpose of the present text is to provide additional insight into the design of coarse spaces in case of unstructured finite element meshes. We focus on geometric multilevel techniques (as opposed to algebraic ones); both fine and coarse level spaces are always associated with a mesh in one way or another. Consequently, we put emphasis on geometrically motivated or inspired transfer concepts.

The proposed multilevel iterations, i. e., multigrid methods and multilevel preconditioners, rely on a variational approach based on a hierarchy of non-nested coarse meshes. The assumptions on the ingredients of our framework are particularly weak. We analyze the presented preconditioners and multigrid methods in the context of additive and multiplicative Schwarz methods to obtain preconditioning and convergence results independent of the mesh size, respectively. Our careful analysis offers a clear view of the requirements the geometric interlevel transfer needs to satisfy. In addition, we relate this approach to other geometry-based multilevel techniques.

Throughout this text, we highlight the actual information transfer between finite element spaces associated with nested or non-nested meshes. This means we also direct our attention to the information transfer as such, in contrast to its distinguished role in the constructed multilevel methods. We are convinced that a deep insight into diverse transfer concepts is very helpful for both the construction of specific operators and their application in multilevel methods. Accordingly, this thesis comprises new results regarding geometric information transfer between finite element spaces associated with non-nested meshes.

Several practical considerations complement the indicated theoretical achievements. In this regard, we address implementation aspects of both the abstract multilevel framework based on non-nested meshes and the concrete realization of diverse transfer concepts of the requisite interlevel transfer. Various numerical experiments constitute an integral part of this thesis. We demonstrate the performance of the introduced multilevel iterative methods and investigate the fundamental characteristics of information transfer between non-nested meshes in practice.

Therefore, we are confident that this text is indeed a comprehensive elaboration of the topic "Multilevel Methods Based on Non-Nested Meshes". We will suggest several open questions in the conclusion.

Multilevel methods based on non-nested meshes

We study the approach of preconditioning based on non-nested mesh hierarchies in full detail. In view of the abstract concept being somewhere between geometric and algebraic multigrid methods but clearly leaning towards the geometric side, the term "semi-geometric" would certainly be justified to name the proposed multilevel methods.

The strategy for the development of the semi-geometric framework to be presented here is to use a family of completely unrelated meshes with associated finite element spaces to construct a hierarchy of nested spaces by a variational approach. This can be achieved by recursively considering the ranges of suitable prolongation operators in the next finer spaces. For the additive variant, two different possibilities are considered.

The core of our analysis is the existence proof of suitable fine-to-coarse mappings, which allow for a stable subspace splitting as introduced in [32, 33, 151, 193] and thus quasi-optimal convergence and preconditioning results. To this end, we identify relevant properties for the information transfer between non-nested finite element spaces, namely the H^1-stability and the L^2-approximation property. In our technical elaboration, we carefully distinguish between the different domains which are represented by the separate meshes. Besides, working with local measures for the mesh sizes, we achieve results that hold true for non-quasi-uniform meshes provided that the interlevel transfer concept is chosen appropriately.

The semi-geometric concept reflects a rather weak setting (yet still variational) compared with other geometry-based approaches; the relations between the employed meshes are particularly loose. The concrete benefit of using non-nested meshes is that the coarse meshes can be chosen quite freely, e. g., generated independently by standard mesh generators. The approach has the additional advantage that for the coarse level problems no approximation of the boundary data is necessary.

The performance and the flexibility of the devised methods are demonstrated by numerical results. We report on a number of experiments carried out in various different ways to study the practical properties of the semi-geometric concept. Finally, we do not only investigate the methodology based on non-nested meshes in detail but also extend our algorithm to a semi-geometric monotone multigrid method for variational inequalities and briefly study an application to Signorini's problem.

Information transfer between non-nested meshes

We believe that a comprehensive analysis of the studied topic needs to include a thorough investigation of the actual information transfer. In this context, it is not at all clear from the start which transfer concepts are best suited for embedding a non-nested coarse space into a finite element space associated with a finer mesh. This is true despite the availability of some obvious candidates.

Our research is in part motivated by the early work of Clément [58] on quasi-interpolation and then Scott and Zhang [171]. We also learned about advanced techniques for

the construction of transfer operators from Wohlmuth [190, 191] and Kim et al. [118] in the context of non-conforming domain decomposition methods. Other interesting studies giving basic insights into the analysis of approximation operators in finite element spaces, which influenced our work, can be found in [8, 30, 31, 43, 49, 174, 175, 177, 193], partially in completely different contexts.

In this thesis, both intuitive and more elaborate mappings are examined in a unique study combining theoretical, practical and experimental considerations. We discuss locally and globally defined operators including well-known quasi-interpolation concepts and also focus on their algorithmic structures. Our considerations cover a respectable range of geometric ideas. On the one hand, we aim to assess the suitability of the transfer concepts for the use in the semi-geometric framework. On the other hand, we are interested in determining qualitative and quantitative differences of the generated operators as such. At this point, we introduce and analyze an operator called "pseudo-L^2-projection", which is based on a global Petrov–Galerkin variational formulation with a discontinuous test space. This generalized projection operator turns out to be the closest to the actual L^2-orthogonal projection. We confirm this by a sophisticated experiment investigating the mutual relations of different prolongation operators.

Outline

This thesis is organized as follows. In Chapter 1, we briefly present the frame in which the various multilevel methods will be discussed. This includes linear and non-linear model problems associated with elliptic partial differential equations. A short description of finite element methods and a motivation for the search of efficient preconditioners for the discrete problems clarify basic notions, which are used throughout this text. Chapter 2 is devoted to standard geometric multilevel methods for finite element discretizations of elliptic equations. We explain the fundamental ideas and algorithmical structures; special attention is paid to the information transfer between nested finite element spaces. Moreover, we outline the framework of parallel and successive subspace correction methods. All this is meant to prepare the subsequent discussion of the generalized multilevel algorithms. In Chapter 3, we introduce the semi-geometric multilevel framework. The flexible approach based on a hierarchy of non-nested meshes is analyzed in full detail including the quasi-optimal convergence and preconditioning results. This chapter also contains the first part on the implementation. Here, essential aspects for the practical realization of the semi-geometric ideas are described. In Chapter 4, several geometry-based multilevel techniques are reviewed. We point out relevant connections to the semi-geometric approach and draw interesting comparisons. Chapter 5 contains the study of the actual information transfer between finite element spaces associated with non-nested meshes. Both qualitative and quantitative properties of a broad range of different transfer concepts are put forward. Moreover, we discuss implementation aspects for specific prolongation operators in practical finite element codes. In Chapter 6, we present numerical results demonstrating the performance of the devised methods. This chapter also includes the extension to a semi-geometric monotone multigrid method for variational inequalities. Chapter 7 serves as an excursus on a parametric multilevel approach with nested spaces, which is in fact a method with an adjusted discretization.

1 Derivation of the model problems

The issues elaborated in this thesis appear in the numerical simulation of phenomena described by elliptic partial differential equations. We discuss the topic of "multilevel methods based on non-nested meshes" in the frame defined by the series of model problems pictured in this introductory chapter. Both the analysis and practical considerations of the novel multilevel preconditioners are carried out in the present context.

As our own contributions concentrate on the efficient iterative solution of the equations rather than advanced modeling aspects, perhaps with the exception of some thoughts in Chapter 7 about parametric discretizations, we keep this part short. To achieve a sufficiently self-contained presentation, we state the considered model problems in Section 1.1 and Section 1.2. Scalar and vector-valued boundary value problems and variational equations are derived, complemented by non-linear obstacle and contact problems, respectively, associated with variational inequalities. As taking the possible time dependence of the modeled physical systems into account does not provide any additional insight for the topic of this thesis, we only consider stationary problems. Naturally, the developed methods may be employed to solve the elliptic systems, which are obtained in each time step by implicit time discretization schemes.

Section 1.3 is concerned with a finite element discretization of the variational problems. The introduced finite-dimensional approximation spaces in one form or another play a fundamental role in all parts of this thesis. We turn our attention to the corresponding operator and matrix notations in Section 1.4. Here, we also briefly motivate the research of efficient iterative solvers and preconditioners for the considered types of problems.

1.1 Elliptic partial differential equations

In this section, we briefly state basic models leading to boundary value problems of elliptic partial differential equations. Besides, a certain amount of notation is introduced to keep the subsequent sections and chapters more compact.

Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain of dimension $d \in \{2, 3\}$, i. e., an open, bounded, connected subset of the Euclidean space with Lipschitz boundary $\Gamma := \partial\Omega$. For a subset $U \subset \mathbb{R}^d$, we denote the closure and the interior of U with respect to the standard topology of $\mathbb{R}^d$ by $\overline{U}$ and $\mathrm{int}(U)$, respectively. We denote vector quantities, matrices and fourth order tensors by bold symbols, e. g., $\boldsymbol{v}$, $\boldsymbol{D}$, $\boldsymbol{C}$, and their components by v_i, D_{ij}, C_{ijlm} for suitable indices i, j, l, m. Analogously, we use bold symbols for vector-valued function spaces, i. e., $\boldsymbol{V} := (V)^d$. Given a real normed vector space V, we denote its dual space consisting of all bounded (or continuous) linear functionals by $V' = \mathrm{Lin}(V, \mathbb{R})$. For a function depending on the variable $\boldsymbol{x} \in \mathbb{R}^d$, the partial derivative with respect to x_j, $1 \leq j \leq d$, is abbreviated with ∂_j. Furthermore, we enforce the summation convention on all repeated indices ranging from 1 to d.

The standard basis of $\mathbb{R}^n$, $n \in \mathbb{N}$, will be referred to as $(\boldsymbol{e}_i)_{1 \leq i \leq n}$. The Euclidean inner product is $\boldsymbol{v} \cdot \boldsymbol{w} := \sum_i v_i w_i$ for $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ and the Euclidean norm of $\boldsymbol{v} \in \mathbb{R}^n$ is $|\boldsymbol{v}| := \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$. Then, the $(d-1)$-sphere is the set $\mathbb{S}^{d-1} := \{\boldsymbol{v} \in \mathbb{R}^d \mid |\boldsymbol{v}| = 1\}$. The symbol $\mathbb{R}^{m \times n}$ denotes

the set of $m \times n$ matrices with real entries. For a matrix $\boldsymbol{A}$, the symbol $\boldsymbol{A}^T$ denotes its transpose. If $m = n$, the index $_{\mathrm{sym}}$ selects the subset of symmetric matrices. We write $\boldsymbol{I}$ for the identity matrix and define the Kronecker delta for some indices i, j by $\delta_{ij} := 1$ if $i = j$ and $\delta_{ij} := 0$ if $i \neq j$.

We will make use of the relation symbols $\lesssim$, $\gtrsim$, and $\eqsim$. Here, $a \lesssim b$ and $f \gtrsim g$, respectively, mean that there are some constants c_1 and c_2, which are independent of the meshes and the considered functions, such that $a \leq c_1 b$ and $f \geq c_2 g$. If $a \lesssim b$ and $a \gtrsim b$, we write $a \eqsim b$.

1.1.1 Diffusion equation

The following boundary value problem plays a fundamental role in the mathematical description of diffusion processes. It commonly appears if the flux of some quantity is proportional to its gradient; see [80, 82].

The boundary Γ may consist of two disjoint, Lebesgue measurable parts, namely the closed Dirichlet boundary Γ_D and the open Neumann boundary Γ_N. Given a sufficiently smooth, symmetric diffusion tensor $\boldsymbol{D} : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and data $f : \Omega \to \mathbb{R}$, $p : \Gamma_N \to \mathbb{R}$, find a function $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\mathrm{div}(\boldsymbol{D}\nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_D, \\ \nabla u \cdot \boldsymbol{n} &= p \quad \text{on } \Gamma_N. \end{aligned} \tag{1.1}$$

Homogeneous media are associated with constant tensors $\boldsymbol{D}$. If $\boldsymbol{D}$ is a multiple of the identity matrix, the equation is called isotropic. Provided $\int_\Omega f \, d\boldsymbol{x} + \int_{\Gamma_N} p \, d\boldsymbol{s} = 0$, this equation may be considered the stationary limit of the heat equation derived as a consequence of Fourier's law of heat conduction and conservation of energy. In this case, u represents the temperature and the thermal conductivity $\boldsymbol{D}$ of the material in the domain Ω is assumed to be temperature-independent. The data f and p model external heat sources and the heat flux across the Neumann boundary Γ_N, respectively.

Similarly, the diffusion equation also models saturated steady-state flow of an incompressible fluid through a porous medium, such as groundwater flow. This relation is derived from Darcy's law, which states that the flux is proportional to the pressure gradient and is supposed to hold true for small Reynolds numbers, and conservation of mass; see [186]. Here, u describes the pressure of the fluid and $\boldsymbol{D} = \frac{1}{\mu}\boldsymbol{\kappa}$ with the symmetric permeability tensor $\boldsymbol{\kappa}$ and the dynamic viscosity $\mu > 0$ of the medium. However, boundary value problems of the form (1.1) arise in many other applications.

We recall several standard notations from functional analysis; see, e. g., [1, 75, 82]. Let $\int \cdot \, d\boldsymbol{x}$ be the Lebesgue integral; for a subset $U \subset \mathbb{R}^d$ we denote the d-dimensional Lebesgue measure by $\mathrm{meas}_d(U)$. Let $L^2(\Omega)$ be the Hilbert space of square integrable functions on Ω with inner product $(v, w)_{L^2(\Omega)} := \int_\Omega vw \, d\boldsymbol{x}$ and norm $\| \cdot \|_{L^2(\Omega)} := (\cdot, \cdot)^{1/2}_{L^2(\Omega)}$. The symbol $L^\infty(\Omega)$ represents the space of essentially bounded functions with norm $\|v\|_{L^\infty(\Omega)} := \mathrm{ess\,sup}_{\boldsymbol{x}\in\Omega} |v(\boldsymbol{x})|$. By $H^m(\Omega)$, as customary, we denote the Sobolev space of functions with $m \geq 0$ square integrable weak derivatives on Ω; in particular, $H^0(\Omega) = L^2(\Omega)$. Let $\alpha \in \mathbb{N}^d$ be a multi-index of order $|\alpha| := \sum_{1\leq i\leq d} \alpha_i$. Then, ∂^α denotes the weak differentiation and

the corresponding norm and semi-norm in $H^m(\Omega)$ are

$$\|v\|_{H^m(\Omega)} := \left(\sum_{|\alpha|\le m} \|\partial^\alpha v\|^2_{L^2(\Omega)} \right)^{\frac{1}{2}} \quad \text{and} \quad |v|_{H^m(\Omega)} := \left(\sum_{|\alpha|= m} \|\partial^\alpha v\|^2_{L^2(\Omega)} \right)^{\frac{1}{2}}.$$

Moreover, the subspace of $H^1(\Omega)$ with vanishing image of the usual trace operator to the Dirichlet boundary Γ_D is called $H^1_D(\Omega)$; we have $H^1_D(\Omega) = \{v \in H^1(\Omega) \mid v|_{\Gamma_D} = 0 \text{ in } H^{\frac{1}{2}}(\Gamma_D)\} = \{v \in H^1(\Omega) \mid v = 0 \text{ a. e. on } \Gamma_D\}$.

The Sobolev space $H^s(\Omega)$ of fractional order $s = m + \sigma$, $m \in \mathbb{N}$, $\sigma \in (0,1)$, is defined as the closure (or completion) of the space of infinitely differentiable functions with respect to the Slobodeckij norm

$$\|v\|_{H^s(\Omega)} := \left(\|v\|^2_{H^m(\Omega)} + \sum_{|\alpha|=m} \int_\Omega \int_\Omega \frac{|\partial^\alpha v(\boldsymbol{x}) - \partial^\alpha v(\boldsymbol{y})|^2}{|\boldsymbol{x}-\boldsymbol{y}|^{d+2\sigma}} \, d\boldsymbol{x}\, d\boldsymbol{y} \right)^{\frac{1}{2}}.$$

For a sufficiently smooth $(d-1)$-dimensional submanifold $S \subset \mathbb{R}^d$, usually a subset of Γ, the Lebesgue integral and the $(d-1)$-dimensional measure of a subset $U \subset S$ are $\int \cdot \, d\boldsymbol{s}$ and $\mathrm{meas}_{d-1}(U)$, respectively. In particular, we make use of the analogues of the above defined function spaces on the considered submanifolds.

The variational formulation of problem (1.1) is: Find $u \in H^1_D(\Omega)$ such that

$$a(u,v) = \mathcal{F}(v), \quad \forall\, v \in H^1_D(\Omega), \tag{1.2}$$

with the continuous and symmetric bilinear form

$$a : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}, \quad a(v,w) := (\boldsymbol{D}\nabla v, \nabla w)_{L^2(\Omega)} = (D_{ij}\partial_i v, \partial_j w)_{L^2(\Omega)}.$$

We assume that the right hand side of equation (1.2) is represented by a bounded linear functional, namely $\mathcal{F} \in H^1_D(\Omega)'$. If the data is sufficiently smooth, we have, e. g., $\mathcal{F}(v) = (f,v)_{L^2(\Omega)} + (p,v)_{L^2(\Gamma_N)}$.

Let $D_{ij} \in L^\infty(\Omega)$, $1 \le i,j \le d$, and $\boldsymbol{D}$ be uniformly elliptic, i. e., there is a constant $\alpha > 0$ such that

$$D_{ij}(\boldsymbol{x})\xi_i\xi_j \ge \alpha|\boldsymbol{\xi}|^2, \quad \text{for a. e. } \boldsymbol{x} \in \Omega,\ \forall\, \boldsymbol{\xi} \in \mathbb{R}^d.$$

Then, the bilinear form a is elliptic if $\mathrm{meas}_{d-1}(\Gamma_D) > 0$. By the Lax–Milgram theorem, problem (1.2) has a unique solution.

1.1.2 Linear elasticity

The second problem class to be considered comprises continuum mechanical models of elastostatics. A deformation of a solid body can be described by kinematic quantities, equilibrium conditions and a constitutive equation. The displacement field induced by applied forces is the solution of an elliptic boundary value problem. We assume small displacements and obtain this system of partial differential equations by linearizing both the strain–displacement relations and the stress–strain relations. Our presentation follows mainly [57]; see also the monographs [80, 135, 192]. We only deal with the full-dimensional case $d = 3$.

Kinematics

In this paragraph, we consider measures of the strains caused by a deformation of a solid. The deformable body is represented by the domain $\Omega \subset \mathbb{R}^3$. We call the closure $\overline{\Omega}$ with respect to the standard topology of $\mathbb{R}^3$ the reference configuration. Then, a deformation of $\overline{\Omega}$ is an orientation-preserving, sufficiently smooth mapping

$$\boldsymbol{\varphi} : \overline{\Omega} \to \mathbb{R}^3$$

which is injective up to the boundary. With the notation $\boldsymbol{\varphi} = \varphi_i \boldsymbol{e}_i$ the deformation gradient is the matrix

$$\boldsymbol{\nabla\varphi} := (\partial_j \varphi_i)_{1\le i,j\le 3}\,.$$

As the deformation is assumed to preserve the orientation, the determinant of the deformation gradient is positive, i. e.,

$$\boldsymbol{\nabla\varphi}(\boldsymbol{x}) \in \mathbb{R}^{3\times 3}_{+} := \{\boldsymbol{M} \in \mathbb{R}^{3\times 3} \mid \det \boldsymbol{M} > 0\}, \quad \forall\, \boldsymbol{x} \in \overline{\Omega}.$$

We call the mapping $\boldsymbol{u} : \overline{\Omega} \to \mathbb{R}^3$ defined by $\boldsymbol{\varphi} = \mathbf{id} + \boldsymbol{u}$ displacement field and its derivative $\boldsymbol{\nabla u} = \boldsymbol{\nabla\varphi} - \boldsymbol{I}$ displacement gradient. The set $\overline{\Omega}^{\varphi} := \boldsymbol{\varphi}(\overline{\Omega})$ is the deformed configuration with points $\boldsymbol{x}^{\varphi} := \boldsymbol{\varphi}(\boldsymbol{x})$.

A comparison of the infinitesimal length elements in the reference and the deformed configuration motivates the definition of the (right) Cauchy–Green strain tensor $\boldsymbol{\nabla\varphi}^T \boldsymbol{\nabla\varphi}$. This tensor is in fact an adequate indicator of the strain emerging inside the deformed body because a deformation is a rigid-body motion, i. e., it is of the form

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{a} + \boldsymbol{Q}\boldsymbol{x}, \quad \boldsymbol{a} \in \mathbb{R}^3, \quad \boldsymbol{Q} \in \mathbb{O}^3_{+}, \quad \forall\, \boldsymbol{x} \in \overline{\Omega},$$

if and only if $\boldsymbol{\nabla\varphi}(\boldsymbol{x})^T \boldsymbol{\nabla\varphi}(\boldsymbol{x}) = \boldsymbol{I}$, $\forall\, \boldsymbol{x} \in \Omega$; see [57, Theorem 1.8-1]. Here, $\mathbb{O}^3_{+}$ denotes the set of orthogonal matrices with determinant $+1$.

To quantify the local deviation of the considered deformation $\boldsymbol{\varphi}$ and a rigid-body motion, one usually introduces the Green–St. Venant strain tensor $\boldsymbol{E}(\boldsymbol{\varphi}) := \frac{1}{2}(\boldsymbol{\nabla\varphi}^T \boldsymbol{\nabla\varphi} - \boldsymbol{I})$. Therefore, one can express the strains in terms of the displacement gradient,

$$\boldsymbol{E}(\boldsymbol{u}) = \frac{1}{2}\left(\boldsymbol{\nabla u}^T + \boldsymbol{\nabla u} + \boldsymbol{\nabla u}^T \boldsymbol{\nabla u}\right). \tag{1.3}$$

In the linear elastic setting, we assume the displacements to be small and neglect the quadratic term in (1.3). This gives rise to the linearized strain tensor

$$\boldsymbol{\epsilon}(\boldsymbol{u}) = \frac{1}{2}\left(\boldsymbol{\nabla u}^T + \boldsymbol{\nabla u}\right). \tag{1.4}$$

Equilibrium conditions

The static equilibrium of the solid body is characterized by a system of partial differential equations. We assume that the deformed body is subjected to volume forces and surface tractions given by the force densities $\boldsymbol{f}^{\varphi} : \Omega^{\varphi} \to \mathbb{R}^3$ and $\boldsymbol{p}^{\varphi} : \Gamma^{\varphi}_N \to \mathbb{R}^3$, respectively. Here,

$\Gamma_N^{\boldsymbol{\varphi}}$ is an open, measurable subset of $\Gamma^{\boldsymbol{\varphi}} := \partial\Omega^{\boldsymbol{\varphi}}$. The corresponding force densities in the reference configuration are, see [57, Section 2.6],

$$\boldsymbol{f} : \Omega \to \mathbb{R}^3, \quad \boldsymbol{f}(\boldsymbol{x}) = (\det \boldsymbol{\nabla}\boldsymbol{\varphi}(\boldsymbol{x}))\boldsymbol{f}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}), \quad \forall\, \boldsymbol{x}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi}(\boldsymbol{x}),\ \boldsymbol{x} \in \Omega, \tag{1.5}$$

and

$$\boldsymbol{p} : \Gamma_N \to \mathbb{R}^3, \quad \boldsymbol{p}(\boldsymbol{x}) = (\det \boldsymbol{\nabla}\boldsymbol{\varphi}(\boldsymbol{x}))|\boldsymbol{\nabla}\boldsymbol{\varphi}(\boldsymbol{x})^{-T}\boldsymbol{n}|\boldsymbol{p}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}), \quad \forall\, \boldsymbol{x}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi}(\boldsymbol{x}),\ \boldsymbol{x} \in \Gamma_N, \tag{1.6}$$

with $\boldsymbol{n}$ the unit outer normal of $\Gamma_N := \boldsymbol{\varphi}^{-1}(\Gamma_N^{\boldsymbol{\varphi}})$.

The stress principle of Euler and Cauchy is the foundation of continuum mechanics for time-independent problems. If the deformable body is in static equilibrium, this axiom asserts the existence of a vector field

$$\boldsymbol{t}^{\boldsymbol{\varphi}} : \overline{\Omega}^{\boldsymbol{\varphi}} \times \mathbb{S}^2 \to \mathbb{R}^3$$

which represents the elementary surface forces inside the deformed configuration. More precisely, the Cauchy stress vector $\boldsymbol{t}^{\boldsymbol{\varphi}}$ for the oriented surface element with normal $\boldsymbol{n}^{\boldsymbol{\varphi}}$ coincides with the given surface tractions, i. e., $\boldsymbol{t}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}, \boldsymbol{n}^{\boldsymbol{\varphi}}) = \boldsymbol{p}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}})$ for $\boldsymbol{x}^{\boldsymbol{\varphi}} \in \Gamma_N^{\boldsymbol{\varphi}}$, and satisfies an axiom of force balance and an axiom of moment balance for any subdomain of $\overline{\Omega}^{\boldsymbol{\varphi}}$; see [57, Axiom 2.2-1].

Now, the following theorem yields a linear dependence of the Cauchy stress vector on $\boldsymbol{n}^{\boldsymbol{\varphi}}$ by the Cauchy stress tensor $\boldsymbol{T}^{\boldsymbol{\varphi}}$ which may be characterized by a system of partial differential equations.

Theorem 1.1 (Cauchy's theorem)**.** *Assume that the applied body force density $\boldsymbol{f}^{\boldsymbol{\varphi}} : \overline{\Omega}^{\boldsymbol{\varphi}} \to \mathbb{R}^3$ is continuous and that the Cauchy stress vector field $\boldsymbol{t}^{\boldsymbol{\varphi}} : \overline{\Omega}^{\boldsymbol{\varphi}} \times \mathbb{S}^2 \to \mathbb{R}^3$, $(\boldsymbol{x}^{\boldsymbol{\varphi}}, \boldsymbol{n}) \mapsto \boldsymbol{t}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}, \boldsymbol{n})$ is continuously differentiable with respect to the variable $\boldsymbol{x}^{\boldsymbol{\varphi}}$ for each $\boldsymbol{n} \in \mathbb{S}^2$ and continuous with respect to the variable $\boldsymbol{n}$ for each $\boldsymbol{x}^{\boldsymbol{\varphi}} \in \overline{\Omega}^{\boldsymbol{\varphi}}$. Then, there exists a continuously differentiable, symmetric tensor field*

$$\boldsymbol{T}^{\boldsymbol{\varphi}} : \overline{\Omega}^{\boldsymbol{\varphi}} \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}, \quad \boldsymbol{x}^{\boldsymbol{\varphi}} \mapsto \boldsymbol{T}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}),$$

such that the Cauchy stress vector satisfies

$$\boldsymbol{t}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}, \boldsymbol{n}) = \boldsymbol{T}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}})\boldsymbol{n}, \quad \forall\, \boldsymbol{x}^{\boldsymbol{\varphi}} \in \overline{\Omega}^{\boldsymbol{\varphi}},\ \boldsymbol{n} \in \mathbb{S}^2,$$

and such that

$$\begin{array}{rcll} -\mathbf{div}^{\boldsymbol{\varphi}}\boldsymbol{T}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}) & = & \boldsymbol{f}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}), & \forall\, \boldsymbol{x}^{\boldsymbol{\varphi}} \in \Omega^{\boldsymbol{\varphi}}, \\ \boldsymbol{T}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}})\boldsymbol{n}^{\boldsymbol{\varphi}} & = & \boldsymbol{p}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}}), & \forall\, \boldsymbol{x}^{\boldsymbol{\varphi}} \in \Gamma_N^{\boldsymbol{\varphi}}, \end{array} \tag{1.7}$$

where $\boldsymbol{n}^{\boldsymbol{\varphi}}$ is the unit outer normal vector along $\Gamma_N^{\boldsymbol{\varphi}}$ and $\mathbf{div}^{\boldsymbol{\varphi}}\boldsymbol{T}^{\boldsymbol{\varphi}} := \partial_j^{\boldsymbol{\varphi}} T_{ij}^{\boldsymbol{\varphi}} \boldsymbol{e}_i$ denotes the divergence of the tensor field $\boldsymbol{T}^{\boldsymbol{\varphi}}$ with respect to the variable $\boldsymbol{x}^{\boldsymbol{\varphi}}$.

Proof. See [57, Theorem 2.3-1]. □

To transform the equations (1.7) to the reference configuration, we define the (second) Piola–Kirchhoff stress tensor as

$$\boldsymbol{\Sigma}(\boldsymbol{x}) := (\det \boldsymbol{\nabla}\boldsymbol{\varphi}(\boldsymbol{x}))\boldsymbol{\nabla}\boldsymbol{\varphi}(\boldsymbol{x})^{-1}\boldsymbol{T}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}})\boldsymbol{\nabla}\boldsymbol{\varphi}(\boldsymbol{x})^{-T}, \quad \boldsymbol{x}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi}(\boldsymbol{x}).$$

Finally, the properties of the Piola transform [57, Theorem 1.7-1] and (1.5), (1.6) yield the

Theorem 1.2 (Equilibrium conditions in the reference configuration)**.** *The Piola–Kirchhoff stress tensor is symmetric and satisfies*

$$\begin{aligned} -\mathbf{div}(\boldsymbol{\nabla\varphi}(\boldsymbol{x})\boldsymbol{\Sigma}(\boldsymbol{x})) &= \boldsymbol{f}(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \Omega, \\ \boldsymbol{\nabla\varphi}(\boldsymbol{x})\boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{n} &= \boldsymbol{p}(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \Gamma_N. \end{aligned}$$

Proof. See [57, Theorem 2.6-2]. □

In the continuum mechanical framework, these are the fundamental equations to characterize a deformation induced by applied forces.

Constitutive equations

We assume that the material is homogeneous and elastic, i. e., there is a response function $\hat{\boldsymbol{\Sigma}} : \mathbb{R}^{3\times3}_+ \to \mathbb{R}^{3\times3}_{\mathrm{sym}}$ for the Piola–Kirchhoff stress tensor such that

$$\boldsymbol{\Sigma}(\boldsymbol{x}) = \hat{\boldsymbol{\Sigma}}(\boldsymbol{\nabla\varphi}(\boldsymbol{x})), \quad \forall\, \boldsymbol{x} \in \overline{\Omega}.$$

In addition, let the material be isotropic, namely the response function may satisfy

$$\hat{\boldsymbol{\Sigma}}(\boldsymbol{FQ}) = \boldsymbol{Q}^T\hat{\boldsymbol{\Sigma}}(\boldsymbol{F})\boldsymbol{Q}, \quad \forall\, \boldsymbol{F} \in \mathbb{R}^{3\times3}_+,\ \boldsymbol{Q} \in \mathbb{O}^3_+.$$

Then, a fundamental result of the mathematical theory of elasticity [57, Theorem 3.8-1] asserts the existence of constants $\lambda > 0$ and $\mu > 0$ such that the material law near a reference configuration with $\hat{\boldsymbol{\Sigma}}(\boldsymbol{I}) = \boldsymbol{0}$ is of the form

$$\hat{\boldsymbol{\Sigma}}(\boldsymbol{F}) = \lambda(\mathrm{tr}\boldsymbol{E})\boldsymbol{I} + 2\mu\boldsymbol{E} + \mathcal{O}(\boldsymbol{E}), \quad \forall\, \boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I}),\ \boldsymbol{F} \in \mathbb{R}^{3\times3}_+, \tag{1.8}$$

where $\mathrm{tr}\boldsymbol{E} = E_{ii}$ is the trace of the matrix $\boldsymbol{E}$. The numbers λ and μ are called Lamé constants of the material. For small deformations and small strains, we neglect terms of higher order in the deformation gradient $\boldsymbol{\nabla u}$ as well as the strain tensor $\boldsymbol{E}$ in (1.8) and obtain a linear material law for the stress tensor, which is denoted by $\boldsymbol{\sigma}$ in this case,

$$\sigma_{ij}(\boldsymbol{u}) = C_{ijlm}\epsilon_{lm}(\boldsymbol{u}) = \lambda\delta_{ij}\epsilon_{kk}(\boldsymbol{u}) + 2\mu\epsilon_{ij}(\boldsymbol{u}). \tag{1.9}$$

Here, $\boldsymbol{\epsilon}$ is the linearization from (1.4) and $\boldsymbol{C} := (C_{ijlm})_{1\le i,j,l,m\le3}$ is Hooke's tensor with the symmetries $C_{ijlm} = C_{lmij} = C_{jilm}$.

Finally, assuming that the force densities $\boldsymbol{f}$ and $\boldsymbol{p}$ are given in the reference configuration and neglecting higher order terms in $\boldsymbol{\nabla u}$ again, we obtain the following boundary value problem with Dirichlet conditions on $\Gamma_D = \Gamma \setminus \Gamma_N$ from Theorem 1.2. Find a displacement field $\boldsymbol{u}$ such that

$$\begin{aligned} -\partial_j(\sigma_{ij}(\boldsymbol{u})) &= f_i \quad \text{in } \Omega, \\ u_i &= 0 \quad \text{on } \Gamma_D, \\ \sigma_{ij}(\boldsymbol{u})n_j &= p_i \quad \text{on } \Gamma_N, \end{aligned} \tag{1.10}$$

for $1 \le i \le 3$; see [57, Section 6.2].

Variational formulation

Wherever appropriate, we use the same symbols as for the scalar diffusion equation in Section 1.1.1. Let the vector-valued function spaces $\boldsymbol{L}^2(\Omega) := \big(L^2(\Omega)\big)^3$ and $\boldsymbol{H}^m(\Omega) := (H^m(\Omega))^3$ be equipped with the respective product topology; the (semi-)norms of these Hilbert spaces are still denoted by $\|\cdot\|_{L^2(\Omega)}$, $\|\cdot\|_{H^m(\Omega)}$ and $|\cdot|_{H^m(\Omega)}$ to avoid bold indices. Moreover, the Dirichlet conditions on the part Γ_D lead to $\boldsymbol{H}^1_D(\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \mid \boldsymbol{v} = \boldsymbol{0}, \text{ a. e. on } \Gamma_D\}$.

We introduce the continuous and symmetric bilinear form

$$a : \boldsymbol{H}^1(\Omega) \times \boldsymbol{H}^1(\Omega) \to \mathbb{R}, \quad a(\boldsymbol{v}, \boldsymbol{w}) := (\sigma_{ij}(\boldsymbol{v}), \partial_j w_i)_{L^2(\Omega)} = (C_{ijlm}\epsilon_{lm}(\boldsymbol{v}), \epsilon_{ij}(\boldsymbol{w}))_{L^2(\Omega)}$$

and the linear form

$$\mathcal{F} : \boldsymbol{H}^1(\Omega) \to \mathbb{R}, \quad \mathcal{F}(\boldsymbol{v}) := (f_i, v_i)_{L^2(\Omega)} + (p_i, v_i)_{L^2(\Gamma_N)}.$$

Then, the weak formulation of problem (1.10) reads as: Find $\boldsymbol{u} \in \boldsymbol{H}^1_D(\Omega)$ satisfying the variational equation

$$a(\boldsymbol{u}, \boldsymbol{v}) = \mathcal{F}(\boldsymbol{v}), \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}^1_D(\Omega). \tag{1.11}$$

As the tensor $\boldsymbol{C}$ from (1.9) is elliptic, i. e., there exists a constant $\alpha > 0$ such that

$$C_{ijlm}\epsilon_{lm}\epsilon_{ij} \geq \alpha \epsilon_{ij}\epsilon_{ij}, \quad \forall\, \boldsymbol{\epsilon} \in \mathbb{R}^{3\times 3}_{\text{sym}},$$

the (uniform) ellipticity of the form a is a consequence of the following

Theorem 1.3 (Korn's inequality)**.** *Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain and may* $\text{meas}_2(\Gamma_D) > 0$*. Then, there is a constant $c > 0$ such that*

$$\int_\Omega \epsilon_{ij}(\boldsymbol{v})\epsilon_{ij}(\boldsymbol{v})\, d\boldsymbol{x} \geq c\|\boldsymbol{v}\|^2_{H^1(\Omega)}, \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}^1_D(\Omega).$$

Proof. See [117, Lemma 6.2]. □

Therefore, if $\text{meas}_2(\Gamma_D) > 0$, problem (1.11) has a unique solution in $\boldsymbol{H}^1_D(\Omega)$ by the Lax–Milgram theorem.

1.2 Variational inequalities

In this section, we outline two types of non-linear problems which can be formulated as variational inequalities. Both problem classes involve one-sided, pointwise constraints either in the domain Ω or on a subset of the boundary Γ which the respective solution has to satisfy. This gives rise to a free boundary value problem; the active set where the constraints are binding has to be determined as part of the solution process.

We consider a scalar obstacle problem and Signorini's problem, a linearized model of elastic contact with a rigid and fixed foundation, and briefly summarize results from calculus of variations about existence and regularity of weak solutions.

1.2.1 Scalar obstacle problems

To obtain a classical obstacle problem, note that the solution of problem (1.2) can be characterized as the (unique) minimum of the energy functional $\mathcal{J} : H^1(\Omega) \to \mathbb{R}$, $\mathcal{J}(v) := \frac{1}{2}a(v,v) - \mathcal{F}(v)$ in $H^1_D(\Omega)$. Next, let the solution be constrained from below by a sufficiently smooth obstacle function $g : \Omega \to \mathbb{R}$, $g \leq 0$ almost everywhere in Ω. In this case, the problem reads as

$$\mathcal{J}(u) = \min_{v \in \mathcal{K}} \mathcal{J}(v), \quad \mathcal{K} := \left\{ v \in H^1_D(\Omega) \mid v \geq g, \text{ a. e. in } \Omega \right\}. \tag{1.12}$$

In fact, the set of admissible functions $\mathcal{K} \subset H^1_D(\Omega)$ is closed and convex.

For $d = 2$, $\Gamma_D = \Gamma$ and $\boldsymbol{D} = \alpha \boldsymbol{I}$ isotropic with $\alpha > 0$, a solution of the above problem is usually interpreted as the deformation of an elastic membrane subjected to a vertical force (with density f) and constrained by an obstacle. In this case, the functional $\frac{1}{2}|v|_{H^1(\Omega)}$ commonly called Dirichlet energy represents, up to higher order terms, the increase of the surface area which is assumed to be proportional to the tension in the membrane.

However, problems of this kind also appear in many other applications; see [119, Chapter VII] and [80, 121]. In some cases, a transformation may be necessary to obtain an auxiliary problem whose solution is characterized by a variational inequality.

1.2.2 Elastic contact problems

For the description of Signorini's problem, we subdivide the boundary Γ into three pairwise disjoint, measurable parts: Γ_D closed, Γ_N open and the possible contact boundary Γ_C closed. Additionally, the condition $\Gamma_C \cap \Gamma_D = \emptyset$ may hold, which appears to be quite natural in most applications.

Let the (non-linear) geometric contact condition, i. e., non-penetration of the deformed solid Ω^φ and the rigid obstacle, be approximated by the following linearization. Assume that an initial gap function $g : \Gamma_C \to \mathbb{R}_+$ modeling the distance in outer normal direction $\boldsymbol{n}$ between the reference configuration and the obstacle is given. Then, the closed and convex set of admissible displacements is

$$\boldsymbol{\mathcal{K}} := \left\{ \boldsymbol{v} \in \boldsymbol{H}^1_D(\Omega) \mid \boldsymbol{v} \cdot \boldsymbol{n} \leq g, \text{ a. e. on } \Gamma_C \right\}. \tag{1.13}$$

Therefore, in the linearized setting with small deformations, a solution of the contact problem has to satisfy linear inequality constraints pointwisely almost everywhere on Γ_C.

Like in the scalar case, the solution of the vector-valued Signorini problem is a minimizer of an energy functional, here of the elastic energy $\mathcal{J}$, in the set $\boldsymbol{\mathcal{K}}$; we have

$$\mathcal{J}(\boldsymbol{u}) = \min_{\boldsymbol{v} \in \boldsymbol{\mathcal{K}}} \mathcal{J}(\boldsymbol{v}), \quad \mathcal{J} : \boldsymbol{H}^1(\Omega) \to \mathbb{R}, \quad \mathcal{J}(\boldsymbol{v}) := \frac{1}{2}a(\boldsymbol{v},\boldsymbol{v}) - \mathcal{F}(\boldsymbol{v}). \tag{1.14}$$

For more about the modeling of elastic contact problems, we refer to the monographs [117, 135, 192] as well as our own paper [71] and the references therein.

1.2.3 Existence of weak solutions and regularity

The solution of the constrained minimization problem (1.14) is also characterized by the variational inequality: Find $\boldsymbol{u} \in \boldsymbol{\mathcal{K}}$ such that

$$a(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) \geq \mathcal{F}(\boldsymbol{v} - \boldsymbol{u}), \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{K}}. \tag{1.15}$$

The respective characterization for problem (1.12) holds with scalar u, $v \in \mathcal{K}$; see, e. g., [80, 82, 117].

We may cite the Lions–Stampacchia theorem verifying the well-posedness of the problems (1.12) and (1.14).

Theorem 1.4 (Lions, Stampacchia)**.** *Let $\boldsymbol{X}$ be a Hilbert space and $\boldsymbol{\mathcal{K}} \subset \boldsymbol{X}$ non-empty, closed and convex. If the bilinear form $a : \boldsymbol{X} \times \boldsymbol{X} \to \mathbb{R}$ is continuous and elliptic and the linear form $\mathcal{F} : \boldsymbol{X} \to \mathbb{R}$ is continuous, then the variational inequality (1.15) has a unique solution which depends continuously on $\mathcal{F}$.*

Proof. See [138] or [119, Theorem II.2.1]. □

For sufficiently smooth data and domain boundary, an interior regularity theorem holds for the linear problems; see [82, Section 6.3] and [57, Section 6.3]. More precisely, the weak solutions are locally H^2-regular, i. e., for all compactly embedded subsets $U \subset\subset \Omega$, the restriction of the solution u of (1.2) or $\boldsymbol{u}$ of (1.11) to U is in $H^2(U)$ or $\boldsymbol{H}^2(U)$, respectively. The same results can be proved for the solutions of the variational inequalities under additional assumptions on the regularity of the obstacle or gap function; see [41, 119] and [117, Section 6.4] and the references therein. Note that, due to potential lack of smoothness at the change from Dirichlet to Neumann conditions, $\Gamma_D \cap \overline{\Gamma}_N$, one can usually not expect the solution to be as regular close to the boundary. If Ω is a polygonal or polyhedral domain, Γ is only piecewise smooth and full regularity is also lost at corner points or edges.

1.3 Finite element approximation

This section is devoted to a brief description of finite element discretizations of the presented variational problems. In particular, several important notations used throughout this thesis are introduced. For a finite-dimensional approximation of the solution, one employs continuous, piecewise polynomial functions with respect to an unstructured mesh. Then, the best approximation with respect to the energy norm is obtained by an a-orthogonal projection. The discretization error is essentially analyzed by Céa's lemma and an interpolation error estimate. We refer to the monographs [25, 39, 56].

A non-overlapping decomposition $\mathcal{T}_\ell$ of Ω into finitely many open polytopes (triangles or quadrilaterals for $d = 2$ and tetrahedra, pyramids, prisms or hexahedra for $d = 3$) is called mesh if the intersection $\overline{T}_1 \cap \overline{T}_2$ is a common vertex, a common edge or, if $d = 3$, a common face for different $T_1, T_2 \in \mathcal{T}_\ell$. Let $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}}$ be a family of shape regular meshes of Ω, i. e., there is a constant c such that

$$\sup_{\ell \in \mathbb{N}} \max_{T \in \mathcal{T}_\ell} \frac{h_T}{r_T} \leq c. \tag{1.16}$$

Here, for an element $T \in \mathcal{T}_\ell$, let $h_T := \operatorname{diam}(T)$ be the diameter of T; besides, r_T denotes the radius of the largest ball inscribed in T. We say $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ is quasi-uniform if, in addition, there is another constant c, independent of ℓ, such that

$$\max_{T\in\mathcal{T}_\ell} h_T \leq c \min_{T\in\mathcal{T}_\ell} h_T, \quad \forall\, \ell \in \mathbb{N}. \tag{1.17}$$

Only in this case it is reasonable to speak about a properly specified mesh parameter $h_\ell > 0$ associated with $\mathcal{T}_\ell$. However, we define the global quantity $h_\ell := \max_{T\in\mathcal{T}_\ell} h_T$ for the moment. During our presentation we will repeatedly return to the question whether quasi-uniformity is needful for the particular point.

Remark 1.5. *We assume that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_\ell} \overline{T}$, $\ell \in \mathbb{N}$. In general, the meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ do not need to represent the domain Ω exactly. They should constitute a sequence $(\Omega_\ell)_{\ell\in\mathbb{N}}$ of polygonal or polyhedral approximations, though. Then, the estimation of the additional consistency error requires sufficiently fast convergence of this sequence to Ω in a certain sense.*

In the present section, each mesh $\mathcal{T}_\ell$, $\ell \in \mathbb{N}$, is used to construct a finite element discretization. More precisely, the sequence of the respective Galerkin projections is supposed to converge to the solution of the continuous problem as described below. Let us, as early as now, clarify that a hierarchy of meshes may also serve another purpose. In the subsequent sections, when we consider multilevel preconditioners for discrete elliptic equations, a "fine" level $L \in \mathbb{N}$ will be specified. Then, the "coarse" meshes $(\mathcal{T}_\ell)_{\ell<L}$ and the associated finite element spaces are used as a means to construct auxiliary problems for the fast iterative solution of the discrete problem at level L. Consequently, the coarse finite element spaces do not need to be proper approximation spaces in the strict sense of the present section. We will distinguish between these two notions carefully.

We denote the set of nodes of $\mathcal{T}_\ell$ which do not lie on Γ_D by $\mathcal{N}_\ell$ and abbreviate its cardinality by $n_\ell := |\mathcal{N}_\ell|$. Let $X_\ell \subset H^1_D(\Omega)$ be the space of Lagrange conforming finite elements of first order with incorporated Dirichlet boundary conditions on Γ_D and denote its nodal basis as $\Lambda_\ell = (\lambda_p^\ell)_{p\in\mathcal{N}_\ell}$ with $\lambda_p^\ell(q) = \delta_{pq}$, $p, q \in \mathcal{N}_\ell$. Note that, as usual, the symbol p is used synonymously to indicate a node in $\mathcal{N}_\ell$ and its position in $\mathbb{R}^d$. Let $\omega_p := \operatorname{supp}(\lambda_p^\ell)$ be the support of the basis function at node $p \in \mathcal{N}_\ell$ commonly called patch. Here, the index ℓ is dropped as it is clear from the choice of p. To approximate the vector-valued space $\boldsymbol{H}^1_D(\Omega)$, we use the finite element space $\boldsymbol{X}_\ell := (X_\ell)^3$ with basis $\boldsymbol{\Lambda}_\ell = (\lambda_p^\ell \boldsymbol{e}_i)_{p\in\mathcal{N}_\ell, 1\leq i\leq 3}$.

The Galerkin discretizations of the linear problems (1.2) and (1.11) are

$$u_\ell \in X_\ell : \;\; a(u_\ell, v) = \mathcal{F}(v), \;\; \forall\, v \in X_\ell \quad \text{and} \quad \boldsymbol{u}_\ell \in \boldsymbol{X}_\ell : \;\; a(\boldsymbol{u}_\ell, \boldsymbol{v}) = \mathcal{F}(\boldsymbol{v}), \;\; \forall\, \boldsymbol{v} \in \boldsymbol{X}_\ell. \tag{1.18}$$

For the finite element discretizations of the variational inequalities, we assume that g and $\boldsymbol{n}$ are sufficiently smooth. Then, the discrete approximations of the constrained minimization problems (1.12) and (1.14) are to find $u_\ell \in \mathcal{K}_\ell$ such that

$$a(u_\ell, v - u_\ell) \geq \mathcal{F}(v - u_\ell), \quad \forall\, v \in \mathcal{K}_\ell := \{v \in X_\ell \mid v(p) \geq g(p), \; \forall\, p \in \mathcal{N}_\ell\},$$

and to find $\boldsymbol{u}_\ell \in \boldsymbol{\mathcal{K}}_\ell$ such that

$$a(\boldsymbol{u}_\ell, \boldsymbol{v} - \boldsymbol{u}_\ell) \geq \mathcal{F}(\boldsymbol{v} - \boldsymbol{u}_\ell), \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{K}}_\ell := \{\boldsymbol{v} \in \boldsymbol{X}_\ell \mid \boldsymbol{v}(p) \cdot \boldsymbol{n}(p) \leq g(p), \; \forall\, p \in \mathcal{N}_\ell \cap \Gamma_C\}, \tag{1.19}$$

respectively. The unique solvability of the discrete equations and inequalities follows just as in the continuous case. Here, as usual, the non-penetration constraints are only enforced at the respective nodes. We briefly return to the issue of modeling discrete inequality constraints in Remark 7.1 in the context of a parametric finite element method.

For the standard discretization error estimates for the linear problems we refer to [25, 39, 56]. If the continuous solution is sufficiently smooth, i. e., $u \in H^{1+\sigma}(\Omega)$ or $\boldsymbol{u} \in \boldsymbol{H}^{1+\sigma}(\Omega)$ for a constant $\sigma \in (0, 1]$, then $\|u - u_\ell\|_{H^1(\Omega)} \lesssim h_\ell^\sigma |u|_{H^{1+\sigma}(\Omega)}$ or $\|\boldsymbol{u} - \boldsymbol{u}_\ell\|_{H^1(\Omega)} \lesssim h_\ell^\sigma |\boldsymbol{u}|_{H^{1+\sigma}(\Omega)}$, respectively. The analogous results hold true for the approximations of the variational inequalities; see [41, 84, 117].

In any case, the boundary Γ needs to be sufficiently smooth. If the mesh $\mathcal{T}_\ell$ does not represent the domain Ω exactly, as indicated in Remark 1.5, one may have to replace the Galerkin projection u_ℓ (or $\boldsymbol{u}_\ell$) by a suitable extension; see, e. g., [117, Section 6.4].

A previous research project of the author was about the numerical simulation of multi-body contact problems on complex three-dimensional geometries. In the case of warped contact boundaries and non-matching finite element meshes, particular emphasis has to be put on the discretization of the transmission of forces and the non-penetration conditions at the contact interfaces between the involved elastic bodies; see [69, 70, 71, 72].

1.4 The need for preconditioning

In this section, a motivation for the search of efficient preconditioners is given. Here, in this introductory chapter, we only explain the concepts of iterative solution and preconditioning for linear problems. Non-linear methods for the discrete variational inequalities are discussed in more detail in Section 6.2. We will also need to establish some more notation. But note that most terminologies are well-known and may be found in many books on numerical mathematics.

From now on, let us consider the scalar case only. As the elaboration of the current section for the vector-valued spaces is straightforward, we do not lose any generality. However, this will appreciably simplify the notation.

As a finite element function $v_\ell \in X_\ell$ has a unique representation

$$v_\ell = \sum_{p \in \mathcal{N}_\ell} v_\ell(p) \lambda_p^\ell,$$

it can be identified with the vector $\boldsymbol{v}_\ell = (v_\ell(p))_{p \in \mathcal{N}_\ell} \in \mathbb{R}^{n_\ell}$. This identification is formalized by the coordinate isomorphism Φ_ℓ given by

$$\Phi_\ell : \mathbb{R}^{n_\ell} \to X_\ell, \quad \Phi_\ell(\boldsymbol{v}) := \sum_{p \in \mathcal{N}_\ell} v_p \lambda_p^\ell. \tag{1.20}$$

Exactly at this point, the "restriction" to the scalar problems pays. This is because, given a basis Λ_ℓ, we may now use bold symbols for vector and matrix representations of finite element functions and operators, respectively; see also below.

Operator and matrix notations

Next, we introduce an operator notation. Let V and W be normed vector spaces with norm $\|\cdot\|_V$ and $\|\cdot\|_W$, respectively. Recall the definition $V' := \mathrm{Lin}(V, \mathbb{R})$ of the dual of a normed vector space V. Generally, $\mathrm{Lin}(V, W) := \{A : V \to W \mid \sup_{\|v\|_V \leq 1} \|Av\|_W < \infty\}$ denotes the space of bounded (or continuous) linear operators from V to W and $\mathrm{Lin}(V) := \mathrm{Lin}(V, V)$. By $\langle \cdot, \cdot \rangle_V$ we denote the duality pairing in V, i. e., $\langle v', w \rangle_V := v'(w)$ for $v' \in V'$ and $w \in V$. For an operator $A \in \mathrm{Lin}(V, W)$ we define the adjoint operator $A' \in \mathrm{Lin}(W', V')$ by

$$\langle A'w', v \rangle_V = \langle w', Av \rangle_W, \quad \forall\, v \in V,\, w' \in W'.$$

If $V = W$ is a Hilbert space with inner product $(\cdot, \cdot)_V$, one may use the Riesz representation theorem to characterize the adjoint operator $A' \in \mathrm{Lin}(V)$ by

$$(A'v, w)_V = (v, Aw)_V, \quad \forall\, v, w \in V.$$

An operator $A \in \mathrm{Lin}(V)$ is called self-adjoint if $A' = A$.

Now, let us return to the elliptic problems described in the previous sections. Note that the bilinear form a induces an operator $A \in \mathrm{Lin}(H^1_D(\Omega), H^1_D(\Omega)')$ via

$$(Av)(w) = \langle Av, w \rangle_{H^1_D(\Omega)} = a(v, w), \quad \forall\, v, w \in H^1_D(\Omega).$$

Then, the variational problem (1.2) is equivalent to: Find $u \in H^1_D(\Omega)$ such that the equation $Au = \mathcal{F}$ holds in the space $H^1_D(\Omega)'$.

Hereafter, for a (not necessarily finite-dimensional) Hilbert space V, we prefer to identify the dual V' with V itself via the Riesz representation theorem. With this in mind, the Galerkin approximation from (1.18) of A in the finite-dimensional spaces X_ℓ yields the discrete operators $A_\ell \in \mathrm{Lin}(X_\ell)$ given by

$$(A_\ell v, w)_{L^2(\Omega)} = a(v, w), \quad \forall\, v, w \in X_\ell,$$

for $\ell \in \mathbb{N}$. The corresponding properties of the bilinear form a, namely symmetry and ellipticity, imply that each operator A_ℓ is self-adjoint and positive definite.

Given a function $f_\ell \in X_\ell$ such that $(f_\ell, v)_{L^2(\Omega)} = \mathcal{F}(v)$ for all $v \in X_\ell$, the discrete problem is to solve the equation

$$A_\ell u_\ell = f_\ell \quad \text{in } X_\ell \tag{1.21}$$

for u_ℓ. To obtain an algebraic representation with respect to the basis Λ_ℓ, we denote the system matrix (or stiffness matrix) by $\boldsymbol{A}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ with the entries $(\boldsymbol{A}_\ell)_{pq} := a(\lambda^\ell_p, \lambda^\ell_q)$ for $p, q \in \mathcal{N}_\ell$. The right hand side $\boldsymbol{F}_\ell \in \mathbb{R}^{n_\ell}$ of the linear system of equations is the vector with the entries $(\boldsymbol{F}_\ell)_p := (f_\ell, \lambda^\ell_p)_{L^2(\Omega)}$ for $p \in \mathcal{N}_\ell$. Then, the algebraic formulation of the discrete problem is the equation

$$\boldsymbol{A}_\ell \boldsymbol{u}_\ell = \boldsymbol{F}_\ell \quad \text{in } \mathbb{R}^{n_\ell}. \tag{1.22}$$

Note that $\boldsymbol{F}_\ell$ is not the algebraic representation of f_ℓ. In fact, for $\boldsymbol{f}_\ell := \Phi_\ell^{-1}(f_\ell) \in \mathbb{R}^{n_\ell}$ we have $\boldsymbol{F}_\ell = \boldsymbol{M}_\ell \boldsymbol{f}_\ell$ where $\boldsymbol{M}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ is the mass matrix with the entries $(\boldsymbol{M}_\ell)_{pq} :=$

$(\lambda_p^\ell, \lambda_q^\ell)_{L^2(\Omega)}$ for $p, q \in \mathcal{N}_\ell$. Similarly, if $\boldsymbol{A}_\ell^\Phi \in \mathbb{R}^{n_\ell \times n_\ell}$ is the algebraic representation of A_ℓ, namely the matrix defined via $\boldsymbol{A}_\ell^\Phi \boldsymbol{v} = \Phi_\ell^{-1}(A_\ell \Phi_\ell(\boldsymbol{v}))$ for all $\boldsymbol{v} \in \mathbb{R}^{n_\ell}$, then the relation $\boldsymbol{A}_\ell = \boldsymbol{M}_\ell \boldsymbol{A}_\ell^\Phi$ holds. Both the matrix $\boldsymbol{A}^\Phi$ and the vector $\boldsymbol{f}_\ell$ do not need be computed.

Naturally, the matrices $\boldsymbol{A}_\ell$ and $\boldsymbol{M}_\ell$ are symmetric, i. e., self-adjoint with respect to the Euclidean scalar product, and positive definite.

Condition number

For any matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $n \in \mathbb{N}$, and any operator $A \in \mathrm{Lin}(V)$ in the space $V \subset L^2(\Omega)$, we define the operator norm induced by the Euclidean norm and the L^2-norm, respectively, by

$$|\boldsymbol{A}| := \sup_{\boldsymbol{v} \in \mathbb{R}^n,\, 0 < |\boldsymbol{v}| \le 1} \frac{|\boldsymbol{A}\boldsymbol{v}|}{|\boldsymbol{v}|} \quad \text{and} \quad \|A\| := \sup_{v \in V,\, 0 < \|v\|_{L^2(\Omega)} \le 1} \frac{\|Av\|_{L^2(\Omega)}}{\|v\|_{L^2(\Omega)}}. \tag{1.23}$$

Later, for the numerical study of certain prolongation operators between non-nested finite element meshes, we will also need other, more general operator norms; see Section 5.8.

Let $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ or $A \in \mathrm{Lin}(V)$ be a symmetric positive definite matrix or operator. Then, the (energy) scalar product $(\cdot, \cdot)_A$ is defined by

$$(\boldsymbol{v}, \boldsymbol{w})_A := \boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{w}, \quad \forall\, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n, \qquad \text{or} \qquad (v, w)_A := (v, Aw)_{L^2(\Omega)}, \quad \forall\, v, w \in V.$$

In both cases, the symbol $\|\cdot\|_A := \sqrt{(\cdot,\cdot)_A}$ denotes the induced (energy) norm. The corresponding operator norm may be defined analogously to (1.23). If $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is positive definite, the condition number $\kappa(\boldsymbol{A})$ with respect to the norm $|\cdot|$ is

$$\kappa(\boldsymbol{A}) := |\boldsymbol{A}|\, |\boldsymbol{A}^{-1}| = \frac{\lambda_{\max}(\boldsymbol{A})}{\lambda_{\min}(\boldsymbol{A})}, \tag{1.24}$$

where

$$\lambda_{\max}(\boldsymbol{A}) := \max_{\boldsymbol{0} \neq \boldsymbol{v} \in \mathbb{R}^n} \frac{(\boldsymbol{v}, \boldsymbol{v})_A}{|\boldsymbol{v}|^2}, \qquad \lambda_{\min}(\boldsymbol{A}) := \min_{\boldsymbol{0} \neq \boldsymbol{v} \in \mathbb{R}^n} \frac{(\boldsymbol{v}, \boldsymbol{v})_A}{|\boldsymbol{v}|^2}$$

are the maximum and minimum eigenvalues of $\boldsymbol{A}$, respectively. Analogously, for positive definite operators $A \in \mathrm{Lin}(V)$ in finite-dimensional spaces $V \subset L^2(\Omega)$, we have

$$\kappa(A) := \|A\|\, \|A^{-1}\| = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \tag{1.25}$$

with the extreme eigenvalues

$$\lambda_{\max}(A) := \max_{0 \neq v \in V} \frac{(v, v)_A}{\|v\|_{L^2(\Omega)}^2}, \qquad \lambda_{\min}(A) := \min_{0 \neq v \in V} \frac{(v, v)_A}{\|v\|_{L^2(\Omega)}^2}.$$

We denote the spectral radius of an operator A and a matrix $\boldsymbol{A}$, which coincides with $\lambda_{\max}$ in the finite-dimensional setting, by the symbols $\rho(A)$ and $\rho(\boldsymbol{A})$, respectively.

The condition number is a measure of the sensitivity of the solution to perturbations of the data. More precisely, the expansion up to first order of the relative error in the solution u of a linear equation $Au = f$ is bounded by $\kappa(A)$ times the relative error in the operator

A and the right hand side f; see, e. g., [92] for more details. However, the issue that the speed of most iterative solvers depends strongly on the condition number is more relevant for the topic of this thesis.

Now, we return to the specific elliptic problem from above. An analysis of the spectrum of $\boldsymbol{A}_\ell$ or A_ℓ yields the following well-known result; see, e. g., [177]. For a family of quasi-uniform meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$, the condition number of the system matrix and the discrete operator is $\kappa(\boldsymbol{A}_\ell) = \kappa(A_\ell) \lesssim h_\ell^{-2}$ for $h_\ell \to 0$. This estimate is sharp. In contrast, the mass matrix is well-conditioned, we have $\kappa(\boldsymbol{M}_\ell) \lesssim 1$.

Convergence behavior of iterative solvers

Finally, let us consider the solution of the discrete equation (1.21) or (1.22) for a fixed finite element space X_ℓ. We may omit the index ℓ and deal with the operator equation: Find $u \in X$ such that $Au = f$.

Let the iterative method start at the initial guess $u^0 \in X$. Constructing a sequence which is supposed to converge to u, we denote the k-th iterate by $u^k \in X$, $k \geq 1$. Then, the error of the k-th iterate is $e^k := u - u^k \in X$ with the residual $r^k := Ae^k = f - Au^k$. The correction of step k is denoted by $c^k := u^k - u^{k-1}$.

To measure the convergence behavior, one examines the asymptotic rate of convergence ρ, also called contraction rate, defined by

$$\rho := \lim_{k\to\infty} \rho^k \quad \text{with} \quad \rho^k := \frac{\|e^k\|}{\|e^{k-1}\|}, \quad k \geq 1,$$

with respect to a suitable norm $\|\cdot\|$ in X. As customary, we set this fraction and the ones considered in the following to 0 if the denominator vanishes.

As the finite element solution u and, thus, the algebraic error $\|e^k\|$ is unknown, in general, alternative quantities have to be used to describe the reduction of the error in the k-th step of the generic iteration. Two common indicators are the L^2-norm of the residual $\|r^k\|_{L^2(\Omega)}$ and the energy norm of the (next) correction $\|c^{k+1}\|_A$ leading to the rates

$$\rho^k_{\text{res}} := \frac{\|r^k\|_{L^2(\Omega)}}{\|r^{k-1}\|_{L^2(\Omega)}} \quad \text{and} \quad \rho^k_A := \frac{\|c^{k+1}\|_A}{\|c^k\|_A}, \quad k \geq 1, \tag{1.26}$$

respectively. Note that, for ill-conditioned problems, the measures from (1.26) may not reflect the actual error decay very well.

In the context of preconditioning, there is another natural, but still less frequently used choice for an estimator of ρ^k. Let $C \in \mathrm{Lin}(X)$ be a symmetric positive definite operator, a preconditioner, with $(C^{-1}v, v)_{L^2(\Omega)} \eqsim (Av, v)_{L^2(\Omega)}$ for all $v \in X$. Further, the k-th correction may be computed as the preconditioned residual, namely $c^k = Cr^{k-1}$. Then, the fraction

$$\rho^k_C := \frac{\sqrt{(c^{k+1}, r^k)_{L^2(\Omega)}}}{\sqrt{(c^k, r^{k-1})_{L^2(\Omega)}}}, \quad k \geq 1,$$

is equivalent to the algebraic error reduction rate because

$$(c^{k+1}, r^k)_{L^2(\Omega)} = (Cr^k, r^k)_{L^2(\Omega)} = (CAe^k, Ae^k)_{L^2(\Omega)} = (ACAe^k, e^k)_{L^2(\Omega)} \eqsim \|e^k\|_A^2.$$

Here, the last equivalence is a well-known result in the analysis of preconditioners; see, e. g., [194, Lemma 2.1]. Naturally, the quality of the estimator for the iteration error $\|e^k\|_A$ depends on the quality of the preconditioner C, more precisely, on the constants in the equivalence $(C^{-1}v, v)_{L^2(\Omega)} \eqsim (Av, v)_{L^2(\Omega)}$.

Finally, for a sufficiently large number of iterations K, the geometric mean

$$\bar{\rho}^K := \left(\prod_{k=1}^{K} \rho^k \right)^{\frac{1}{K}} \tag{1.27}$$

is a reasonable approximation of the asymptotic convergence rate ρ. If the early stage of the iteration appears very differently from the asymptotic behavior, one may want to neglect the first couple of steps in the definition (1.27). In practice, ρ^k needs to be replaced by one of the approximate variants from above.

We consider some corresponding quantities for finite element vectors in the context of geometric multigrid methods in Section 2.1 and Section 2.2.1.

Preconditioning

Iterative solvers suffer from their heavy dependence on the condition number. This is a critical point for our finite element model problems since the condition numbers of the representing operators or matrices grow like $\mathcal{O}(h_\ell^{-2})$ for $h_\ell \to 0$.

For symmetric positive definite problems, the conjugate gradient method (cg) proposed in [111] is one of the most important solvers. However, its convergence speed depends crucially on the condition number as specified below. The idea of preconditioning is to transform the original system to another one with smaller condition number. We call a symmetric positive definite operator $C \in \mathrm{Lin}(X)$ a preconditioner for the operator A, if the condition number of the preconditioned operator is relatively small, namely $\kappa(CA) \ll \kappa(A)$. In addition, the action of the operator C, which can be interpreted as an approximate inverse of A, should be easy to evaluate.

Then, for the iteration error of the preconditioned conjugate gradient method (pcg), the (sharp) estimate

$$\|e^k\|_A \leq 2 \left(\frac{\sqrt{\kappa(CA)} - 1}{\sqrt{\kappa(CA)} + 1} \right)^k \|e^0\|_A$$

is well-known. Of course, the operators C and A do not need to commute. This is because the pcg method is equivalent to the cg method applied to the operator $C^{1/2}AC^{1/2}$ where $C^{1/2}$ is the uniquely defined positive square root of C such that $C^{1/2}C^{1/2} = C$ and symmetric. Still, only the action of C has to be computed in the algorithm. This implies that the efficiency of the overall iterative method principally depends on the size of $\kappa(C^{1/2}AC^{1/2}) = \kappa(CA)$ and the cost of the assembly and the evaluation of C. Note that, for a preconditioner C, the linear iteration

$$u^{k+1} = u^k + \omega C(f - Au^k) \tag{1.28}$$

converges for $\omega \in (0, \frac{2}{\lambda_{\max}(CA)})$.

Therefore, the quest for an efficient iterative solution strategy for the (ill-conditioned) symmetric positive definite linear problem (1.21) amounts to finding an operator $C_\ell \in \mathrm{Lin}(X_\ell)$ such that $\kappa(C_\ell A_\ell) \ll \kappa(A_\ell)$. In fact, the goal is to obtain robust convergence behavior of the iterative method with respect to the mesh size h_ℓ. More precisely, both the condition number $\kappa(C_\ell A_\ell)$ of the preconditioned operator and the cost of the preconditioning routine should be bounded as far as possible. At the best, one can show a result like $\kappa(C_\ell A_\ell) \lesssim 1$ with a bound independent of the discretization parameter. In this case, the pcg method should yield robust convergence with respect to the problem size.

This is exactly the objective of the multilevel methods investigated in the last decades. The fundamental ideas and a fairly detailed overview of the literature are presented in Chapter 2. We develop a generalized concept in Chapter 3.

2 Multilevel methods for elliptic equations

In this chapter, the main features of multilevel methods for elliptic partial differential equations are presented in a setting which is to a certain extent standard. We describe the fundamental ideas and the algorithmic structures as well as outline the convergence analysis from the theory of additive and multiplicative Schwarz methods. This is meant to prepare the subsequent discussion of a generalized framework based on non-nested meshes. Therefore, during the introduction of the multilevel algorithms, we put emphasis on the information transfer between the involved finite element spaces.

To start with, Section 2.1 reviews the classical linear iterative methods giving a motivation of the development of multilevel techniques to overcome the determined deficiencies. Section 2.2 is about geometric multigrid methods. Here, we present the standard algorithms for elliptic partial differential equations based on multilevel finite elements associated with a hierarchy of nested meshes. In Section 2.3, it is shown how the framework of parallel and successive subspace correction methods (Schwarz methods) can be used to prove optimal preconditioning and convergence results. Note that the Schwarz theory, while not yielding quantitative estimates of contraction rates, is capable of treating very general problem classes with only minor regularity assumptions. We conclude with some remarks about the applicability of multigrid methods to other problems settings in Section 2.4.

2.1 Standard linear iterative methods

To understand the fundamental concepts of the multilevel methods this thesis is concerned with, it may be considered essential to briefly recall the basic properties of classical iterative methods. This is done in the present section. At the same time, we motivate the development of multilevel techniques by observing the inability of standard linear iterations to reduce low-frequency contributions of the algebraic error. We also use this opportunity to introduce several notations which will be used in the algorithms.

Let us first state some notations concerning the residual during an iteration. For an index ℓ, let $u_\ell^k = \Phi_\ell(\boldsymbol{u}_\ell^k) \in X_\ell$ be the k-th iterate. Then, we have

$$\boldsymbol{R}_\ell^k := \boldsymbol{F}_\ell - \boldsymbol{A}_\ell \boldsymbol{u}_\ell^k = \Big((f_\ell, \lambda_p^\ell)_{L^2(\Omega)} - (A_\ell u_\ell^k, \lambda_p^\ell)_{L^2(\Omega)} \Big)_{p \in \mathcal{N}_\ell} = \Big((r_\ell^k, \lambda_p^\ell)_{L^2(\Omega)} \Big)_{p \in \mathcal{N}_\ell} . \quad (2.1)$$

Further, with $\boldsymbol{r}_\ell^k := \Phi_\ell^{-1}(r_\ell^k)$, it is $\boldsymbol{R}_\ell^k = \boldsymbol{M}_\ell \boldsymbol{r}_\ell^k$ and $\|r_\ell^k\|_{L^2(\Omega)} = (\boldsymbol{M}_\ell^{-1} \boldsymbol{R}_\ell^k) \cdot \boldsymbol{R}_\ell^k$. Therefore, one usually considers $|\boldsymbol{R}_\ell^k|$ instead of $\|r_\ell^k\|_{L^2(\Omega)}$ to measure the convergence behavior of an iterative method in practice, although the two quantities are only equivalent up to scaling with the local mesh size.

Generally, the considered standard iterations are of the form (1.28). The methods are called linear or stationary because the error propagation is controlled by the fixed linear operator $E := \mathrm{id} - \omega CA$; the error update of the k-th step is $e^{k+1} = Ee^k$. Naturally, the method converges if and only if the spectral radius of E satisfies $\rho(E) < 1$.

Given some initial residual $r_\ell \in X_\ell$ and a current iterate $v_\ell \in X_\ell$, let the next iterate

for the corresponding defect problem in X_ℓ be defined by the step

$$S_\ell(v_\ell, r_\ell) := v_\ell + \omega C_\ell(r_\ell - A_\ell v_\ell). \tag{2.2}$$

We also use a common product notation, which is recursively defined for $\nu > 1$ by

$$S_\ell^\nu(v_\ell, r_\ell) := S_\ell(S_\ell^{\nu-1}(v_\ell, r_\ell), r_\ell), \tag{2.3}$$

starting with $S_\ell^1(v_\ell, r_\ell) := S_\ell(v_\ell, r_\ell)$ given by (2.2). Then, it is $u_\ell^{k+\nu} = S_\ell^\nu(u_\ell^k, f_\ell)$ if $u_\ell^0 = 0$ for the discrete problem in equation (1.21). In matrix notation, for $\boldsymbol{v}_\ell, \boldsymbol{R}_\ell \in \mathbb{R}^{n_\ell}$, we have

$$\boldsymbol{S}_\ell(\boldsymbol{v}_\ell, \boldsymbol{R}_\ell) := \boldsymbol{v}_\ell + \omega \boldsymbol{C}_\ell(\boldsymbol{R}_\ell - \boldsymbol{A}_\ell \boldsymbol{v}_\ell), \tag{2.4}$$

which gives the next iterate in $\mathbb{R}^{n_\ell}$, and $\boldsymbol{S}_\ell^\nu(\boldsymbol{v}_\ell, \boldsymbol{R}_\ell)$ defined analogously to (2.3). In (2.2) and (2.4), one needs a suitable operator $C_\ell \in \mathrm{Lin}(X_\ell)$ and a suitable matrix $\boldsymbol{C}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$, respectively. Actually, the method given by the formula (2.4) can be derived and analyzed as a pure matrix iteration without referring to the finite element context; see, e. g., the standard references [104, 196]. The notation of the residual vector $\boldsymbol{R}_\ell$ shall still indicate that this input is obtained according to (2.1) here.

The most common iterations employed for smoothing in multigrid methods are constructed by a splitting of the system matrix $\boldsymbol{A}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ into three $(n_\ell \times n_\ell)$-matrices via

$$\boldsymbol{A}_\ell = \boldsymbol{D}_\ell - \boldsymbol{L}_\ell - \boldsymbol{U}_\ell.$$

Here, $\boldsymbol{D}_\ell$ is the diagonal, $-\boldsymbol{L}_\ell$ is the lower and $-\boldsymbol{U}_\ell$ is the upper triangular part. Then, the Jacobi method is specified by $\boldsymbol{C}_\ell^{\mathrm{Jac}} = \boldsymbol{D}_\ell^{-1}$; the Gauß–Seidel method has the representation $\boldsymbol{C}_\ell^{\mathrm{GS}} = (\boldsymbol{D}_\ell - \boldsymbol{L}_\ell)^{-1}$. Both linear methods are Richardson iterations with relaxation parameter ω; the latter depends on the ordering of the degrees of freedom. It is important to note that the cost of one step grows like $\mathcal{O}(n_\ell)$ for sparse matrices such as the ones coming from the finite element discretization in Section 1.3.

To illustrate the correlation between operator and matrix notations in the finite element setting employed throughout this thesis, we remark that, for instance, the operator S_ℓ^{Jac} in the iteration

$$u_\ell^{k+1} = S_\ell^{\mathrm{Jac}}(u_\ell^k, f_\ell) = u_\ell^k + \omega \sum_{p \in \mathcal{N}_\ell} a(\lambda_p^\ell, \lambda_p^\ell)^{-1} (r_\ell^k, \lambda_p^\ell)_{L^2(\Omega)} \lambda_p^\ell \tag{2.5}$$

represents the Jacobi method. Both characterizations (2.4) and (2.5) indicate that only the data $\boldsymbol{R}_\ell^k = \boldsymbol{M}_\ell \boldsymbol{r}_\ell^k$ is required to evaluate the above pointwise relaxation methods.

It is well-known that the performance of the basic iterative methods for large systems is rather poor and strongly deteriorates with an increasing number of equations. For instance, the square of the error reduction factor of the Gauß–Seidel method (with $\omega = 1$) grows like $1 - \mathcal{O}(h_\ell^2)$. As a matter of fact, this may be verified by the theory of subspace correction methods outlined in Section 2.3; see [194].

Consequently, these simple iterations cannot be used as stand-alone methods in practice. Many classical iterative schemes show a strong smoothing effect on the algebraic error, though. The oscillatory parts of the error are damped relatively fast, mostly in very few

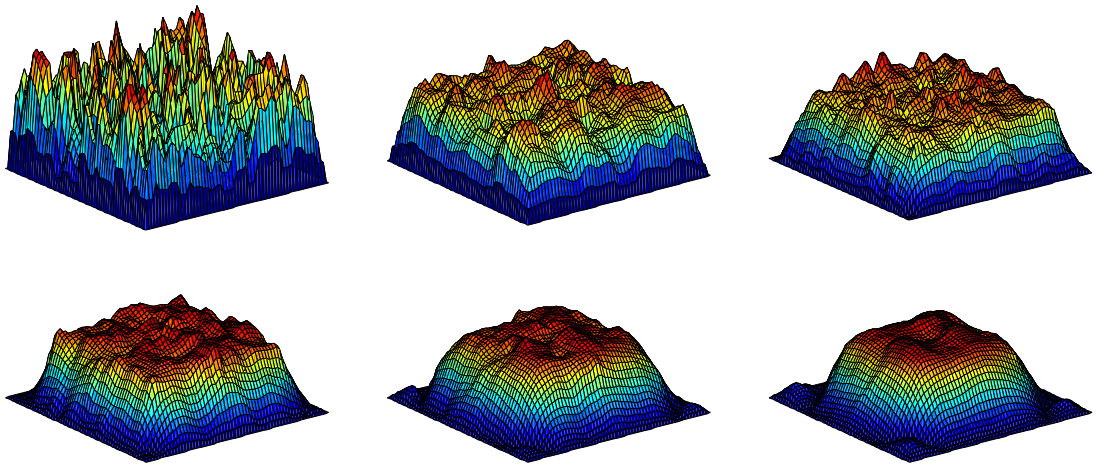


Figure 2.1. Illustration of the smoothing effect of the Gauß–Seidel method: absolute error of the initial guess (random, normalized) and the first five iterations.

iteration steps, whereas the overall convergence speed indeed turns out to be as poor as the above pessimistic estimate predicts. This behavior is illustrated in Figure 2.1*.

Here, the conception of the term "smooth error" relies on geometric considerations related to the finite element discretization. This notion is different in the context of algebraic multigrid methods; sometimes the term "slow-to-converge error" is preferred. Further, the smoothing property is relative. More precisely, at each level $\ell > 0$, a smoothing operator acting on the finite element space X_ℓ manages to reduce the high-frequency contributions of the error with respect to the current mesh such that the remaining error may be approximated accurately in the subspace $X_{\ell-1}$ in a certain sense. The qualitative behavior does not depend on the actual mesh size h_ℓ. This is the fundamental observation which led to the development of multigrid methods in the first place. We do not state this formally here but return to this issue later.

A lot of effort has been spent to investigate both quantitative and qualitative properties of smoothing operators for multigrid methods. However, the focus of our own research has been directed to an application of the abstract Schwarz theory which is beneficial to cover very general problem classes with minor regularity assumptions. As we treat the multilevel iterations considered in Section 2.2 as subspace correction methods in Section 2.3, no quantitative characterization of the smoothing properties enters the analysis. This already indicates that the entire theory is a qualitative one which aims at robust convergence statements for general problems instead of quantitative estimates of contraction numbers. To obtain the latter, one usually needs rather restrictive assumptions on the problem settings. Therefore, we neither go into detail about Brandt's local mode analysis [37] (or local Fourier analysis, for short LFA) nor investigate the classical smoothing properties identified by Hackbusch [103]. For a discussion of these issues, we refer to [178].

*Certainly, although I have created the graphs from my own computational results, I do not deserve credit for this nice demonstration of the smoothing behavior of standard relaxation methods. One should be able to find similar figures in every noteworthy multigrid book; see, e. g., [178].

Let us, as early as now, remark that the classical iterative methods work well in the new context to be developed in Chapter 3. This thesis does not include new results on smoothing operators. But the numerical examples suggest that the smoothed error is well represented by functions in the range of the prolongation from the coarser level.

2.2 Geometric multigrid methods

In this section, we discuss the major ideas of multilevel methods for elliptic partial differential equations. In particular, we introduce the (standard) multigrid algorithm usually implemented in a recursive fashion and BPX-type preconditioners. This is the class of algorithms that will be extended to a semi-geometric method in Chapter 3. Particular attention is paid to the formal description of the ordinary information transfer between nested finite element spaces.

It seems generally accepted that multigrid methods have first been described by Fedorenko [86, 87] and Bakhvalov [12]. An important milestone was reached with the work by Brandt summarized in [37]. Furthermore, essential contributions are due to Hackbusch; see the monograph [103]. We also briefly comment on the historical developments in the context of the convergence analysis in Section 2.3.

It is well-known that the performance of the classical linear methods of Section 2.1 is not satisfactory. In rather general settings, they have smoothing properties, though. The paradigm of multigrid methods is to perform only very few steps of a relaxation method to obtain a defect problem where the remaining error is relatively smooth, at least with respect to a weighted energy norm at the current level. Then, the defect problem or, more precisely, the corresponding error may be well represented in a space with less degrees of freedom. An essential ingredient is a methodology of how to realize such a coarse approximation.

The power of multilevel iterations results from a sophisticated combination of smoothing iterations and coarse level corrections. These ingredients should be complementary in the sense that they reduce different components of the error; at each level a different section of the spectrum should be processed.

2.2.1 Information transfer between nested finite element spaces

First, we describe a standard way to obtain a coarse level approximation in geometric multigrid methods. Because in this thesis we are concerned with the study of diverse interpolating or projecting operators, we do this in a little more detail than absolutely necessary.

To obtain a representation of the defect problem in a space with less degrees of freedom, we directly work with finite element spaces of different scales. Note that the applicability of multigrid methods is not restricted to this class of problems. Other variants may be developed more conveniently by introducing certain stencil notations; see, e. g., [103, 178]. Still, aiming at efficient multilevel methods for unstructured finite element meshes, we do not consider this point of view as it does not seem to provide any additional insight into the structure of the studied problems. The presentation in this section is, moreover, closer to

our investigation of geometrically inspired transfer concepts between finite element spaces associated with non-nested meshes.

Recall that we have already introduced a scale of finite element spaces $(X_\ell)_{\ell\in\mathbb{N}}$ associated with meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ in Section 1.3. For the considerations of the present chapter, assume that the spaces are nested, i. e.,

$$X_\ell \subset X_{\ell+1}, \quad \forall\, \ell \in \mathbb{N}.$$

We consider the variational approach in the following. When dealing with a sequence of nested finite element spaces, there is a very natural way to realize the information transfer between two successive spaces. An important ingredient is the fact that the problem is set in the Hilbert space $L^2(\Omega)$.

The index $\ell > 0$ may be fixed. For the coarse-to-fine transfer, let

$$\mathcal{I}_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell$$

be the natural embedding of $X_{\ell-1}$ into X_ℓ. Its matrix representation $\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ is defined via $\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell}\boldsymbol{v} = \Phi_\ell^{-1}(\mathcal{I}_{\ell-1}^{\ell}\Phi_{\ell-1}(\boldsymbol{v}))$ for all $\boldsymbol{v} \in \mathbb{R}^{n_{\ell-1}}$. Evidently, the entries of $\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell}$ are the coefficients in the unique linear combination representing a basis function of $X_{\ell-1}$ with respect to the basis of X_ℓ. This is the mapping which is used as prolongation operator to transfer a coarse level correction to the next finer level.

For finite element problems set in the space $L^2(\Omega)$, the natural way to transfer the residual to a coarser level is the orthogonal projection with respect to $(\cdot,\cdot)_{L^2(\Omega)}$. To understand this, it is reasonable to consider the residual as a dual object, namely a functional in the respective dual space. Let $r_\ell \in X_\ell$. Then, as $X_{\ell-1} \subset X_\ell$, it is possible to represent the corresponding functional $(r_\ell, \cdot) \in X_\ell'$ exactly in $X_{\ell-1}'$. More precisely, there is a unique function $r_{\ell-1} =: \mathcal{Q}_\ell^{\ell-1} r_\ell \in X_{\ell-1}$ such that

$$(r_{\ell-1}, v)_{L^2(\Omega)} = (r_\ell, v)_{L^2(\Omega)}, \quad \forall\, v \in X_{\ell-1}.$$

In other words, $r_{\ell-1}$ is computed by the L^2-projection of r_ℓ to the subspace $X_{\ell-1} \subset X_\ell$. This gives a fine-to-coarse mapping

$$\mathcal{Q}_\ell^{\ell-1} : X_\ell \to X_{\ell-1}$$

which is used as restriction operator yielding the coarse level representation of the residual.

A matrix representation of $\mathcal{Q}_\ell^{\ell-1}$ with respect to the bases $\Lambda_{\ell-1}$ and Λ_ℓ is obtained by use of the one of the embedding $\mathcal{I}_{\ell-1}^{\ell}$; we have

$$\boldsymbol{\mathcal{Q}}_\ell^{\ell-1} = \boldsymbol{M}_{\ell-1}^{-1}(\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell})^T \boldsymbol{M}_\ell. \tag{2.6}$$

In particular, for a residual $\boldsymbol{r}_\ell = \Phi_\ell^{-1}(r_\ell)$, we have the following relation between the residual vectors at two successive levels,

$$\boldsymbol{R}_{\ell-1} := \boldsymbol{M}_{\ell-1}\boldsymbol{r}_{\ell-1} = \boldsymbol{M}_{\ell-1}\boldsymbol{\mathcal{Q}}_\ell^{\ell-1}\boldsymbol{r}_\ell = \boldsymbol{M}_{\ell-1}\boldsymbol{M}_{\ell-1}^{-1}(\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell})^T\boldsymbol{M}_\ell\boldsymbol{r}_\ell = (\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell})^T\boldsymbol{R}_\ell.$$

Therefore, the inversion of the mass matrix $\boldsymbol{M}_{\ell-1}$ in the evaluation of the operator $\mathcal{Q}_\ell^{\ell-1}$ can be avoided in a very natural way as the relevant input of the smoothing operator at level

$\ell - 1$ is the quantity $\boldsymbol{R}_{\ell-1}$ rather than $\boldsymbol{r}_{\ell-1}$; see Section 2.1. Consequently, the transpose of the prolongation matrix appears as restriction matrix in multigrid algorithms like the one below. As a matter of fact, this is one of the most important reasons why standard multigrid methods can be implemented in an efficient manner. A further discussion of how to replace the fine-to-coarse L^2-projections by other operators that are computable more easily can be found in [31]. This is only necessary in case the structure of the smoothers or subspace solvers cannot be exploited in the above fashion.

Now, let the operator A_ℓ and the matrix $\boldsymbol{A}_\ell$, respectively, be given by a suitable finite element discretization as specified in Section 1.3. To obtain a coarse representation of equation (1.21) at level $\ell - 1$, consider the operator

$$A_{\ell-1} = \mathcal{Q}_\ell^{\ell-1} A_\ell \mathcal{I}_{\ell-1}^\ell. \tag{2.7}$$

This immediately implies that

$$(A_{\ell-1} v, w)_{L^2(\Omega)} = (A_\ell v, w)_{L^2(\Omega)}, \quad \forall\, v, w \in X_{\ell-1}. \tag{2.8}$$

For the matrix equation (1.22), we have

$$\boldsymbol{A}_{\ell-1} = \boldsymbol{M}_{\ell-1} \boldsymbol{A}_{\ell-1}^\Phi = \boldsymbol{M}_{\ell-1} \boldsymbol{M}_{\ell-1}^{-1} (\boldsymbol{\mathcal{I}}_{\ell-1}^\ell)^T \boldsymbol{M}_\ell \boldsymbol{A}_\ell^\Phi \boldsymbol{\mathcal{I}}_{\ell-1}^\ell = (\boldsymbol{\mathcal{I}}_{\ell-1}^\ell)^T \boldsymbol{A}_\ell \boldsymbol{\mathcal{I}}_{\ell-1}^\ell.$$

Due to the above relations, the concept outlined here is called Galerkin or variational approach.

2.2.2 Coarse level correction

Assume that we have linear coarse-to-fine operators $(\mathcal{I}_\ell^{\ell+1})_{\ell\in\mathbb{N}}$ and fine-to-coarse operators $(\mathcal{Q}_{\ell+1}^\ell)_{\ell\in\mathbb{N}}$ between the spaces $(X_\ell)_{\ell\in\mathbb{N}}$. According to the previous section, the most reasonable choices, if the equations come from finite element discretizations and allow for natural (nested) coarse scales, are the canonical inclusion and the L^2-projection. Then, in the variational setting, the operators $(A_\ell)_{\ell\in\mathbb{N}}$ are actually representations of the operator A at each level ℓ as stated in (2.8). In particular, they are symmetric positive definite.

Suppose the high-frequency components of the error have been reduced at some specified level $\ell > 0$ by a standard relaxation method. Then, the remaining error (i. e., the solution of the defect equation) is expected to be relatively smooth and may be well approximated at a coarser level. This is done in a finite element space with larger mesh size. Now, the idea of computing a coarse level correction is to reduce the low-frequent error contributions by a suitable routine at level $\ell - 1$.

We illustrate this concept by the variational principle in the following two-level setting. Let the coarse level operator $A_{\ell-1}$ be defined by the Galerkin approach (2.7). For a guess $v_\ell \in X_\ell$, let

$$c_\ell = \mathcal{I}_{\ell-1}^\ell A_{\ell-1}^{-1} \mathcal{Q}_\ell^{\ell-1} (f_\ell - A_\ell v_\ell) \tag{2.9}$$

be the (exact) coarse level correction. Then, the variational principle ensures that the associated error update operator is an orthogonal projection to the range of $\mathcal{I}_{\ell-1}^\ell$ with

respect to inner product $(\cdot,\cdot)_A$; the coarse level correction is a minimizer of the energy norm $\|\cdot\|_A$, i. e.,

$$\|u_\ell - v_\ell - c_\ell\|_A = \min_{c_{\ell-1}\in X_{\ell-1}} \|u_\ell - v_\ell - \mathcal{I}_{\ell-1}^{\ell} c_{\ell-1}\|_A. \tag{2.10}$$

In other words, the variational construction implies that $\mathcal{I}_{\ell-1}^{\ell} c_{\ell-1}$ is the Galerkin approximation of the exact solution of the defect equation at the fine level ℓ in the space $X_{\ell-1}$.

Naturally, it is only practicable to compute the coarse level correction c_ℓ exactly, namely via (2.9), if the dimension of the respective space $X_{\ell-1}$ is sufficiently small. Otherwise it needs to be approximated. In the context of multigrid methods, this is done by recursion such that the inverse of the matrix $\boldsymbol{A}_\ell$ is only computed at level $\ell = 0$, as described in the next paragraph. In this case, the characterization (2.10) does not hold true any more.

Still, at this point, one realizes that the two-level method obtained by a simple combination of a fine level smoothing procedure and an exact coarse level correction cannot diverge if the smoother converges. Under certain conditions, this also holds true if the exact coarse level correction is replaced by a suitable approximation; see [176]. Note that adding the coarse level correction does not converge as a stand-alone method. This is because the restriction operator generally has a non-trivial kernel.

2.2.3 The standard algorithm

In this section, we state a standard form of the multigrid method which comprises the ideas outlined in the previous paragraphs.

The efficiency of the geometric multigrid cycle arises from an effective treatment of the different frequencies of the error by a sophisticated combination of smoothing iterations and coarse level corrections. The parameters $\nu_1, \nu_2 \geq 0$ denote the number of pre-smoothing and post-smoothing steps, respectively. By $\gamma \geq 1$ we denote the number of recursive calls of the coarse level correction routine.

We prefer to write down the multigrid cycles as preconditioners such that the respective algorithms provide corrections rather than updated iterates. Note that this is in contrast to the notation of the smoothing operators $(S_\ell)_{\ell\in\mathbb{N}}$. As indicated in Section 1.4, a preconditioner has to be evaluated for a given residual. It is a linear mapping in $\mathrm{Lin}(X_\ell)$ that may be represented by a matrix in $\mathbb{R}^{n_\ell\times n_\ell}$. Then, the overall iterative solver is a preconditioned Richardson or a preconditioned conjugate gradient method to be considered afterwards. We only state the operator formulation here.

Algorithm 2.1 (Multigrid cycle)**.** *For (the residual) $r \in X_\ell$ compute the value*

$$C_\ell r = \mathrm{MG}_\ell^{\gamma,\nu_1,\nu_2}(r) = \mathrm{MG}\ (\ell,\ \gamma,\ \nu_1,\ \nu_2,\ r) \in X_\ell$$

as follows.

```
MG (ℓ, γ, ν1, ν2, r) {
    if (ℓ = 0) {
            Solve exactly: x ← A_0^{-1} r
    }
```

```
    else {
            Pre-smoothing steps: x ← S_ℓ^{ν1}(0, r)
            Coarse level correction:
                Restriction: r' ← Q_ℓ^{ℓ−1}(r − A_ℓ x)
                Initialize: x' ← 0
                for (i = 1, ..., γ) do {
                    Recursive call: x' ← x' + MG (ℓ − 1, γ, ν1, ν2, r')
                }
                Prolongation: x ← x + I_{ℓ−1}^{ℓ} x'
            Post-smoothing steps: x ← S_ℓ^{ν2}(x, r)
    }
    return x
}
```

As indicated before, x and x' are primal objects whereas r and r' are dual objects. For $\gamma = 1$ and $\gamma = 2$, this is the $\mathcal{V}(\nu_1, \nu_2)$-cycle and the $\mathcal{W}(\nu_1, \nu_2)$-cycle, respectively. Other combinations of smoothing and coarse level correction are possible. For simplicity, we have not incorporated an additional relaxation parameter for the coarse level correction. We will comment on the usefulness of an over-relaxation of the coarse level correction in certain multigrid variants later.

To obtain the actual multigrid method, assume that the problem (1.21) to be solved is set in a space corresponding to the fixed level $L > 0$. This means that the nested finite element spaces $(X_\ell)_{\ell=0,\dots,L}$ and the associated transfer operators and coarse level operators discussed in Section 2.2.1 are given. Note that all this data essentially comes from a hierarchy of nested meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L}$. Suppose that the dimension of the space X_0 at the coarsest level is sufficiently small such that efficient methods for the computation of an exact solution exist. Finally, let $(S_\ell)_{\ell=1,\dots,L}$ be suitable symmetric smoothing operators. Then, the desired multigrid method is a Richardson iteration with the preconditioner $C_L \in \mathrm{Lin}(X_L)$ defined by Algorithm 2.1. For the sake of completeness, we state the multigrid iteration in Algorithm 2.2 below. In practice, a sufficiently large reduction of the algebraic error is achieved by employing a suitable (absolute or relative) stopping criterion according to the estimates in Section 1.4. However, symmetric multigrid cycles are for linear problems more commonly used as preconditioners inside pcg iterations.

Algorithm 2.2 (Multigrid iteration)**.** *For chosen parameters ν_1, ν_2, γ, compute an approximate solution of the equation $A_L u = f_L$ in X_L by the following iteration.*

```
Initialize: u^0 ← 0
for (k = 1, 2, ...) do {
    Multigrid cycle: u^k ← u^{k−1} + MG (L, γ, ν1, ν2, f_L − A_L u^{k−1})
    Check stopping criterion
}
```

2.2.4 Full multigrid or nested iteration

Let us briefly discuss the cost of the multigrid methods, i. e., the number of operations that have to be carried out asymptotically. The arguments put forward in this section are well-known and may be found, e. g., in [103, 178]. However, the advantages of multilevel techniques to obtain an asymptotically optimal method for the solution of elliptic boundary value problems have first been investigated systematically and to the full extent in [37].

First of all, if n_0 is sufficiently small, the cost to solve the coarse level equation exactly is considered negligible; thus, the cost of one multigrid cycle grows like the number of unknowns at the finest level, namely $\mathcal{O}(h_L^{-d})$ in case of quasi-uniform meshes. In the present context, the finite element spaces represent discretizations of a boundary value problem for a given partial differential equation. Therefore, it is natural to compute an approximate solution of the equation, e. g., by the multigrid iteration of Algorithm 2.2, such that the algebraic error is of the size of the discretization error. Even if the multigrid method is optimal for any fixed size of the algebraic error, i. e., the number of necessary steps to reach a specified accuracy remains bounded, a direct application of Algorithm 2.2 is not sufficient to obtain an optimal method in this case. This is because the tolerance has to be decreased as $h_L \to 0$.

Assume that the finite element approximation described in Section 1.3 yields a discretization error (in the energy norm) of the order $\mathcal{O}(h_L)$ for the Galerkin solution in the space X_L. In the next section about the fundamental analysis of multilevel methods for elliptic equations (Section 2.3), we discuss the dependence of the convergence behavior of the multigrid algorithms on the dimension of the finite element space. If the convergence is independent of the discretization parameter h_L, i. e., the contraction number of each single multigrid step is uniformly bounded away from one, the number of multigrid cycles needed for the above accuracy is $\mathcal{O}(|\log(h_L)|)$ asymptotically. Therefore, the number of operations to reach an iterate with an algebraic error of the size of the discretization error grows like $\mathcal{O}(h_L^{-d}|\log(h_L)|)$.

Nested iteration is a technique to reduce the cost to optimal order $\mathcal{O}(h_L^{-d})$. It is motivated by the well-known fact that iterative solvers profit from good initial iterates. The paradigm of the following procedure is to ensure that the iteration error at each level is proportional to the discretization error at that same level.

Algorithm 2.3 (Full multigrid iteration)**.** *For chosen parameters ν_1, ν_2, γ, compute an approximate solution of the equation $A_L u = f_L$ in X_L by the following iteration.*

Solve exactly: $\widetilde{u}^0 \leftarrow A_0^{-1} f_0$
for $(\ell = 1, \dots, L)$ **do** {
    Interpolation: $u^0 \leftarrow \widetilde{\mathcal{I}}_{\ell-1}^{\ell} \widetilde{u}^0$
    for $(k = 1, \dots, K_\ell)$ **do** {
        Multigrid cycle: $u^k \leftarrow u^{k-1} + \mathrm{MG}\ (\ell, \gamma, \nu_1, \nu_2, f_\ell - A_\ell u^{k-1})$
        Check stopping criterion
    }
    Final iterate: $\widetilde{u}^0 \leftarrow u^k$
}

Here, the operators $(A_\ell)_{\ell=0,\dots,L}$ and the right hand side data $(f_\ell)_{\ell=0,\dots,L}$ are as specified in Section 1.3. Special attention has to be paid to the stopping criteria (or the number of multigrid steps) at the different levels. Certain assumptions on the problem settings guarantee the existence of small numbers $(K_\ell)_{\ell=1,\dots,L}$ such that the iterates at each level are sufficiently accurate in the above sense. Naturally, these numbers depend on the convergence of the multigrid iteration itself. For a sufficiently small contraction rate (a traditional assumption is $\rho^k < 1/4$), one cycle may be enough. Usually, no more than two or three steps are necessary. Moreover, one carries out the same number of steps at each level because the convergence rate is bounded independently of the level.

In this case, the total cost is still bounded by a constant times the computational work at the final level L; thus, the method is of optimal complexity. Note that, for the above considerations to be valid, the operators $(\widetilde{\mathcal{I}}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ need to satisfy an additional approximation property, e. g., formulated as consistency condition; see [103, Chapter 5]. Evidently, in the variational setting of the present section, the canonical inclusion $(\mathcal{I}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ is sufficiently accurate.

Cascadic multigrid methods

A different multigrid-like method to obtain optimal complexity for the task to reduce the iteration error to the size of the discretization error is the so-called cascadic multigrid method. This algorithm has first been proposed by Deuflhard in [68] in a slightly different form.

Here, the call of the multigrid cycle MG $(\ell,\, \gamma,\, \nu_1,\, \nu_2,\, f_\ell - A_\ell u^{k-1})$ in Algorithm 2.3 is replaced by a non-recursive (one-level) iterative scheme, e. g., one of the standard relaxation methods from Section 2.1 or the conjugate gradient method. Therefore, no coarse level correction in the strict sense of the word is computed. We refer to [21, 22] for further details.

2.2.5 BPX-like preconditioners

We also state a class of additive multilevel preconditioners, which work directly with the generating system $(\Lambda_\ell)_{\ell=0,\dots,L}$ usually called multilevel nodal basis in this context. These methods have been developed to achieve preconditioning with a highly parallel algorithmic structure. Note that they do generally not converge as stand-alone solvers, though.

For a residual $r \in X_L$, one computes a correction via

$$C_L r := A_0^{-1} \mathcal{Q}_0 r + \sum_{\ell=1}^{L} \sum_{p \in \mathcal{N}_\ell} \frac{(r, \lambda_p^\ell)_{L^2(\Omega)}}{d_p^\ell}\, \lambda_p^\ell \tag{2.11}$$

with a nodewise scaling by $d_p^\ell = 4^\ell (\lambda_p^\ell, \lambda_p^\ell)_{L^2(\Omega)}$ or $d_p^\ell = 4^\ell (\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}$ or $d_p^\ell = a(\lambda_p^\ell, \lambda_p^\ell)$; see [34, 198, 204]. Analogously to before, $\mathcal{Q}_0 : X_L \to X_0$ is the L^2-projection; here and in the following, if the domain of a mapping is clearly the finest finite element space, we only indicate the target subspace (in this case, X_0) and drop the index L. As equation (2.11) is formulated in the space X_L and each term in the double sum lies in some subspace $X_\ell \subset X_L$,

we have omitted the inclusion operators $(\mathcal{I}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ for clarity, as common in this field. If the meshes are adaptively refined, there may appear basis functions at level $\ell > 0$ with $\lambda_p^\ell = \lambda_q^{\ell-1}$ for some $p \in \mathcal{N}_\ell$ and $q \in \mathcal{N}_{\ell-1}$. In this case, the corresponding correction terms need to be excluded from the sums. Note that the direct solution of the coarse problem at level 0 may be replaced by another method, for instance, also nodewise scaling. This is done in the genuine BPX-preconditioner developed by Bramble, Pasciak and Xu in [34]. We do not consider the hierarchical basis preconditioner and the corresponding multigrid method developed in [197, 198] and [13], respectively, here.

Again, one notes that only the residual vectors $\boldsymbol{R}_\ell$, $\ell \in \{0, \dots, L\}$ are required to evaluate the preconditioner. This is because the quantity $c_0 := A_0^{-1}\mathcal{Q}_0 r \in X_0$ is characterized by

$$a(c_0, v) = (r, v)_{L^2(\Omega)}, \quad \forall\, v \in X_0;$$

thus, we have $\boldsymbol{c}_0 = \boldsymbol{A}_0^{-1}\boldsymbol{R}_0$.

Remark 2.4. *The influential research published in [95, 96] gives an interpretation of multilevel methods (including the multigrid methods and the BPX-type preconditioners) as levelwise block Gauß–Seidel or Jacobi methods with respect to the associated semi-definite system. This point of view allows, for instance, to regroup the degrees of freedom in a nodewise fashion to obtain a point-oriented multilevel algorithm; see [95, Chapter 6]. This has major consequences for the (parallel) implementation of the resulting block iterative methods.*

2.3 Convergence analysis

In this section, we briefly explain the convergence analysis of the multilevel algorithms in the framework of parallel and successive subspace correction methods. Relevant characteristics are highlighted to facilitate the subsequent discussion of a generalized methodology based on non-nested meshes. More precisely, for the analysis of the semi-geometric multilevel methods to be developed in Chapter 3, we refer to some of the main results particularly summarized here.

The theory to be presented in the following is a qualitative as opposed to a quantitative one. It is fair to say that among all available approaches this is the most robust one as it allows for the most general statements. No quantitative results concerning convergence factors are achievable in this framework, though. The main advantage is that no regularity assumptions are required to prove asymptotically optimal preconditioning and convergence results.

Based on a formulation of the multigrid method in an abstract Hilbert space setting developed in [27], the decisive steps were taken by Bramble, Pasciak, Wang and Xu in [32, 33, 193]; see also the review paper [194]. The result was not optimal, namely independent of the mesh size h_L and the number of levels L, before the importance of a general norm equivalence in multilevel finite element approximation theory was completely understood by Oswald [151] and others.

The earlier approach of combining stronger smoothing and approximation properties, which relies on a regularity result for the underlying partial differential equation, is presumably presented most elaborately in [103]; we also refer to this monograph and the references

therein for an overview of the historical development of this particular theory. Finally, the paper [199] is a notable review of different techniques and also includes many historical comments.

2.3.1 Subspace splitting and subspace correction

As a matter of fact, the fundamental idea of constructing convergent sequences in a function space by a subspace correction technique traces back to the work by Schwarz [169] published as early as in the year 1870, although in a purely analytic context. Here, we outline how the multilevel methods presented in Section 2.2 fit into the framework of additive and multiplicative Schwarz methods acting on the residual by parallel and successive subspace correction, respectively. The presentation follows [151, 177].

Here and in the following, we make exclusive use of the operator notations. As before, the letter C stands for an (overall) preconditioner and E represents an iteration operator for an error update. Moreover, the character B denotes certain approximations of the operator A in some subspaces whereas T stands for a specified subspace correction.

Let the finite-dimensional space V be decomposed into a usually not direct sum of subspaces $(V_\ell)_{\ell=0,\dots,L}$, $L < \infty$, i. e.,

$$V = \sum_{\ell=0}^{L} V_\ell.$$

Assume that the subspaces are equipped with suitable auxiliary forms $b_\ell(\cdot,\cdot) = (B_\ell \cdot, \cdot)_{L^2(\Omega)}$, $\ell \in \{0,\dots,L\}$, defined by the symmetric positive definite operators $B_\ell \in \mathrm{Lin}(V_\ell)$, which are supposed to represent approximations of the restrictions A_ℓ of the operator A to the subspaces V_ℓ. Then, the subspace splitting reads as

$$(V, a) = \sum_{\ell=0}^{L} (V_\ell, b_\ell). \tag{2.12}$$

This notation emphasizes the importance of the bilinear forms $b_\ell(\cdot,\cdot)$ for the following discussion. A norm $|\!|\!| \cdot |\!|\!|$ associated with the splitting (2.12) may be defined by

$$|\!|\!| u |\!|\!|^2 := \inf_{u_\ell \in V_\ell,\ u = \sum_{\ell=0}^{L} u_\ell} \sum_{\ell=0}^{L} b_\ell(u_\ell, u_\ell), \quad u \in V. \tag{2.13}$$

In the next paragraph (Section 2.3.2), we introduce the notion of stable subspace splittings by considering the relation between the energy norm and the splitting norm $|\!|\!| \cdot |\!|\!|$. This facilitates the development of a very general concept to analyze additive preconditioners by investigating specific norms of the above form. An extension to the respective multiplicative preconditioners is also achievable; see [97, 151, 155].

But first, let us present several subspace correction methods directly associated with the splitting (2.12). Considering the operators $T_\ell \in \mathrm{Lin}(V, V_\ell)$, $\ell \in \{0,\dots,L\}$, given by

$$u \mapsto T_\ell u \in V_\ell : \quad b_\ell(T_\ell u, v) = a(u, v), \quad \forall\, v \in V_\ell,$$

and the right hand side $\widetilde{f} = \sum_{\ell=0}^{L} \widetilde{f}_\ell$ defined by the variational formulation

$$b_\ell(\widetilde{f}_\ell, v) = \mathcal{F}(v), \quad \forall\, v \in V_\ell,$$

we obtain a preconditioned system $P_L u = \widetilde{f}$ with an operator $P_L \in \mathrm{Lin}(V)$ via

$$P_L = \sum_{\ell=0}^{L} T_\ell.$$

Note that in operator notation T_ℓ can be written as $B_\ell^{-1} \mathcal{Q}_\ell A_L$ with the L^2-projections

$$\mathcal{Q}_\ell : V \to V_\ell, \quad (\mathcal{Q}_\ell v, w)_{L^2(\Omega)} = (v, w)_{L^2(\Omega)}, \quad \forall\, v \in V,\, w \in V_\ell;$$

thus, we have $P_L = C_L A_L$ with $C_L = \sum_{\ell=0}^{L} B_\ell^{-1} \mathcal{Q}_\ell$.

Recall that $e^k \in V$ denotes the algebraic error of the k-th iterate, $k \geq 0$. Then, the error reduction of the above parallel subspace correction method is

$$e^{k+1} = E_L e^k := (\mathrm{id} - P_L) e^k = (\mathrm{id} - \sum_{\ell=0}^{L} T_\ell) e^k.$$

We will see that the BPX-like methods considered in Section 2.2.5 may be formulated as additive Schwarz methods in the above fashion.

The multiplicative algorithms from the previous section, namely the geometric multigrid cycles defined by Algorithm 2.1, also fit into this framework. To illustrate this, let us examine preconditioners where the approximate subspace solvers operate one after another. The corresponding error equation of the (non-symmetric) successive subspace correction method induced by the splitting (2.12) and the auxiliary operators $(B_\ell)_{\ell=0,\ldots,L}$ is given by

$$e^{k+1} = \widetilde{E}_L e^k := (\mathrm{id} - T_0) \cdots (\mathrm{id} - T_\ell) \cdots (\mathrm{id} - T_L)\, e^k. \tag{2.14}$$

Here, we use the customary convention that products (or chains) of operators are evaluated in sequential order from right to left. A symmetric multiplicative variant can be defined by

$$e^{k+1} = \widetilde{E}_L' \widetilde{E}_L e^k$$

with $(\widetilde{E}_L' v, w)_A = (v, \widetilde{E}_L w)_A$ for all $v, w \in V$; see [33]. In this case, the preconditioned operator is $(\mathrm{id} - \widetilde{E}_L' \widetilde{E}_L)$. Further, we have

$$\widetilde{E}_L' = (\mathrm{id} - T_L) \cdots (\mathrm{id} - T_\ell) \cdots (\mathrm{id} - T_0).$$

The action of this multiplicative preconditioner on a residual may be implemented as a common multigrid cycle (here, a $\mathcal{V}$-cycle) if the subspaces and the subspace operators are chosen appropriately. Note that the coarse level operator $(\mathrm{id} - T_0)$ is a projection, namely $(\mathrm{id} - T_0)^2 = (\mathrm{id} - T_0)$, if $B_0 = A_0$, i. e., if the exact subspace correction in V_0 is computed. Often, each single operator B_ℓ, $\ell \in \{1, \ldots, L\}$ represents an additive or a multiplicative subspace correction method associated with a splitting $(V_\ell, a) = \sum_{p \in \mathcal{N}_\ell} (V_{\ell,p}, b_{\ell,p})$. In case

the subspace solver in $V_{\ell,p}$ is exact, i. e., $B_{\ell,p} = A_\ell$ for $p \in \mathcal{N}_\ell$, this corresponds to "smoothing" in V_ℓ with the Jacobi or the Gauß–Seidel method, respectively; see, e. g., [194] for an elaboration of this connection.

We need to point out that the application of the above concepts is not limited to multilevel preconditioning or multigrid methods. On the contrary, the strategy of space decomposition and parallel or successive subspace correction originates from the development of the abstract Schwarz theory and is equally popular in domain decomposition methods; see, in particular, [151, 177, 194].

2.3.2 A relevant norm equivalence

A basic tool for the study of the additive preconditioner is to consider the relation between the energy norm and the splitting norm. Of course, as the space V is of finite dimension, the norms $\|\!|\cdot|\!\|$ and $\|\cdot\|_A$ are equivalent. On top of that, the fundamental observation is that the values

$$0 < \lambda_{\min} := \inf_{v\in V,\, v\neq 0} \frac{a(v,v)}{\|\!|v|\!\|^2} \le \lambda_{\max} := \sup_{v\in V,\, v\neq 0} \frac{a(v,v)}{\|\!|v|\!\|^2} < \infty,$$

which are the best possible constants in the norm equivalence

$$a(v,v) \eqsim \|\!|v|\!\|^2, \quad \forall\, v \in V, \tag{2.15}$$

are also the minimal and maximal eigenvalue $\lambda_{\min}(P_L)$ and $\lambda_{\max}(P_L)$, respectively, of the operator P_L. Thus, it is sufficient to estimate the condition number of the splitting, namely $\kappa = \lambda_{\max}/\lambda_{\min}$. This is a celebrated result from the theory of domain decomposition methods or Schwarz methods; see [151, 177, 187, 204].

As an illustration, for the BPX-preconditioner [34, 194], each $V_\ell = X_\ell$, $\ell \in \{0,\dots,L\}$, is again split into the subspaces $V_{\ell,p} = \operatorname{span}\{\lambda_p^\ell\}$, $p \in \mathcal{N}_\ell$. Then, the norm $\|\!|\cdot|\!\|$ reads as

$$\|\!|v|\!\|^2 = \inf_{v_{\ell,p}\in V_{\ell,p},\, v=\sum_{\ell=0}^{L}\sum_{p\in\mathcal{N}_\ell} v_{\ell,p}} \left(\sum_{\ell=0}^{L} \sum_{p\in\mathcal{N}_\ell} b_{\ell,p}(v_{\ell,p}, v_{\ell,p}) \right), \quad v \in V, \tag{2.16}$$

with $b_{\ell,p}(\cdot,\cdot) = 4^\ell(\cdot,\cdot)_{L^2(\Omega)}$. Here, the subspace correction in V_0 is not treated any differently; cf. (2.11) and the following comments. For more general smoothers, i. e., more complicated auxiliary forms $b_\ell(\cdot,\cdot)$, one may not have a compact representation of the form (2.11) and (2.16), respectively. If in each subspace V_ℓ, e. g., a symmetric Gauß–Seidel method is to be applied, which allows for a representation via two triangular matrices with respect to the chosen basis, a forward-backward substitution scheme needs to be incorporated.

In the early nineties, Oswald established a fundamental connection of multilevel finite elements to approximation theory. Introducing abstract scales of approximation spaces (see [150]), one can prove that certain Besov spaces coincide with the Sobolev space $H^1(\Omega)$. Applied to the above subspace splitting, the theory implies that, if adequate Jackson and Bernstein inequalities are satisfied in the respective subspaces, the norm equivalence (2.15)

holds true with constants independent of the mesh size and the number of levels. The relevant results can be found in [151, Theorem 15].

Note that the validity of the required inequalities is guaranteed for nested sequences of standard finite element spaces as considered in Section 2.2. An inequality of Jackson-type asserts suitable approximation properties whereas a Bernstein-type inequality states an inverse estimate. This immediately implies that the BPX-method and most of its variants are optimal preconditioners; see [151, Section 4.2]. In general, if one can show L-independent Jackson and Bernstein inequalities for a sequence of nested subspaces, Oswald's considerations yield a condition number bound which is independent of the mesh parameters and the number of levels of the employed hierarchy.

We also refer to [23, 34, 64, 197, 198, 199], especially for an analysis of the case of adaptively refined scales of finite element spaces. General norm equivalences have also been derived for problems with jumping coefficients in [78, 154] (if the jumps align with the coarsest mesh) and, more recently, in [98, 167, 168].

2.3.3 The theory of Schwarz methods

In this section, we briefly outline the abstract theory of Schwarz methods. For the specified additive methods, one basically needs to work on the norm equivalence (2.15); the task is to find assumptions which guarantee good constants. For the derivation of an energy estimate for the error reduction of the multiplicative methods, we follow [32, 33, 194] and [177]. See also the review paper [199].

First, we state an assumption on the auxiliary forms $(B_\ell)_{\ell=0,\dots,L}$. Let $\omega_1 > 0$ be the smallest constant such that

$$(A_\ell v, v)_{L^2(\Omega)} \leq \omega_1 (B_\ell v, v)_{L^2(\Omega)}, \quad \forall\, v \in V_\ell, \quad \forall\, \ell \in \{0, \dots, L\}. \tag{2.17}$$

In fact, $\omega_1 = \max_{0\leq\ell\leq L} \rho(B_\ell^{-1} A_\ell)$. A more classical way to formulate this "smoothing property" is

$$\rho(A_\ell)^{-1} \|v\|^2_{L^2(\Omega)} \leq \widetilde{\omega}_1 (B_\ell v, v)_{L^2(\Omega)}, \quad \forall\, v \in V_\ell \quad \forall\, \ell \in \{0, \dots, L\},$$

with another constant $\widetilde{\omega}_1 > 0$.

Second, one requires a stability estimate for the subspace splitting (2.12). Assume that there is a constant $K_0 > 0$ such that, for $v \in V$, there is a representation $v = \sum_{\ell=0}^{L} v_\ell$ with $v_\ell \in V_\ell$, $\ell \in \{0, \dots, L\}$, satisfying

$$\sum_{\ell=0}^{L} (B_\ell v_\ell, v_\ell)_{L^2(\Omega)} \leq K_0 (Av, v)_{L^2(\Omega)}. \tag{2.18}$$

In the context of multigrid methods, this may be illustrated as follows. To guarantee the existence of small upper bounds for K_0 and ω_1 at the same time, the subspace solvers B_ℓ^{-1} should be spectrally equivalent to $(\rho(A_\ell)^{-1}\,\mathrm{id})$.

Third, the interaction between the subspaces and the involved auxiliary forms is measured by strengthened Cauchy–Schwarz inequalities. We define the constant $K_1 := \rho(\boldsymbol{\epsilon}) > 0$

where $\boldsymbol{\epsilon} \in \mathbb{R}^{(L+1)\times(L+1)}$ is a symmetric matrix with positive entries such that

$$a(T_i v, T_j w) \leq \omega_1 \epsilon_{ij} a(T_i v, v)^{\frac{1}{2}}\, a(T_j w, w)^{\frac{1}{2}}, \quad \forall\, v, w \in V. \tag{2.19}$$

Slightly weaker assumptions may be found in the literature; see [199] and the references therein. Note that $\epsilon_{ij} \leq 1$ holds for all $i, j \in \{0, \ldots, L\}$. Therefore, Gershgorin's circle theorem implies at least the upper bound $\rho(\boldsymbol{\epsilon}) \leq L + 1$.

As a result, for the parallel subspace correction method, the condition number estimate

$$\kappa(P_L) \leq \omega_1 K_0 K_1$$

holds. This is because the assumptions (2.17), (2.18) and (2.19) imply the norm equivalence

$$K_0^{-1} |\!|\!| v |\!|\!|^2 \leq a(v, v) \leq \omega_1 K_1 |\!|\!| v |\!|\!|^2, \quad \forall\, v \in V.$$

Now, assume that $\omega_1 < 2$. This is equivalent to the convergence of the Richardson iteration

$$u^{k+1} = u^k + B_\ell^{-1}(f_\ell - A_\ell u^k), \quad k \geq 0,$$

for the subspace equation in V_ℓ with preconditioner B_ℓ^{-1} for all $\ell \in \{0, \ldots, L\}$. Then, the convergence behavior of the multiplicative subspace correction method given by (2.14) is estimated with respect to the energy norm by

$$\|\widetilde{E}_L\|_A^2 \leq 1 - \frac{2 - \omega_1}{K_0 (1 + \omega_1 K_1)^2}; \tag{2.20}$$

see [194, Theorem 4.4]. The result does not depend on the order of the subspace corrections.

Another way to analyze the multiplicative Schwarz algorithms can be found in [97, 151]. The authors show that the convergence rate of the multiplicative method can be directly linked to that of the additive method. Thus, it is sufficient to find the best possible constants $\lambda_{\min}$ and $\lambda_{\max}$ in the norm equivalence (2.15).

2.3.4 Convergence estimates for multigrid algorithms

Having established the basic concepts of the Schwarz theory, let us now return to the setting of Section 2.2. In particular, we consider the sequence

$$X_0 \subset \ldots \subset X_L =: V.$$

The key to the successful application of the abstract results concerning the multiplicative Schwarz methods as described in Section 2.3.3, which has been developed by Bramble, Pasciak, Wang and Xu in [32, 33], is to find suitable subspaces $V_\ell \subset X_\ell$, $\ell \in \{0, \ldots, L\}$, and establish the required assumptions. For this purpose, the following L^2-orthogonal decomposition was first considered in [33]. Using the L^2-projections $\mathcal{Q}_\ell : X_L \to X_\ell$ to the finite element subspaces $X_\ell \subset X_L$, $\ell \in \{0, \ldots, L-1\}$, the splitting reads as

$$V_\ell := (\mathcal{Q}_\ell - \mathcal{Q}_{\ell-1})\, V := \{\mathcal{Q}_\ell v - \mathcal{Q}_{\ell-1} v \mid v \in V\} \subset X_\ell, \quad \ell \in \{0, \ldots, L\}, \tag{2.21}$$

with $\mathcal{Q}_L := \mathrm{id}$ and $\mathcal{Q}_{-1} := 0$. Then, under the assumption that the operators $(S_\ell)_{\ell=1,\dots,L}$ satisify the smoothing properties

$$(S_\ell v, v)_{L^2(\Omega)} \eqsim \|h_\ell v\|^2_{L^2(\Omega)}, \quad \forall\, v \in V_\ell, \tag{2.22}$$

the uniform convergence of the multigrid method (Algorithm 2.2) by the rate given in (2.20) may be proved; see [33, Section 4] and [199, Section 7]. In particular, the multigrid cycle (Algorithm 2.1) yields a preconditioner with a condition number estimate independent of h_ℓ. Earlier, a-orthogonal decompositions have been used in combination with additional regularity assumptions; see [199] for a comparison.

In fact, a main result of [33] is that it is sufficient to find a set of (fine-to-coarse) operators $\mathcal{Q}^X_\ell : X_L \to X_\ell$ such that

$$|\mathcal{Q}^X_\ell v|_{H^1(\Omega)} \lesssim |v|_{H^1(\Omega)}, \quad \forall\, v \in X_L, \tag{2.23}$$

and

$$\|h_\ell^{-1}(v - \mathcal{Q}^X_\ell v)\|_{L^2(\Omega)} \lesssim |v|_{H^1(\Omega)}, \quad \forall\, v \in X_L, \tag{2.24}$$

for $\ell \in \{0, \dots, L-1\}$. Then, one considers a splitting analogous to the L^2-orthogonal one in (2.21) to establish the assumptions of the previous discussion from (2.22), (2.23), (2.24); this is formulated in [33, Theorem 1]. We study the above properties in detail in Section 3.4.

At this point, let us remark that the theory also covers hierarchies of finite element spaces generated by adaptive refinement. This follows essentially from the fact that, by (2.21) and (2.22), it is sufficient that the subspace solvers act on the subdomains of refinement. We do not go into detail here but take the liberty of referring to [33] and the references given in the comments at the end of Section 2.3.2.

2.4 Remarks on robustness

Multigrid methods are widely used and adaptations for various problem classes exist in the literature. A modern and comprehensive monograph presenting many advanced problems and possible cures is [178]. Let us indicate only very few points here.

The standard prolongation operator, namely the natural embedding acting on the finite element functions, for instance, by linear or trilinear interpolation, is very simple. The overall algorithm can be expected to be efficient only if the smoothing procedure effectively makes the error geometrically smooth; the remaining error should vary slowly in space, i. e., be low-frequent with respect to the current finite element mesh. As indicated in Section 2.1, for isotropic problems with moderately varying coefficients, standard iterations like the damped Jacobi or Gauß–Seidel method are very successful at this. However, one may need more sophisticated ingredients in case of other problem classes.

In principle, there are two different approaches. On the one hand, an adaptation of the smoothing operators keeping the standard multilevel finite element hierarchy may lead to an efficient method if one achieves that the remaining error is indeed geometrically smooth. On the other hand, one might attempt to adapt the hierarchy itself keeping a standard relaxation scheme such that the error which is “slow-to-converge” is well approximated in

the range of the prolongation. The latter is the paradigm of algebraic multigrid methods; we refer to [159, 176] for an introduction.

For instance, in case of jumping coefficients, appropriate techniques may involve a matrix- or operator-dependent prolongation; see, e. g., [7, 67]. An adequate coarse level hierarchy for anisotropic problems can be constructed by semi-coarsening strategies, at least in case the structure is sufficiently simple. In this thesis, alongside the development of a semi-geometric framework (Chapter 3), we cannot investigate every robustness issue described above. In particular, we do not aspire to cover the entire spectrum of possible applications. But at least one interesting topic is addressed which is beyond the pure semi-geometric multilevel methods. As indicated in Section 1.2, we will demonstrate that a semi-geometric monotone multigrid method for variational inequalities may be constructed in a rather straightforward way. This circumstance and ongoing experiments make us confident that the semi-geometric concepts developed in this thesis are capable of being extended to other problem classes and combined with various techniques from the literature.

Systems of partial differential equations

Naturally, the theory described in this chapter also applies to systems of partial differential equations. To illustrate this, let $\boldsymbol{V} := (X_L)^d$ be a vector-valued space. Then, for instance, a suitable splitting for the BPX-type preconditioner may be realized as decomposition into d-dimensional subspaces, namely

$$(\boldsymbol{V}, a) = \sum_{\ell=0}^{L} \sum_{p \in \mathcal{N}_\ell} (\boldsymbol{V}_p, b_p)$$

with

$$\boldsymbol{V}_p = \operatorname{span}\{\lambda_p^\ell \boldsymbol{e}_i \,|\, 1 \le i \le d\}, \quad p \in \mathcal{N}_\ell,\ \ell \in \{0, \dots, L\},$$

or by considering the one-dimensional subspaces in the triple sum

$$(\boldsymbol{V}, a) = \sum_{\ell=0}^{L} \sum_{p \in \mathcal{N}_\ell} \sum_{i=1}^{d} (\boldsymbol{V}_{p,i}, b_{p,i})$$

with

$$\boldsymbol{V}_{p,i} = \operatorname{span}\{\lambda_p^\ell \boldsymbol{e}_i\}, \quad 1 \le i \le d,\ p \in \mathcal{N}_\ell,\ \ell \in \{0, \dots, L\}.$$

For the standard smoothing operators in multigrid cycles such as the damped Jacobi or Gauß–Seidel methods, the first approach is called collective relaxation whereas the second one is referred to as decoupled relaxation. Note that much more general block relaxation schemes are possible; see, e. g., [178].

3 Semi-geometric multilevel preconditioners

This chapter is about a novel multilevel preconditioning strategy for linear problems arising from finite element discretizations. We introduce a class of semi-geometric methods which are based on a hierarchy of unrelated, especially non-nested meshes. We investigate the characteristics of the approach in full detail and provide a complete analysis.

In numerical simulations based on finite elements, one may take advantage of unstructured meshes to adapt the discretization to complicated geometries particularly appearing in many important applications in computational engineering. As such meshes in general do not allow for straightforward multilevel hierarchies, the disadvantage usually experienced is that iterative solvers may be less efficient or even not applicable at all in this case. The idea of accelerating an iteration by a correction from non-nested meshes has probably first been used in practice by aerospace engineers as early as the year 1986; see [143, 144] and [139]. It has also found its way into domain decomposition methods in [43, 48, 49, 51, 52, 195].

Some related techniques are studied in Chapter 4. In a way, the semi-geometric concept reflects the presumably weakest setting compared with other geometry-based approaches; the relations between the employed meshes are particularly loose. We still achieve a proof that the constructed multigrid methods based on non-nested meshes converge uniformly with respect to the discretization parameter. In contrast to, e. g., the methods to be reviewed in Section 4.1, we wish to keep the original fine level discretization, with respect to which the problem is stated, and do not change the corresponding mesh or the finite element space. In this case, it is generally a challenging task to construct suitable coarse level approximations. However, the proposed method allows for a rather direct control of the additional coarse degrees of freedom; the little geometric information entering the setup leads to a very efficient multilevel hierarchy.

At this point, the construction of the semi-geometric space hierarchy employs generic prolongation operators which are specified in Chapter 5. Fundamental requirements for the information transfer between non-nested finite element meshes are yet derived by the present analysis.

We start with some introductory remarks to advance from the standard nested setting to a non-nested multilevel framework. An outline of the rest of the chapter is given at the end of Section 3.1.

3.1 Introduction into the semi-geometric framework

In this section, we carefully prepare the derivation of the multilevel preconditioners based on non-nested meshes. For this purpose, some of the concepts and a few notations from the previous sections (especially, Section 1.3) need to be adjusted. There are several important differences to the standard multigrid setting.

Let $\mathcal{T}_L$ be a shape regular mesh of Ω such that the problems introduced in Chapter 1 are set in the associated finite element space X_L. Now, we consider a set of non-nested shape regular meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$. We emphasize that the meshes do not need to be quasi-uniform for both the construction of the multilevel methods and their analysis. However, we will return to this question when investigating possible concepts for the information transfer between non-nested meshes in more detail.

It is important to remark that each mesh represents a possibly different domain. More precisely, the coarse meshes $(\mathcal{T}_\ell)_{\ell<L}$ do not necessarily represent Ω; see Remark 1.5. However, at the finest level L, we keep the computational domain $\Omega_L := \Omega$ on which the problem to be solved has initially been defined. At the coarser levels, the domains $(\Omega_\ell)_{\ell<L}$ may be introduced by

$$\Omega_\ell := \operatorname{int}\left(\bigcup\{\overline{T} \mid T \in \mathcal{T}_\ell\}\right), \quad \ell \in \{0,\dots,L-1\}.$$

According to Section 1.3, the space $X_L \subset H^1_D(\Omega_L)$ has the Dirichlet boundary conditions on $\Gamma_D \subset \partial\Omega_L$ incorporated. As, in general, $\Omega_\ell \neq \Omega_L$ for $\ell < L$, not only the Dirichlet values but even the $(d-1)$-dimensional set Γ_D cannot be described exactly at each level. Therefore, let $(X_\ell)_{\ell=0,\dots,L-1}$ be the standard finite element spaces associated with the meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ without any boundary modifications such that $\mathcal{N}_\ell$ is the set of all nodes including all boundary nodes. Consequently, the functions which are constant in Ω_ℓ are contained in X_ℓ for $\ell \in \{0,\dots,L-1\}$.

The goal is to develop a multilevel method for the iterative solution of the discrete problems in X_L based on the non-nested coarse meshes. For the semi-geometric preconditioning framework built from the auxiliary spaces $(X_\ell)_{\ell<L}$, the recursive structure implies that it is most reasonable to assume

$$\Omega_0 \supset \dots \supset \Omega_L = \Omega. \tag{3.1}$$

We discuss the treatment of the boundaries and especially the boundary conditions in full detail in Section 3.3 and Section 3.4.

Note that, in case of non-nested meshes, the finite element spaces are non-nested, too. As a consequence, a basis function λ_p^ℓ associated with a node $p \in \mathcal{N}_\ell$, in general, cannot be written as a linear combination of the form

$$\sum_{q\in\mathcal{N}_{\ell+1}} \alpha_{pq}\lambda_q^{\ell+1},$$

which is an essential ingredient in standard multigrid methods. Despite this issue, the heuristics which led to the invention of multigrid methods in the first place, namely the role of the individual spaces as a decomposition of frequencies, may not be affected. In other applications, a non-nested hierarchy of spaces may result from a non-conforming (finite element) approximation. We do not consider this kind of non-nestedness here.

Outline

The rest of this chapter is organized as follows. We develop a suitable space hierarchy and the associated coarse level approximations by a variational approach in Section 3.2. The multilevel methods considered in Chapter 2 are reformulated with respect to the novel

spaces. This includes a setup phase as the construction of the coarse level operators is not standard. In Section 3.3, we state basic properties of the presented approach concerning the resolution of boundaries and boundary conditions by the coarse level spaces. Section 3.4 is the main part of this chapter containing the convergence analysis of the proposed methods. In Section 3.5, we comment on the application of the semi-geometric concept to two-level overlapping Schwarz methods in the context of the domain decomposition paradigm. Finally, Section 3.6 is concerned with the implementation of the developed framework.

As indicated before, the information transfer between non-nested finite element spaces, both the issue as such and the application to the present framework, is studied in detail in Chapter 5.

3.2 Multilevel preconditioners based on non-nested meshes

This section aims at a semi-geometric preconditioning framework. We introduce additive and multiplicative multilevel preconditioners in the fashion of the methods described in Section 2.2. The delicate point, though, is the construction of an appropriate hierarchy of spaces from the originally unrelated spaces $(X_\ell)_{\ell=0,\dots,L}$. This has to be done in a way which allows for a powerful convergence analysis as well as an efficient implementation.

3.2.1 Construction of a space hierarchy with multilevel bases

For this purpose, let the spaces $(X_\ell)_{\ell=0,\dots,L}$ be connected by the prolongation operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$, namely

$$\Pi_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell, \quad \forall\, \ell \in \{1,\dots,L\}.$$

As before, we carefully distinguish between operators and matrices. The analysis of a set of suitable linear operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ will be a key issue of this thesis. This is because the chosen transfer concept is a major ingredient of the construction of the space hierarchy $(V_\ell)_{\ell=0,\dots,L}$ which is presented in this section. The generated spaces are significantly influenced by this data.

As we will use the notion of the "range of an operator" repeatedly throughout this thesis, let us make this perfectly clear.

Definition 3.1 (Range or image)**.** *Let X, Y be sets and $\Pi : X \to Y$ a mapping. For a (not necessarily strict) subset $Z \subset X$, the output of Π restricted to Z is formally denoted by*

$$\Pi Z := \Pi(Z) := \{y \in Y \mid \exists\, z \in Z,\, y = \Pi z\} \subset Y.$$

If the domain, i. e., the argument Z, is clear from the context, $\Pi(Z)$ may be called in words the range (or image) of Π.

Furthermore, we introduce a product notation for the composition of linear operators formally by induction. For suitable data v, let

$$\Pi_{k'-1}^{k'} \cdots \Pi_{k-1}^{k} v := \Pi_{k'-1}^{k'} (\Pi_{k'-2}^{k'-1} \cdots \Pi_{k-1}^{k} v) \tag{3.2}$$

for $k' > k$; the initial step is

$$\Pi_{k-1}^{k} \cdots \Pi_{k-1}^{k} v := \Pi_{k-1}^{k} v. \tag{3.3}$$

Now, let $V_L := X_L$; we emphasize that the fine space will not be touched in the present framework. This may be different for some of the approaches reviewed in Section 4.1. We construct a nested sequence of spaces $(V_\ell)_{\ell=0,\dots,L}$ via

$$V_\ell := \Pi_{L-1}^{L} \cdots \Pi_{\ell}^{\ell+1} X_\ell, \quad \forall\, \ell \in \{0, \dots, L-1\}.$$

That way, the images of certain compositions of the given operators determine the space hierarchy. We have, indeed, that

$$V_L \supset V_{L-1} \supset \ldots \supset V_0$$

because by construction

$$\Pi_{\ell-1}^{\ell} X_{\ell-1} \subset X_\ell, \quad \forall\, \ell \in \{1, \dots, L\},$$

and, thus,

$$V_{\ell-1} = (\Pi_{L-1}^{L} \cdots \Pi_{\ell}^{\ell+1})(\Pi_{\ell-1}^{\ell} X_{\ell-1}) \subset (\Pi_{L-1}^{L} \cdots \Pi_{\ell}^{\ell+1})(X_\ell) = V_\ell, \quad \forall\, \ell \in \{1, \dots, L-1\}.$$

With the nodal bases of the finite element spaces $X_{\ell-1}$ and X_ℓ a matrix representation $\mathbf{\Pi}_{\ell-1}^{\ell} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ of $\Pi_{\ell-1}^{\ell}$ can be computed for $\ell \in \{1, \dots, L\}$ via $\mathbf{\Pi}_{\ell-1}^{\ell} \boldsymbol{v} := \Phi_\ell^{-1}(\Pi_{\ell-1}^{\ell} \Phi_{\ell-1}(\boldsymbol{v}))$ for all $\boldsymbol{v} \in \mathbb{R}^{n_{\ell-1}}$. Here, $\Phi_\ell : \mathbb{R}^{n_\ell} \to X_\ell$ is the coordinate isomorphism introduced in (1.20). Assume that these matrices have full rank. Then, a basis of V_{L-1} can be written as

$$\widetilde{\lambda}_q^{L-1} := \sum_{p \in \mathcal{N}_L} (\mathbf{\Pi}_{L-1}^{L})_{pq} \lambda_p^L, \quad \forall\, q \in \mathcal{N}_{L-1}, \tag{3.4}$$

and, finally, a basis of V_ℓ for $\ell \in \{0, \dots, L-2\}$ can recursively be defined by

$$\widetilde{\lambda}_q^{\ell} := \sum_{p \in \mathcal{N}_{\ell+1}} (\mathbf{\Pi}_{\ell}^{\ell+1})_{pq} \widetilde{\lambda}_p^{\ell+1}, \quad \forall\, q \in \mathcal{N}_\ell. \tag{3.5}$$

For convenience, we set $\widetilde{\lambda}_q^L := \lambda_q^L$ for $q \in \mathcal{N}_L$. Moreover, the new coordinate isomorphisms with respect to the bases $\widetilde{\Lambda}_\ell := (\widetilde{\lambda}_p^\ell)_{p \in \mathcal{N}_\ell}$, $\ell \in \{0, \dots, L\}$, will be denoted by $\widetilde{\Phi}_\ell : \mathbb{R}^{n_\ell} \to V_\ell$.

In this manner, basis functions at level $\ell - 1$ are nothing but linear combinations of basis functions at level ℓ induced by the operator $\Pi_{\ell-1}^{\ell}$. For a simple one-dimensional setting, the constructed bases are exemplarily illustrated in Figure 3.1. Note that the mapping $\Pi_{\ell-1}^{\ell}$ between the given spaces $X_{\ell-1}$ and X_ℓ usually does not act on $V_{\ell-1}$ directly. Still, the matrix $\mathbf{\Pi}_{\ell-1}^{\ell}$ determines an operator $\widetilde{\Pi}_{\ell-1}^{\ell} : V_{\ell-1} \to V_\ell$ by

$$v \mapsto \widetilde{\Pi}_{\ell-1}^{\ell} v := \widetilde{\Phi}_\ell(\mathbf{\Pi}_{\ell-1}^{\ell} \widetilde{\Phi}_{\ell-1}^{-1}(v)), \quad \forall\, v \in V_{\ell-1}, \quad \forall\, \ell \in \{1, \dots, L\}. \tag{3.6}$$

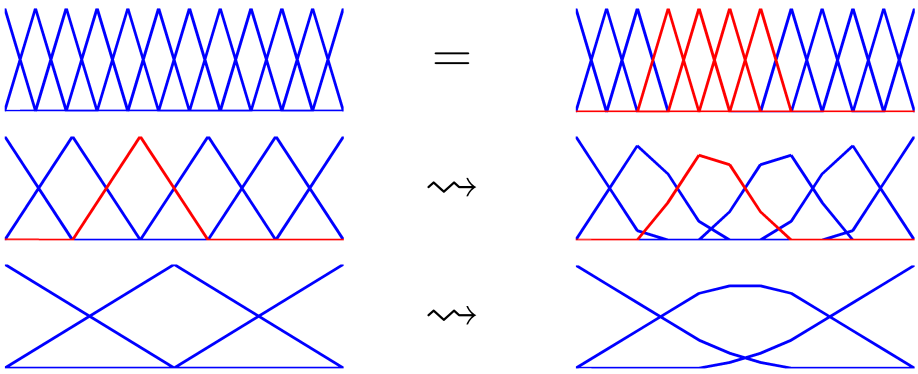


Figure 3.1. One-dimensional example: Bases of the original non-nested spaces $X_L \not\supset X_{L-1} \not\supset X_{L-2}$ (left) and of the constructed nested spaces $V_L \supset V_{L-1} \supset V_{L-2}$ (right). A nodal function λ_p^{L-1}, the respective function $\widetilde{\lambda}_p^{L-1}$, and the fine space functions involved in the corresponding linear combination are highlighted.

One can easily see that $\widetilde{\Pi}_{\ell-1}^{\ell}$ is the natural embedding because it interpolates the respective basis exactly. More precisely, by the definitions in (3.4) and (3.5), we have

$$\begin{aligned}\widetilde{\Pi}_{\ell-1}^{\ell}\widetilde{\lambda}_q^{\ell-1} &= \widetilde{\Phi}_\ell(\boldsymbol{\Pi}_{\ell-1}^{\ell}\widetilde{\Phi}_{\ell-1}^{-1}(\widetilde{\lambda}_q^{\ell-1})) = \widetilde{\Phi}_\ell(\boldsymbol{\Pi}_{\ell-1}^{\ell}\boldsymbol{e}_q) = \widetilde{\Phi}_\ell \begin{pmatrix} (\boldsymbol{\Pi}_{\ell-1}^{\ell})_{1q} \\ \vdots \\ (\boldsymbol{\Pi}_{\ell-1}^{\ell})_{n_\ell q} \end{pmatrix} \\ &= \sum_{p\in\mathcal{N}_\ell} (\boldsymbol{\Pi}_{\ell-1}^{\ell})_{pq}\widetilde{\lambda}_p^{\ell} = \widetilde{\lambda}_q^{\ell-1}, \quad \forall\, q \in \mathcal{N}_{\ell-1},\end{aligned}$$

for $\ell \in \{1,\dots,L\}$. Thus, we can regard the matrix $\boldsymbol{\Pi}_{\ell-1}^{\ell} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ as an algebraic representation of the natural embedding of $V_{\ell-1}$ into V_ℓ. Consequently, the L^2-projection from V_ℓ to $V_{\ell-1}$ is, as before, represented by the matrix

$$\boldsymbol{M}_{\ell-1}^{-1}(\boldsymbol{\Pi}_{\ell-1}^{\ell})^T\boldsymbol{M}_\ell. \tag{3.7}$$

In Section 2.2.1, we have seen that this holds true for any imaginable set of operators between the original non-nested spaces $(X_\ell)_{\ell=0,\dots,L}$; no special structure is required.

As we have chosen not to prescribe any boundary conditions at coarser levels, the basis Λ_ℓ is a partition of unity in the domain Ω_ℓ, namely

$$\sum_{q\in\mathcal{N}_\ell} \lambda_q^{\ell}(\boldsymbol{x}) = 1, \quad \forall\, \boldsymbol{x} \in \Omega_\ell,$$

for $\ell \in \{0,\dots,L-1\}$. The fine level space X_L approximates the space $H_D^1(\Omega)$ and has the Dirichlet values on Γ_D incorporated. Therefore, the basis Λ_L is a partition of unity on all elements not meeting the Dirichlet boundary. In certain circumstances, this property carries over to the new bases $(\widetilde{\Lambda}_\ell)_{\ell<L}$ as stated in the following

Lemma 3.2. *Let the row totals of the matrices* $(\mathbf{\Pi}_\ell^{\ell+1})_{\ell=0,\dots,L-1}$ *be equal to one, i. e.,* $\sum_{q\in\mathcal{N}_\ell}(\mathbf{\Pi}_\ell^{\ell+1})_{pq}=1$ *for all* $p\in\mathcal{N}_{\ell+1}$, $\ell\in\{0,\dots,L-1\}$. *Then, the new basis* $(\widetilde{\lambda}_p^\ell)_{p\in\mathcal{N}_\ell}$ *is a partition of unity in* $\Omega=\Omega_L$ *away from the Dirichlet boundary, namely*

$$\sum_{q\in\mathcal{N}_\ell}\widetilde{\lambda}_q^\ell(\boldsymbol{x})=1,\quad\forall\,\boldsymbol{x}\in\Omega\setminus\Omega_D,$$

for $\ell\in\{0,\dots,L-1\}$. *Here,*

$$\Omega_D:=\operatorname{int}\left(\bigcup\{\overline{T}\in\mathcal{T}_L\mid\overline{T}\cap\Gamma_D\neq\emptyset\}\right).$$

In any case, for $v\in V_\ell$, *one has that* $v(\boldsymbol{x})=0$ *if* $\boldsymbol{x}\notin\overline{\Omega}$.

Proof. The assertion follows by induction from the relation

$$\sum_{q\in\mathcal{N}_\ell}\widetilde{\lambda}_q^\ell=\sum_{q\in\mathcal{N}_\ell}\sum_{p\in\mathcal{N}_{\ell+1}}(\mathbf{\Pi}_\ell^{\ell+1})_{pq}\widetilde{\lambda}_p^{\ell+1}=\sum_{p\in\mathcal{N}_{\ell+1}}\left(\sum_{q\in\mathcal{N}_\ell}(\mathbf{\Pi}_\ell^{\ell+1})_{pq}\right)\widetilde{\lambda}_p^{\ell+1},$$

which is valid for $\ell\in\{0,\dots,L-1\}$, and the fact that the basis at level L itself is a partition of unity away from the Dirichlet boundary, that is in $\Omega\setminus\Omega_D$. In particular, each basis function $\widetilde{\lambda}_p^\ell$, $p\in\mathcal{N}_\ell$, is zero outside $\overline{\Omega}$. □

The condition on the row totals means that the operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$ need to preserve constants in $\Omega\setminus\Omega_D$. Note that every non-trivial projection operator clearly satisfies this assumption if the functions which are constant in $\Omega\setminus\Omega_D$ are contained in its domain. Further examples are given in Chapter 5. But, in general, the constructed coarse level bases $\widetilde{\Lambda}_\ell=(\widetilde{\lambda}_p^\ell)_{p\in\mathcal{N}_\ell}$ are not interpolatory. Although $\lambda_p^\ell(q)=\delta_{pq}$ for $p,q\in\mathcal{N}_\ell$, this property does not persist for the new basis functions of the nested spaces $(V_\ell)_{\ell=0,\dots,L}$. By construction the functions at the coarser levels are piecewise linear with respect to the finest mesh $\mathcal{T}_L$; cf. Figure 3.1.

With this information we can summarize our efforts as follows. From the completely unrelated finite element spaces $(X_\ell)_{\ell=0,\dots,L}$ we have constructed a sequence of nested spaces $(V_\ell)_{\ell=0,\dots,L}$ such that the given prolongation operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$ induce the natural embeddings $(V_{\ell-1}\hookrightarrow V_\ell)_{\ell=1,\dots,L}$ by their matrix representations $(\mathbf{\Pi}_{\ell-1}^\ell)_{\ell=1,\dots,L}$ with respect to the original bases $(\Lambda_\ell)_{\ell=0,\dots,L}$. In particular, the coarse level matrices for the nested spaces with the respective bases $\widetilde{\Lambda}_\ell$, as customary in a variational approach, can be written as

$$\boldsymbol{A}_{\ell-1}=(\mathbf{\Pi}_{\ell-1}^\ell)^T\boldsymbol{A}_\ell\,\mathbf{\Pi}_{\ell-1}^\ell,\quad\forall\,\ell\in\{1,\dots,L\};$$

see Section 2.2.1. If $\boldsymbol{A}_L$ is symmetric positive definite and if $\mathbf{\Pi}_{\ell-1}^\ell$ has full rank for all $\ell\in\{1,\dots,L\}$, the respective coarse level matrices $(\boldsymbol{A}_\ell)_{\ell=0,\dots,L-1}$ are symmetric positive definite, too. Note that the bandwidth of the coarse matrices highly depends on the transfer concept employed to obtain the prolongation operators.

As the coarse level equations are merely set up as auxiliary problems, similar to (purely) algebraic multigrid approaches, the operators $(A_\ell)_{\ell<L}$ may usually not be considered discretizations or Galerkin approximations in proper finite element spaces of the original problem set in some Sobolev space. However, under certain conditions, one might be able to

prove approximation results in terms of a priori discretization error estimates for the constructed spaces $(V_\ell)_{\ell<L}$.

3.2.2 Semi-geometric multigrid methods

Let us now proceed to the prospective multilevel preconditioners. We point out that, again for simplicity, we state one of the simplest multiplicative Schwarz preconditioners with suitable smoothing operators $(S_\ell)_{\ell=1,\dots,L}$. Naturally, the constructed spaces $(V_\ell)_{\ell=0,\dots,L}$ and the prolongation operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ can also be used in more sophisticated methods.

In Section 2.2, a multigrid algorithm has been written purely in operator notation. We have always carefully distinguished between operators and the representing matrices with respect to selected bases and will continue doing so. Here, we prefer the matrix notations.

To start with, we adopt the notion of a setup phase, which is a paraphrase for the construction of a multilevel space hierarchy common in the area of algebraic multigrid methods; see, e. g., [176]. At this point, we state a rather general frame to merely illustrate that a certain setup is an integral part of the presented semi-geometric methods. The intermediate steps of this procedure (the "other factors") are elaborated more precisely later, as noted below.

Algorithm 3.3 (Setup semi-geometric multigrid method)**.** *Choose* type *of prolongation operator according to Chapter 5.*

```
setupSGMG (type, (T_ℓ)_{ℓ=0,...,L}) {
    for (ℓ = L, ..., 1) do {
        Compute the prolongation matrix Π_{ℓ-1}^ℓ
        Take some other factors into account
        Compute the coarse level matrix A_{ℓ-1} = (Π_{ℓ-1}^ℓ)^T A_ℓ Π_{ℓ-1}^ℓ
    }
}
```

The algorithm implicitly comprises a generation of coarse level nodes or degrees of freedom. The choice of these variables is facilitated by the usage of the nodesets $(\mathcal{N}_\ell)_{\ell<L}$. However, the coarse meshes are just auxiliary devices and one might disregard some of the nodes, i. e., some of the basis functions in $(\Lambda_\ell)_{\ell<L}$, when computing the new bases $(\widetilde{\Lambda}_\ell)_{\ell<L}$. We return to this in Section 3.3. One may consider the nodes of the given meshes as "suggested degrees of freedom". The span of a basis function λ_p^ℓ or rather of $\widetilde{\lambda}_p^\ell$ is not necessarily included in the coarse space V_ℓ, though. Moreover, the selected "transfer concept" or "type of prolongation" may require some modifications of the meshes, nodes, matrices and so forth; see Section 3.6.

Ultimately, the semi-geometric multigrid cycle appears as standard multigrid cycle with respect to the spaces $(V_\ell)_{\ell=0,\dots,L}$. Therefore, we obtain the following multilevel preconditioner in matrix notation by an application of the ideas of Section 2.2.3 to the novel space hierarchy.

Algorithm 3.4 (Multigrid cycle)**.** *For (the residual vector)* $\boldsymbol{R} \in \mathbb{R}^{n_\ell}$ *compute the correction*

$$\boldsymbol{C}_\ell \boldsymbol{R} = \mathrm{SGMG}_\ell^{\gamma,\nu_1,\nu_2}(\boldsymbol{R}) = \mathrm{SGMG}\ (\ell,\ \gamma,\ \nu_1,\ \nu_2,\ \boldsymbol{R}) \in \mathbb{R}^{n_\ell}$$

by the following procedure.

SGMG $(\ell,\, \gamma,\, \nu_1,\, \nu_2,\, \boldsymbol{R})$ {
  if $(\ell = 0)$ {
    Solve exactly: $\boldsymbol{x} \leftarrow \boldsymbol{A}_0^{-1}\boldsymbol{R}$
  }
  else {
    Pre-smoothing steps: $\boldsymbol{x} \leftarrow \boldsymbol{S}_\ell^{\nu_1}(\boldsymbol{0}, \boldsymbol{R})$
    Coarse level correction:
      Restriction: $\boldsymbol{R}' \leftarrow (\boldsymbol{\Pi}_{\ell-1}^{\ell})^T(\boldsymbol{R} - \boldsymbol{A}_\ell \boldsymbol{x})$
      Initialize: $\boldsymbol{x}' \leftarrow \boldsymbol{0}$
      for $(i = 1, \ldots, \gamma)$ **do** {
        Recursive call: $\boldsymbol{x}' \leftarrow \boldsymbol{x}' + \mathrm{SGMG}\ (\ell - 1,\, \gamma,\, \nu_1,\, \nu_2,\, \boldsymbol{R}')$
      }
      Prolongation: $\boldsymbol{x} \leftarrow \boldsymbol{x} + \boldsymbol{\Pi}_{\ell-1}^{\ell}\boldsymbol{x}'$
    Post-smoothing steps: $\boldsymbol{x} \leftarrow \boldsymbol{S}_\ell^{\nu_2}(\boldsymbol{x}, \boldsymbol{R})$
  }
  return $\boldsymbol{x}$
}

Note that it is straightforward to rewrite the above algorithms in operator notation. In fact, Algorithm 3.4 for the semi-geometric multigrid cycle is basically equivalent to Algorithm 2.1 for the geometric one. The only aspect worth mentioning is that the action of the matrix $\boldsymbol{\Pi}_{\ell-1}^{\ell}$ on vectors representing finite element functions in $V_{\ell-1}$ corresponds to the operator $\widetilde{\Pi}_{\ell-1}^{\ell}$ from (3.6). Moreover, the action of $(\boldsymbol{\Pi}_{\ell-1}^{\ell})^T$ on residual vectors in $\mathbb{R}^{n_\ell}$ is part of the fine-to-coarse L^2-projection as discussed before in (3.7) and more detailed in Section 2.2.1.

As all operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\ldots,L}$ employed in the construction of the space hiearchy are linear, one can always find proper matrices $(\boldsymbol{\Pi}_{\ell-1}^{\ell})_{\ell=1,\ldots,L}$. But note that, in case the transfer concept involves inverting local or global (mass) matrices, prolongation and restriction may also be realized differently. In this case, it may be advantageous to keep the transfer routines separate such that a multigrid cycle does not comprise just one matrix-vector multiplication but rather suitable local or global forward-backward substitution procedures in each transfer step. Although it is difficult to make a clear statement about what is the more efficient way to implement the prolongation and restriction routines for extraordinary transfer concepts, we believe that it will usually be the precomputation of the matrices $(\boldsymbol{\Pi}_{\ell-1}^{\ell})_{\ell=1,\ldots,L}$.

3.2.3 Additive semi-geometric preconditioners

The following parallel subspace correction methods in the spirit of Section 2.2.5 are interesting in our context because the mappings of the non-nested finite element spaces into the finest space X_L can be realized in two different (likewise straightforward) ways. For an additive variant, one needs prolongation operators

$$\Pi_\ell^L : X_\ell \to X_L, \quad \forall\, \ell \in \{0, \ldots, L-1\},$$

mapping straight into the space X_L. Then, the corresponding subspaces $(W_\ell)_{\ell=0,\dots,L-1} \subset W_L := X_L$ read as

$$W_\ell := \Pi_\ell^L X_\ell, \quad \forall\, l \in \{0, \dots, L-1\}.$$

Given the original bases $(\Lambda_\ell)_{\ell=0,\dots,L}$, the coarse level bases and matrices with respect to the spaces $(W_\ell)_{\ell=0,\dots,L}$ are defined as before in Section 3.2.1 with the obvious adaptations. Then, an additive preconditioner associated with the multilevel hierarchy $(W_\ell)_{\ell=0,\dots,L}$ can be written as a sum.

Algorithm 3.5. *For (the residual vector) $\boldsymbol{R} \in \mathbb{R}^{n_L}$ compute the correction $\boldsymbol{C}_L\boldsymbol{R}$ as*

$$\boldsymbol{C}_L\boldsymbol{R} = \mathrm{SGMG}_{\mathrm{add}}(L,\, \nu,\, \boldsymbol{R}) = \boldsymbol{\Pi}_0^L \boldsymbol{A}_0^{-1} (\boldsymbol{\Pi}_0^L)^T \boldsymbol{R} + \sum_{\ell=1}^{L-1} \boldsymbol{\Pi}_\ell^L \boldsymbol{S}_\ell^\nu(\boldsymbol{0}, (\boldsymbol{\Pi}_\ell^L)^T \boldsymbol{R}) + \boldsymbol{S}_L^\nu(\boldsymbol{0}, \boldsymbol{R}).$$

Here, S_ℓ is a suitable symmetric smoothing operator in the subspace W_ℓ, $\ell \in \{0, \dots, L\}$, and thus $\boldsymbol{S}_\ell$ acts on $\mathbb{R}^{n_\ell}$. For academic reasons, one may also consider a modification of Algorithm 3.5 where the fine smoothing (at level L) is performed before the restriction and after the prolongation. The modified algorithm should be expected to yield better preconditioning results at the cost of a less parallel structure. This expectation is due to the fact that the coarse level spaces yield better approximations of the defect problems in case of less oscillating errors.

To illustrate the appearance of the new basis functions, let us consider a BPX-like preconditioner in operator notation. If the coarsest auxiliary problem in Algorithm 3.5 is not solved exactly and, in addition, the smoothers $(S_\ell)_{\ell=0,\dots,L}$ have a natural parallel structure such as diagonal scaling or the Jacobi method, the subspaces $(W_\ell)_{\ell=0,\dots,L}$ can further be split allowing for even more parallelism. In this case, the additive preconditioner can have, e. g., the following form,

$$C_L r = \sum_{\ell=0}^{L} \sum_{p \in \mathcal{N}_\ell} \frac{(r, \widetilde{\lambda}_p^\ell)_{L^2(\Omega)}}{a(\widetilde{\lambda}_p^\ell, \widetilde{\lambda}_p^\ell)} \widetilde{\lambda}_p^\ell.$$

This is the additive (diagonal scaling) preconditioner with respect to the multilevel nodal basis $\{\widetilde{\lambda}_p^\ell \mid p \in \mathcal{N}_\ell,\, \ell \in \{0, \dots, L\}\}$; see Section 2.2.5.

In the rest of this paragraph, we comment on the setup phase of the additive semi-geometric method. The central observation is that the prolongation operators $(\Pi_\ell^L)_{\ell=0,\dots,L-1}$ can be realized in two different ways.

Given some operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$, one can choose to define the operator Π_ℓ^L via $\Pi_{L-1}^L \cdots \Pi_\ell^{\ell+1}$. But this sort of construction is not mandatory. In fact, if the operator Π_ℓ^L is constructed directly, namely in a similar algorithmic manner as the various operators to be presented in Chapter 5, we do not have $W_\ell = V_\ell$. An equivalence like this can only hold true in special cases, e. g., if the meshes are nested. This is because without any relation between the meshes one cannot prohibit a loss of information when interpolating from one space to another. Although the considered finite element spaces $(X_\ell)_{\ell=0,\dots,L}$ and the connecting operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$ are assumed to satisfy some relevant approximation properties, this does not mean that every bit of local information can be transferred to the

next space. This stronger property is supposably only valid for nested meshes. But even in the nested case, this depends on the choice of the transfer concept; it is only guaranteed for special operator types.

Evidently, the setup of the first variant is the same recursive one as stated in Algorithm 3.3. In contrast, the obvious non-recursive ("immediate") routine for the computation of the matrices $(\mathbf{\Pi}_\ell^L)_{\ell=0,\dots,L-1}$ and $(\boldsymbol{A}_\ell)_{\ell=0,\dots,L-1}$ from the meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L}$ may be called setupSGMG$_{\text{imm}}$.

In practical computations, even if the coarse level matrices obtained by Galerkin assembly have very similar structures, the approach with "immediate" mappings $(\Pi_\ell^L)_{\ell=0,\dots,L-1}$ tends to need a little more memory. This is due to the fact that the total number of entries of the prolongation matrices $(\mathbf{\Pi}_\ell^L)_{\ell=0,\dots,L-1}$ may be considerably larger, depending on the chosen transfer concept. However, the advantage of this approach may be a little extra flexibility of the generation of the coarse meshes. This is because it obviates the need for ensuring reasonable relations between two consecutive meshes.

In the successive approach, the operators are assembled in a recursive fashion. Only the single matrices $(\mathbf{\Pi}_\ell^{\ell+1})_{\ell=0,\dots,L-1}$ need to be computed. Then, the evaluation of prolongation and restriction is done by applying the operators one after the other. In fact, admittedly, this disagrees with the additive structure of the overall algorithm. We examine whether the difference of the two approaches is significant in practical computations in Chapter 6.

3.3 Coarse representation of boundaries and boundary conditions

In this section, we explain the semi-geometric multilevel hierarchy in more detail, complementing the introductory remarks from Section 3.1. The treatment of boundary values for the coarse level problems is also reconsidered. Here, the focus lies on an illustrative presentation and useful practical considerations. We take a more analytic viewpoint next.

For the development of the multilevel preconditioning framework, we have chosen to use a quite general set of non-nested meshes. The desire to detach the generation of coarse meshes (and spaces) as far as possible has several reasons. First of all, it allows for employing independent mesh generation processes, possibly of different nature, at each level. The coarse domains $(\Omega_\ell)_{\ell<L}$ can be considerably different from the fine domain Ω_L. This may be beneficial, for example, to acquire coarse meshes of simpler shapes, perhaps even meshes with some regular structures. In other cases, if the computational domain Ω_L varies due to (pseudo-)time stepping or changes slightly during (shape) optimization, it may be of practical use to reuse the coarse level meshes. Finally, we emphasize that the analysis of the presented multilevel methods, which is carried out in the next section, does not require very restrictive assumptions on the particular interaction of two successive meshes. Of course, the case that the coarse meshes are nested is included.

To retain the capability to capture the behavior of the functions under consideration in the whole computational domain, each domain in the sequence $(\Omega_\ell)_{\ell=0,\dots,L}$ should be covered by all other domains which are used to provide coarse level information. For the methods with a recursive structure of the information transfer, this is reflected in (3.1).

For the additive method with immediate mappings $(\Pi_\ell^L)_{\ell=0,\dots,L-1}$, this condition becomes weaker as it is sufficient to assume $\Omega_\ell \supset \Omega_L$ for $\ell \in \{0,\dots,L-1\}$.

It is important to note that the constructed coarse level spaces resolve the boundary of the computational domain in a certain sense. This is an immediate consequence of the fact that the bases $(\widetilde{\Lambda}_\ell)_{\ell=0,\dots,L-1}$ are defined by linear combinations of basis functions in Λ_L. For quite a few transfer concepts, each basis $\widetilde{\Lambda}_\ell$ is a partition of unity in Ω_ℓ as asserted in Lemma 3.2. In any case, the equality

$$\Omega = \operatorname{int}\Big(\bigcup\{\boldsymbol{x} \in \mathbb{R}^d \mid \exists\, v \in V_\ell,\, v(\boldsymbol{x}) \neq 0\}\Big)$$

holds true for all $\ell \in \{0,\dots,L\}$.

As the domains $(\Omega_\ell)_{\ell=0,\dots,L}$ are principally independent of each other, the Dirichlet boundary $\Gamma_D \subset \partial\Omega_L$ of the mixed boundary value problem is assumed to be resolved by the finest mesh $\mathcal{T}_L$ only. Still, by the assumption that the range of the operator Π_{L-1}^L is in the space $X_L \subset H_D^1(\Omega_L)$, the Dirichlet conditions are incorporated into all coarse spaces in a very natural way. As a general rule, as we only study the variational approach in this thesis, namely the operators at the coarser levels are entirely defined by a Galerkin relation, the coarse space problems do not need any special considerations of the respective boundaries. This means that all possible boundary conditions only have to be treated in the finest space X_L. In particular, the concept used to obtain the matrices $(\mathbf{\Pi}_{\ell-1}^\ell)_{\ell=1,\dots,L}$ may be rather general; no boundary modifications of the employed prolongation operators are necessary, in contrast to [171].

This paradigm (to consider coarse level basis functions as linear combinations of fine level basis funtions and then cancel undesired contributions) reminds of monotone multigrid methods for variational inequalities. There, the coarse level spaces are modified by a recursive truncation of basis functions depending on the active constraints of the current fine level iterate. This approach ensures that the coarse level correction does not violate the active constraints; see [121, 122, 124]. We explain this in a little more detail in Section 6.2. The resolution of the boundary at coarser levels by design, namely by means of suitable linear combinations of fine level functions, also resembles the composite finite element method [105, 106]; see Section 4.1.3.

For most transfer concepts, especially for the ones which yield "local" operators according to Definition 3.8 to be considered in the next section, it is reasonable to neglect the elements at level ℓ which lie completely outside of $\Omega_{\ell+1}$, $\ell \in \{0,\dots,L-1\}$. Otherwise, if the domain Ω_ℓ constituted by $\mathcal{T}_\ell$ is considerably larger than Ω_L, it can happen that $(\widetilde{\lambda}_q^\ell)_{q\in\mathcal{N}_\ell}$ is not a basis but merely a spanning set of the constructed subspace V_ℓ. In this case, the discrete representation of the composed operator $\Pi_{L-1}^L \cdots \Pi_\ell^{\ell+1}$ has zero columns. Indeed, one may also realize this procedure as a modification of the nodes depending on the transfer operators. Assuming that the matrices $(\mathbf{\Pi}_\ell^{\ell+1})_{\ell=0,\dots,L-1}$ have been computed according to some formula to be specified later, we can reduce the sets of nodes via

$$\mathcal{N}_\ell \longmapsto \{p \in \mathcal{N}_\ell \mid \exists\, q \in \mathcal{N}_{\ell+1},\, (\mathbf{\Pi}_\ell^{\ell+1})_{pq} \neq 0\}$$

for descending $\ell \in \{0,\dots,L-1\}$. Naturally, in a practical implementation the redundant degrees of freedom associated with neglected nodes are never created. The above procedure

needs to be included in the setup of the semi-geometric multilevel hierarchy setupSGMG (Algorithm 3.3).

Finally, let us remark that it is possible to slightly relax the conditions on the coarse domains such that they only need to cover the interior nodes. For the Dirichlet part of the boundary, this can be achieved without additional assumptions on the transfer operators. However, for the Neumann boundary, one needs to construct special interpolation operators [47, 91] because plain extension by zero to the part of the fine domain lying outside of the coarse domain is not sufficiently accurate. We postpone the technical discussion to the end of Section 3.4.3; see Remark 3.11 and Remark 3.13.

3.4 Quasi-optimality of the semi-geometric multilevel methods

In this section, we examine the convergence and preconditioning properties of the multilevel algorithms. We aim to investigate the new spaces and more importantly the involved operators more closely to establish conditions under which the multilevel methods from Algorithm 3.4 and Algorithm 3.5 yield good preconditioners. The constants in the norm equivalence discussed in Section 2.3.2 are not derived directly. Instead, the approach by Bramble, Pasciak, Wang, and Xu [33] as specified in Section 2.3.4 is applied to the semi-geometric situation. Ultimately, we check whether a specific choice of the spaces $(X_\ell)_{\ell=0,\dots,L}$ and of the prolongation operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ affirms the existence of suitable fine-to-coarse operators $(\mathcal{Q}_\ell^V)_{\ell=0,\dots,L-1}$ allowing for a stable splitting of the space V_L into a sum of the subspaces $(V_\ell)_{\ell=0,\dots,L}$. The truth of the matter is that the existence of such (non-trivial) fine-to-coarse mappings provides fundamental information about the smoothness of the coarse level functions.

We consider this a very natural context to examine the convergence of the methods because one directly sees where the various operators and their stability and approximation properties enter the proofs. We emphasize that the relevant estimates follow from assumptions on the original spaces $(X_\ell)_{\ell=0,\dots,L}$ and the prolongation operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ rather than on the spaces $(V_\ell)_{\ell=0,\dots,L}$. The issue of how to find transfer concepts satisfying the determined assumptions is addressed subsequently.

As we have to overcome additional technical difficulties, the proof is split into several parts. This will be stated more precisely at the end of Section 3.4.1. A special feature of our analysis is that we carefully distinguish between the different domains $(\Omega_\ell)_{\ell=0,\dots,L}$. In this, especially in the Sections 3.4.2 and 3.4.3, we cannot avoid some technical considerations. The analysis is finalized in Section 3.4.4 where we complete the proof by an application of the techniques mentioned above and make some concluding comments.

3.4.1 Stability and approximation properties

First, we introduce local mesh size functions. Recall that a mesh $\mathcal{T}_\ell$ is a non-overlapping decomposition of the computational domain into open polytopes $\{T \in \mathcal{T}_\ell\}$ satisfying certain regularity assumptions specified in Section 1.3. In case of non-quasi-uniform meshes, one

usually introduces functions $h_{\mathcal{T}_\ell} : \Omega_\ell \to \mathbb{R}_+$ reflecting the local mesh size of $\mathcal{T}_\ell$. For this purpose, for any subset $U \subset \mathbb{R}^d$, let

$$L^\infty_>(U) := \{v \in L^\infty(U) \mid \exists\, \alpha > 0, \text{ such that } v(\boldsymbol{x}) > \alpha \text{ for a. e. } \boldsymbol{x} \in U\}$$

be the space of positive functions in $L^\infty(U)$ which are bounded away from zero. Then, one may consider the standard, piecewise constant functions in $L^\infty_>(\Omega_\ell)$ defined almost everywhere by

$$h_{\mathcal{T}_\ell}(\boldsymbol{x}) := h_T, \quad \text{if } \boldsymbol{x} \in T.$$

Naturally, the values at the boundaries of the elements do not need to be specified. If pointwise evaluation of $h_{\mathcal{T}_\ell}$ is desired, continuous functions may be used, e. g., in the finite element space X_ℓ defined by

$$h_{\mathcal{T}_\ell} := \sum_{p \in \mathcal{N}_\ell} h_p \lambda_p^\ell \tag{3.8}$$

with $h_p := \max_{p \in \overline{T}} h_T$ or $h_p := \frac{\sum_{p \in \overline{T}} h_T}{|\{T \,|\, p \in \overline{T}\}|}$; see, e. g., [63]. After all, we abbreviate $h_{\mathcal{T}_\ell} =: h_\ell$ and assume that local mesh size functions $h_\ell \in L^\infty_>(\Omega_\ell)$ are given for $\ell \in \{0, \dots, L\}$.

As illustrated during the derivation of the algorithms in the previous sections, the coarse meshes shall in principle be independent of the fine mesh and of each other. However, it is convenient for the analysis and very reasonable in practical algorithms to consider the following coarsening assumption. We assume that the local relation of the mesh sizes is designed such that a constant $\sigma < 1$ exists satisfying

$$h_\ell(\boldsymbol{x}) \leq \sigma h_{\ell-1}(\boldsymbol{x}), \quad \text{for a. e. } \boldsymbol{x} \in \Omega_\ell,\ \forall\, \ell \in \{1, \dots, L\}. \tag{3.9}$$

In other words, the mesh $\mathcal{T}_{\ell-1}$ mimics a coarsening of the mesh $\mathcal{T}_\ell$. Note that the function $h_{\ell-1}$ is well-defined in Ω_ℓ in case $\Omega_0 \supset \ldots \supset \Omega_L$.

The following general definition of stability and approximation properties of operators between certain function spaces introduces notions which are crucial to the analysis of the semi-geometric multilevel methods. Here, we work with localized estimates instead of global L^2-approximation inequalities, the latter of which involving the quantity $\max_{T \in \mathcal{T}_\ell} h_T$.

Definition 3.6. *Let $\Omega_Y \subset \Omega_X \subset \mathbb{R}^d$ be domains. Given a subspace $X \subset H^1(\Omega_X)$ and a (target) finite element space $Y \subset H^1(\Omega_Y)$ with discretization parameter $h_Y \in L^\infty_>(\Omega_Y)$, an operator $\Pi : X \to Y$ is called H^1-stable in X if*

$$|\Pi v|_{H^1(\Omega_Y)} \lesssim |v|_{H^1(\Omega_X)}, \quad \forall\, v \in X.$$

We say that the operator Π satisfies the (or an) L^2-approximation property if

$$\|h_Y^{-1}(v - \Pi v)\|_{L^2(\Omega_Y)} \lesssim |v|_{H^1(\Omega_X)}, \quad \forall\, v \in X.$$

Note that the inclusions (of sets) in the above definitions do not need to be strict. Moreover, no relation between X and Y has been specified other than the fact that functions from X are also well-defined in the domain Ω_Y. Definition 3.6 constitutes a quite general concept. We will use the notions for both coarse-to-fine operators and fine-to-coarse operators. The term H^1-stability is slightly stronger than H^1-continuity, especially in the

finite-dimensional case where every linear operator is continuous (with respect to every equivalent norm), because the latter notion includes mappings with a continuity constant dependent on the mesh.

Certainly, one will not be able to find operators which are H^1-stable and satisfy the L^2-approximation property according to Definition 3.6 if Dirichlet boundary conditions are incorporated into the target space Y unless they are already satisfied in the space X. For example, consider the function $\mathbb{1} \in H^1(\Omega_X)$ which is constant with value 1 in the entire domain. As $|\mathbb{1}|_{H^1(\Omega_X)} = 0$, the only "stable" image is $\Pi\mathbb{1} = 0$ in case there are Dirichlet conditions in Y. But this violates the approximation property as $\|h_Y^{-1}(\mathbb{1} - 0)\|_{L^2(\Omega_Y)} > 0$.

This important observation also influences our analysis of the spaces $(V_\ell)_{\ell=0,\dots,L}$ as we need to prove the existence of suitable (fine-to-coarse) mappings $\mathcal{Q}_\ell^V : V_L \to V_\ell$. To make the following steps perfectly clear, we split the analysis into two parts. In Section 3.4.2, we state a preliminary result and carry out the proof assuming that the prolongation operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$ are H^1-stable and possess L^2-approximation properties according to Definition 3.6 with respect to the spaces $(X_\ell)_{\ell=0,\dots,L}$. This may be considered as an analysis of the special case without any Dirichlet conditions such that all bases $(\Lambda_\ell)_{\ell=0,\dots,L}$ are partitions of unity on their respective domains $(\Omega_\ell)_{\ell=0,\dots,L}$ because suitable operators are available (only) in this case. In Section 3.4.3, we relax the assumptions to cover the general semi-geometric setting.

3.4.2 Existence proof of suitable fine-to-coarse mappings

Let us proceed to the proof. We need to make some effort because, in general, the constructed coarse spaces $(V_\ell)_{\ell<L}$ are not standard finite element spaces.

The assumptions in the following lemma are a little too restrictive to cover the general semi-geometric setting described in Section 3.1 and Section 3.2. One may basically view this as the case without any Dirichlet boundary conditions prescribed in the finite element spaces, namely $\Gamma_D = \emptyset$. We do not aim at stating this explicitly, though. Although some additional difficulties have to be overcome in the next section, the present technical considerations give the main structure of the final proof.

We state the required inequalities explicitly to stress the different domains. Note that the estimates (3.12) and (3.13) asserted in the lemma are with respect to $\Omega_L = \Omega$. The coarse domains $(\Omega_\ell)_{\ell<L}$ appear in the assumptions (3.10) and (3.11).

Lemma 3.7. *Let $\Pi_{\ell-1}^\ell : X_{\ell-1} \to X_\ell$, $\ell \in \{1, \dots, L\}$, be H^1-stable prolongation operators, i. e.,*

$$|\Pi_{\ell-1}^\ell v|_{H^1(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_{\ell-1})}, \quad \forall\, v \in X_{\ell-1}, \tag{3.10}$$

with the L^2-approximation properties

$$\|h_\ell^{-1}(v - \Pi_{\ell-1}^\ell v)\|_{L^2(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_{\ell-1})}, \quad \forall\, v \in X_{\ell-1}. \tag{3.11}$$

Then, there are mappings $\mathcal{Q}_\ell^V : V_L \to V_\ell$, $\ell \in \{0, \dots, L-1\}$, which are also H^1-stable and have an L^2-approximation property, i. e.,

$$|\mathcal{Q}_\ell^V v|_{H^1(\Omega_L)} \lesssim |v|_{H^1(\Omega_L)}, \quad \forall\, v \in V_L, \tag{3.12}$$

and

$$\|h_\ell^{-1}(v - \mathcal{Q}_\ell^V v)\|_{L^2(\Omega_L)} \lesssim |v|_{H^1(\Omega_L)}, \quad \forall\, v \in V_L. \tag{3.13}$$

Proof. First of all, recall the product notation from equations (3.2) and (3.3). Now, let $\ell \in \{0, \dots, L-1\}$ be fixed. To construct a suitable mapping from V_L to the subspace V_ℓ, we employ the recursive structure of the definition of the multilevel space hierarchy in Section 3.2.1. For this purpose, let $E_\ell : V_L \subset H^1(\Omega_L) \to H^1(\Omega_\ell)$ be an extension operator satisfying $(E_\ell v)|_{\Omega_L} = v$ in $H^1(\Omega_L)$ and

$$|E_\ell v|_{H^1(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_L)}. \tag{3.14}$$

Such an operator exists if $\partial\Omega_L$ is Lipschitz; see, e. g., [1, Theorem 4.32]. Note that the extended function does not need to have a special form.

As X_ℓ is a standard finite element space associated with Ω_ℓ, it is not difficult to find an H^1-stable mapping $\mathcal{Q}_\ell^X : H^1(\Omega_\ell) \to X_\ell$ which satisfies the relevant L^2-approximation property

$$\|h_\ell^{-1}(v - \mathcal{Q}_\ell^X v)\|_{L^2(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_\ell)}, \quad \forall\, v \in H^1(\Omega_\ell).$$

For instance, one may employ a Clément-type quasi-interpolation operator discussed in full detail in Section 5.2. Then, we choose $\mathcal{Q}_\ell^V$ as $\Pi_{L-1}^L \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell$ which is indeed a mapping from V_L to V_ℓ. The H^1-stability of $\mathcal{Q}_\ell^V$ follows directly from the respective inequalities for the operators in the composition; for $v \in V_L$, we have

$$\begin{aligned} |\mathcal{Q}_\ell^V v|_{H^1(\Omega_L)} &= |\Pi_{L-1}^L \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v|_{H^1(\Omega_L)} \\ &\lesssim |\Pi_{L-2}^{L-1} \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v|_{H^1(\Omega_{L-1})} \lesssim \dots \\ &\lesssim |\mathcal{Q}_\ell^X E_\ell v|_{H^1(\Omega_\ell)} \lesssim |E_\ell v|_{H^1(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_L)}. \end{aligned} \tag{3.15}$$

This confirms the estimate (3.12).

Let $v \in V_L$. To prove the approximation property (3.13), we use the triangle inequality to calculate

$$\begin{aligned} \|h_\ell^{-1}(v - \mathcal{Q}_\ell^V v)\|_{L^2(\Omega_L)} &= \|h_\ell^{-1}(E_\ell v - \mathcal{Q}_\ell^V v)\|_{L^2(\Omega_L)} \\ &\le \|h_\ell^{-1}(E_\ell v - \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_L)} \\ &\quad + \|h_\ell^{-1}(\mathcal{Q}_\ell^X E_\ell v - \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_L)} \\ &\quad + \sum_{k=\ell+1}^{L-1} \|h_\ell^{-1}(\Pi_{k-1}^k \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v - \Pi_k^{k+1} \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_L)}. \end{aligned}$$

Adjusting the domains, we continue the above estimate with

$$\begin{aligned} &\le \|h_\ell^{-1}(E_\ell v - \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_\ell)} + \|h_\ell^{-1}(\mathcal{Q}_\ell^X E_\ell v - \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_{\ell+1})} \\ &\quad + \sum_{k=\ell+1}^{L-1} \|h_\ell^{-1}(\Pi_{k-1}^k \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v - \Pi_k^{k+1} \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_{k+1})}. \end{aligned}$$

Then, we rearrange the mesh size functions and make use of the estimate (3.9) to obtain

$$\begin{aligned}
&= \|h_\ell^{-1}(E_\ell v - \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_\ell)} + \|h_\ell^{-1} h_{\ell+1} h_{\ell+1}^{-1}(\mathcal{Q}_\ell^X E_\ell v - \Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_{\ell+1})} \\
&\quad + \sum_{k=\ell+1}^{L-1} \|h_\ell^{-1} h_{k+1} h_{k+1}^{-1}(\Pi_{k-1}^k \cdots \Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v - \Pi_k^{k+1}\cdots\Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_{k+1})} \\
&\leq \|h_\ell^{-1}(E_\ell v - \mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_\ell)} + \sigma\, \|h_{\ell+1}^{-1}(\mathcal{Q}_\ell^X E_\ell v - \Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_{\ell+1})} \\
&\quad + \sum_{k=\ell+1}^{L-1} \sigma^{k+1-\ell}\, \|h_{k+1}^{-1}(\Pi_{k-1}^k \cdots \Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v - \Pi_k^{k+1}\cdots\Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v)\|_{L^2(\Omega_{k+1})}.
\end{aligned}$$

Finally, the L^2-approximation properties of the single operators yield

$$\lesssim |E_\ell v|_{H^1(\Omega_\ell)} + \sigma\, |\mathcal{Q}_\ell^X E_\ell v|_{H^1(\Omega_\ell)} + \sum_{k=\ell+1}^{L-1} \sigma^{k+1-\ell}\, |\Pi_{k-1}^k \cdots \Pi_\ell^{\ell+1}\mathcal{Q}_\ell^X E_\ell v|_{H^1(\Omega_k)}.$$

As in (3.15), we use the H^1-stability properties of the operators $(\Pi_k^{k+1})_{k=\ell,\dots,L-1}$ and of the mapping $\mathcal{Q}_\ell^X$ for the terms in the sum and then the boundedness (3.14) of the extension operator to obtain

$$\lesssim \sum_{k=\ell}^{L} \sigma^{k-\ell}\, |E_\ell v|_{H^1(\Omega_\ell)} \lesssim \sum_{k=\ell}^{L} \sigma^{k-\ell}\, |v|_{H^1(\Omega_L)} \lesssim |v|_{H^1(\Omega_L)}. \tag{3.16}$$

In the very last step of (3.16), the sum over k is bounded by means of a geometric series. This concludes the proof of the approximation property of $\mathcal{Q}_\ell^V$. □

We needed to find an H^1-regular extension (from $H^1(\Omega_L)$ to the potentially larger domain Ω_ℓ) such that the L^2-approximation property of the operator $\mathcal{Q}_\ell^X$ may be exploited. Plain extension by zero may not be H^1-regular. But mere evaluation is still well-defined for, e. g., Clément-type operators.

The derivation of the space hierarchy with multilevel bases in Section 3.2.1 clarifies that all coarse level functions are merely associated with the computational domain $\Omega_L = \Omega$; see, in particular, Lemma 3.2. However, in the proof of Lemma 3.7, the coarse domains $(\Omega_\ell)_{\ell<L}$ appear and the treatment of functions which are defined with respect to different domains becomes crucial. Note that this is only because we exploit the product structure of the definition of the space hierarchy when looking for suitable mappings $\mathcal{Q}_\ell^V$. If one had a means to work directly with the basis functions in $(\widetilde{\Lambda}_\ell)_{\ell<L}$, such a recursive access to the transfer operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$ and thus to the domains $(\Omega_\ell)_{\ell<L}$ could possibly be avoided.

3.4.3 Relaxation of the assumptions

As indicated before, we need to modify the proof of Lemma 3.7 because it is in general impossible to find suitable prolongation operators satisfying the assumptions (3.10) and (3.11) for all function in $(X_\ell)_{\ell<L}$ in case $V_L \subset H_D^1(\Omega_L)$ with $\mathrm{meas}_{d-1}(\Gamma_D) > 0$. We will

relax these requirements in the following. More precisely, it is sufficient that the operators are stable and possess approximation properties in subspaces of functions which already satisfy the zero boundary conditions in a certain sense. We essentially show that mappings $\mathcal{Q}_\ell^V : V_L \to V_\ell$ also exist in this case.

Let us emphasize that the following construction, which recursively excludes basis functions which "couple with Γ_D", is not necessary for the practical algorithms. This is because the Dirichlet boundary conditions are immediately incorporated into the coarse level functions in a very natural manner; see Section 3.3. However, for the sake of the analysis to be carried out here, we need to be careful that the single operators constituting $\mathcal{Q}_\ell^V$ respect the boundary conditions. This requires, again, a recursive technique which will be discussed in full detail here.

Let the Dirichlet nodes of $\mathcal{T}_L$, namely the nodes on Γ_D, be denoted by $\mathcal{N}_L^D$. To avoid confusion, let us remark that these nodes have not appeared so far; they are not contained in $\mathcal{N}_L$. For $\ell < L$, recall that the set $\mathcal{N}_\ell$ consists of all nodes of $\mathcal{T}_\ell$ including all boundary nodes. We recursively define for $\ell \in \{0, \dots, L-1\}$ the sets

$$\mathcal{N}_\ell^D := \{p \in \mathcal{N}_\ell \mid \exists\, q \in \mathcal{N}_{\ell+1}^D,\ \mathrm{int}(\mathrm{supp}(\lambda_p^\ell)) \cap \mathrm{int}(\mathrm{supp}(\lambda_q^{\ell+1})) \neq \emptyset\}$$

and

$$\widehat{\mathcal{N}}_\ell := \mathcal{N}_\ell \setminus \mathcal{N}_\ell^D.$$

Then, the auxiliary spaces to be used in the following are

$$\widehat{X}_\ell := \mathrm{span}\{\lambda_p^\ell \mid p \in \widehat{\mathcal{N}}_\ell\} \subset H^1(\Omega_\ell)$$

for $\ell \in \{0, \dots, L-1\}$. In other words, if a basis function is removed, one has to remove all functions from coarser spaces which "couple" with this basis function. By this means, we take care that the constructed operator really maps to the coarse space V_ℓ. This will become clear in the following. Note that no new fine space is introduced.

It will be beneficial to require the prolongation operators to satisfy

$$\Pi_\ell^{\ell+1} \widehat{X}_\ell \subset \widehat{X}_{\ell+1}. \tag{3.17}$$

To illustrate this assumption, let us derive a sufficient condition which is easy to verify. We introduce the notion of "local operators" mapping to a finite element space in the following

Definition 3.8. *Let X be a finite element space associated with a mesh of a domain Ω_X with nodes $\mathcal{N}_X$, $n_X := |\mathcal{N}_X|$, and coordinate isomorphism $\Phi_X : \mathbb{R}^{n_X} \to X$. An operator $\Pi_X : H^1(\Omega_X) \to X$ is called local if*

$$\Phi_X^{-1}(\Pi_X v)_p = \Phi_X^{-1}(\Pi_X w)_p, \quad \forall\, p \in \mathcal{N}_X,\ \forall\, v, w \in H^1(\Omega_X),\ v|_{\omega_p} = w|_{\omega_p}.$$

The above definition specifies the operators whose image at a node of the target finite element space only depends on the values of the input in the corresponding patch. Now, the following statement is the promised characterization of a class of operators satisfying the condition stated in (3.17). Its proof is straightforward.

Lemma 3.9. *A local operator satisfies (3.17).*

The above construction is designed to relax the assumptions of Lemma 3.7 by merely requiring the H^1-stability and the L^2-approximation properties of the prolongation operators in the spaces $(\widehat{X}_\ell)_{\ell=0,\dots,L-1}$. Moreover, under certain conditions, one may expect each space $\widehat{X}_\ell$ to have enough approximation power in the sense that a mapping $\mathcal{Q}_\ell^X : H^1(\Omega_\ell) \to \widehat{X}_\ell$ exists such that

$$\|h_\ell^{-1}(v - \mathcal{Q}_\ell^X v)\|_{L^2(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_\ell)}, \quad \forall\, v \in H^1_D(\Omega_\ell). \tag{3.18}$$

As before, $\mathcal{Q}_\ell^X$ needs to operate on functions associated with Ω_ℓ as it will be applied to extended functions $E_\ell v$ with $v \in H^1(\Omega_L)$. The issue that, in general, $\Gamma_D \not\subset \partial\Omega_\ell$ but merely $\Gamma_D \subset \overline{\Omega}_\ell$ is just a technicality. This is because a suitable trace operator from $H^1(\Omega_\ell)$ to $H^{\frac{1}{2}}(\Gamma_D)$ also exists in the present case; see, e. g., [75]. Therefore, the space $H^1_D(\Omega_\ell)$ may be understood as $\{v \in H^1(\Omega_\ell) \mid v|_{\Gamma_D} = 0 \text{ in } H^{\frac{1}{2}}(\Gamma_D)\} = \{v \in H^1(\Omega_\ell) \mid v = 0 \text{ a. e. on } \Gamma_D\}$.

Similar to the proof of Lemma 3.7, we choose the mappings $\mathcal{Q}_\ell^X$ as quasi-interpolation operators. The idea to guarantee (3.18) is to limit the size of the subdomain in which the fine-to-coarse operator approximates by zero due to the reduction from X_ℓ to $\widehat{X}_\ell$. Loosely speaking, this "strip" must not grow asymptotically faster into the domain than the mesh size. This is stated more precisely in the following lemma, which is the main result of this section.

Lemma 3.10. *Let the transfer operators $\Pi_{\ell-1}^\ell : X_{\ell-1} \to X_\ell$, $\ell \in \{1,\dots,L\}$, be local and H^1-stable. Suppose that they satisfy L^2-approximation properties with respect to the auxiliary spaces $(\widehat{X}_{\ell-1})_{\ell=1,\dots,L}$, i. e.,*

$$|\Pi_{\ell-1}^\ell v|_{H^1(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_{\ell-1})}, \quad \forall\, v \in \widehat{X}_{\ell-1}, \tag{3.19}$$

and

$$\|h_\ell^{-1}(v - \Pi_{\ell-1}^\ell v)\|_{L^2(\Omega_\ell)} \lesssim |v|_{H^1(\Omega_{\ell-1})}, \quad \forall\, v \in \widehat{X}_{\ell-1}. \tag{3.20}$$

In addition, assume that

$$\operatorname{dist}(p, \Gamma_D) \lesssim h_\ell(p), \quad \forall\, p \in \mathcal{N}_\ell^D, \tag{3.21}$$

for $\ell \in \{0,\dots,L-1\}$ with suitably chosen (continuous) mesh size functions h_ℓ, e. g., according to (3.8). Then, there are mappings $\mathcal{Q}_\ell^V : V_L \to V_\ell$, $\ell \in \{0,\dots,L-1\}$, which satisfy the same H^1-stability and L^2-approximation properties as previously stated in Lemma 3.7, namely (3.12) and (3.13).

Proof. For $\ell \in \{0,\dots,L-1\}$, choose $\mathcal{Q}_\ell^V$ like before as $\Pi_{L-1}^L \cdots \Pi_\ell^{\ell+1} \mathcal{Q}_\ell^X E_\ell$. In particular, one may employ a standard quasi-interpolation operator $\mathcal{Q}_\ell^X : H^1(\Omega_\ell) \to \widehat{X}_\ell$, which now maps to the modified auxiliary space $\widehat{X}_\ell$. Note that the extension operator E_ℓ preserves the zero values on Γ_D. Anticipating the validity of the approximation property (3.18), we conclude the proof following the lines of the proof of Lemma 3.7. Indeed, $\mathcal{Q}_\ell^V$ maps V_L to V_ℓ. This is a consequence of the recursive construction and the fact that the operators $(\Pi_{\ell-1}^\ell)_{\ell=1,\dots,L}$ are local.

Let us postpone the proof of the existence of an H^1-stable Clément-type operator $\mathcal{Q}_\ell^X$ satisfying the relevant L^2-approximation property to Section 5.2, Lemma 5.4. □

In general, a non-recursive definition of the spaces $(\widehat{X}_\ell)_{\ell<L}$ is inadequate as it is not clear how to enforce $\Pi_{L-1}^{L}\cdots\Pi_{\ell}^{\ell+1}\widehat{X}_\ell \subset V_\ell$, $\ell \in \{0,\dots,L-1\}$, in this case. This property is evidently used in the proof of Lemma 3.10. In particular, it is not enough to make modifications at level $L-1$ only.

Although the assumption (3.21) does not seem to have appeared in the literature in this form before, we consider it a very natural requirement. It is a rough statement about the resolution of Γ_D by the eliminated coarse level nodes, which formally specifies the asymptotic growth condition discussed above. This is further illustrated in the prospective proof of Lemma 5.4. As only the coarse mesh size h_ℓ, $\ell < L$, is involved, the assumption is rather weak.

Remark 3.11. *With the arguments put forward above one may "refine" the analysis and relax the conditions on the coarse domains a little. From our proof of Lemma 3.10, we see that it is not necessary that the coarse domains $(\Omega_\ell)_{\ell<L}$ cover Γ_D. At these parts an extension by zero would be sufficiently accurate in the sense that $\mathcal{Q}_\ell^X$ still possesses the required approximation property. To satisfy a suitable analogon of (3.21), the coarse domains should cover all other nodes, though; this means $\Omega_0 \supset \ldots \supset \Omega_{L-1}$ and $p \in \Omega_{L-1}$ for all $p \in \mathcal{N}_L$.*

Finally, let us state a more convenient form of the conditions the mappings $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ need to satisfy. Naturally, one does not examine single entities but rather considers entire types or classes of approximation operators. Generally speaking, a transfer concept is an instruction which provides a concrete prolongation operator depending on certain data, namely the domains $\Omega_Y \subset \Omega_X \subset \mathbb{R}^d$, the space $X \subset H^1(\Omega_X)$ and the finite element space $Y \subset H^1(\Omega_Y)$. One might formally write this as

$$(X,\, Y) \longmapsto (\Pi : X \to Y). \tag{3.22}$$

Such a concept is usually specified by a simple interpolation formula or some other (to a greater or lesser extent abstract) rule. The following lemma states a sufficient condition, which will be used later when we examine different choices of the coarse-to-fine information transfer.

Lemma 3.12. *Let a transfer concept in the sense of (3.22) generate only local, H^1-stable operators which satisfy L^2-approximation properties for all settings with the following characteristics:*

- *arbitrary domain $\Omega_Y = \Omega_X =: \Omega$,*
- *finite element space Y with Dirichlet boundary conditions on $\Gamma_D \subset \partial\Omega$,*
- *$X = H_D^1(\Omega)$.*

Then, this transfer concept applied to the situation of Lemma 3.10 yields prolongation operators which fulfill the requirements (3.19) and (3.20).

Proof. For $\ell \in \{1, \dots, L\}$, let $\Pi_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell$ originate from a transfer concept with the assumed qualifications. We denote the reduced domains by $\widehat{\Omega}_\ell := \Omega_\ell \setminus \Omega_\ell^D$ where Ω_ℓ^D is the maximal subset of Ω_ℓ such that all functions in $\widehat{X}_\ell$ are zero everywhere; it may happen that $\Omega_\ell^D = \emptyset$. In addition, let $\Gamma_D^\ell \subset \partial\widehat{\Omega}_\ell$ be maximal such that $v|_{\Gamma_D^\ell}$ is zero everywhere for all $v \in \widehat{X}_\ell$. The spaces $H_D^1(\widehat{\Omega}_\ell)$ may accordingly be defined with Dirichlet boundary conditions on Γ_D^ℓ.

Let $v \in \widehat{X}_{\ell-1}$. We have that $\Pi_{\ell-1}^{\ell} v = \Pi_{\ell-1}^{\ell}(v|_{\widehat{\Omega}_\ell})$ because the operators are local. Moreover, $v|_{\widehat{\Omega}_\ell} \in H_D^1(\widehat{\Omega}_\ell)$ as $\Omega_{\ell-1}^D \supset \Omega_\ell^D$ by construction. Consequently, we can exploit the properties of the transfer concept to conclude the stated estimates. □

As we have proved that the above conditions are sufficient for the direct applicability of the considered class of operators in Lemma 3.10, it remains to check one generic case later.

Remark 3.13. *A further relaxation of the conditions on the overlap of the domains is discussed in [47, 91]. It is possible to allow coarse meshes without the relation* $\Omega_{\ell-1} \supset \Omega_\ell$*; the size of the uncovered part needs to be bounded, though. This comes at the price of designing special interpolation operators. If the Neumann boundaries or the respective portions of the coarse domains are not completely covered, both analysis and numerical experiments show that a naive extension by zero in* $\Omega_\ell \setminus (\Omega_\ell \cap \Omega_{\ell-1})$ *is only sufficiently accurate near the Dirichlet boundary. To overcome this, interpolation formulas which are modified near the Neumann boundary have been considered in [47, 91] in case* $d = 2$*.*

This observation also relates to the fact that the approach of [123] is not suitable for problems other than pure Dirichlet problems; see Section 4.1.1 for more detailed comments.

3.4.4 Convergence theorem

The proofs of Lemma 3.7 and Lemma 3.10 are essential steps of the convergence analysis of the semi-geometric multilevel methods. For general finite element spaces $(X_\ell)_{\ell=0,\dots,L}$ satisfying the coarsening assumption (3.9), we have shown the existence of mappings $(\mathcal{Q}_\ell^V)_{\ell=0,\dots,L-1}$ with the required H^1-stability and L^2-approximation properties. The analysis has been carried out for the multiplicative and also the first additive variant as both use the recursive procedure setupSGMG. Note that the adaptation of Lemma 3.10 and its proof to the second additive variant with setupSGMG$_{\text{imm}}$ (see Section 3.2.3) is straightforward.

Now, we are in a position to apply the theoretical achievements mainly due to Bramble, Pasciak, Wang and Xu [32, 33] as outlined in Section 2.3 to the semi-geometric setting. For this purpose, assume that the smoothing operators $(S_\ell)_{\ell=1,\dots,L}$ satisfy the elementary properties stated in (2.22). We do not aim to put forward more explicit results concerning the smoothers in this thesis. Instead, we take the liberty of referring to, e. g., [28]. Then, our considerations result in the following

Theorem 3.14. *Let the assumptions of Lemma 3.10 hold and the smoothing property (2.22) be satisfied. Then, the semi-geometric multilevel methods Algorithm 3.4 and Algorithm 3.5 yield preconditioning uniformly with respect to the mesh size. Further, the semi-geometric multigrid method, i. e., the Richardson iteration preconditioned by Algorithm 3.4, converges uniformly with respect to the mesh size.*

Again, we remark that the result holds for adaptively refined meshes in the sense that the smoothing property (2.22) only needs to be satisfied in a suitable subspace; no relaxation operations are necessary outside of the subdomain of refinement. This line of reasoning has been put forward in comparably straightforward manners by numerous authors in the literature; in particular, see [34]. Note that such a generalization is possible only if the stability and approximation properties of the constructed operators also hold true in the non-quasi-uniform case. In addition, if two successive meshes locally match, the respective prolongation operator should coincide with the identity in the corresponding parts of the domain. If this is not achieved naturally, namely following from the basic properties of the employed transfer concept, it should be enforced in a suitable other way.

On the way to *L*-independent convergence

Theorem 3.14 states uniform convergence of the semi-geometric multigrid methods with respect to the mesh size. Let us comment on the dependence of the result on the number of levels L. As, in theory, the number of levels needs to grow to infinity with decreasing fine mesh size to retain coarse level problems which are small enough to be solved very accurately, the convergence theorem is only quasi-optimal in case it worsens with increasing L.

Note that the stability constant of $\mathcal{Q}_\ell^V$ can only be proved to be bounded by the product of the stability constants of the single operators. This is because the proof directly exploits the product structure of the constructed space hierarchy. Thus, the possible dependence of the constants in (3.12) and (3.13) on the number of levels is not ruled out. However, we obtain a convergence result which is optimal if the number of levels is fixed. In our context, as we have the advantage to choose the coarser meshes quite freely, we can easily ensure that the number of degrees of freedom roughly reduces by one order of magnitude per level for $d = 3$; thus, this restriction is not too relevant for practical purposes. An analysis providing not only quasi-optimal but L-independent convergence results for the developed multilevel methods with the specially constructed space hierarchy has to be carried out in a different way. In principle, one should examine the norm equivalence put forward in Section 2.3.2. If one can show that an inverse inequality holds uniformly in each new space V_ℓ, the norm equivalence holds true without a dependence on the number of levels L.

The following special case is of particular practical importance. If the coarse meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ are nested, e. g., resulting from a regular refinement routine of an initial coarse mesh $\mathcal{T}_0$, the operators connecting the associated finite element spaces $(X_\ell)_{\ell=0,\dots,L-1}$ may be chosen as the natural inclusions (because the spaces are nested). In this case, the stability constants of the mappings $\mathcal{Q}_\ell^V$ from the previous lemmas are estimated merely by the product of the stability constants of the two mappings $\mathcal{Q}_\ell^X$ and Π_{L-1}^L. Therefore, the (constants in the) inequalities (3.12) and (3.13) do not depend on the number of levels L because neither the estimate of $\mathcal{Q}_\ell^X$ nor the one of Π_{L-1}^L do. This holds true under the general assumptions of this section. Although this case resembles the auxiliary space method analyzed in [195], note that the latter suggests a non-variational approach to keep standard (multilevel) iterations applicable in the respective auxiliary space. In addition, the auxiliary spaces considered in [112, 113, 195], in part in a quite different context, though, need to be almost as fine as the original spaces. This is in contrast to the framework developed

in this chapter.

A further aspect, which has not been investigated in the literature so far, concerns the magnitude of the stability constants of suitable operators for the information transfer between non-nested finite element spaces. They appear to be rather small in practice. We will examine this in Chapter 5; especially, see the numerical studies in Section 5.8.

Remark 3.15. *Finally, we need to comment on a series of papers about non-nested multigrid methods which do not fall into the category of variational methods. In part these methods are specially designed for certain non-conforming discretizations. Starting with [35] and then [172, 201, 202, 203], several assumptions have been formulated concerning the relation between the operators at the single levels, the most intuitive being*

$$\|\Pi_{\ell-1}^{\ell} v\|_{A_\ell} \leq \|v\|_{A_{\ell-1}}, \quad \forall\, v \in X_{\ell-1},$$

for $\ell \in \{1, \ldots, L\}$. Here, the operators $(A_\ell)_{\ell=0,\ldots,L}$ do not stem from a Galerkin restriction and are therefore not related to each other by the variational equation (2.7).

The more sophisticated analyses need an approximation property which can, in general, only be guaranteed by additional regularity assumptions, though. This is because orthogonal projections with respect to the corresponding energy inner products $(a_\ell(\cdot,\cdot))_{\ell=0,\ldots,L-1}$ are used in all mentioned approaches. For the development of additive preconditioners for some non-conforming discretizations, we refer to [152, 153]. Further progress for the non-nested $\mathcal{V}$-cycle has been made more recently by [38, 79].

3.5 A coarse space for overlapping Schwarz methods

This is a convenient point to comment briefly on the application of the above concept to two-level overlapping Schwarz methods with a global coarse space. For a detailed introduction into this field, see, e. g., [177, Chapter 3]. We also mention the enormous number of references in this monograph. The idea of this class of methods is to decompose the computational domain into overlapping subdomains, preferably of simple structure, on which local (e. g., exact) solvers can be successfully applied. The information transfer between the single problems is realized by the local overlaps and a global component.

Without going into great detail, we point out some significant connections to the previous considerations. The notations are a little different here. As there are only two distinct levels, the fine and coarse quantities may be labeled with the indices $_h$ and $_H$, respectively. In this context, let us mark a decomposition into subspaces by the index $_i$ (instead of $_\ell$). Assume that $V = X_h$ is a given finite element space associated with a mesh $\mathcal{T}_h$ of the computational domain $\Omega_h = \Omega$ to approximate the solution as described in Section 1.3. Let the local finite element spaces

$$(V_i)_{i=1,\ldots,N} \quad \text{with} \quad V_i \subset V,\ i \in \{1, \ldots, N\},$$

be associated with a "horizontal" overlapping decomposition $(\Omega_i)_{i=1,\ldots,N}$ of Ω. In particular, $v = 0$ in $\Omega \setminus \Omega_i$ for all $v \in V_i$. Naturally, one will not have access to the global space V

but only to the local spaces $(V_i)_{i=1,\dots,N}$ in practice, which are defined with respect to local meshes $(\mathcal{T}_i)_{i=1,\dots,N}$. As most proof techniques involve a coloring argument, one usually assumes that the number of adjacent subdomains is bounded (independently of h and N).

A Schwarz preconditioner, which acts as (successive or parallel) subspace correction method corresponding to the above decomposition, is in general not scalable with respect to the number of subdomains N. To prove preconditioning results independent of the number of subdomains, one usually introduces a global coarse space; see, e. g., [177]. For a historical overview of the role of coarse spaces in domain decomposition methods, we refer to [189]. In many cases, bounds on the condition number of the preconditioned operator may be proved, which essentially depend on the ratio of the maximum size of the subdomains to the minimum size of the overlaps. Coarse spaces associated with elementary coarse meshes are ready to hand as long as the structure of the decomposition is simple. To achieve the mentioned result, in general, one needs to ensure enough global information transfer by choosing the coarse mesh size sufficiently small, e. g., comparable to the size of the subdomains. In addition, an (almost) exact solution of the subproblems is often required.

Two-level overlapping Schwarz methods have also been developed for problems with unstructured meshes. In this case, as we have seen in the multilevel setting, it may be difficult to construct nested coarse spaces as the local fine meshes do typically not allow for a proper global coarse mesh. Advanced coarse spaces for this problem class include partition of unity spaces [160, 161, 162] and spaces obtained by smoothed aggregation [134, 164]. A paper on recent progress of the theory of domain decomposition methods with irregular subdomains is [188]; see also the references therein. Here, we focus on an approach which is close to the one presented before. If X_H is a global finite element space associated with a mesh $\mathcal{T}_H$ of $\Omega_H \supset \Omega$ which is not related to the space X_h, a coarse space can be constructed by means of a suitable prolongation operator $\Pi_H^h : X_H \to X_h$ correlating the global coarse mesh with the local fine meshes. Similarly to the multilevel case, a nested space is defined via

$$V_0 := \Pi_H^h X_H \tag{3.23}$$

like in [43, 49, 51, 52]. Then, the analysis of the decomposition

$$V = V_0 + \sum_{i=1}^{N} V_i,$$

namely the proof of a partition lemma, requires respective H^1-stability and L^2-approximation properties of the applied operator Π_H^h. This may indicate that the considerations of the present chapter and the research of geometrically inspired transfer concepts in Chapter 5 are useful for other purposes, too.

Naturally, in this two-level setting, we do not need to worry about the fact that a composed mapping is used in the proof; the composition always consists of as few as two mappings. This means that the stability and approximation results for the respective fine-to-coarse mapping $\mathcal{Q}_0^V : V \to V_0$ are optimal if the coarse space is constructed via (3.23) and if the operator Π_H^h satisfies the meantime well-known properties.

In the context of the semi-geometric multilevel methods, not only for the analysis but also for practical purposes, the locality of the information transfer is crucial. Here, the

situation is a little different. Whereas, in a multilevel setting, the reduction of the number of the degrees of freedom per level is only by a factor typically in the order of 2^{-d}, one may be able to manage with very few coarse degrees of freedom per subdomain in case of the overlapping two-level method. Therefore, it might not be mandatory to use a local transfer operator. A global transfer concept might not affect the overall efficiency as only small dense systems need to be solved corresponding to the subspace V_0. Admittedly, this does not hold true for the asymptotic range as the size of the subdomains and, thus, the coarse mesh size needs to decrease to retain spaces $(V_i)_{i=1,\dots,N}$ with sufficiently small dimension.

3.6 Implementation aspects

In this section, we outline some essential aspects concerning the implementation of the abstract semi-geometric framework. The developed software concepts are quite close to the description of the algorithmic structure given in the current chapter. Here, we focus on the realization of the basic principles of the semi-geometric multilevel methods. Both analysis and practical aspects of the linear operators for the information transfer between non-nested finite element spaces are presented separately in Chapter 5. We have realized our concrete implementation as a new module `nnmglib` in the package `obslib++`, which is maintained by the author of [124] and his work group. The software uses fundamental components of the finite element toolbox `ug`; see [17].

3.6.1 Bounding the complexity of the multilevel hierarchy

In this paragraph, we consider an important point concerning the complexity of the studied multilevel algorithms. As a measure of the efficiency of the multilevel hierarchy itself, in addition to iteration counts or convergence rates, we put forward the notions of grid complexity $\mathcal{C}_{\mathrm{gr}}$ and operator complexity $\mathcal{C}_{\mathrm{op}}$ defined by

$$\mathcal{C}_{\mathrm{gr}} := \frac{\sum_{\ell=0}^{L} n_\ell}{n_L}, \quad \mathcal{C}_{\mathrm{op}} := \frac{\sum_{\ell=0}^{L} n_\ell^A}{n_L^A}, \tag{3.24}$$

which are common in the literature on algebraic multigrid methods. Here, n_ℓ^A is the number of non-zero entries in the sparse matrix $\boldsymbol{A}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$, $\ell \in \{0, \dots, L\}$.

As an illustration, let us elaborate on the above concept of complexity in the case of the geometric multigrid method with uniform refinement. This will also come in useful for a later comparison with the semi-geometric approach. Note that a priori estimates of the quantities $\mathcal{C}_{\mathrm{gr}}$ and $\mathcal{C}_{\mathrm{op}}$ are much more difficult to establish for (purely) algebraic multigrid methods. This is because the determination of the coarse degrees of freedom and the corresponding coarse level operators is part of the algorithm in this case; see, e. g., [176].

To start with, a uniform refinement of a mesh consisting of triangles and quadrilaterals for $d = 2$ or tetrahedra and hexahedra for $d = 3$ results in a multiplication of the number of elements with the factor 2^d. Asymptotically this also holds true for the number of nodes, namely $\frac{n_{\ell+1}}{n_\ell} =: c_\ell \to 2^d$ for $\ell \to \infty$. As the sequence $(c_\ell)_{\ell=0,1,\dots}$ is strictly increasing, we

may estimate*

$$\mathcal{C}_{\mathrm{gr}} = \frac{\sum_{\ell=0}^{L}(\prod_{\ell'=0}^{\ell-1} c_{\ell'})n_0}{(\prod_{\ell'=0}^{L-1} c_{\ell'})n_0} = \sum_{\ell=0}^{L} \frac{1}{\prod_{\ell'=\ell}^{L-1} c_{\ell'}} > \sum_{\ell=0}^{L} \frac{1}{2^{\ell d}} \to \frac{2^d}{2^d-1} \quad \text{for } L \to \infty. \tag{3.25}$$

This gives a precise idea of the magnitude of $\mathcal{C}_{\mathrm{gr}}$ for geometric multigrid methods. In fact, the limit $\frac{2^d}{2^d-1}$ is not a lower bound of the grid complexity but at least a good approximation. For example, in case $d = 3$, and a structured simplicial mesh of a cube with 96 elements in $\mathcal{T}_0$, the difference of $\mathcal{C}_{\mathrm{gr}}$ to the limit $\frac{8}{7}$ is less than 0.0004 for a hierarchy of four levels with three coarse meshes ($L = 3$). Neglecting the boundary nodes, whose influence on the measure decreases exponentially with the refinement index ℓ, one obtains the same estimate for the operator complexity $\mathcal{C}_{\mathrm{op}}$.

A prevalent technique to keep $\mathcal{C}_{\mathrm{gr}}$ and $\mathcal{C}_{\mathrm{op}}$ small (and the application of both Algorithm 3.4 and Algorithm 3.5 efficient) is truncation of the prolongation operators by deleting the matrix entries which are smaller than a parameter $\varepsilon_{\mathrm{tr}} > 0$ times the maximal entry in the respective row. Afterwards, the modified rows are rescaled such that the row totals remain unchanged; see [176]. In the semi-geometric framework, it is absolutely necessary to perform a truncation procedure to retain the optimality of the algorithms. Otherwise, one can in general not prevent the appearance of very small and thus irrelevant entries in the prolongation matrices. Therefore, let the outlined truncation be included in the setup of the semi-geometric multilevel hierarchy, setupSGMG (Algorithm 3.3). At level $\ell \in \{1, \dots, L\}$, it is carried out after the computation of the matrix $\mathbf{\Pi}_{\ell-1}^{\ell}$ and before the Galerkin product $\boldsymbol{A}_{\ell-1}$. The incorporation into setupSGMG$_{\mathrm{imm}}$ is done analogously. Note that there is no straightforward way to prevent the entries of a final prolongation matrix from being very small as early as during its assembly. This is because it is usually built from local contributions, which have to be summed up first. We study this in more detail in Chapter 5; see, in particular, Section 5.7. Regardless of the precise transfer concept, at least the entries smaller than 1% of the maximum entry must be removed for a minimum degree of sparseness of the prolongation matrices.

Consequently, one has at least two different options to influence the bandwidth of the prolongation matrices. On the one hand, we may control the structure of the matrices $(\mathbf{\Pi}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ a priori by the choices of the meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ and the type of information transfer between the associated non-nested finite element spaces. On the other hand, the sparsity pattern of the computed matrices may be controlled a posteriori by suitable modifications, e. g., the described truncation with rescaling. The first point is special to the semi-geometric framework whereas the second one is more common. In principle, intermediate approaches are conceivable, e. g., modifying the setup procedure (Algorithm 3.3) by including local adaptations of the coarse meshes. Notable methodologies aiming at this direction are reviewed in Chapter 4.

As a general rule, to construct efficient multilevel methods, one needs to trade off the approximation properties of the coarse level problems against the band structures of the representing coarse level matrices. Especially, in the context of non-nested coarse

*The product symbol $\prod$ in (3.25) must not be confused with the notation of the generic prolongation operators. The former only appears here whereas the latter is used throughout this thesis.

meshes, one should expect better approximation properties of the coarse spaces in case of a variational approach, such as the one presented in Section 3.2.1, whereas the sparsity patterns of the matrices $(\boldsymbol{A}_\ell)_{\ell=0,\dots,L-1}$ and, accordingly, the complexity $\mathcal{C}_{\mathrm{op}}$ are generally more favorable with a non-variational approach.

3.6.2 Information transfer between non-nested meshes

As indicated before, we keep the transfer concept still abstract in this section. The implementation of concrete prolongation and restriction operators is discussed, in particular, in Section 5.7. However, let us remark that the routines for the information transfer between non-nested meshes are usually built from local coupling contributions such as global or local integrals or function evaluations. This holds true whether or not the transfer is actually local in the sense of Definition 3.8.

As the set of non-nested meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L}$ does not come with natural parent–child relations stemming from a regular refinement procedure, one needs to compute suitable neighborhood relations between elements in successive meshes to ensure that the eventual assembly routines of the prolongation matrices are local. For this purpose, we have incorporated the quadtree/octree implementation of [5] into `obslib++`. Suitable advancing front techniques exploiting the connectivities of the single meshes can be applied instead; see, e. g., [89] in a related context. Although a hierarchical structure is an adequate choice to treat general problems in a flexible fashion, note that a plain sorting variant may be more efficient for certain cases in the present context. This holds particularly true if the elements of the used meshes are distributed rather evenly because in this case the overhead of the hierarchical quadtree/octree structure might not be negligible.

Naturally, any reasonable choice of coarse meshes satisfies $\sum_{\ell=0}^{L} |\mathcal{T}_\ell| \lesssim |\mathcal{T}_L|$. This implies that the number of operations to compute the desired relations of the elements in the pairs $(\mathcal{T}_{\ell-1}, \mathcal{T}_\ell)_{\ell=1,\dots,L}$ (or in the pairs $(\mathcal{T}_\ell, \mathcal{T}_L)_{\ell=0,\dots,L-1}$ if the variant setupSGMG$_{\mathrm{imm}}$ is employed in case of the additive semi-geometric preconditioners) grows at most like $\mathcal{O}(n_L \log n_L)$. In Section 5.7, it turns out that we achieve optimal complexity $\mathcal{O}(n_L)$ for the actual assembly of the sequence of prolongation matrices, having these relations ready to hand.

Coarse spaces from non-nested coarse meshes

Our practical implementation is indeed as flexible as the theoretical considerations in this chapter indicate. The coarse meshes can be unstructured; they do not need to be nested. The user provides the fine level mesh $\mathcal{T}_L$ with respect to which the boundary value problem to be solved is actually set. In addition, coarse level meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ are imported using the extended geometry handling incorporated into `obslib++` by [76, 100], too.

We have implemented a new module `nnmglib` in `obslib++` to manage the setup of the semi-geometric (monotone) multigrid methods and additive preconditioners. This library also includes the methods for the computation of a variety of concrete transfer concepts (see Chapter 5 and in particular Section 5.7) as well as the methods for the elaborate study concerning the information transfer between non-nested meshes as such (see Section 5.8). Let us point out that, at an intermediate stage, a part of the basic data structures, which

are derived from related transfer classes in ug, has been developed during the preparation of a student research project [76].

Note that the mesh $\mathcal{T}_L$ is directly equipped with a set of degrees of freedom as the finite element space $V_L = X_L$ and a basis Λ_L are known from the start. In contrast, the nodes and the elements of the meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ merely represent auxiliary geometric entities and are thus not yet included in the algebraic structure. Their intended purpose is to supply the bases $(\Lambda_\ell)_{\ell=0,\dots,L-1}$ which are not used in the eventual semi-geometric algorithm. In other words, we proceed as customary in algebraic multigrid methods; the coarse level degrees of freedom are not created before the setup.

As indicated in the discussion of the algorithms in Section 3.2.2 and Section 3.2.3, once the discrete operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ or $(\Pi_\ell^L)_{\ell=0,\dots,L-1}$ and $(A_\ell)_{\ell=0,\dots,L-1}$ are known, i. e., once the matrices $(\mathbf{\Pi}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ or $(\mathbf{\Pi}_\ell^L)_{\ell=0,\dots,L-1}$ and $(\boldsymbol{A}_\ell)_{\ell=0,\dots,L-1}$ are computed by setupSGMG or setupSGMG$_{\text{imm}}$, the presented multilevel iterations are nothing but algebraic operations involving those matrices. In principle, one does not need the prolongation matrices to be given explicitly. It is sufficient to have routines performing the respective evaluations for given residual or correction vectors ready to hand.

The efficiency of the multigrid method relies on the effectivity of the individual smoothing iterations. More precisely, at each level $\ell \in \{1, \dots, L\}$, one needs to be able to reduce the oscillating error components with respect to the space X_ℓ sufficiently fast, i. e., using very few (Gauß–Seidel) iterations. The remaining error has to be sufficiently smooth, namely its representation at the coarser level in $X_{\ell-1}$ is sufficiently accurate. In standard geometric multigrid methods, one relies on sequences of nested meshes with $h_{\ell-1} = 2h_\ell$. For our purposes, the coarse meshes need to be chosen appropriately such that a similar coarsening assumption holds. Conversely, an increased number of local relaxations in certain regions, possibly carried out in a special order or as block relaxation, could compensate for a "locally bad choice" of the coarser mesh. There are indeed first approaches, although not in the present context, which try to control or optimize the amount of local work performed by the smoother during the algorithm. One example is an a priori redistribution of the total number of relaxation steps towards regions with badly shaped elements. We learned about this methodology from [90].

As a matter of fact, for the above conditions on an effective interplay of smoothing and coarse level correction to be satisfied in practice, one only needs to guarantee that the coarsening factor is in a rather generous range. Our numerical examples include results on the basic robustness of the semi-geometric approach with respect to the choice of the coarse meshes. However, for adaptively refined meshes, we cannot eliminate the possibility that an undesirable local relation between coarse and fine mesh affects the convergence behavior, unless a robust method for readjusting the coarse mesh is applied. In case a highly non-uniform mesh originates from an adaptive refinement procedure (presumably based on suitable error estimators), one might exploit this additional information for an adaptation of the coarse level meshes. A careful elaboration of these issues is beyond the scope of this thesis, though. The utility of an automatic coarse mesh construction, in other words of coarsening procedures, is discussed in Section 4.2.

4 Other geometry-based multilevel techniques

In the past years, several methodologies have been developed for the application of basic multilevel algorithms to problems with complicated boundaries of the computational domain. In this chapter, we describe some of the accomplishments of the research on multigrid methods since the very first algorithms have been recorded with respect to finite difference schemes in the unit square. Some of the efforts which have been made to improve the applicability of general multilevel ideas are explained; relevant connections to the introduction of the semi-geometric framework of the previous chapter are established. In part, the presented approaches aim at constructing coarse approximations of finite element spaces associated with unstructured meshes. Others employ structured meshes coming with tailored discretizations which allow for a (to some extent) straightforward coarse level hierarchy.

Certainly, most of the developments to be reviewed in this chapter are in a sense interwoven. One cannot overlook that the research activities of the cited authors have influenced each other in some form or another. Some ideas immediately build upon the theoretical and algorithmical achievements in multigrid and domain decomposition methods presented in Chapter 2, whereas it seems that others have been developed from a somewhat different point of view.

In agreement with the overall concept of this thesis, we focus on geometric techniques and discuss some important methodologies. Multigrid methods based on adjusted discretizations, which are mostly built from structured meshes, are reviewed in Section 4.1. Next, in Section 4.2, we turn our attention to geometric coarsening techniques for unstructured meshes. We always point out relevant connections and draw comparisons.

As we prefer to put our emphasis on a most thorough study of the properties of the semi-geometric framework, it is beyond the scope of this thesis to evaluate all of the algorithmic ideas described below in detail. We believe that the present chapter provides an adequately deep insight into the development of multilevel methods for problems with complicated boundaries, though. Moreover, we select one method working with a special discretization and present a (monotone) multigrid method based on parametric finite elements in Chapter 7. Note that the overall structure of the paradigm put forward in Chapter 3 is a rather general one compared with other concepts which have been investigated in this or a related context.

4.1 Geometric multigrid methods with adjusted discretizations

In this section, we report on several geometric multilevel techniques based on adjustments of the (fine level) discretizations. Each of the specified approaches employs a discretization scheme of a special nature to construct a suitable coarse space hierarchy. So, for the methods in this section, one has to be willing to give up some of the advantages that finite element discretizations associated with unstructured meshes have. We have also considered another

way to adjust the discretization for the above purpose, namely a parametric finite element approach. A brief study of a (monotone) multigrid method based on a parametric concept is provided as an excursus in Chapter 7.

The goal of our presentation is to highlight structural similarities and differences and relate the methodologies to the semi-geometric framework of Chapter 3 and to each other. In part, the basic idea is not at all difficult and, e. g., immediately builds on Cartesian (auxiliary) meshes. Both the analysis and the algorithmic realization may be more involved, though. A major difference of the approaches to be described in this section, compared to the semi-geometric concept, is the fact that a new discretization of the problems set up in Chapter 1 is introduced. The formulation of the respective finite element spaces is designed to allow for natural coarse scales. However, just as before, one may need fine level information to evaluate coarse level functions, too.

In this thesis, we do not consider meshfree discretizations such as the partition of unity or generalized finite element methods. Instead of using finite element spaces associated with proper meshes, such an approach basically employs a partition of unity associated with an overlapping decomposition of the computational domain and local approximation spaces; see [10, 145] and, for a more recent overview, [11]. Multilevel methods for partition of unity discretizations of elliptic partial differential equations have been developed, e. g., in [99, 170]. We refer to [55, 62] for recent analytical results.

Before going into detail about the single approaches, we consider it particularly important to point out that here the relation of two successive meshes is generally much closer than in the semi-geometric setting; the families of meshes exhibit some additional structure. This will be easy to see for the methods described in Section 4.1.1, Section 4.1.2 and Section 4.1.4, but it especially holds true for the composite finite element method reviewed in Section 4.1.3. Still, the latter is indeed a "multilevel method based on non-nested meshes", too. The research of this and the other techniques has been driven by the demand to construct multilevel finite element splittings for efficient iterative solvers. In contrast to the more flexible semi-geometric framework, the techniques considered here cannot be applied to given unstructured meshes. We emphasize that, among all studied methods, the variant of composite finite element spaces described in Section 4.1.4 is the only one which is able to resolve the domain without changing the gradients of the multilevel finite element functions. Unfortunately, the straightforward applicability is affected by the disadvantage that the Dirichlet boundary conditions require special attention.

As in the previous chapter, auxiliary spaces and the spaces which are finally used in the multilevel algorithms are denoted by X and V, respectively, with some indices.

4.1.1 Filling the domain gradually

Let us start the overview with the conceptually simplest method—at least from an algorithmical point of view—to determine multilevel hierarchies for domains which are not easily resolved by coarse meshes. This approach has been analyzed in [123]. The idea is to embed the computational domain Ω into a larger square ($d = 2$) and to employ a family of structured meshes generated by a regular (uniform) refinement routine. This yields a crude successive approximation of Ω by a finite element mesh from the inside in the following

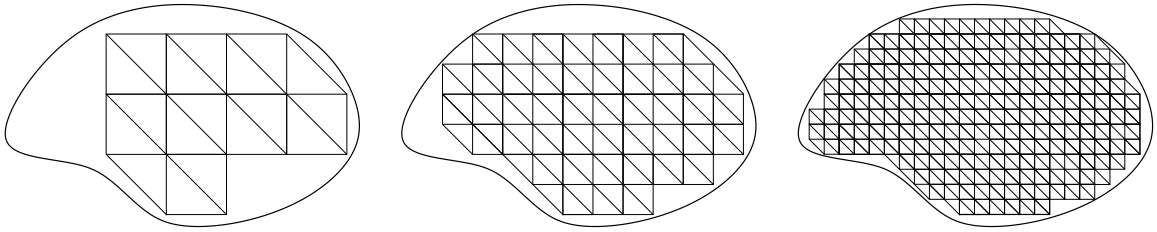


Figure 4.1. Generation of a multilevel hierarchy by uniform refinement of a structured mesh as proposed in [123]. As only those basis functions are taken into account whose supports are completely contained in the computational domain $\overline{\Omega}$, one obtains a sequence of nested finite element spaces approximating $H_0^1(\Omega)$.

way; see Figure 4.1 for an illustration of such a sequence of triangulations.

Let $(\mathcal{T}_\ell^\square)_{\ell\in\mathbb{N}}$ be a family of (uniform) meshes of a surrounding square, not of the presumably much more complicated domain Ω itself. Then, using the standard nodal bases $(\Lambda_\ell)_{\ell\in\mathbb{N}}$ of the finite element spaces $(X_\ell)_{\ell\in\mathbb{N}}$ associated with the considered meshes, namely $\Lambda_\ell = (\lambda_p^\ell)_{p\in\mathcal{N}_\ell^\square}$ for $\ell \in \mathbb{N}$, a hierarchy of nested finite element spaces can be defined by

$$V_\ell := \operatorname{span}\{\lambda_p^\ell \,|\, \omega_p \subset \overline{\Omega}\}, \quad \forall\, \ell \in \mathbb{N}.$$

This means that one fills the domain bit by bit with all elements which fit into it. But only those basis functions whose supports are entirely contained in $\overline{\Omega}$ are included in the definition of V_ℓ as only interior nodes, denoted by $\mathcal{N}_\ell$, may bear degrees of freedom. Consequently, at level ℓ, only those elements which belong to at least one $\omega_p := \operatorname{supp}(\lambda_p^\ell)$, $p \in \mathcal{N}_\ell$, are depicted in Figure 4.1. Note that a similar technique of approximation from the inside has also been analyzed in [29, Section 6] for a refinement application.

By this procedure, one obtains a sequence $V_0 \subset \ldots \subset V_\ell \subset \ldots \subset H_0^1(\Omega)$. Thus, the method seems only reasonable for Dirichlet boundary conditions on $\partial\Omega$; see also Remark 3.13. Moreover, from an application point of view, the discretization V_L at a fixed (finest) level L is not expected to be appropriate in case boundary effects are of particular interest. For the analysis of multilevel preconditioners, the authors of [123] consider an L^2-orthogonal decomposition as in (2.21) associated with the constructed subspaces. Under suitable regularity assumptions on the boundary $\partial\Omega$, a quasi-optimal result on the relevant equivalence between the energy norm and the corresponding splitting norm is achieved; see [123, Theorem 6.1]. For this purpose, the cited reference involves a careful investigation of how the fact that the discrete domain "grows" during the refinement affects the stability of the constructed subspace splitting. This yields preconditioning uniformly with respect to the mesh size for problems set in $H_0^1(\Omega)$. But, as in our analysis in Section 3.4, the condition number estimate depends on the number of refinements.

Naturally, the above ideas transfer to the case $d = 3$ although the analysis may be more involved; see [123]. Note that the methods to be described in Section 4.1.3 and Section 4.1.4 make use of a sequence of structured meshes to derive a tailored discretization scheme for the function space associated with the computational domain Ω, too. However, those approaches are based on much more sophisticated adaptations at the boundary $\partial\Omega$.

Remark 4.1 (Finite elements associated with tree structures)**.** *In recent years, tree structures have become more and more important for high performance computing. In the context of large-scale numerical simulations based on finite element discretizations, the obvious advantage of regular meshes, e. g., generated from the leaves of a quadtree or octree decomposition of the computational domain, is that the stiffness matrix does not need to be assembled. This is usually accompanied by a substantial speed-up of the matrix-vector multiplications involving the precomputed local stiffness matrices. The real benefit of a tree structure, however, is that it may be exploited for adaptive mesh refinement in parallel and efficient load-balancing. We take the liberty of referring to [42] and the references therein.*

As the composite finite element methods described in Section 4.1.3 and Section 4.1.4 are also based on regular meshes, they may take advantage of some of the mentioned points, too. Note that we also employ a tree structure for a fast and flexible determination of "neighboring" elements from unrelated meshes. This is explained within the implementation aspects of, first, the semi-geometric framework (Section 3.6) and, second, the concrete prolongation and restriction operators (Section 5.7).

4.1.2 Boundary fitted elements

In this section, we present a method which is different in many respects. Also for academic reasons, one may be interested in the so-called boundary fitted finite element spaces introduced in [200]. Such a space is associated with a mesh $\mathcal{T}$ of Ω with the property that there is a constant c such that

$$\operatorname{dist}(T, \partial\Omega) \leq c\, h_T, \quad \forall\, T \in \mathcal{T}. \tag{4.1}$$

In other words, the size of the elements $T \in \mathcal{T}$ increases at least linearly with the distance to $\partial\Omega$. This is illustrated by the left sketch in Figure 4.2*. Here, the term "boundary fitted" expresses the fact that the relatively coarse mesh $\mathcal{T}$ still resolves the computational domain exactly, provided that this is possible. Note that the meshes depicted in Figure 4.2 contain hanging nodes only to keep the illustration as simple as possible. As we do not aspire to address non-conforming (multilevel) discretizations, such nodes should be either avoided during the mesh generation process or eliminated by algebraic constraints as usual.

Let the boundary fitted finite element space be denoted by $X_{\mathcal{T}}$. Obviously, Dirichlet boundary conditions on a portion $\Gamma_D \subset \partial\Omega$ may be incorporated as in the standard case of Section 1.3 such that $X_{\mathcal{T}} \subset H^1_D(\Omega)$. Further, consider the weighted L^2-norm $\|\cdot\|_{L^2(\Omega),\mathcal{T}}$ given by a weighted L^2-inner product via

$$\|v\|_{L^2(\Omega),\mathcal{T}} := (v,v)^{\frac{1}{2}}_{L^2(\Omega),\mathcal{T}} \quad \text{and} \quad (v,w)_{L^2(\Omega),\mathcal{T}} := \sum_{T\in\mathcal{T}} \frac{1}{h_T^2}(v,w)_{L^2(T)}, \quad \forall\, v,w \in L^2(\Omega).$$

Then, for families of triangulations ($d = 2$) satisfying (4.1) with a universal constant c and sufficiently regular domains, Yserentant showed in [200] that the semi-norm $|\cdot|_{H^1(\Omega)}$ is equivalent to $\|\cdot\|_{L^2(\Omega),\mathcal{T}}$, namely

$$|v|_{H^1(\Omega)} \eqsim \|v\|_{L^2(\Omega),\mathcal{T}}, \quad \forall\, v \in X_{\mathcal{T}}.$$

*This sketch resembles the one in the original paper; cf. [200, Figure 1]. The straight shape of the boundary has been chosen for simplicity.

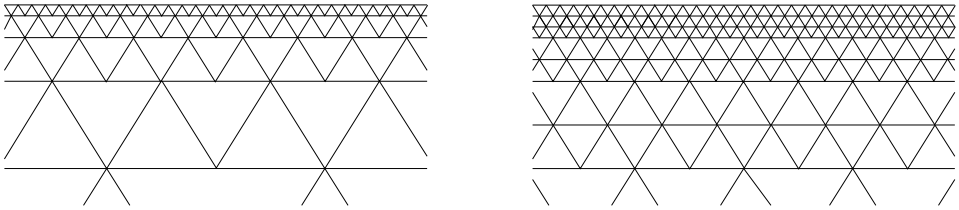


Figure 4.2. Detail of a boundary fitted multilevel triangulation. The actual boundary fitted finite element space is defined with respect to the coarsest mesh on the left. The mesh on the right comes from a refinement of all triangles except the ones at the boundary.

As a consequence, the condition number of the discrete representation of the operator A with respect to the standard basis in $X_{\mathcal{T}}$, see Section 1.4, is bounded independently of the size of the triangles near the boundary.

Now, a space of boundary fitted elements shall be used as coarse space in a geometric multigrid method, i. e., $V_0 := X_{\mathcal{T}}$. In fact, the theoretical considerations in the cited reference may be considered a justification of the paradigm to construct a multilevel hierarchy by coarsening away from the boundary (starting from a fine mesh). This appears natural in case the given fine mesh exhibits a layer-like structure. Naturally, a boundary fitted mesh will be created by a converse procedure, namely by a refinement towards the boundary (starting from a coarse mesh) followed by a simple post-processing; see [200].

To obtain the next finer level in a multilevel hierarchy $(\mathcal{T}_\ell)_{\ell=0,\dots,L}$, every triangle except for the ones at the boundary is refined. This process is illustrated in Figure 4.2. Note that the coarse space associated with the boundary fitted mesh $\mathcal{T} =: \mathcal{T}_0$ can still have a lot of degrees of freedom. But, as the special discretization makes the coarse level problem well-conditioned, relatively few steps of any decent iterative solver should suffice to reduce the error by orders of magnitude.

One should remark that a similar idea is the basis of the "boundary concentrated finite element method" [116]. This discretization scheme is designed for elliptic problems with varying regularity of the solution in the domain, which is due to non-smooth boundary data or complicated boundaries. Exploiting the inner regularity, the method does not only increase the mesh size but also the polynomial degree away from the boundary in the fashion of an *hp*-method. Then, a moderate increase of the condition number of the system matrix with respect to the mesh size and the polynomial degree can be proved; see [116, Proposition 3.4].

4.1.3 Composite finite elements

In this and the following section, we describe some other geometric multigrid methods which are based on adjustments of the discretization. The development of these methods is also driven by the desire to use simpler meshes in the definition of multilevel finite element hierarchies. However, in contrast to the simple variant presented in Section 4.1.1, the design of the space hierarchies involves a rather sophisticated adaptation of certain auxiliary spaces

to the domain boundaries. We present two different variants of the composite finite element method, the technique originally introduced for triangulations ($d = 2$) in [105, 106] and a variation motivated by image based computing [137].

A logically regular mesh hierarchy

A new discretization of the space $H^1_D(\Omega)$ has been introduced in [105, 106]. Similar to the method discussed in Section 4.1.1, the finite element scheme is based on a sequence of regular meshes. But instead of merely neglecting all elements or degrees of freedom which are not admissible in a certain sense, the meshes are adapted to the computational domain as described in the following. This yields more flexibility when it comes to the incorporation of essential and natural boundary conditions.

Here, we do not aspire to give a most precise definition of composite finite element spaces but rather summarize the essential ideas brought forward in [105, 106]. For example, it is sufficient to treat only bounded domains for our purposes.

As before, let $(\mathcal{T}_\ell^\square)_{\ell\in\mathbb{N}}$ be a family of structured meshes with mesh size $h_\ell = 2^{-\ell} h_0$. The nodes of $\mathcal{T}_\ell^\square$ are denoted by $\mathcal{N}_\ell^\square$, $\ell \in \mathbb{N}$. To derive multilevel finite elements to be used in computations, one fixes a maximal level index L. Then, a very specific mesh hierarchy is constructed from the geometry of the domain Ω and the given meshes. Later, an associated hierarchy of nested finite element spaces may be defined similar to the abstract concept already analyzed in Chapter 3. In fact, the created spaces highly depend on the algorithm employed for the adaptation of the fine mesh to Ω and, thus, the adaptation of the coarse meshes, too.

First, let us create a fine mesh $\mathcal{T}_L$ which resolves the computational domain but remains logically regular, i. e., the connectivity of the nodes is unchanged. For this purpose, the mesh size h_L has to be sufficiently small such that the mesh $\mathcal{T}_L^\square$ is capable of resolving all geometric details of $\partial\Omega$. At this specified fine level, the elements near the boundary are adapted by small distortions moving a node $p \in \mathcal{N}_L^\square$ to $\partial\Omega$ if $\operatorname{dist}(p, \partial\Omega) \ll h_L$.

We assume that the boundary of the computational domain is represented exactly by this procedure, although in practice this will only be done to sufficient accuracy; see also Remark 1.5. Then, only interior elements are maintained in the new mesh $\mathcal{T}_L$; the set of nodes is called $\mathcal{N}_L$. Simple examples of such an adapted mesh and the domain, respectively, are illustrated in the right part of Figure 4.3.

As approximation of $H^1_D(\Omega)$ at level L, we use the standard space of Lagrange conforming finite elements of first order associated with the constructed fine mesh $\mathcal{T}_L$ and denote it by X_L, as usual. For the derivation of coarse approximation spaces, the meshes $(\mathcal{T}_\ell)_{\ell<L}$ are generated from the regular meshes $(\mathcal{T}_\ell^\square)_{\ell<L}$ via small distortions by moving the nodes according to the corresponding fine nodes. This means that a node in $\mathcal{N}_\ell^\square$ is relocated if and only if its counterpart in $\mathcal{N}_L^\square$ is; see Figure 4.3.

Moreover, an element $T \in \mathcal{T}_\ell^\square$, $\ell < L$, potentially adapted by a movement of some of its nodes as explained above, is only included in $\mathcal{T}_\ell$, if at least one of its (logical) children is an element in $\mathcal{T}_{\ell+1}$, namely has not been discarded as "outside element", or if T (physically) contains a node from $\mathcal{N}_{\ell+1}$. On the whole, the outlined procedure may be considered a

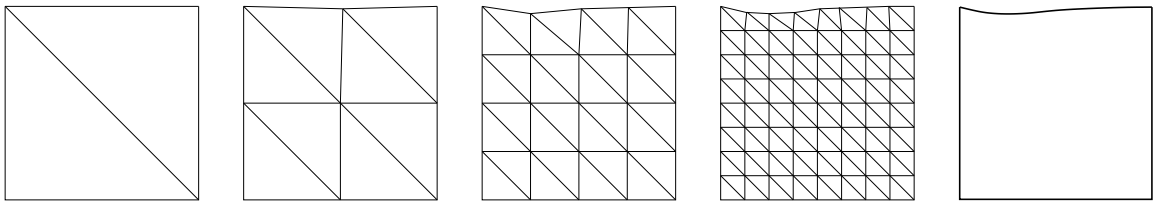


Figure 4.3. Detail of a composite finite element mesh hierarchy $\mathcal{T}_{L-3}, \ldots, \mathcal{T}_L$ and the respective portion of the domain. The regular meshes are distorted by moving some of the fine mesh nodes with their associated coarse mesh nodes to the boundary of Ω. As the displacement of the nodes is of a size significantly smaller than h_L, the influence on the much coarser meshes is rather small; here, the distortion is almost invisible in $\mathcal{T}_{L-2}$.

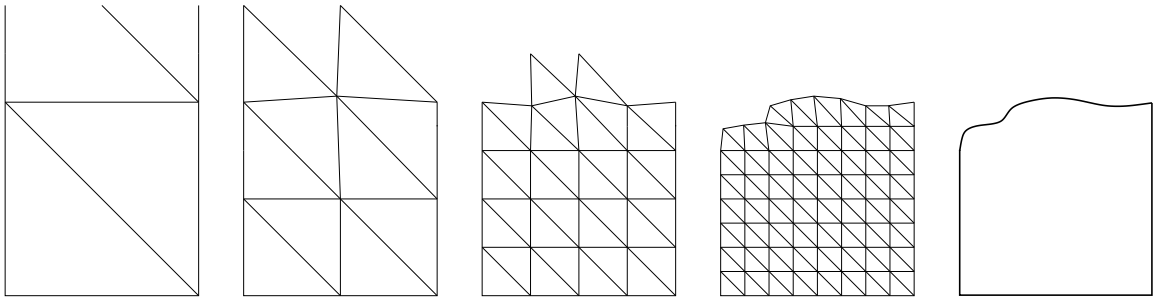


Figure 4.4. Here, the situation is more complex than in the previous sketch. A parent element is included in the coarse mesh if at least one of its (logical) children is present in the fine mesh or at least one fine mesh node is contained in it.

formal mapping

$$(\mathcal{T}_\ell^\square, \mathcal{N}_\ell^\square)_{\ell=0,\ldots,L} \quad \longmapsto \quad (\mathcal{T}_\ell, \mathcal{N}_\ell)_{\ell=0,\ldots,L}.$$

We refer to the provided illustrations, which both represent details of composite finite element triangulations ($d = 2$). The part of the coarse meshes depicted in Figure 4.3 is very simple as every element in $\mathcal{T}_\ell$, $\ell < L$, has exactly four children in the next finer mesh $\mathcal{T}_{\ell+1}$. Note that no triangle has a node lying outside its (logical) parent triangle. This is different in Figure 4.4, where further coarse level elements need to be considered.

Obviously, the mesh hierarchy $(\mathcal{T}_\ell)_{\ell=0,\ldots,L}$ becomes non-nested but remains logically nested as, on the one hand, no new elements are created during the adaptation and, on the other hand, only elements lying (almost) completely outside Ω are discarded. In other words, there are well-defined parent–child relations stemming from the regular meshes, but the portion of "physical space" covered by the children may be slighly different from that of the parent.

As all nodes which are present at coarser levels are moved according to the fine level adaptation, the procedure essentially yields a node-nested sequence of meshes $(\mathcal{T}_\ell)_{\ell=0,\ldots,L}$, namely coarse nodes are also fine nodes. Only coarse nodes lying outside the domain and thus corresponding to zero values may not have a fine level analogon. Other node-nested

hierarchies obtained by geometric coarsening will be studied in Section 4.2.

Remark 4.2 (Mesh adaptation versus mesh generation)**.** *Certainly, the issue of mesh adaptation has not been treated carefully enough. As we have investigated a more flexible way to construct coarse level spaces using general non-nested meshes in Chapter 3, we do not go into detail here. However, let us point out that the development of robust adaptation methods which produce meshes of high quality is a demanding task and may require quite elaborate techniques, especially for $d = 3$. In fact, one needs to proceed as carefully as in general mesh generation algorithms.*

As a rule, unless more sophisticated techniques are used, the mesh size h_L has to be relatively small to avoid degeneration of elements. The user also has to balance the desired accuracy of the resolution of the domain boundary by the meshes or spaces of the hierarchy to be constructed. We refer to [105, 106] for more details on possible adaptation procedures.

The composite finite element method can in principle exploit some of the advantages of regular meshes summarized in Remark 4.1. For the intergrid transfer, this holds true outright as long as no special post-processing is necessary which destroys the simple connectivity. In any case, one needs to find a way to locally process the additional information generated during the adaptation procedure to retain the overall regular structure. Then, it turns out that the local stiffness matrices only need to be computed in the neighborhoods of the boundary in the sense of Remark 4.3. We return to this issue in the next paragraph.

Construction of nested coarse spaces

Now, having the coarse meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ ready to hand, we may define a space hierarchy by a recursive construction. For this purpose, let X_ℓ be the finite element space associated with the adapted mesh $\mathcal{T}_\ell$, $\ell \in \{0, \dots, L-1\}$. Recall that these spaces are non-nested as the meshes are merely node-nested.

As we have already seen in full detail in Chapter 3, one needs to make use of suitable transfer operators to obtain a sequence of nested spaces. Here, following [105], the nodal interpolation $\mathcal{I}_\ell^{\ell+1} : X_\ell \to X_{\ell+1}$ is employed as prolongation operator for $\ell \in \{0, \dots, L-1\}$. Then, owing to the abstract formulation developed in Chapter 3, the coarse spaces are immediately defined by

$$V_\ell := \mathcal{I}_{L-1}^{L} \cdots \mathcal{I}_\ell^{\ell+1} X_\ell, \quad \forall\, \ell \in \{0, \dots, L-1\}. \tag{4.2}$$

Again, we have $V_0 \subset \ldots \subset V_L := X_L$ and bases of the new spaces may be written down via linear combinations of fine level basis functions in a straightforward manner.

A major motivation for constructing the composite finite element spaces is the desire to represent complicated domains by few degrees of freedom, similar to the boundary fitted finite element method from Section 4.1.2. Here, this comes at the price of more complicated basis functions which are certainly no longer piecewise linear with respect to any reasonable coarse mesh. Note that, if one is primarily interested in such a coarse space, it may be possible to avoid the construction of large portions of the meshes $(\mathcal{T}_\ell)_{\ell=0,\dots,L}$; see [106]. Nevertheless, a successive refinement of a given regular mesh in a sufficiently large neighborhood of the potentially complicated boundary is necessary such that the presented recursive definition of the spaces in (4.2) makes sense at least locally.

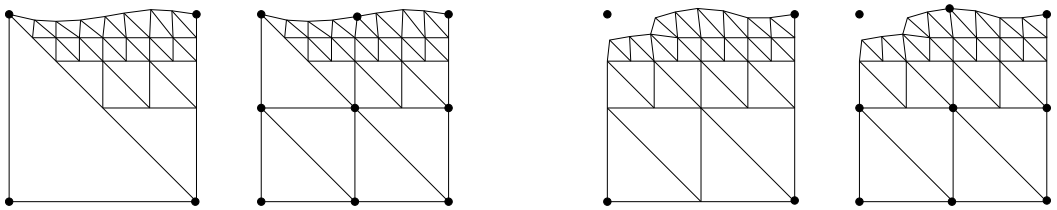


Figure 4.5. Details of composite meshes (levels $L-3$ and $L-2$) for the examples from Figure 4.3 (left) and Figure 4.4 (right). As explained in Remark 4.3, elements from different levels are gathered to illustrate the composition of the coarse level functions near the boundary. Note that, for the second example, the depicted meshes (but not the coarse degrees of freedom marked by dots) happen to be identical for these two levels.

Remark 4.3 (Composite meshes)**.** *To illustrate the term "composite finite element method" a little more, we point out the dependence of a coarse level function on the fine level mesh, more precisely on the adaptation process. This point of view actually comprehends the idea of constructing a "composite mesh" which is not a proper finite element mesh but merely indicates the hierarchical composition of the coarse level spaces; see [106].*

We have seen that only in a neighborhood of the boundary the finite element spaces X_ℓ are modified to "composite spaces" V_ℓ consisting of suitable linear combinations of fine level functions from $V_L := X_L$. Thus, gathering all elements which are necessary to evaluate a function at level $\ell < L$, one may imagine a "composite mesh" where, recursively, an element is replaced by its logical children if one of them has been adapted before. Figure 4.5 illustrates these purely logical meshes for the examples presented in the previous two sketches (Figure 4.3 and Figure 4.4), in each case for the levels $L-3$ and $L-2$. To avoid ambiguity, the dots represent the positions the coarse degrees of freedom are associated with, namely the nodes in $\mathcal{N}_\ell$, $\ell < L$.

This highlights the fact that, near the boundary, basis functions at coarse levels are "composed" of certain fine level functions. Note that the notion of being close to the boundary depends on the level. The coarser the mesh $\mathcal{T}_\ell$, the more likely an element $T \in \mathcal{T}_\ell$ is close to the boundary in the sense that one of its children has been adapted by moving one of its nodes. In this case, one of the mappings $\mathcal{I}_\ell^{\ell+1}, \dots, \mathcal{I}_{L-1}^{L}$ is not the identity on T. For example, in Figure 4.5, the node on the bottom right should be considered "close to the boundary" at level $L-3$ but "far away from the boundary" at level $L-2$ in both the left and the right case.

Finally, we notice that an equivalent way to construct the multilevel spaces is achieved by interpolating the original functions from X_ℓ on this composite meshes at all levels $\ell < L$; see [106] for a precise definition. We prefer the more algebraic description in (4.2) as it is closer to our own formulation.

The previous remark indicates that one can take advantage of the hierarchical structure for the implementation of the presented multilevel finite elements; see also Remark 4.1. In the semi-geometric framework of Chapter 3, this means that we may exploit certain a priori knowledge of the generated mesh hierachy for the computation of the level transfer

and thus for the assembly of the coarse level operators. More precisely, for an efficient evaluation of coarse level (basis) functions, it is beneficial that the prolongation operators locally coincide with the identity mapping on certain parts of the domain where successive meshes are locally nested. Therefore, depending on the complexity of $\partial\Omega$, at different levels $\ell \in \{0, \dots, L-1\}$, there are usually a lot of nodes $p \in \mathcal{N}_\ell$ and elements $T \in \mathcal{T}_\ell$ such that

$$\mathcal{I}_\ell^{\ell+1} \lambda_p^\ell(\boldsymbol{x}) = \lambda_p^\ell(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in T. \tag{4.3}$$

In case (4.3) holds locally on $\omega_p \cap \omega_q$ for some nodes $p, q \in \mathcal{N}_\ell$, the corresponding stiffness matrix entry may even be precomputed due to the regularity of the mesh.

Approximation properties

It is true that the coarse level spaces are primarily determined by the resolution of the fine level boundary and are thus no finite element spaces in the proper sense. Still, the relation to the standard multilevel finite element setting is rather close. After all, the constructed meshes are not just logically nested but also physically "almost nested". In the adaptation of the fine mesh $\mathcal{T}_L$, only a small modification of some nodes is admissible. Therefore, the displacement of the nodes or the distortion of the elements is in the same order, namely significantly smaller than h_L, also at the coarser levels. Simply put, the influence of the relocated nodes becomes more negligible with decreasing level index ℓ. This is the main ingredient of the a priori error analysis in [105].

Let us emphasize that this is in contrast to the "loose" coupling of successive meshes in the semi-geometric setting of Chapter 3. Nevertheless, the weak assumptions elaborated in full detail in Section 3.4 allow for the proof of a quasi-optimal result. The semi-geometric preconditioners may be created in a slightly more general way as both the fine and the coarse level meshes are not confined to such a specific structure. In addition, we will investigate a wide range of possible operators for the information transfer between non-nested meshes.

It is the restrictive structure of the mesh hierarchy which allows the authors of [105] to obtain optimal approximation properties of the constructed coarse spaces. More precisely, assuming that for each element the sum of certain distortion measures of all its logical children over all levels remains bounded, see [105, Assumption 3 (f)], one may prove the H^1-stability of the required compositions of the prolongation operators $(\mathcal{I}_\ell^{\ell+1})_{\ell=0,\dots,L-1}$ and also L^2- and H^1-approximation properties of the spaces $(V_\ell)_{\ell=0,\dots,L}$ for H^2-regular problems, in each case with constants which are independent of the number of levels L.

Note that the above assumption constitutes not just a requirement on the relation between two successive meshes but indeed a condition on the adaptation process and thus on the entire mesh hierarchy. It can be satisfied by using the routines for the generation of composite finite element meshes proposed in [105, 106].

4.1.4 A variant with cut off finite element functions

A related technique using structured meshes is motivated by image based computing and resolves the potentially complicated boundaries of the computational domain by means of a level set function. Here, the abstract formulation of a multilevel space hierarchy is almost

as simple as the one put forward in Section 4.1.1 for "gradually filling the domain". The algorithmic realization may be more involved, though. In this section, we briefly review this variant of the composite finite element method introduced in [137]. We do not explore the approach, which is particularly favorable for some applications, in full detail but rather extract a couple of key aspects.

As before, we consider regular (triangular or tetrahedral) meshes $(\mathcal{T}_\ell^\square)_{\ell=0,\dots,L}$ of a surrounding square or cube symbolically denoted by $\square \subset \mathbb{R}^d$. Different from the previous paragraph, the auxiliary spaces $(X_\ell)_{\ell=0,\dots,L}$ are the linear Lagrange finite element spaces directly associated with these undistorted meshes. The impressively simple idea for the construction of a sequence of nested approximation spaces $(V_\ell)_{\ell=0,\dots,L}$ is merely to restrict the above functions to the domain Ω. Thus, the multilevel discretization may be defined all at once via

$$V_\ell := X_\ell \cap \mathcal{C}^0(\Omega) = \{v|_\Omega \mid v \in X_\ell\}, \quad \forall\, \ell \in \{0,\dots,L\}. \tag{4.4}$$

Evidently, the characterization (4.4) is not inherently recursive although the algorithmical realization usually is. This is due to the need for a suitable approximation of $\partial\Omega$; we will return to this point below.

Naturally, the restriction in (4.4) can be understood as a resolution of Ω by cutting off the basis functions of $X_\ell \subset H^1(\square)$. For this purpose, let χ_Ω be the characteristic function of the computational domain, i. e.,

$$\chi_\Omega : \mathbb{R}^d \to \mathbb{R}, \quad \chi_\Omega(\boldsymbol{x}) := \begin{cases} 1, & \text{if } \boldsymbol{x} \in \overline{\Omega} \\ 0, & \text{otherwise.} \end{cases}$$

As before, let $\Lambda_\ell = (\lambda_p^\ell)_{p\in\mathcal{N}_\ell}$ be the nodal basis of X_ℓ for $\ell \in \{0,\dots,L\}$. Then, one obtains a basis $\widetilde{\Lambda}_\ell = (\widetilde{\lambda}_p^\ell)_{p\in\widetilde{\mathcal{N}}_\ell}$ of V_ℓ with

$$\widetilde{\mathcal{N}}_\ell = \{p \in \mathcal{N}_\ell \mid \operatorname{supp}(\lambda_p^\ell) \cap \Omega \neq \emptyset\}, \quad \forall\, \ell \in \{0,\dots,L\},$$

and

$$\widetilde{\lambda}_p^\ell = \lambda_p^\ell\, \chi_\Omega, \quad \forall\, p \in \widetilde{\mathcal{N}}_\ell,\ \ell \in \{0,\dots,L\}. \tag{4.5}$$

Note that, in general, no Dirichlet conditions may be prescribed at $\partial\Omega$. More precisely, one needs a proper mesh at those portions $\Gamma_D \subset \partial\Omega$ where Dirichlet conditions shall be enforced. Unless treated in a more sophisticated manner, Γ_D needs to align the regular mesh $\mathcal{T}_L^\square$. Although this poses difficulties for the general applicability, the method is effective for plenty of problems, for example in case $\Gamma_D \subset \partial\square$; see [137, 157]. Let us assume that the Dirichlet boundaries are treated appropriately and formally replace V_ℓ by $V_\ell \cap H_D^1(\Omega)$. Then, the nestedness of the constructed spaces

$$V_0 \subset \ldots \subset V_L \subset H_D^1(\Omega) \tag{4.6}$$

is trivial.

As indicated before, some efforts have to be made to attain the cut off basis functions conceptionally defined in (4.5). We emphasize that $\partial\Omega$ needs to be approximated only once

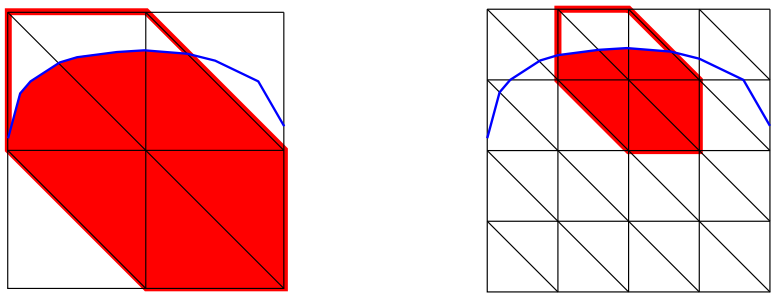


Figure 4.6. Coarse and fine level basis functions are cut off depending on the representation of the boundary at the finest level (depicted on the right). The cut off support is filled with red whereas the original one is indicated by the red frame.

for the entire space hierarchy as the resolution at the fine level will determine the resolution at all coarser levels. For this purpose, following [137], suppose that the boundary of the computational domain is given as the zero level set of a function $\phi : \square \to \mathbb{R}$ such that $\partial\Omega = \{\boldsymbol{x} \in \square \mid \phi(\boldsymbol{x}) = 0\}$ and $\Omega = \{\boldsymbol{x} \in \square \mid \phi(\boldsymbol{x}) < 0\}$. Neglecting potential additional geometry error, we assume that $\phi \in X_L$. In practice, if the level set function is generated by an image processing algorithm, its finite element interpolation in X_L should be sufficiently accurate in case the mesh size h_L is in the order of the pixel or voxel size.

Figure 4.6 illustrates an example of the cut off basis functions. One needs to point out that the requirement $\phi \in X_L$ is only one of many possible variants. Basically, the support of each basis function needs to be sufficiently simple, namely allow for efficient quadrature. This may require the introduction of local auxiliary meshes of the simplices intersected by $\partial\Omega$. In addition, stability issues may occur if $\mathrm{meas}_d(\mathrm{supp}(\widetilde{\lambda}_p^L))$ becomes very small for some nodes $p \in \widetilde{\mathcal{N}}_L$. We refer to [137] for a detailed discussion. A similar local remeshing procedure for the integration of basis functions over complicated portions of structured meshes is used in the different context of discontinuous Galerkin methods in [81].

Remark 4.4. *For the sake of clarity, let us point out that the "cut off functions" in the present composite finite element variant must not be confused with the "truncated functions" employed in the monotone multigrid methods as explained in Section 6.2.*

For the definition of the multilevel hierarchy (4.6), the presumably most important point is that the difficulties only lie in the representation of the interface at the finest level. This is because the coarse level basis functions are simple linear combinations of the fine level basis functions entirely determined by the connectivity of the regular meshes $(\mathcal{T}_\ell^\square)_{\ell=0,\dots,L}$. Once the basis $\widetilde{\Lambda}_L$ is constructed, e. g., via the procedure developed in [137], the bases $(\widetilde{\Lambda}_\ell)_{\ell<L}$ are explicitely given by the standard prolongation operators between the original spaces $(X_\ell)_{\ell=0,\dots,L}$. Thus, the construction of a convenient representation of the basis $\widetilde{\Lambda}_L$ is a key issue. As in the previous case of composite finite elements, one usually needs to take coarse level basis functions into account which are associated with nodes lying far outside of Ω. But, unlike before, the present variant also incorporates fine level degrees of freedom associated with outside nodes.

This variant of the composite finite element method is very attractive for multigrid

methods not only from an algorithmical point of view but also when it comes to the required multilevel approximation properties. It has the advantage, as opposed to all other approaches, that the computational domain in resolved exactly without changing the gradients of the (coarse level) basis functions. As long as one does not introduce artificial couplings of components which are physically too far away from each other, see [137, 157], the coarse spaces have perfect approximation properties provided that the stability issues concerning the minimum size of the supports of the basis functions are taken care of. As indicated before, this is subject to the restriction that the part Γ_D of the boundary is resolved by the mesh or the Dirichlet boundary conditions have been treated in another appropriate way.

Remark 4.5 (Immersed interface multigrid method)**.** *Finally, let us briefly mention another approach which naturally fits into this context. In the fashion of the previously described methods, it may be considered a very special multigrid method associated with a very special discretization also based on Cartesian meshes.*

The immersed interface method developed in [136] is a discretization for elliptic equations which is specifically designed for problems whose solutions have jumps across complicated interfaces through the computational domain. Employing a standard finite difference scheme in the rest of the domain, one resolves the interfaces by modified stencils near the jump. Now, the application of geometric multigrid ideas essentially requires similar techniques for the derivation of suitable (weighted) prolongation operators as introduced earlier in [7, 67]. Note that, just as the discretization itself, the immersed interface multigrid method [2, 3, 4] takes the local shape of the interface into account.

For their operator-dependent prolongation, the authors of [4] experience a loss of the diagonal dominance of the coarse level matrices if the restriction is chosen as the transpose of the prolongation. Alternative (non-variational) choices are studied in [2]. See also [185] for a similar approach.

4.2 Geometric coarsening

In this section, we consider another concept for the construction of coarse level approximations of finite element spaces. This methodology does not introduce a special fine level discretization which is more amenable to the purpose of multilevel algorithms. Instead, the main input is the original unstructured mesh. The procedures we bring into focus in the following are geometric coarsening and agglomeration algorithms which attempt to construct a space hierarchy based on the given fine mesh.

Let us emphasize that the coarsening techniques reviewed in Section 4.2.1 and Section 4.2.2 produce non-nested but node-nested mesh hierarchies. Similar to our approach, one needs an additional mesh generation for the coarse levels. However, the procedures described in the following enforce special structures on the constructed coarsened meshes, which may lead to degenerate meshes unless treated with great care; see also the comprehensive paper [146].

In a way, the semi-geometric concept constitutes the "weakest" of all conceivable coarsening strategies. Indeed, it seems that the theoretical and algorithmical aspects developed

in Chapter 3 may be applied to the mesh hierarchies derived in the present section. This is because the algorithms are explicitly designed to satisfy a local coarsening assumption comparable to (3.9), at least following some heuristic. Still, it would be surprising if any of the following coarsening or agglomeration schemes allowed for a superior theoretical result without requiring strong additional assumptions. Compare also with the discussion of the composite finite element method in Section 4.1.3.

Most geometric coarsening strategies are in fact graph-based algorithms; geometric agglomeration may also fall into this category. A central notion in graph theory is the one of independent sets. To avoid any conflicts with the notation in the other parts of this thesis, we use calligraphic symbols in the context of graph theoretical techniques. As it will be useful at several points, let us recall the following definition without digressing too much.

Definition 4.6 (Maximal independent set)**.** *Let $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ be a graph with vertex set $\mathscr{V}$ and edge set $\mathscr{E}$. Besides, (v_1, v_2) denotes the edge between $v_1, v_2 \in \mathscr{V}$. Then, a subset $\mathscr{U} \subset \mathscr{V}$ is called maximal independent set (MIS) if*

- *it is independent: $(u_1, u_2) \notin \mathscr{E}$ for all $u_1, u_2 \in \mathscr{U}$ and*
- *it is maximal: for all $v \in \mathscr{V} \setminus \mathscr{U}$ there is a vertex $u \in \mathscr{U}$ such that $(u, v) \in \mathscr{E}$.*

In the following, maximal independent sets either of graphs essentially representing the mesh connectivity or of some conflict graphs are employed in the geometric coarsening strategies. Although it is not difficult to obtain such a set, in general, an MIS of a graph is not unique. However, a feasible set may be computed using both well-known and simple search algorithms; see, e. g., [102]. An example is the following

Algorithm 4.7. *For a given graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, determine a maximal independent set $\mathscr{U}$ via the following greedy algorithm.*

```
greedyMIS (𝒱, ℰ) {
    Initialize: 𝒰 = ∅, 𝒲 = ∅
    for (v ∈ 𝒱) do {
                if (v ∉ 𝒲) {
                        Update:
                        𝒰 ← 𝒰 ∪ {v}
                        𝒲 ← 𝒲 ∪ {w ∈ 𝒱 | (v, w) ∈ ℰ}
                }
    }
    return 𝒰
}
```

Evidently, the resulting set highly depends on the sorting of the vertices in $\mathscr{V}$. Note that it is not easy to keep the MIS particularly small or particularly large at reasonable costs; one cannot expect it to satisfy any additional assumptions. The fact that both Definition 4.6 and Algorithm 4.7 are of local nature implies that the greedy search may be implemented efficiently as an advancing front technique.

4.2.1 Basic node-nested coarsening with remeshing

For the basic geometric coarsening techniques, regard a given unstructured mesh $\mathcal{T}$ as graph $\mathcal{G}$ based on the mesh connectivity in the usual sense, i. e., the nodes $\mathcal{N}$ of $\mathcal{T}$ are the vertices of $\mathcal{G}$ which are connected by an edge if and only if they share an edge of an element in $\mathcal{T}$. We comment on other possibilities later.

Now, a node-nested coarse mesh is constructed in [48, 102] by a suitable remeshing of a selected subset of the nodes $\mathcal{N}$. The required set of coarse nodes may be obtained as an MIS of the graph $\mathcal{G}$, for instance, by Algorithm 4.7. Then, it is suggested to compute a (generally not unique) Delaunay triangulation of the just determined set of points; see [102]. The Delaunay algorithm is supposed to generate one of the best possible meshes with respect to the aspect ratio. The authors of [48] prefer algorithms of Cavendish-type [46] for the generation of the coarse meshes. All mentioned papers deal with the case $d = 2$.

Usually, it is beneficial to treat the boundary nodes in a special way. Heuristic approaches to roughly retain the overall shape of the computational domain are either to consider the boundary separately, i. e., form two different graphs for the boundary nodes and the interior nodes, or to perform Algorithm 4.7 with the boundary nodes appearing first in the vertex list; see [48, 102].

A multilevel hierarchy is generated naturally by recursion. Unfortunately, the repeated application of the above strategies based on remeshing maximal independent sets may lead to a degrading mesh quality, namely a decreasing aspect ratio. This issue is analyzed in full detail in [146]. In any case, if the multilevel spaces are constructed by a variational or a non-variational approach using the created mesh hierarchy, one should expect the quality of the constructed coarse meshes to matter. Therefore, every coarsening algorithm should have some instruments to guarantee for a minimum quality of the constructed elements. The same is, in principle, true of the agglomeration techniques reported on below.

The authors of [48] also experiment with a closely related coarsening scheme which uses an MIS of the dual graph of the mesh. Here, in the dual graph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$, the elements $\mathcal{T}$ are the vertices $\mathcal{V}'$ which are connected by an edge in $\mathcal{E}'$ if and only if they share an edge in $\mathcal{E}$. Then, a remeshing of the centroids of the determined elements yields a non-nested mesh which is not node-nested any more.

4.2.2 Advanced coarsening algorithms

The coarsening of an unstructured mesh via a remeshing of an MIS is motivated by the following basic considerations. On the one hand, it is reasonable to pick an independent set to prevent regions from being excluded in the overall coarsening procedure. On the other hand, it is mandatory to pick a maximal set as otherwise it is likely that too many degrees of freedoms are removed in certain regions.

In computational geometry, more sophisticated techniques exist which do not just consider the standard mesh connectivity graph. Instead, a so-called conflict graph is established, which may include additional geometric information about the given mesh. The fundamental idea proposed in [146] is to connect "conflicting" nodes, namely the ones which are physically too close together, by an edge in the new graph. For this purpose, suitable local length scales need to be determined from the given mesh; see [146, 149].

Moreover, it is possible to utilize special routines to retain certain characteristic features such as sharp corners or folds. We refer to [149] and the references therein.

Instead of globally remeshing an MIS, one may also perfom an incremental vertex deletion, e. g., by edge contraction. However, note that coarsening of three-dimensional meshes is much more demanding than that of two-dimensional meshes. This is due to the fact that the connectivity is more involved in case $d = 3$; see [149]. Some of the above techniques have been applied, e. g., in [6] to build multilevel preconditioners associated with triangular surface meshes.

Another coarsening algorithm for unstructured triangular meshes ($d = 2$) has been developed in [14, 15] for the construction of hierarchical basis preconditioners. The idea is to regard the given mesh as a perturbation of an adaptively refined structured mesh. Assuming certain knowledge of the underlying refinement procedure, the authors attempt to mimic a "reverse" process. As a result, a non-nested but node-nested coarse mesh is obtained which is connected to the fine mesh by a particular logical structure similar to the one considered for the composite finite element methods in Section 4.1.3. For instance, each fine node which does not coincide with a coarse node lies logically, but not necessarily physically, on a coarse edge. This special relation between the successive meshes makes it possible to prove stability estimates for certain composed interpolation operators immediately exploiting the logical structure. We refer to [15] for a detailed analysis.

4.2.3 Element agglomeration

In the early nineties, multigrid-like methods based on volume or element agglomeration techniques have been developed and employed in practice by engineers in computational fluid dynamics; see, e. g., [120, 130, 183]. In fact, the idea to construct coarse level approximations by simply forming agglomerates appears most natural for finite volume discretizations. The technique of element agglomeration has also been applied in the context of algebraic multigrid methods [115]. On the (purely) algebraic side, related considerations led to the development of aggregation-based algebraic multigrid methods [24]; see, e. g., [20, 148] for an overview and the series of papers [180, 181, 182] for the more sophisticated variant of smoothed aggregation.

Applied to the finite element method, instead of constructing proper coarse meshes for an unstructured fine mesh, as is done in all previously described approaches, geometric agglomeration procedures generate macroelements, i. e., sets of polygons or polyhedra consisting of plain unions of fine level elements. This may be done, e. g., by a direct greedy search [120, 130, 183] comparable to Algorithm 4.7, essentially fusing neighboring elements which have not yet been handled.

In their proceedings contribution [50], the authors identify the coarse nodes and, more importantly, the surrounding macroelements by an advanced graph-based method. We skip the details but remark that a basic ingredient is the removal of certain edges of the dual graph of the given mesh (the ones surrounding the coarse nodes, which have been chosen as subset of the fine nodes). See Figure 4.7 for an examplary illustration of the agglomeration process. Then, the coarse level spaces are defined by suitable prolongation operators, which is similar to the construction in Section 3.2.1. More precisely, a basis

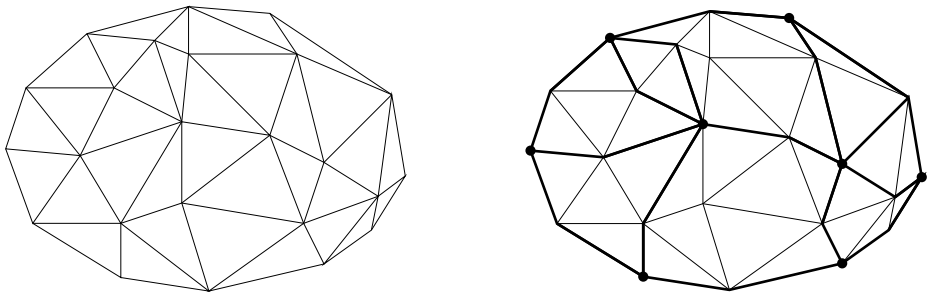


Figure 4.7. Agglomeration of elements of an unstructured mesh to macroelements according to [50]. The coarse nodes marked by dots form an MIS of the graph of the fine mesh.

function associated with a selected coarse node is a linear combination of the fine level basis functions associated with the nodes of the adjacent macroelements. In particular, this linear combination generally may depend on the locations of the fine nodes. However, it seems that very simple interpolation weights solely coming from combinatorical considerations are sufficient; see [50]. One could call the traditional choice in [120], where only the numbers one and zero appear as coefficients, "plain agglomeration".

In general, during the recursion to obtain a multilevel space hierarchy, even if a straightforward procedure is available to further coarsen an agglomerated mesh in a reasonable fashion, special care has to be taken to prohibit a degeneration of the formed macroelements by a repeated agglomeration. The same issue arises in the context of the coarsening algorithms with remeshing in Section 4.2.1 and Section 4.2.2.

To obtain a quasi-optimal convergence result for the derived multilevel algorithms employing the particular agglomeration strategy of [50], the existence of suitable fine-to-coarse mappings (as described in detail in Section 3.4) has been proved. As a matter of fact, we need to point out that the critical assumption in [50, Section 5], namely the uniform boundedness of the diameter of the macroelements by a constant times the fine mesh size, is nothing else but a condition on the regularity of the agglomeration process. Considering the practical examples presented in [120, 130, 183], this condition seems to be rather difficult to satisfy. The ideas of [146], i. e., to establish a suitable conflict graph as briefly outlined in Section 4.2.2, might be a remedy.

5 Prolongation and restriction operators between non-nested meshes

The practical suitability of the semi-geometric framework developed in Chapter 3 relies on an effective choice of the employed prolongation operators. In this chapter, we analyze the information transfer between finite element spaces originating from non-nested meshes. We are convinced that a deep insight into diverse transfer concepts is very helpful for both the construction of specific operators and their application in multilevel methods.

It seems that, until now, virtually all development of approximation operators in this context was for the purpose of numerical analysis exclusively, neglecting practical concerns. Not enough attention has been paid to the feasible usage of geometrically inspired transfer operators between non-nested finite element spaces in numerical algorithms. The author believes that, principally, this is due to the fact that the potential capability of algorithms based on non-nested meshes have not been studied to the full extent. Mostly, the opportunity was missed to include an adequate discussion of the properties relevant for the present purposes and to draw conclusions for the development of algorithms. However, we need to point out that in the context of a meshfree method possible replacements for the L^2-projection between non-nested partition of unity spaces have been studied in [99, 170]. The only other notable exception seems to be the formulation of suitable coupling conditions at interior interfaces in non-conforming domain decomposition methods [190, 191].

To understand similarities and differences between possible prolongation operators for non-nested meshes, we review basic properties of some selected transfer concepts in a theoretical part (Section 5.1 to Section 5.5). We are confident that our considerations cover a respectable range of geometric ideas. A part of our observations reflects the current state of the literature. However, it seems that other points made in this chapter, especially concerning the practical relevance of certain stability and projection properties, have not been examined so far. In particular, we consider a generalized projection operator called "pseudo-L^2-projection", which is based on a Petrov–Galerkin variational formulation with respect to biorthogonal test spaces, and provide a complete analysis. In Section 5.6, we comment on the applicability of the transfer concepts in the semi-geometric framework. The practical implementation of the respective operators in a customary finite element code is addressed in Section 5.7.

In an equally important experimental part (Section 5.8), we investigate the fundamental characteristics of information transfer between non-nested meshes in practice. The properties of the analyzed operators are assessed and the differences between them are measured with respect to suitable norms. In particular, we draw a map showing the mutual relations of the transfer concepts. Indeed, the performed studies provide valuable insight into the nature of mappings between distinct and non-nested finite element spaces; they also facilitate some conclusions for the application in the semi-geometric multilevel methods.

We point out that the quest for suitable prolongation operators which allow for the construction of efficient preconditioners is related to the development of coarsening strategies in algebraic multigrid methods. Such a coarsening is, in general, not free to access geomet-

ric information except for the matrix entries, though. For most strategies, this makes it difficult to find operators comparable to the ones we analyze in the following. This further illustrates that a major difference between the semi-geometric and the algebraic approach is how the approximation properties of the constructed coarse spaces are achieved, a fact which has already been mentioned in the previous two chapters.

Outline of the theoretical part

In the following, beside a new generalized projection operator, which is derived and analyzed in Section 5.5, we consider quite a few different, geometrically inspired transfer concepts. This is done in the first four sections starting with Section 5.1. Both intuitive and more elaborate mappings are examined, partly from the literature on multigrid methods and domain decomposition with unstructured meshes. Most of the operators to be studied comprise some kind of weighting (with discrete test functions, particularly in $L^\infty(\Omega)$) and are thus well-defined in $L^1(\Omega)$. But some need more regularity such as continuity of the input function. We discuss locally and globally defined operators including well-known quasi-interpolation concepts and also focus on their algorithmic structure. Above all, we aim to assess the suitability of the transfer concepts for the use in the semi-geometric framework.

From now on we assume the current level index ℓ to be chosen appropriately. By Lemma 3.12 we know that it is sufficient to consider a fixed domain called Ω in the following, without any index. We have denoted a generic or unspecific prolongation operator by Π. To every concrete operator we will assign a different calligraphic symbol ($\mathcal{I}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$), sometimes varied by a tilde or a hat. As before, if an operator maps between the two non-nested spaces $X_{\ell-1}$ and X_ℓ or between the two nested spaces $V_{\ell-1}$ and V_ℓ, this will be indicated by, e. g., $\Pi_{\ell-1}^{\ell}$. Similarly, an operator mapping some other space, such as a Lebesgue or Sobolev space, to the finite element space X_ℓ will be denoted by, e. g., Π_ℓ. This shall suggest that a mesh $\mathcal{T}_\ell$ with a local mesh size function $h_\ell \in L^\infty_>(\Omega)$ in the spirit of Section 3.4.1 is always involved.

Before going into detail about the single concepts, let us recall Lemma 3.2 which implies that the property of an operator to preserve constants, except for the domain boundaries, ensures that each of the constructed bases $(\widetilde{\Lambda}_\ell)_{\ell=0,\dots,L-1}$ is a partition of unity there. Note that every projection clearly satisfies this condition.

5.1 Standard finite element interpolation

First, we consider the most elementary operator. The standard finite element interpolation or nodal interpolation in case of first order Lagrange elements is defined by

$$\mathcal{I}_\ell : \mathcal{C}^0(\Omega) \to X_\ell, \quad u \mapsto \mathcal{I}_\ell u := \sum_{p \in \mathcal{N}_\ell} u(p)\lambda_p^\ell.$$

The operator is surjective, namely $\mathcal{I}_\ell(\mathcal{C}^0(\Omega)) = X_\ell$. The interpolation in X_ℓ with the domain restricted to the finite element space $X_{\ell-1}$ is called $\mathcal{I}_{\ell-1}^{\ell}$. One has to note that, in

general, $\mathcal{I}_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell$ is no L^2-orthogonal projection if the meshes $\mathcal{T}_{\ell-1}$ and $\mathcal{T}_\ell$ are non-nested. But the interpolation operator $\mathcal{I}_\ell$ is obviously a projection, i. e., for every $v \in \mathcal{C}^0(\Omega)$ we have $\mathcal{I}_\ell \mathcal{I}_\ell v = \mathcal{I}_\ell v$. Evidently, the operator is local according to Definition 3.8. Moreover, for shape regular meshes, it fortunately possesses the H^1-stability and L^2-approximation properties given in Section 3.4.1, at least when restricted to finite element spaces. This result can be found in several papers; see, e. g., [43, 49, 177].

It is surprising that an operator as simple as the nodal interpolation satisfies the requirements of the analysis in Section 3.4. This means that the sequence $(\mathcal{I}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ can actually be used in algorithms to realize the information transfer between the non-nested finite element spaces. Note that the fact that its domain is just the (strict) subset $\mathcal{C}^0(\Omega)$ of $H^1(\Omega)$ has no major influence on the applicability of the standard interpolation operator in the proofs of Section 3.4.2 and Section 3.4.3. This is because the operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ in the employed compositions are always applied to functions in standard finite element spaces. One needs to have a fine-to-coarse operator ready to hand to map from $H^1(\Omega_\ell)$ to X_ℓ in the first place, though. This may be done by means of a quasi-interpolation operator to be defined in Section 5.2.

Remark 5.1. *Let a transfer concept be such that (given two admissible spaces, the latter one being a finite element space X_ℓ) each generated operator Π_ℓ acts as the identity mapping on the target space. This is true if, for instance, Π_ℓ is a surjective projection. Then, if the meshes $\mathcal{T}_{\ell-1}$ and $\mathcal{T}_\ell$ are nested, the fact that the corresponding spaces $X_{\ell-1} \subset X_\ell$ are also nested implies immediately that the restricted mapping $\Pi_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell$ is the natural embedding and thus coincides with the standard interpolation $\mathcal{I}_{\ell-1}^{\ell}$. In any case, even if the meshes are non-nested, the induced operator $\widetilde{\Pi}_{\ell-1}^{\ell} : V_{\ell-1} \to V_\ell$ from (3.6) is the natural embedding and thus an L^2-projection between the specially constructed spaces.*

Otherwise (if there exists an element $v \in X_{\ell-1}$ such that $\Pi_\ell \Pi_\ell v \neq \Pi_\ell v$) the semi-geometric construction of the space hierarchy in Section 3.2, which directly employs the operators $(\Pi_{\ell-1}^{\ell})_{\ell=1,\dots,L}$, does not reduce to the standard scheme with the usual interpolation in case of nested meshes. Thus, a considerable part of structure is unnecessarily disregarded.

From a computational point of view, the standard nodal interpolation is very attractive. Given an arbitrary function in $\mathcal{C}^0(\Omega)$, the computation of the interpolant is very cheap with one function evaluation per node in $\mathcal{N}_\ell$, i. e., per basis function in Λ_ℓ. It is without any doubt the least expensive way to transfer information to a finite element space in a reasonable way. For the computation of the matrix $\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell}$, this amounts to the evaluation of $\lambda_q^{\ell-1}(p)$ for all $q \in \mathcal{N}_{\ell-1}$ and $p \in \mathcal{N}_\ell$. Naturally, one may neglect the combinations with $p \notin \omega_q$. This is straightforward if successive meshes are nested and parent–child relations are known. In the non-nested setting, such neighborhood relations have to be computed; see Section 3.6.2.

On the H^1-stability of the nodal interpolation

In the literature, several different proofs have been brought forth for the stability and approximation property of the standard finite element interpolation in case the domain is restricted to a (coarser) finite element space; see [43, 49, 177]. All techniques seem to have in common the dependence of the stability estimates on certain mesh properties such as

the shape regularity. This leads, in general, to constants greater than one. However, our numerical experiments suggest that the H^1-stability constant is bounded by one; see Section 5.8. Indeed, we found an elementary proof in case $d = 1$ for the fact that "interpolation smoothes". Note that the notion of shape regularity (1.16) only makes sense for $d \geq 2$. For one-dimensional problems no such assumption needs to be considered.

Only for the following reasoning, we investigate a one-dimensional setting. Let X_ℓ be a finite element space as introduced in Section 3.1 without any incorporated Dirichlet boundary conditions. In the present case $d = 1$, this approximation space is associated with a subdivision of the domain Ω_ℓ into a set of intervals $\mathcal{T}_\ell$. Then, the following lemma basically states that the linear interpolation operator $\mathcal{I}_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell$ has an H^1-stability constant less or equal one whether or not Dirichlet conditions come into play.

Lemma 5.2. *In the case $d = 1$, let the set of intervals $\mathcal{T}_{\ell-1}$ be coarser than $\mathcal{T}_\ell$ in the sense that there is no pair of elements $T_{\ell-1} \in \mathcal{T}_{\ell-1}$, $T_\ell \in \mathcal{T}_\ell$ such that $T_{\ell-1} \subset T_\ell$. Then, the nodal interpolation operator $\mathcal{I}_{\ell-1}^{\ell} : X_{\ell-1} \to X_\ell$ satisfies the following optimal H^1-stability estimates.*

- *If $\Omega_\ell \subset \Omega_{\ell-1}$,*

$$|\mathcal{I}_{\ell-1}^{\ell} v|_{H^1(\Omega_\ell)} \leq |v|_{H^1(\Omega_\ell)} \leq |v|_{H^1(\Omega_{\ell-1})}, \quad \forall\, v \in X_{\ell-1}, \tag{5.1}$$

- *otherwise,*

$$|\mathcal{I}_{\ell-1}^{\ell} v|_{H^1(\Omega_\ell)} \leq |v|_{H^1(\Omega_\ell \cap \Omega_{\ell-1})}, \quad \forall\, v \in X_{\ell-1} \cap H_D^1(\Omega_{\ell-1}). \tag{5.2}$$

Moreover, the interpolation operator $\mathcal{I}_{\ell-1}^{\ell,D} : X_{\ell-1} \to X_\ell$ enforcing zero function values at the boundary $\partial\Omega_\ell$ satisfies

$$|\mathcal{I}_{\ell-1}^{\ell,D} v|_{H^1(\Omega_\ell)} \leq |v|_{H^1(\Omega_\ell \cap \Omega_{\ell-1})}, \quad \forall\, v \in X_{\ell-1} \cap H_D^1(\Omega_\ell) \cap H_D^1(\Omega_{\ell-1}). \tag{5.3}$$

Remark. *In statement (5.3), note that $H_D^1(\Omega_\ell) \cap H_D^1(\Omega_{\ell-1}) = H_D^1(\Omega_\ell)$ if $\Omega_\ell \subset \Omega_{\ell-1}$. This is because we treat functions defined on the smaller domain as extended to the larger one by zero here. Moreover, the symbols $\subset$ and $\supset$ always include the case of equality.*

Proof. First, consider the case $\Omega_\ell \subset \Omega_{\ell-1}$. Let $v \in X_{\ell-1}$ be arbitrary. If $T_\ell \in \mathcal{T}_\ell$ is completely contained in an element of $\mathcal{T}_{\ell-1}$, we have locally exact interpolation, i. e., $(\mathcal{I}_{\ell-1}^{\ell} v)|_{T_\ell} = v|_{T_\ell}$; thus, $|\mathcal{I}_{\ell-1}^{\ell} v|_{H^1(T_\ell)} = |v|_{H^1(T_\ell)}$.

The assumptions imply that all other elements in $\mathcal{T}_\ell$ have intersections of positive (one-dimensional) measure with exactly two elements in $\mathcal{T}_{\ell-1}$. For such an element $T_\ell \in \mathcal{T}_\ell$, let us introduce some notations in Figure 5.1. Then, one has

$$|v|_{H^1(T_\ell)}^2 = \int_{S_1} \left(\frac{v_p - v_s}{|S_1|} \right)^2 dx + \int_{S_2} \left(\frac{v_t - v_p}{|S_2|} \right)^2 dx = \frac{(v_p - v_s)^2}{|S_1|} + \frac{(v_t - v_p)^2}{|S_2|}$$

as the first derivative of v is piecewise constant. Besides, we estimate

$$\begin{aligned} |\mathcal{I}_{\ell-1}^{\ell} v|_{H^1(T_\ell)}^2 = \int_{T_\ell} \left(\frac{v_t - v_s}{|T_\ell|} \right)^2 dx = \frac{(v_t - v_s)^2}{|T_\ell|} &= \frac{(v_t - v_p + v_p - v_s)^2}{|S_1| + |S_2|} \\ &< \frac{(v_t - v_p)^2}{|S_1|} + \frac{(v_p - v_s)^2}{|S_2|} + \frac{2(v_t - v_p)(v_p - v_s)}{|S_1| + |S_2|}. \end{aligned}$$

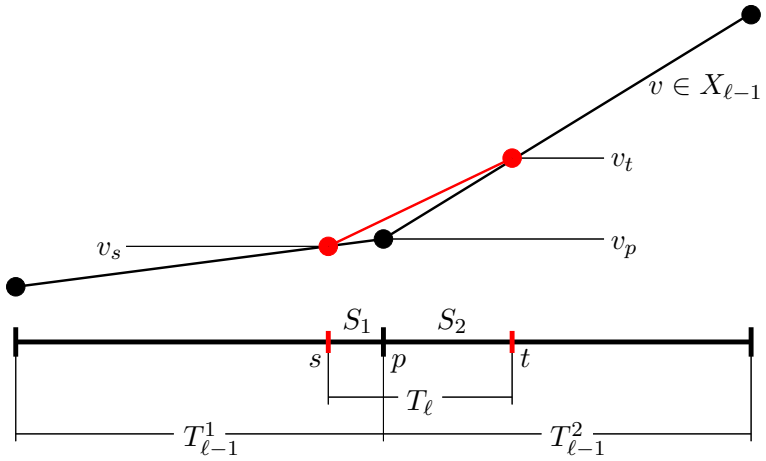


Figure 5.1. This sketch for the proof of Lemma 5.2 shows a function $v \in X_{\ell-1}$ on the coarse elements $T^1_{\ell-1}, T^2_{\ell-1} \in \mathcal{T}_{\ell-1}$ and its interpolant on the fine element $T_\ell \in \mathcal{T}_\ell$. We have the intersections $S_1 := T_\ell \cap T^1_{\ell-1}$ and $S_2 := T_\ell \cap T^2_{\ell-1}$. The values of v at the nodes $s, t \in \mathcal{N}_\ell$ and $p \in \mathcal{N}_{\ell-1}$ are denoted by v_s, v_t and v_p, respectively.

If the last term on the right hand side is less or equal zero, we have $|\mathcal{I}^\ell_{\ell-1} v|^2_{H^1(T_\ell)} < |v|^2_{H^1(T_\ell)}$. This holds in each of the following cases:

- $v_p = v_s$ or $v_p = v_t$,
- $v_p > v_s$ and $v_p > v_t$,
- $v_p < v_s$ and $v_p < v_t$.

It remains to consider the elements where $v_s < v_p < v_t$ (cf. Figure 5.1) or $v_s > v_p > v_t$. It suffices to study the first case. We will show in the following that the linear interpolant on the fine level interval T_ℓ for given values v_s and v_t at the left and right endpoint, respectively, has minimal H^1-semi-norm out of all piecewise linear functions.

Without loss of generality, let $T_\ell = (0,1)$ and $v_s = 0$, $v_t = 1$. This leaves as variables the coordinate $p \in (0,1)$ and the intermediate value $v := v_p \in (0,1)$. We introduce the functional

$$\mathcal{H} : (0,1)^2 \to \mathbb{R}, \quad \mathcal{H}(p,v) := \frac{v^2}{p} + \frac{(1-v)^2}{1-p}$$

describing the square of the H^1-semi-norm of the piecewise linear function connecting the three points $(0,0)$, (p,v) and $(1,1)$. It is not difficult to see that, given any $p \in (0,1)$, the functional $\mathcal{H}$ attains its minimum if the measured function is linear, namely if $v = p$. Indeed, the calculation of the first partial derivatives shows that the gradient $\nabla\mathcal{H}(p,v)$ vanishes if and only if $v = p$. In addition, the Hessian at these points

$$\nabla^2\mathcal{H}(p,p) = \frac{2}{p(1-p)} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

is positive semi-definite. Therefore, the functional $\mathcal{H}$ has a set of minima (with value 1) on the diagonal line $\{p = v\}$.

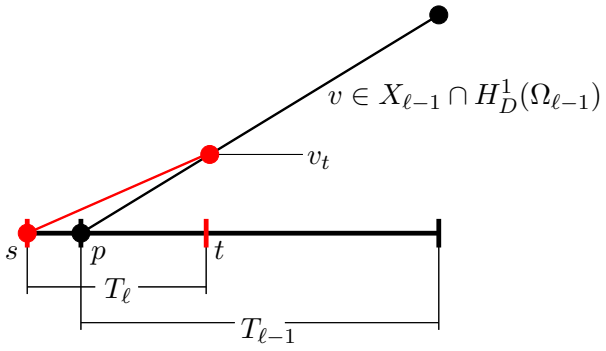


Figure 5.2. In this sketch for the proof of Lemma 5.2, the fine element $T_\ell \in \mathcal{T}_\ell$ contains a boundary node $p \in \mathcal{N}_{\ell-1} \cap \partial\Omega_{\ell-1}$. It shows a function $v \in X_{\ell-1} \cap H^1_D(\Omega_{\ell-1})$ on the coarse element $T_{\ell-1} \in \mathcal{T}_{\ell-1}$ and its interpolant on T_ℓ. We have the values 0 at the nodes $s \in \mathcal{N}_\ell$ and $p \in \mathcal{N}_{\ell-1}$ and v_t at the node $t \in \mathcal{N}_\ell$.

Applied to our problem, this means that $|\mathcal{I}^\ell_{\ell-1} v|^2_{H^1(T_\ell)} \leq |v|^2_{H^1(T_\ell)}$ also holds true for the elements $T_\ell \in \mathcal{T}_\ell$ with $v_s < v_p < v_t$ or $v_s > v_p > v_t$. Now, summation over all $T_\ell \in \mathcal{T}_\ell$ concludes the proof of (5.1).

For the second assertion (5.2), we realize that the nodal interpolation is identically zero on the exterior elements $T_\ell \subset \Omega_\ell \setminus \Omega_{\ell-1}$. On the interior elements $T_\ell \subset \Omega_{\ell-1}$, the local estimates hold as before. One of the two remaining cases is illustrated in Figure 5.2; the other one is the respective situation at the right boundary of $\Omega_{\ell-1}$. We obviously have that

$$|\mathcal{I}^\ell_{\ell-1} v|^2_{H^1(T_\ell)} = \frac{v_t^2}{|T_\ell|} \leq \frac{v_t^2}{|T_\ell \cap T_{\ell-1}|} = |v|^2_{H^1(T_\ell \cap T_{\ell-1})}.$$

Consequently, the estimate (5.2) follows as before by adding up the local contributions.

Finally, let us consider the operator $\mathcal{I}^{\ell,D}_{\ell-1}$ with the property $(\mathcal{I}^{\ell,D}_{\ell-1} v)(p) = 0$ if $p \in \mathcal{N}_\ell \cap \partial\Omega_\ell$. As indicated in (5.3), we show the stability estimate for the functions in the space $X_{\ell-1} \cap H^1_D(\Omega_\ell) \cap H^1_D(\Omega_{\ell-1})$. Therefore, the input v vanishes on the elements $T_{\ell-1} \in \mathcal{T}_{\ell-1}$ with $T_{\ell-1} \cap \partial\Omega_\ell \neq \emptyset$ or $T_{\ell-1} \subset \Omega_{\ell-1} \setminus \Omega_\ell$, if any. This implies that $\mathcal{I}^{\ell,D}_{\ell-1}$ coincides with the standard operator on the considered subspace, namely

$$\mathcal{I}^{\ell,D}_{\ell-1} v = \mathcal{I}^\ell_{\ell-1} v, \quad \forall\, v \in X_{\ell-1} \cap H^1_D(\Omega_\ell) \cap H^1_D(\Omega_{\ell-1}).$$

Therefore, the third assertion (5.3) follows from the previous two and the proof of the lemma is concluded. □

It seems not feasible to use the same elementary techniques to prove analogous results for space dimension $d \in \{2, 3\}$. This is because the constant of the shape regularity assumption (1.16) comes into play by a local inverse inequality and the sum over neighboring elements.

5.2 The concept of quasi-interpolation

Next, we deal with a class of approximation operators which have originally been introduced to generalize the nodal interpolation in finite element spaces. Such quasi-interpolation

operators are a necessary tool if the considered functions are discontinuous.

5.2.1 Clément interpolation

The Clément interpolation operator, first introduced in [58], is defined by

$$\mathcal{R}_\ell : L^2(\Omega) \to X_\ell, \quad u \mapsto \mathcal{R}_\ell u := \sum_{p \in \mathcal{N}_\ell} (Q_p u)(p) \lambda_p^\ell, \tag{5.4}$$

with the L^2-projections Q_p onto the local polynomial spaces $\mathbb{P}_r(\omega_p)$ of degree $r \in \mathbb{N}$, i. e.,

$$u \mapsto Q_p u \in \mathbb{P}_r(\omega_p) : \quad (Q_p u, v)_{L^2(\omega_p)} = (u, v)_{L^2(\omega_p)}, \quad \forall\, v \in \mathbb{P}_r(\omega_p),\ p \in \mathcal{N}_\ell. \tag{5.5}$$

For instance, each projection Q_p simply acts as local averaging by $Q_p u = \frac{1}{|\omega_p|} \int_{\omega_p} u$ if $r = 0$. The class of Clément-type interpolation operators is probably most famous for its frequent usage in proofs of the reliability of a posteriori error estimators; see [45, 184] for a detailed review.

Whereas the original results in [58] have been derived for triangular meshes in case $d = 2$, the relevant properties can indeed be proved for finite element spaces associated with general, not necessarily affine meshes and $d \in \{2, 3\}$. This is the content of the following

Lemma 5.3 (Clément)**.** *The operator $\mathcal{R}_\ell : L^2(\Omega) \to X_\ell$ is H^1-stable and has the L^2-approximation property for $r \in \mathbb{N}$.*

Note that the assertion holds true for non-quasi-uniform meshes; we refer the reader to the discussion in [8]. The technical ideas of the proof are perhaps most clearly elaborated in [30, Lemma 3.1], although in a slightly different context. Both the required estimates, namely stability and approximation property, are valid even if $r = 0$, as asserted in the above lemma originating from the work by Clément [58]. At this point, there seems to be no profound reason why the authors in [47, 49, 52] insist on the usage of the space of linear polynomials as local trial space instead. They do not comment on the choice at all.

By definition, the Clément interpolation acts as the nodal interpolation on polynomials of degree r, namely $\mathcal{R}_\ell v = \mathcal{I}_\ell v$ for all $v \in \mathbb{P}_r(\Omega)$. For the purpose of information transfer between both nested and non-nested finite element spaces which are built from piecewise polynomials, this cannot be exploited, though.

Let us now complete the analysis of Section 3.4 by showing the applicability of the just defined local operator $\mathcal{R}_\ell$. For the proof of Lemma 3.7, this has been managed by the above Lemma 5.3 by Clément. In the more complex case of Lemma 3.10, it is still the same operator which may be applied. The H^1-stability is retained. However, as the target space $\widehat{X}_\ell$ contains less functions due to the recursive modification, we need to prove the following

Lemma 5.4. *Let assumption (3.21) of Lemma 3.10 hold. Then, there exists a Clément-type quasi-interpolation operator $\mathcal{Q}_\ell^X : H^1(\Omega_\ell) \to \widehat{X}_\ell$ which satisfies the L^2-approximation property (3.18).*

Proof. We adapt a technique used in [49, Lemma 3] to our considerably more general case. Let $\mathcal{R}_\ell : H^1(\Omega_\ell) \to X_\ell$ be the standard Clément operator defined by (5.4) where the target space is the full finite element space X_ℓ at level $\ell < L$. Then, for any $v \in H^1(\Omega_\ell)$, we choose

$$\mathcal{Q}_\ell^X v = \sum_{p \in \mathcal{N}_\ell \setminus \mathcal{N}_\ell^D} (Q_p v)(p) \lambda_p^\ell = \mathcal{R}_\ell v - \sum_{p \in \mathcal{N}_\ell^D} (Q_p v)(p) \lambda_p^\ell.$$

Recall that $\mathcal{N}_\ell^D$ denotes the recursively excluded nodes in a neighborhood of the Dirichlet boundary at level L. We obtain indeed a mapping to $\widehat{X}_\ell$.

For any node $p \in \mathcal{N}_\ell$, let us first calculate

$$\begin{aligned}\|(Q_p v)(p)\lambda_p^\ell\|_{L^2(\omega_p)}^2 &= \sum_{T \in \mathcal{T}_\ell, T \subset \omega_p} \|(Q_p v)(p)\lambda_p^\ell\|_{L^2(T)}^2 \le \sum_{T \in \mathcal{T}_\ell, T \subset \omega_p} \|Q_p v\|_{L^\infty(T)}^2 \|\lambda_p^\ell\|_{L^2(T)}^2 \\ &\lesssim \sum_{T \in \mathcal{T}_\ell, T \subset \omega_p} h_T^{-d} \|Q_p v\|_{L^2(T)}^2 h_T^d = \|Q_p v\|_{L^2(\omega_p)}^2 \lesssim \|v\|_{L^2(\omega_p)}^2.\end{aligned}$$

Here, a well-known inverse estimate between $\|\cdot\|_{L^\infty(T)}$ and $\|\cdot\|_{L^2(T)}$ for polynomials in the domain T has been employed; see, e. g., [25].

Now, let $v \in H_D^1(\Omega_\ell)$. By assumption (3.21), for $p \in \mathcal{N}_\ell^D$, we can find a connected set $\widetilde{\omega}_p$, $\omega_p \subset \widetilde{\omega}_p \subset \overline{\Omega}_\ell$, with sufficiently smooth boundary such that $\operatorname{diam}(\widetilde{\omega}_p) \eqsim \operatorname{diam}(\omega_p)$ and $\operatorname{meas}_{d-1}(\widetilde{\omega}_p \cap \Gamma_D) > 0$. Moreover, the assumption (3.21) also implies that, for any node $p \in \mathcal{N}_\ell^D$, the number of elements in $\mathcal{T}_\ell$ which are intersected by a straight line from p to a closest point on Γ_D is bounded uniformly in ℓ. Therefore, as the mesh $\mathcal{T}_\ell$ is shape regular, one may construct these sets $(\widetilde{\omega}_p)_{p \in \mathcal{N}_\ell^D}$ with a finite covering property, i. e., there is a universal constant $C \in \mathbb{N}$ such that, for each $\boldsymbol{x} \in \Omega_\ell$, the number of extended patches with $\boldsymbol{x} \in \widetilde{\omega}_p$ is bounded by C.

By the Poincaré inequality, we have

$$\|v\|_{L^2(\omega_p)}^2 \le \|v\|_{L^2(\widetilde{\omega}_p)}^2 \lesssim \operatorname{diam}(\widetilde{\omega}_p)^2 |v|_{H^1(\widetilde{\omega}_p)}^2 \eqsim \operatorname{diam}(\omega_p)^2 |v|_{H^1(\widetilde{\omega}_p)}^2.$$

This amounts to

$$\|h_\ell^{-1}(Q_p v)(p)\lambda_p^\ell\|_{L^2(\omega_p)} \lesssim |v|_{H^1(\widetilde{\omega}_p)}$$

with a mesh size function $h_\ell \in L_>^\infty(\Omega_\ell)$ suitably chosen; see Section 3.4.1. Consequently,

$$\begin{aligned}\|h_\ell^{-1}(v - \mathcal{Q}_\ell^X v)\|_{L^2(\Omega_\ell)} &\le \|h_\ell^{-1}(v - \mathcal{R}_\ell v)\|_{L^2(\Omega_\ell)} + \sum_{p \in \mathcal{N}_\ell^D} \|h_\ell^{-1}(Q_p v)(p)\lambda_p^\ell\|_{L^2(\omega_p)} \\ &\lesssim |v|_{H^1(\Omega_\ell)} + \sum_{p \in \mathcal{N}_\ell^D} |v|_{H^1(\widetilde{\omega}_p)} \lesssim |v|_{H^1(\Omega_\ell)},\end{aligned}$$

where the final step follows from the choice of the extended patches $(\widetilde{\omega}_p)_{p \in \mathcal{N}_\ell^D}$ and the shape regularity of the mesh $\mathcal{T}_\ell$. □

Projection properties

Restricting the attention to the discrete space X_ℓ, one notes that the Clément interpolation $\mathcal{R}_\ell : X_\ell \to X_\ell$ does not keep invariant the basis functions; see (5.6) below. But this information is not sufficient to determine the projection properties of the operator in the spaces $L^2(\Omega)$ and $X_{\ell-1}$, respectively. This is because we do not know whether the functions $\lambda_p^\ell \in \Lambda_\ell$ are contained in the range $\mathcal{R}_\ell(L^2(\Omega))$ or even $\mathcal{R}_\ell(X_{\ell-1})$. However, considering the size of the supports of the images of certain functions, we can prove the following quite universal

Proposition 5.5. *Let the mesh $\mathcal{T}_\ell$ contain at least two interior nodes. Then, the quasi-interpolation $\mathcal{R}_\ell : L^2(\Omega) \to X_\ell$ is not a projection.*

Proof. Let $v \in L^2(\Omega)$ be a non-negative, non-trivial function such that $\operatorname{supp}(\mathcal{R}_\ell v) \neq \overline{\Omega}$. It is easy to see that such a "local" function v exists if the mesh $\mathcal{T}_\ell$ has at least two interior nodes. Then, one can find an element $T_0 \in \mathcal{T}_\ell$ with $T_0 \not\subset \operatorname{supp}(\mathcal{R}_\ell v)$ but $\overline{T}_0 \cap \operatorname{supp}(\mathcal{R}_\ell v) \neq \emptyset$, in other words an element adjacent to the support of $\mathcal{R}_\ell v$. It is obvious that

$$\operatorname{supp}(\mathcal{R}_\ell \lambda_p^\ell) = \bigcup \left\{ \overline{T} \mid T \in \mathcal{T}_\ell,\, \overline{T} \cap \omega_p \neq \emptyset \right\}, \quad \forall\, p \in \mathcal{N}_\ell. \tag{5.6}$$

By definition, we have the linear combination $\mathcal{R}_\ell \mathcal{R}_\ell v = \sum_{p \in \mathcal{N}_\ell} (Q_p v)(p) \mathcal{R}_\ell \lambda_p^\ell$ with numbers $(Q_p v)(p) \geq 0$. Because the functions $\mathcal{R}_\ell \lambda_p^\ell = \sum_{r \in \mathcal{N}_\ell} (Q_r \lambda_p^\ell)(r) \lambda_r^\ell$, $p \in \mathcal{N}_\ell$, are also non-negative, the contributions coming from $\mathcal{R}_\ell \lambda_p^\ell$ and $\mathcal{R}_\ell \lambda_q^\ell$, $p \neq q$, do not cancel out each other in the calculation of the effective coefficients of $\mathcal{R}_\ell \mathcal{R}_\ell v$ with respect to the basis Λ_ℓ. Thus, it follows that $T_0 \subset \operatorname{supp}(\mathcal{R}_\ell \mathcal{R}_\ell v)$ and, consequently, $\mathcal{R}_\ell \mathcal{R}_\ell v \neq \mathcal{R}_\ell v$. This concludes the proof of the proposition. □

There are in fact subspaces $U \subset L^2(\Omega)$ such that $\mathcal{R}_\ell \mathcal{R}_\ell u = \mathcal{R}_\ell u$ for all $u \in U$; for instance, $\mathbb{P}_r(\Omega)$ has this property, as mentioned before. We now investigate to what extent the above considerations also hold true for $\mathcal{R}_{\ell-1}^\ell : X_{\ell-1} \to X_\ell$, namely if the domain of the operator is restricted to the discrete subspace $X_{\ell-1}$. For this purpose, suppose that there is a node $p \in \mathcal{N}_{\ell-1}$ and an element $T_1 \in \mathcal{T}_\ell$ such that

$$\operatorname{int}\left(\bigcup \{ \overline{T} \mid T \in \mathcal{T}_\ell,\, T \cap \operatorname{supp}(\lambda_p^{\ell-1}) \neq \emptyset \} \right) \cap \operatorname{int}\left(\bigcup \{ \overline{T} \mid T \in \mathcal{T}_\ell,\, \overline{T} \cap \overline{T}_1 \neq \emptyset \} \right) = \emptyset. \tag{5.7}$$

Simply put, T_1 needs to be sufficiently far away from the "reach of p". This implies that $\operatorname{int}(\operatorname{supp}(\mathcal{R}_\ell \lambda_p^{\ell-1})) \cap T_1 = \emptyset$. Thus, $\operatorname{supp}(\mathcal{R}_\ell \lambda_p^{\ell-1}) \neq \overline{\Omega}$ and one can find an element $T_0 \in \mathcal{T}_\ell$ which is adjacent to the support of $\mathcal{R}_\ell \lambda_p^{\ell-1}$. Concluding as before, we have the following

Proposition 5.6. *Provided that (5.7) can be fulfilled, the Clément interpolation is not a projection even if its domain is restricted to the discrete subspace $X_{\ell-1}$.*

Note that the relatively weak assumption (5.7) is valid for virtually every pair of meshes $(\mathcal{T}_{\ell-1}, \mathcal{T}_\ell)$ one might handle. Therefore, we have shown that the Clément interpolation operator is practically never a projection.

From Proposition 5.6 and Remark 5.1 we observe the following: Neither does the Clément operator reduce to the standard interpolation in case of nested meshes $\mathcal{T}_{\ell-1}$ and $\mathcal{T}_\ell$ nor is it the identity mapping if the meshes and hence the associated spaces are identical. Evidently, this observation is valid for any polynomial degree $r \in \mathbb{N}$.

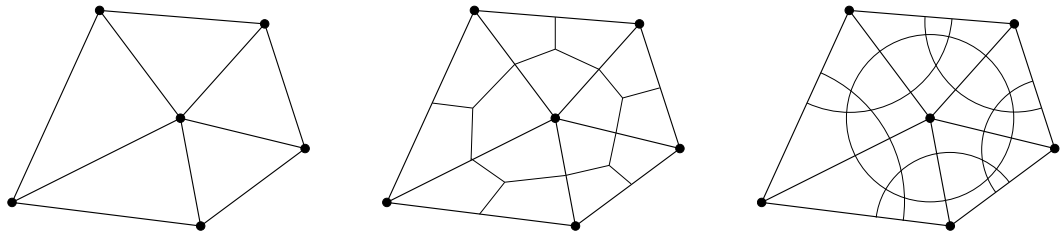


Figure 5.3. Original patch ω_p (left), a non-overlapping decomposition into sub-elements by means of a dual mesh (center), a covering by overlapping circles (right).

5.2.2 On the design of local quasi-interpolation operators

Clément's idea was to employ a local L^2-projection for each basis function and evaluate the resulting polynomial at the respective node. The local trial spaces are defined with respect to the patches $(\omega_p)_{p\in\mathcal{N}_\ell}$. Therefore, the value of the interpolant $\mathcal{R}_\ell v \in X_\ell$ of a function $v \in L^2(\Omega)$ inside an element $T \in \mathcal{T}_\ell$ depends on the values of v in the neighborhood

$$\omega_T := \operatorname{int}\left(\bigcup \left\{\overline{T'} \mid T' \in \mathcal{T}_\ell,\, \overline{T'} \cap \overline{T} \neq \emptyset\right\}\right).$$

Presumably, this is the most reasonable "domain of influence" for an interpolation operator. In general, it seems to be impossible or at least inadvisable to construct an acceptable interpolant just from the values of each element separately, especially in the finite element setting with continuous ansatz spaces.

An interpolation process, namely the computation of the coefficients $((\Pi_\ell v)(p))_{p\in\mathcal{N}_\ell}$ should be symmetric to a certain degree. Usually, it is not favorable to exclude adjacent elements to the node $p \in \mathcal{N}_\ell$ from the weighting procedure. Still, other weighting operators might work with suitably chosen subsets of the elements in ω_p. For instance, this could be realized as a decomposition of each element into sub-elements. For each node $p \in \mathcal{N}_\ell$, this defines a smaller neighborhood $\widetilde{\omega}_p \subset \omega_p$. Then, replacing the local polynomial spaces $\mathbb{P}_r(\omega_p)$ by the spaces $\mathbb{P}_r(\widetilde{\omega}_p)$, we obtain the L^2-projections

$$u \mapsto Q_p u \in \mathbb{P}_r(\widetilde{\omega}_p): \quad (Q_p u, v)_{L^2(\widetilde{\omega}_p)} = (u, v)_{L^2(\widetilde{\omega}_p)}, \quad \forall\, v \in \mathbb{P}_r(\widetilde{\omega}_p),\ p \in \mathcal{N}_\ell.$$

The nature of the decomposition essentially determines the character of the operator. Simply put, larger overlaps tend to result in a smoother interpolant. Naturally, for the corresponding stability and approximation properties to be valid, the subdivision of the single elements has to be designed in such a way that the size of each sub-element decreases by the same rate as the size of the associated element. Simple examples in case $d = 2$ are given in Figure 5.3. The sketch at center originates from a dual mesh used, e. g., in [174]. The decomposition on the right employs overlapping circles centered at the nodes.

However, for computational purposes the above approach seems to be somewhat extravagant; see also Section 5.7 for the implementation of quasi-interpolation operators. Finally, let us remark that making the "domain of influence" smaller cannot change the operator $\mathcal{R}_\ell$ into a projection.

5.2.3 Convergence of approximation operators

Clément [58] included the proposition that the interpolant $\mathcal{R}_\ell v$ converges to v with respect to the norm $\|\cdot\|_{H^r(\Omega)}$ as $h_\ell \to 0$. We state this assertion in a slightly more general way because the convergence result does not only hold for the quasi-interpolation specified by (5.4) but also for other operators satisfying basic approximation properties. Still, the proof is basically extracted from the original work.

Lemma 5.7. *Given the integer $m > 0$, let $X \subset H^m(\Omega)$. Suppose the operator $\Pi : X \to Y$ maps to a finite element space $Y \subset H^{m-1}(\Omega)$ with (global) discretization parameter $h_Y > 0$. Then, if Π is H^{m-1}-stable and satisfies the approximation property*

$$|v - \Pi v|_{H^{m-1}(\Omega)} \lesssim h_Y |v|_{H^m(\Omega)}, \quad \forall\, v \in X, \tag{5.8}$$

the convergence result

$$\lim_{h_Y \to 0} |v - \Pi v|_{H^{m-1}(\Omega)} = 0, \quad \forall\, v \in X,$$

holds for decreasing mesh size h_Y of the target space.

Proof. For every $w \in H^m(\Omega)$, we have

$$\begin{aligned} |v - \Pi v|_{H^{m-1}(\Omega)} &\leq |w - \Pi w|_{H^{m-1}(\Omega)} + |v - w|_{H^{m-1}(\Omega)} + |\Pi(v - w)|_{H^{m-1}(\Omega)} \\ &\leq c_1 h_Y |w|_{H^m(\Omega)} + (1 + c_2)|v - w|_{H^{m-1}(\Omega)} \end{aligned}$$

with constants $c_1, c_2 > 0$. For given $\varepsilon > 0$, we may choose the function w such that $(1 + c_2)|v - w|_{H^{m-1}(\Omega)} < \frac{\varepsilon}{2}$. Now, let $\overline{h}_Y > 0$ such that $c_1 \overline{h}_Y |w|_{H^m(\Omega)} < \frac{\varepsilon}{2}$. Then, $|v - \Pi v|_{H^{m-1}(\Omega)} < \varepsilon$ for $h_Y < \overline{h}_Y$. This completes the proof. □

We have stated the lemma for a global quantity h_Y. Naturally, the assertion may also be proved in case the approximation property (5.8) is given locally, namely with respect to a function $h_Y^{-1} \in H^{m-1}(\Omega)$.

Lemma 5.7 is an interesting observation illustrating the nature of the approximation operators addressed in this thesis a bit. It is rather weak and we do not use it for the theoretical considerations of this and the other sections, though. Also note that we only deal with L^2-approximation properties and will, as a rule, not be able to prove approximation properties in Sobolev spaces of higher order for the operators considered in this chapter.

5.2.4 An alternative quasi-interpolation procedure

As we have seen in the previous paragraphs, the local orthogonal projections are usually defined with respect to polynomial spaces. A variant of Clément's quasi-interpolation operator can be found in [30, 174]. It is defined by

$$\widetilde{\mathcal{R}}_\ell : L^2(\Omega) \to X_\ell, \quad u \mapsto \widetilde{\mathcal{R}}_\ell u := \sum_{p \in \mathcal{N}_\ell} (\widetilde{Q}_p u)(p) \lambda_p^\ell,$$

with the L^2-projections $\widetilde{Q}_p$ onto the local trial spaces X_ℓ^p, i. e.,

$$u \mapsto \widetilde{Q}_p u \in X_\ell^p : \quad (\widetilde{Q}_p u, v)_{L^2(\omega_p)} = (u, v)_{L^2(\omega_p)}, \quad \forall\, v \in X_\ell^p,\ p \in \mathcal{N}_\ell, \tag{5.9}$$

where

$$X_\ell^p := \left\{ v|_{\omega_p} \mid v \in X_\ell \right\}, \quad \forall\, p \in \mathcal{N}_\ell. \tag{5.10}$$

In the original context of [30], $\widetilde{\mathcal{R}}_\ell$ has been employed to prove a new result on the H^1-stability of the standard L^2-projection under weaker assumptions.

The operator $\widetilde{\mathcal{R}}_\ell$ is a projection with domain $L^2(\Omega)$, namely $\widetilde{\mathcal{R}}_\ell \widetilde{\mathcal{R}}_\ell v = \widetilde{\mathcal{R}}_\ell v$ for any $v \in L^2(\Omega)$, because $(\widetilde{Q}_p \lambda_p^\ell)(p) = \lambda_p^\ell(p) = 1$ for any $p \in \mathcal{N}_\ell$ and $(\widetilde{Q}_q \lambda_p^\ell)(q) = \lambda_p|_{\omega_q}(q) = 0$ for $q \neq p$. The same argument implies that the quasi-interpolation $\widetilde{\mathcal{R}}_\ell$ is surjective, i. e., $\widetilde{\mathcal{R}}_\ell(L^2(\Omega)) = X_\ell$. In addition, it satisfies the requested stability and approximation estimates; see [30, Lemma 3.1]. However, to our knowledge, this alternative quasi-interpolation has never been used in practical computations.

In this context, another local operator immediately suggests itself. Replacing the trial spaces X_ℓ^p in (5.9) by $\operatorname{span}\{\lambda_p^\ell\}$, we notice that the local L^2-projections to the one-dimensional spaces are very easy to evaluate. This yields the formula

$$\widetilde{\mathcal{R}}'_\ell : L^2(\Omega) \to X_\ell, \quad u \mapsto \widetilde{\mathcal{R}}'_\ell u := \sum_{p \in \mathcal{N}_\ell} \frac{(\lambda_p^\ell, u)_{L^2(\Omega)}}{(\lambda_p^\ell, \lambda_p^\ell)_{L^2(\Omega)}} \lambda_p^\ell. \tag{5.11}$$

We return to this operator in Remark 5.8. It is not a projection; see Remark 5.10. We will explain by Remark 5.19 in the context of some theoretical considerations and by Remark 5.22 in the context of our numerical experiments that $\widetilde{\mathcal{R}}'_\ell$, although looking quite similar, is qualitatively and quantitatively very different from the other analyzed transfer concepts.

5.3 The L^2-projection

In this section, we comment on the use of an operator in the present context, which has repeatedly appeared in this thesis before. Let $\mathcal{Q}_\ell : L^2(\Omega) \to X_\ell$ be the L^2-projection onto X_ℓ, i. e., the orthogonal projection in the Hilbert space $L^2(\Omega)$ to the subspace X_ℓ characterized by the variational equation

$$u \mapsto \mathcal{Q}_\ell u \in X_\ell : \quad (\mathcal{Q}_\ell u, v)_{L^2(\Omega)} = (u, v)_{L^2(\Omega)}, \quad \forall\, v \in X_\ell.$$

The mapping $\mathcal{Q}_\ell$ is global as opposed to Definition 3.8. This can be understood considering the algebraic representation of the fully discrete operator $\mathcal{Q}_{\ell-1}^\ell$ via a product similar to (2.6),

$$\mathcal{Q}_{\ell-1}^\ell v = \Phi_\ell(\boldsymbol{M}_\ell^{-1} \boldsymbol{B}_\ell \Phi_{\ell-1}^{-1}(v)), \quad \forall\, v \in X_{\ell-1}, \tag{5.12}$$

with the mass matrix $\boldsymbol{M}_\ell$ associated with X_ℓ and a sparse coupling matrix $\boldsymbol{B}_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ with the entries

$$(\boldsymbol{B}_\ell)_{pq} = (\lambda_p^\ell, \lambda_q^{\ell-1})_{L^2(\Omega)}, \quad \forall\, p \in \mathcal{N}_\ell,\ q \in \mathcal{N}_{\ell-1}. \tag{5.13}$$

From a heuristic point of view the L^2-projection might be quite a good choice. But the usage of the L^2-projection "as is" to transfer information from coarser levels to finer levels cannot be expected to result in a computationally efficient multilevel algorithm for non-nested meshes. This is because a mass matrix associated with the finer finite element

space has to be inverted. This inverse $\boldsymbol{M}_\ell^{-1}$ is usually dense; thus, the basis functions of the coarse spaces constructed in (3.4) and (3.5) have large supports. This yields dense coarse matrices $(\boldsymbol{A}_{\ell-1})_{\ell=1,\dots,L}$ whose Galerkin assembly requires the application of the inverse mass matrix $\boldsymbol{M}_\ell^{-1}$ to $2n_\ell$ vectors at level ℓ. Moreover, prolongation and restriction of a coarse level correction and a fine level residual, respectively, are expensive to compute.

Note that the operator attained by simply lumping the matrix $\boldsymbol{M}_\ell$ will be considered in the following section. We want to point out that some of the (to a greater or lesser extent) sophisticated prolongation operators taken from the literature and discussed here are distinctly motivated by the idea to find an L^2-projection-like mapping or a weighted interpolation which is more suitable for computations. Partly, this is achieved by modifying the fine-level mass matrix $\boldsymbol{M}_\ell$ but keeping the coupling matrix $\boldsymbol{B}_\ell$ from (5.13) between coarse and fine space.

To obtain a stability estimate for the L^2-projection, the requirement of quasi-uniformity of the mesh $\mathcal{T}_\ell$ has been considered inevitable for quite a long time. Meanwhile, weaker criteria ensuring the H^1-stability of $\mathcal{Q}_\ell$ are available; see, e. g. [30, 44, 60, 174]. For estimates with respect to other Lebesgue norms, see [77] and the references therein. Two different proofs both using inverse estimates of Bernstein-type, which generally hold true only for quasi-uniform meshes, can be found in [36, Theorem 3.4] and [25, Folgerung II.7.8].

If another suitable approximation operator such as the Clément quasi-interpolation from Section 5.2 is at hand, the L^2-approximation property in case of quasi-uniform meshes can be proved with elementary techniques similar to Lemma 5.15. However, a direct proof is achievable. This proof employs the fact that $\mathcal{Q}_\ell$ is the orthogonal projection with respect to $(\cdot,\cdot)_{L^2(\Omega)}$; thus, it yields the best approximation in X_ℓ with respect to the norm $\|\cdot\|_{L^2(\Omega)}$. Further ingredients are a standard finite element interpolation error estimate and an interpolation technique between Sobolev spaces. See, e. g., [36, Theorem 3.2].

5.4 On L^2-quasi-projections

In this section, we give two transfer concepts from the literature yielding quasi-projection operators. The first one has been proposed in the thesis [193] and is defined for simplicial meshes by the choice of a discrete inner product $(\cdot,\cdot)_\ell$ in X_ℓ, namely

$$(u,v)_\ell := \frac{1}{d+1}\sum_{T\in\mathcal{T}_\ell}|T|\sum_{p\in\mathcal{N}_\ell\cap T}u(p)v(p),\quad \forall\, u,v\in X_\ell. \tag{5.14}$$

Then, consider the operator $\widetilde{\mathcal{Q}}_\ell : L^2(\Omega)\to X_\ell$ specified by

$$u\mapsto \widetilde{\mathcal{Q}}_\ell u:\quad (\widetilde{\mathcal{Q}}_\ell u, v)_\ell = (u,v)_{L^2(\Omega)},\quad \forall\, v\in X_\ell. \tag{5.15}$$

Note that $\widetilde{\mathcal{Q}}_\ell$ is usually not a projection; see the discussion in Remark 5.10 below. This motivates the term quasi-projection. A proof of the H^1-stability and the L^2-approximation property of the operator $\widetilde{\mathcal{Q}}_\ell$ in case of a quasi-uniform mesh $\mathcal{T}_\ell$ is given in [193, Lemma 3.6].

By construction, the mass matrix with respect to the discrete inner product,

$$(\widetilde{\boldsymbol{M}}_\ell)_{pq} = (\lambda_p^\ell,\lambda_q^\ell)_\ell = \delta_{pq}\frac{|\omega_p|}{d+1},\quad \forall\, p,q\in\mathcal{N}_\ell,$$

is diagonal and for $u \in L^2(\Omega)$ the quasi-projection operator reads as

$$\widetilde{\mathcal{Q}}_\ell u = \sum_{p \in \mathcal{N}_\ell} \frac{(\lambda_p^\ell, u)_{L^2(\Omega)}}{(\widetilde{\boldsymbol{M}}_\ell)_{pp}} \lambda_p^\ell. \tag{5.16}$$

For $u \in X_{\ell-1}$, the definition (5.15) yields a fully discrete representation of the operator $\widetilde{\mathcal{Q}}_\ell$, as explained in the previous section, namely

$$\widetilde{\mathcal{Q}}_{\ell-1}^\ell u = \Phi_\ell(\widetilde{\boldsymbol{M}}_\ell^{-1} \boldsymbol{B}_\ell \Phi_{\ell-1}^{-1}(u)), \quad \forall\, u \in X_{\ell-1},$$

with the same coupling matrix $\boldsymbol{B}_\ell$ as in (5.13).

The constant factor $d+1$ in (5.16), which comes from the interpretation of (5.14) as a quadrature rule, does not matter for our purposes. Another discrete inner product also motivated by a quadrature rule (on centroids of faces instead of nodes) can be found in [26]. There, it is used in the fashion of (5.15) to define a prolongation operator between the non-nested spaces associated with a discretization with Crouzeix–Raviart elements on nested meshes.

Remark 5.8. *The operator $\widetilde{\mathcal{R}}_\ell'$ defined by (5.11) may also be obtained in the above fashion by choosing the discrete inner product*

$$(u, v)_\ell' := \sum_{T \in \mathcal{T}_\ell} (\lambda_p^\ell, \lambda_p^\ell)_{L^2(\Omega)} \sum_{p \in \mathcal{N}_\ell \cap T} u(p) v(p), \quad \forall\, u, v \in X_\ell. \tag{5.17}$$

In other words, $\widetilde{\mathcal{R}}_\ell'$ is the orthogonal projection to X_ℓ equipped with this inner product.

The following second quasi-projection operator has been introduced in [31] as approximation operator replacing the L^2-projection from the space $H^1(\Omega)$ to the discrete spaces X_ℓ. It is a mapping directly defined via a representation like (5.16), here,

$$\widehat{\mathcal{Q}}_\ell : L^2(\Omega) \to X_\ell, \quad u \mapsto \widehat{\mathcal{Q}}_\ell u := \sum_{p \in \mathcal{N}_\ell} \frac{(\lambda_p^\ell, u)_{L^2(\Omega)}}{(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}} \lambda_p^\ell, \tag{5.18}$$

where $\mathbb{1}$ denotes the constant function with value 1. After all, we can obtabin a matrix representation of the fully discrete operator $\widehat{\mathcal{Q}}_{\ell-1}^\ell : X_{\ell-1} \to X_\ell$ from the one of the standard L^2-projection in a simple way by lumping the mass matrix $\boldsymbol{M}_\ell$ in (5.12). In the numerical practice, this seems a very natural thing to do. Moreover, for simplicial meshes, it is easy to verify by integration over the reference element that $(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)} = \frac{|\omega_p|}{d+1}$ for all $p \in \mathcal{N}_\ell$; thus, the operators $\widetilde{\mathcal{Q}}_\ell$ and $\widehat{\mathcal{Q}}_\ell$ are equivalent in this case. Again, one needs to notice that there is virtually no experience with the quasi-projection operators in practical computations.

Let us postpone the proof of the following lemma to the next section; see Remark 5.19.

Lemma 5.9. *The operator $\widehat{\mathcal{Q}}_\ell : L^2(\Omega) \to X_\ell$ is H^1-stable and has the L^2-approximation property for all shape regular families of meshes.*

Although we do not study prolongation operators for higher order finite element spaces here, we emphasize that this concept is, in general, only well-defined for first order elements. This is because the denominators in (5.18), coming from lumping the mass matrix $\boldsymbol{M}_\ell$, may vanish for other element types.

Remark 5.10. *We remark a striking fact about the quasi-projection operators sketched in this section. As a rule, neither the operator $\widetilde{\mathcal{Q}}_\ell$ nor the operator $\widehat{\mathcal{Q}}_\ell$ defined in (5.15) and (5.18), repectively, is a projection. A proof of this assertion can be achieved analogously to the ones of Proposition 5.5 and Proposition 5.6, which treat the same issue for the Clément interpolation. This is also true for the operator $\widetilde{\mathcal{R}}'_\ell$ from (5.11) as the main ingredient is merely the property (5.6). In addition, one can easily see that for $p \in \mathcal{N}_\ell$ we have $(\lambda_p^\ell, \lambda_p^\ell)_{L^2(\Omega)} < (\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}$; thus, $\widehat{\mathcal{Q}}_\ell \lambda_p^\ell \neq \lambda_p^\ell$.*

Finally, we notice that certain operators are self-adjoint with respect to the L^2-inner product. For instance,

$$(\widehat{\mathcal{Q}}_\ell u, v)_{L^2(\Omega)} = \sum_{p \in \mathcal{N}_\ell} \frac{(\lambda_p^\ell, u)_{L^2(\Omega)} (\lambda_p^\ell, v)_{L^2(\Omega)}}{(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}} = (u, \widehat{\mathcal{Q}}_\ell v)_{L^2(\Omega)}, \quad \forall\, u, v \in L^2(\Omega).$$

This is also true for $\widetilde{\mathcal{Q}}_\ell$ and $\widetilde{\mathcal{R}}'_\ell$. However, we do not know whether this property may be put to a good use in the analysis or the practical computations at this point. This is because the two involved spaces are usually not identical in applications. Note that the only operator that is self-adjoint and at the same time a projection is the orthogonal projection $\mathcal{Q}_\ell$.

5.5 The pseudo-L^2-projection

We propose a transfer concept which is different in some respects. The new operator will be denoted by the symbol $\mathcal{P}$ with the appropriate indices. Generally speaking, we introduce a Petrov–Galerkin scheme with a discontinuous test space built from a set of functions which are biorthogonal to the standard nodal basis with respect to the L^2-inner product $(\cdot, \cdot)_{L^2(\Omega)}$. In the fully discrete setting, this will yield a band matrix representation of the operator as no mass matrix has to be inverted. The transfer concept based on a biorthogonal system, which is examined in full detail in this section, is in fact a projection from $L^2(\Omega)$ onto the finite element space X_ℓ. But, in general, it is not an orthogonal projection although it resembles one strongly (and coincides with $\mathcal{Q}_\ell$ in case of nested meshes). Additionally, in the author's view, the operator $\mathcal{P}_\ell$ represents a way to get "as close as possible" to the real L^2-projection while at the same time it guarantees an efficient evaluation. This is clearly confirmed in a remarkable manner by the numerical experiments in Section 5.8. Therefore, we call this oblique projection operator "pseudo-L^2-projection". This term is also meant to contrast, e. g., the L^2-quasi-projection concepts of Section 5.4, which yield in actual fact no projections. Moreover, the pseudo-L^2-projection seems to be the only, reasonably straightforward operator in the fashion of the previous ones (5.11), (5.16), (5.18) which is actually a projection.

5.5.1 An operator with a dual test space

In this paragraph, we derive the pseudo-L^2-projection, followed by some historical remarks in Section 5.5.2. Then, the relevant properties of the operators generated by the proposed transfer concept are analyzed in Section 5.5.3. Recall that, by the arguments put forward

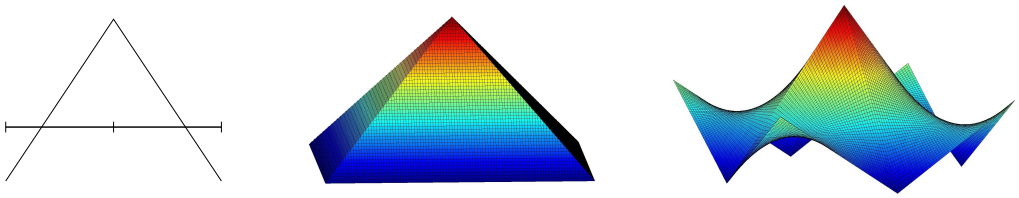


Figure 5.4. Examples for dual basis functions ψ_p^ℓ in case $d = 1, 2$. Outside the depicted domains (two line segments, four triangles, four quadrilaterals, respectively) they are extended discontinuously by 0.

in the introduction of the chapter, it is sufficient to consider one fixed domain Ω. The notational convention of the level indices applies here, too.

Biorthogonal basis functions

For the definition of the pseudo-L^2-projection, choose a set of functions $\Psi_\ell = (\psi_p^\ell)_{p\in\mathcal{N}_\ell}$ with $\psi_p^\ell \in \mathcal{C}^0(\omega_p)$ for all $p \in \mathcal{N}_\ell$ such that

$$(\psi_p^\ell, \lambda_q^\ell)_{L^2(\Omega)} = \delta_{pq}(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}, \quad \forall\, p, q \in \mathcal{N}_\ell, \tag{5.19}$$

and set the discontinuous test space as

$$Y_\ell := \operatorname{span}\{\psi_p^\ell \,|\, p \in \mathcal{N}_\ell\} \not\subset \mathcal{C}^0(\Omega).$$

Note that such a dual basis with respect to $(\cdot,\cdot)_{L^2(\Omega)}$ of the nodal finite element basis $\Lambda_\ell = (\lambda_p^\ell)_{p\in\mathcal{N}_\ell}$ exists. This can also be seen in the explicit construction of the set Ψ_ℓ which is carried out below.

Before proceeding to the actual definition of the projection operator, let us elaborate on the dual test space. First, we show that a biorthogonal system satisfying (5.19) may be constructed from the nodal basis Λ_ℓ in a straightforward manner. More precisely, on its support ω_p, each ψ_p^ℓ can be represented by a linear combination of the nodal basis functions associated with the adjacent elements restricted to ω_p; see Figure 5.4 for an illustration.

Note that, for general elements like bilinear quadrilaterals or trilinear hexahedra, the coefficients in this linear combination depend on the non-affine transformation to the reference element. Nevertheless, it is not difficult to find a suitable set Ψ_ℓ. For this purpose, we denote the number of nodes of an element $T \in \mathcal{T}_\ell$ by n_T. Let $\boldsymbol{M}_T \in \mathbb{R}^{n_T\times n_T}$ be the element mass matrix and $\boldsymbol{D}_T \in \mathbb{R}^{n_T\times n_T}$ a diagonal scaling matrix, for each element $T \in \mathcal{T}_\ell$ given by

$$(\boldsymbol{M}_T)_{pq} = (\lambda_p^\ell, \lambda_q^\ell)_{L^2(T)}, \quad (\boldsymbol{D}_T)_{pq} = \delta_{pq}(\lambda_p^\ell, \mathbb{1})_{L^2(T)}, \quad \forall\, p, q \in \mathcal{N}_\ell \cap \overline{T}.$$

Obviously, $\boldsymbol{M}_T$ is symmetric positive definite. Then, for $p \in \mathcal{N}_\ell \cap \overline{T}$, the local function $\psi_{p,T}$ defined by

$$\psi_{p,T}(x) := \sum_{r\in\mathcal{N}_\ell\cap\overline{T}} (\boldsymbol{D}_T\boldsymbol{M}_T^{-1})_{pr}\lambda_r^\ell(x), \quad \text{if } x \in \overline{T}, \tag{5.20}$$

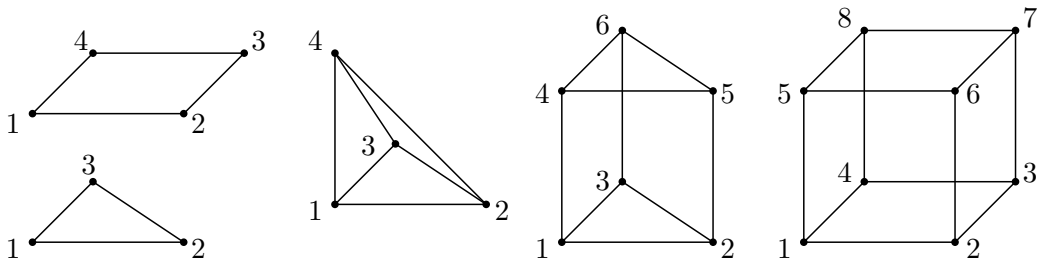


Figure 5.5. The two- and three-dimensional reference elements with the local numbering of the nodes: triangle, quadrilateral, tetrahedron, prism, hexahedron.

and extended by 0 outside T, is biorthogonal to the family Λ_ℓ because for all $q \in \mathcal{N}_\ell$ we have

$$(\psi_{p,T}, \lambda_q^\ell)_{L^2(\Omega)} = \sum_{r \in \mathcal{N}_\ell \cap \overline{T}} (\boldsymbol{D}_T \boldsymbol{M}_T^{-1})_{pr} (\boldsymbol{M}_T)_{rq} = (\boldsymbol{D}_T \boldsymbol{M}_T^{-1} \boldsymbol{M}_T)_{pq} = (\boldsymbol{D}_T)_{pq}.$$

Thus, for the composed function

$$\psi_p^\ell := \sum_{T \in \mathcal{T}_\ell,\, p \in \overline{T}} \psi_{p,T} \notin \mathcal{C}^0(\Omega) \tag{5.21}$$

the desired relation (5.19) holds. Indeed, $\operatorname{supp}(\psi_p^\ell) = \omega_p$ and $\psi_p^\ell \in \mathcal{C}^0(\omega_p)$ for all $p \in \mathcal{N}_\ell$. The above procedure has been proposed in a similar form for functions on two-dimensional interfaces in three-dimensional space in [88]. This reference also includes another way to define the biorthogonal basis by a special transformation from the reference element directly incorporating the non-constant Jacobian. See also [156].

In case of affine elements, i. e., elements with an affine transformation to the respective reference element, the coefficients in the above sum do not depend on the actual node p and element T but can be computed on the reference element in a one-time process. This is due to the scaling with $(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}$ on the right hand side of (5.19). As most unstructured meshes are entirely built from triangles and tetrahedra for $d \in \{2, 3\}$, respectively, the computation of the inverse element mass matrices is not necessary to obtain the test space Y_ℓ in this case.

As an illustration of the affine case, for some $p \in \mathcal{N}_\ell$, consider an element $T \in \mathcal{T}_\ell$ with $T \subset \omega_p$ and rewrite the sum (5.20), which gives the value of ψ_p^ℓ on the chosen element, as

$$\sum_{i=1}^{n_T} m_p^i \lambda_i^\ell. \tag{5.22}$$

Here, the index i is running over the nodes of T in the local numbering given in Figure 5.5; the set $(\lambda_i^\ell)_{i=1,\dots,n_T}$ may contain the basis functions associated with these nodes. Assume that the global index $p \in \mathcal{N}_\ell$ corresponds to the local number $i = 1$. Then, for the one-dimensional line segment and the two- and three-dimensional element types depicted in Figure 5.5, the coefficients $(m_p^i)_{i=1,\dots,n_T}$ are specified in Table 5.1. The dual functions associated with the other nodes are obtained from the given ones by simple permutations.

element type	i =	1	2	3	4	5	6	7	8
line segment		2	−1						
triangle		3	−1	−1					
quadrilateral		4	−2	1	−2				
tetrahedron		4	−1	−1	−1				
prism		6	−2	−2	−3	1	1		
hexahedron		8	−4	2	−4	−4	2	−1	2

Table 5.1. The coefficients $(m_p^i)_{i=1,\dots,n_T}$ in the linear combination (5.22) for the case that $p \in \mathcal{N}_\ell$ corresponds to the local number $i = 1$.

We conclude the study of the dual test space Y_ℓ by an L^∞-estimate for the basis Ψ_ℓ. As a matter of fact, the scaling by $(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}$ on the right hand side of (5.19) does not only provide universal coefficients for affine elements but also implies the following boundedness result.

Lemma 5.11. *The functions $(\psi_p^\ell)_{p\in\mathcal{N}_\ell}$ constructed as stated above are bounded independently of the mesh size, i. e.,*

$$\|\psi_p^\ell\|_{L^\infty(\Omega)} \lesssim 1, \quad \forall\, p \in \mathcal{N}_\ell. \tag{5.23}$$

Proof. It is well-known that $(\lambda_p^\ell, \mathbb{1})_{L^2(T)} = \|\lambda_p^\ell\|_{L^1(T)} \eqsim h_T^d$ and that the local matrices $\boldsymbol{M}_T$ and $\boldsymbol{D}_T$ are spectrally equivalent, namely

$$\boldsymbol{v} \cdot \boldsymbol{M}_T \boldsymbol{v} \eqsim \boldsymbol{v} \cdot \boldsymbol{D}_T \boldsymbol{v}, \quad \forall\, \boldsymbol{v} \in \mathbb{R}^{n_T},$$

independently of the local mesh size h_T of the element $T \in \mathcal{T}_\ell$; see, e. g., [177, Lemma B.31]. Consequently, for any $p \in \mathcal{N}_\ell$, the boundedness of ψ_p^ℓ on T follows from its definition via the equations (5.20) and (5.21) and from the boundedness of the nodal basis Λ_ℓ. □

Remark 5.12. *Evidently, we also have $\|\psi_p^\ell\|_{L^\infty(\Omega)} \gtrsim 1$ for all $p \in \mathcal{N}_\ell$.*

Finally, we note that the system Ψ_ℓ is a partition of unity. This follows immediately from the fact that the row totals of the element mass matrix $\boldsymbol{M}_T$ are equal to the entries of the scaling matrix $\boldsymbol{D}_T$. In particular, $\mathbb{P}_0(\Omega) \subset Y_\ell$; see also [88, Lemma 1].

For a more detailed analysis of biorthogonal bases, carried out in the context of the mortar finite element method, and the construction of such systems for higher order finite element spaces, we refer to [118, 131, 132, 191].

Petrov–Galerkin formulation

Now, we define the pseudo-L^2-projection $\mathcal{P}_\ell : L^2(\Omega) \to X_\ell$ by a Petrov–Galerkin variational formulation with trial space X_ℓ and test space Y_ℓ, i. e.,

$$u \mapsto \mathcal{P}_\ell u : \quad (\mathcal{P}_\ell u, v)_{L^2(\Omega)} = (u, v)_{L^2(\Omega)}, \quad \forall\, v \in Y_\ell. \tag{5.24}$$

This variational problem has a unique solution because $\dim(Y_\ell) = \dim(X_\ell) < \infty$ and for $u \in X_\ell$ it is $(u, v)_{L^2(\Omega)} = 0$ for all $v \in Y_\ell$ if and only if $u = 0$. In particular, the definition yields the obvious representation formula

$$\mathcal{P}_\ell u = \sum_{p \in \mathcal{N}_\ell} \frac{(\psi_p^\ell, u)_{L^2(\Omega)}}{(\lambda_p^\ell, \mathbb{1})_{L^2(\Omega)}} \lambda_p^\ell, \quad \forall\, u \in L^2(\Omega). \tag{5.25}$$

Actually, the operator $\mathcal{P}_\ell$ is well-defined in $L^1(\Omega)$. This follows by Hölder's inequality because $\psi_p^\ell \in L^\infty(\Omega)$ for all $p \in \mathcal{N}_\ell$. The fully discrete representation of the pseudo-L^2-projection $\mathcal{P}_{\ell-1}^\ell : X_{\ell-1} \to X_\ell$ is obtained analogously to the ones of the L^2-projection and the L^2-quasi-projections from the previous sections.

The idea to use a Petrov–Galerkin scheme to define a generalized projection operator can be found in [174] for $d \in \{1, 2\}$, too. There, the test space is constructed differently; more precisely, the local test functions are associated with a dual mesh (see Figure 5.3) often found in finite volume schemes. We establish deeper connections to the literature in the following paragraph.

5.5.2 Historical remarks

Undoubtedly, the root of the class of operators considered in this section lies in the research of quasi-interpolation concepts by Clément [58]; we have described some fundamental ideas in Section 5.2. However, the first appearance of a weighted interpolation operator using a system of biorthogonal test functions was in [171]. We comment on this in more detail below. Note that biorthogonal systems of some form or another are considerably more common in the context of wavelets; see, e. g., [59, 65].

Generalized projections using dual test functions have first been introduced to the area of domain decomposition methods by [190, 191] and then [118]. In this context, one considers a non-conforming discretization associated with a non-overlapping decomposition of the computational domain. In each subdomain, an independent finite element space is used; the meshes meeting at the interior interfaces do not need to match. Then, the mortar finite element method developed in [19] achieves an optimal a priori error estimate by enforcing weak matching conditions in terms of orthogonal projections with respect to the L^2-inner product on the interfaces. This approach was reformulated as saddle point problem in [18] with suitable trace spaces of the standard finite element functions as Lagrange multiplier spaces. A motivation for the usage of discontinuous test functions (or Lagrange multipliers) in [190, 191] was the fact that the normal derivative of the global function is not necessarily continuous at the coupling interfaces. All in all, in practical computations, operators of the type (5.25) have been used to map trace functions between non-matching interfaces; see also [71, 72] and the references therein.

As indicated before, an earlier version of the transfer concept based on biorthogonal test functions was proposed by Scott and Zhang in [171]. The operator is defined by

$$\mathcal{S}_\ell : W_p^m(\Omega) \to X_\ell, \quad u \mapsto \mathcal{S}_\ell u := \sum_{q \in \mathcal{N}_\ell} (\psi_{\sigma_q}^\ell, u)_{L^2(\sigma_q)} \lambda_q^\ell, \tag{5.26}$$

with $m \geq 1$ if $p = 1$ and $m > 1/p$ otherwise. As is customary, $W_p^m(\Omega)$ is the space of p-integrable functions with p-integrable weak derivatives up to order m on Ω defined analogously to the Sobolev spaces $H^m(\Omega)$ in Section 1.1.1. The operator $\mathcal{S}_\ell$ has originally been introduced for simplicial meshes $\mathcal{T}_\ell$. With this in mind, the definition (5.26) employs suitably chosen d- or $(d-1)$-dimensional simplices $(\sigma_q)_{q\in\mathcal{N}_\ell}$ with the additional property that $\sigma_q \subset \partial\Omega$ if $q \in \partial\Omega$. Then, there exists a system of dual functions, namely a set $(\psi^\ell_{\sigma_q})_{q\in\mathcal{N}_\ell}$ with $\psi^\ell_{\sigma_q} \in \mathcal{C}^0(\sigma_q)$ for all $q \in \mathcal{N}_\ell$ which is biorthogonal to the restricted basis $(\lambda^\ell_q|_{\sigma_q})_{q\in\mathcal{N}_\ell}$ with respect to the d- or $(d-1)$-dimensional inner product $(\cdot,\cdot)_{L^2(\sigma_q)}$, i. e.,

$$(\psi^\ell_{\sigma_q}, \lambda^\ell_r)_{L^2(\sigma_q)} = \delta_{qr}, \quad \forall\, q, r \in \mathcal{N}_\ell.$$

The assumptions on m and p ensure that the required trace theorem holds for $u \in W_p^m(\Omega)$; thus, by Hölder's inequality, the integrals on the chosen sub-simplices in (5.26) are well-defined. Note that the mapping $\mathcal{S}_\ell$ depends on the specific choice of these simplices.

The Scott–Zhang operator is designed to preserve Dirichlet boundary conditions on $\partial\Omega$; see also Remark 5.13 below. This makes its definition via the choice of the sub-simplices $(\sigma_q)_{q\in\mathcal{N}_\ell}$ more intricate and unsymmetric in a sense. In contrast, our version is symmetric, namely the test functions $\psi^\ell_q \in \Psi_\ell$ constructed in Section 5.5.1 have the same support as the basis functions $\lambda^\ell_q \in \Lambda_\ell$. As we employ functions which are biorthogonal to the standard d-dimensional nodal basis with respect to the full L^2-inner product $(\cdot,\cdot)_{L^2(\Omega)}$, we do not have to choose any sub-simplices but rather work with the given finite element meshes. A further advantage, which we do not yet exploit here, is the lower requirement for the regularity of the considered functions, namely $L^1(\Omega)$ instead of $W_p^m(\Omega)$ with m, p as specified above.

Remark 5.13. *For the Scott–Zhang operator, by forcing the sub-simplex σ_q to be a subset of $\partial\Omega$ if the corresponding node q lies on the boundary, one obtains the property that $(\mathcal{S}_\ell u)|_{\partial\Omega} = 0$ in case of $u|_{\partial\Omega} = 0$. Note that the trace of u is a well-defined function in $L^1(\partial\Omega)$ for $u \in W_p^m(\Omega)$ with m and p as aforementioned. By contrast, in the semi-geometric framework, we cope with the Dirichlet boundary conditions in a more direct way as described in Section 3.3. Even if the auxiliary coarse spaces $(X_\ell)_{\ell=0,\dots,L-1}$ do not satisfy the desired boundary conditions, the variational approach ensures that all coarse functions in the constructed spaces $(V_\ell)_{\ell=0,\dots,L-1}$ do.*

At this point, the Scott–Zhang operator seems to be a little more general because it preserves vanishing traces without further ado, whereas we need a priori knowledge about the Dirichlet boundary. But this extra flexibility, which involves more regularity assumptions, does not seem to be of any use in the present context. We rather believe our more direct way to treat the boundary conditions to be superior in practice.

5.5.3 On the properties of the pseudo-L^2-projection

In this section, we examine the new operator more closely. First of all, the mapping $\mathcal{P}_\ell$ is surjective, namely $\mathcal{P}_\ell(L^2(\Omega)) = X_\ell$, because (5.19) and (5.25) immediately imply $\mathcal{P}_\ell\lambda^\ell_p = \lambda^\ell_p$ for all $p \in \mathcal{N}_\ell$. Moreover, it is a projection onto X_ℓ. This is a simple consequence of the linearity of the operator and, again, the biorthogonality property (5.19). In addition, it is important to note the following stability property of the constructed operator.

Lemma 5.14. *The pseudo-L^2-projection $\mathcal{P}_\ell$ is continuous or L^2-stable, i. e.,*

$$\|\mathcal{P}_\ell v\|_{L^2(\Omega)} \lesssim \|v\|_{L^2(\Omega)}, \quad \forall\, v \in L^2(\Omega). \tag{5.27}$$

Proof. For a function $v \in L^2(\Omega)$ and an element $T \in \mathcal{T}_\ell$, we have

$$\|\mathcal{P}_\ell v\|_{L^2(T)} \leq \sum_{p \in \mathcal{N}_\ell \cap \overline{T}} \frac{\|\psi_p^\ell\|_{L^2(\omega_p)} \|v\|_{L^2(\omega_p)}}{\|\lambda_p^\ell\|_{L^1(\omega_p)}} \|\lambda_p^\ell\|_{L^2(T)}.$$

The shape regularity of the mesh $\mathcal{T}_\ell$ implies that $\operatorname{diam}(\omega_p) \eqsim \operatorname{diam}(T) =: h_T$ and that the number of terms in the sum is uniformly bounded. Using Hölder's inequality, and the boundedness (5.23), one obtains the following three estimates:

$$\|\psi_p^\ell\|_{L^2(\omega_p)} \lesssim h_T^{d/2}, \quad \|\lambda_p^\ell\|_{L^2(T)} \lesssim h_T^{d/2}, \quad \|\lambda_p^\ell\|_{L^1(\omega_p)} \eqsim h_T^d. \tag{5.28}$$

Therefore,

$$\|\mathcal{P}_\ell v\|_{L^2(T)} \lesssim \|v\|_{L^2(\omega_T)}, \quad \omega_T := \operatorname{int}\left(\bigcup \left\{\overline{T'} \in \mathcal{T}_\ell \,|\, \overline{T'} \cap \overline{T} \neq \emptyset\right\}\right). \tag{5.29}$$

Again, because of the shape regularity of $\mathcal{T}_\ell$, each element of $\mathcal{T}_\ell$ belongs to a uniformly bounded number of patches ω_T; thus, we have

$$\|\mathcal{P}_\ell v\|_{L^2(\Omega)}^2 \lesssim \sum_{T \in \mathcal{T}_\ell} \|v\|_{L^2(\omega_T)}^2 \lesssim \|v\|_{L^2(\Omega)}^2.$$

This completes the proof of the lemma. □

Note that Lemma 5.14 holds true for all shape regular (not necessarily quasi-uniform) meshes because all estimates in (5.28) are local in the sense of Section 3.4.1. Also confer Section 5.2.2.

In the following, we prove the sufficient conditions according to Lemma 3.12 such that the pseudo-L^2-projection is applicable in the semi-geometric multilevel algorithms. Again, all estimates in an element $T \in \mathcal{T}_\ell$ are local, namely only involving the values in ω_T. Therefore, the required properties hold for shape regular meshes.

L^2-approximation property

First, we notice that a preliminary result about the L^2-approximation property of $\mathcal{P}_\ell$ can be proved in a very simple way, similarly to [133]. As a matter of fact, the proof of the estimate is not specially-designed for the pseudo-L^2-projection but holds true for a fairly broad class of operators. This is the assertion of the following

Lemma 5.15. *Let the operator $\Pi_\ell : H^1(\Omega) \to X_\ell$ be L^2-stable and a surjective projection. The mesh $\mathcal{T}_\ell$ may be quasi-uniform with global parameter $h_\ell > 0$. Then, Π_ℓ satisfies the L^2-approximation property.*

Proof. For $v \in H^1(\Omega)$ we employ the Clément quasi-interpolation operator $\mathcal{R}_\ell$ onto X_ℓ and the triangle inequality to calculate

$$\begin{aligned} \|v - \Pi_\ell v\|_{L^2(\Omega)} &\leq \|v - \mathcal{R}_\ell v\|_{L^2(\Omega)} + \|\mathcal{R}_\ell v - \Pi_\ell v\|_{L^2(\Omega)} \\ &= \|v - \mathcal{R}_\ell v\|_{L^2(\Omega)} + \|\Pi_\ell(\mathcal{R}_\ell v - v)\|_{L^2(\Omega)} \\ &\lesssim \|v - \mathcal{R}_\ell v\|_{L^2(\Omega)} \\ &\lesssim h_\ell \, |v|_{H^1(\Omega)}. \end{aligned}$$

By assumption, the operator Π_ℓ is the identity on the image of $\mathcal{R}_\ell$; thus, the equality in the second line holds. The remaining estimates follow directly from the stability of Π_ℓ and the approximation property of $\mathcal{R}_\ell$ stated in Lemma 5.3. □

As a direct consequence one obtains that the pseudo-L^2-projection satisfies the L^2-approximation property in case of quasi-uniform meshes. It seems impossible to generalize the above technique to non-quasi-uniform meshes with a mesh size function $h_\ell \in L^\infty_>(\Omega)$ without requiring an assumption on the stability of Π_ℓ with respect to a weighted L^2-norm. However, the relevant property can be proved directly; the following lemma holds for all shape regular meshes.

Lemma 5.16. *The pseudo-L^2-projection $\mathcal{P}_\ell$ has the L^2-approximation property.*

Proof. For a function $v \in H^1_D(\Omega)$ and an element $T \in \mathcal{T}_\ell$ such that $\overline{T} \cap \Gamma_D = \emptyset$, let $\bar{v} = \frac{1}{|\omega_T|} \int_{\omega_T} v$ be the local average. As the projection operator $\mathcal{P}_\ell$ reproduces constants apart from Γ_D, we have

$$\begin{aligned} \|v - \mathcal{P}_\ell v\|_{L^2(T)} &\leq \|v - \bar{v}\|_{L^2(T)} + \|\mathcal{P}_\ell v - \bar{v}\|_{L^2(T)} \\ &= \|v - \bar{v}\|_{L^2(T)} + \|\mathcal{P}_\ell(v - \bar{v})\|_{L^2(T)} \\ &\lesssim \|v - \bar{v}\|_{L^2(T)} + \|v - \bar{v}\|_{L^2(\omega_T)} \\ &\lesssim \|v - \bar{v}\|_{L^2(\omega_T)} \\ &\lesssim \operatorname{diam}(\omega_T) |v|_{H^1(\omega_T)}. \end{aligned}$$

This follows from the local L^2-stability of the operator $\mathcal{P}_\ell$, see (5.29), and the Poincaré inequality. If $\overline{T} \cap \Gamma_D \neq \emptyset$, the above calculations hold true for the choice $\bar{v} = 0$. As $\operatorname{diam}(\omega_T) \eqsim h_T$, this implies

$$\|h_T^{-1}(v - \mathcal{P}_\ell v)\|_{L^2(T)} \lesssim |v|_{H^1(\omega_T)}.$$

The global version of the above estimate, with the piecewise constant mesh size function $h_\ell \in L^\infty_>(\Omega)$ considered in Section 3.4.1, follows from the shape regularity of the meshes $\mathcal{T}_\ell$; confer the proof of Lemma 5.14. □

H^1-stability

Finally, let us study the stability of the new operator in the following lemmas. This concludes the analysis of the properties which are relevant for the application of the transfer concept in the semi-geometric framework.

Lemma 5.17. *The pseudo-L^2-projection $\mathcal{P}_\ell$ is H^1-stable.*

Proof. Let $v \in H^1_D(\Omega)$. For an element $T \in \mathcal{T}_\ell$ such that $\overline{T} \cap \Gamma_D = \emptyset$, consider the local average $\bar{v} = \frac{1}{|\omega_T|}\int_{\omega_T} v$. As the projection operator $\mathcal{P}_\ell$ reproduces constants away from Γ_D, we have

$$\begin{aligned} |\mathcal{P}_\ell v|_{H^1(T)} &= |\mathcal{P}_\ell v - \bar{v}|_{H^1(T)} = |\mathcal{P}_\ell(v-\bar{v})|_{H^1(T)} \\ &\leq \sum_{p \in \mathcal{N}_\ell \cap \overline{T}} \frac{\|\psi_p^\ell\|_{L^2(\omega_p)}\|v-\bar{v}\|_{L^2(\omega_p)}}{\|\lambda_p^\ell\|_{L^1(\omega_p)}} |\lambda_p^\ell|_{H^1(T)}. \end{aligned}$$

The Poincaré inequality yields

$$\|v-\bar{v}\|_{L^2(\omega_p)} \leq \|v-\bar{v}\|_{L^2(\omega_T)} \leq \operatorname{diam}(\omega_T)|v|_{H^1(\omega_T)}.$$

If $\overline{T} \cap \Gamma_D \neq \emptyset$, the above calculations hold true for the choice $\bar{v} = 0$. With the inverse estimate $|\lambda_p^\ell|_{H^1(T)} \lesssim \operatorname{diam}(T)^{-1}\|\lambda_p^\ell\|_{L^2(T)}$ and the three local estimates in (5.28), we conclude

$$|\mathcal{P}_\ell v|_{H^1(T)} \lesssim |v|_{H^1(\omega_T)}$$

because $\operatorname{diam}(\omega_T) \eqsim h_T$ due to the shape regularity of the mesh. The proof is completed by the same argument as in Lemma 5.14. □

Note that the proof of Lemma 5.17 holds for all shape regular meshes. We give an alternative proof for the stability estimate for arbitrary operators satisfying an L^2-approximation property in case of quasi-uniform meshes.

Lemma 5.18. *Let $\Pi_\ell : H^1(\Omega) \to X_\ell$ satisfy the L^2-approximation property and assume that the mesh $\mathcal{T}_\ell$ is quasi-uniform. Then, Π_ℓ is H^1-stable.*

Proof. To prove this statement, note that Lagrange conforming finite element spaces associated with quasi-uniform meshes come with global inverse inequalities (and mesh size parameters $h_\ell = \max_{T\in\mathcal{T}_\ell} h_T$). As is known, this implies the H^1-stability of the L^2-projection $\mathcal{Q}_\ell$. Moreover, $\mathcal{Q}_\ell$ is the identity on X_ℓ. Thus, for $v \in H^1_D(\Omega)$, we estimate

$$\begin{aligned} |\Pi_\ell v|_{H^1(\Omega)} &\leq |\Pi_\ell v - \mathcal{Q}_\ell v|_{H^1(\Omega)} + |\mathcal{Q}_\ell v|_{H^1(\Omega)} \\ &\lesssim h_\ell^{-1}\|\Pi_\ell v - \mathcal{Q}_\ell v\|_{L^2(\Omega)} + |v|_{H^1(\Omega)} \\ &= h_\ell^{-1}\|\mathcal{Q}_\ell(\Pi_\ell v - v)\|_{L^2(\Omega)} + |v|_{H^1(\Omega)} \\ &\lesssim h_\ell^{-1}\|\Pi_\ell v - v\|_{L^2(\Omega)} + |v|_{H^1(\Omega)} \\ &\lesssim |v|_{H^1(\Omega)}. \end{aligned}$$

This is the assertion, which also holds true for $\mathcal{P}_\ell$. □

As indicated before, we complete the analysis of the L^2-quasi-projections described in Section 5.4 now. More precisely, a proof of Lemma 5.9 about the properties of the operator $\widehat{\mathcal{Q}}_\ell$ is given in the following

Remark 5.19 (Proof of Lemma 5.9)**.** *First of all, note that the operator $\widehat{\mathcal{Q}}_\ell$ has the representation formula (5.18) which is very similar to the one of the pseudo-L^2-projection considered in this section. Besides, constant functions are reproduced locally (again, apart from Γ_D) as*

$$(\widehat{\mathcal{Q}}_\ell \bar{v})|_T = \sum_{p \in \mathcal{N}_\ell \cap \overline{T}} \frac{(\lambda_p^\ell, \bar{v})_{L^2(\omega_p)}}{(\lambda_p^\ell, \mathbb{1})_{L^2(\omega_p)}} \lambda_p^\ell|_T = \sum_{p \in \mathcal{N}_\ell \cap \overline{T}} \bar{v} \lambda_p^\ell|_T = \bar{v}$$

for $\bar{v} \in \mathbb{R}$. Then, for both the L^2-norm and the H^1-semi-norm, the terms of the sum over the nodes $p \in \mathcal{N}_\ell$ are estimated in a straightforward manner, as was done for $\mathcal{P}_\ell$ in the present section. The L^2-approximation property is derived analogously to Lemma 5.16.

We remark that, even though the formula (5.11) is quite similar, the analysis of the simplified quasi-interpolation operator $\widetilde{\mathcal{R}}'_\ell$ may not be carried out in this way. This is because the property to preserve constant functions is missing.

5.6 Application to the semi-geometric multilevel methods

We have indeed found a selection of geometrically inspired transfer concepts which may be employed in the semi-geometric framework of Chapter 3. So, the final step of the theoretical considerations concerning the information transfer between non-nested meshes, which have been made in full detail at least for the pseudo-L^2-projection, is to summarize the applicability of the different approaches.

If a transfer concept has proved to generate local operators satisfying the required stability and approximation properties for all shape regular meshes, we obtain the final result by Theorem 3.14. The multilevel methods based on non-nested meshes yield quasi-optimal preconditioners for a variety of choices of the prolongation operators as stated in the following

Theorem 5.20. *Let the assumption (3.21) of Lemma 3.10 hold and the smoothing property (2.22) be satisfied. Besides, the transfer operators may be generated by one of the following concepts.*

- *Standard nodal finite element interpolation $\mathcal{I}$ (Section 5.1)*
- *Classical Clément quasi-interpolation $\mathcal{R}$ (Section 5.2.1)*
- *Alternative quasi-interpolation $\widetilde{\mathcal{R}}$ (Section 5.2.4)*
- *The L^2-quasi-projections $\widetilde{\mathcal{Q}}$, $\widehat{\mathcal{Q}}$ (Section 5.4)*
- *Pseudo-L^2-projection $\mathcal{P}$ (Section 5.5)*

Then, the semi-geometric multilevel methods Algorithm 3.4 and Algorithm 3.5 yield preconditioning uniformly with respect to the mesh size. Further, the semi-geometric multigrid method converges uniformly with respect to the mesh size.

Note that only the concepts $\mathcal{I}$, $\widetilde{\mathcal{R}}$, $\mathcal{P}$ produce projection operators. But this is not a requirement of the analysis of Section 3.4. We remark that the relevant properties have not been proved for general shape regular families of meshes for the simplified quasi-interpolation $\widetilde{\mathcal{R}}'$ (Section 5.2.4). The main reason that the techniques used for some of the other operators cannot be applied here is the fact that the constant functions are not preserved at least locally, as discussed in Remark 5.19. However, we still intend to assess the performance of these operators in practice. Certainly, the actual L^2-projection $\mathcal{Q}$ (Section 5.3) may not be employed as the generated operators are not local.

To further assess the practical properties of the transfer operators both qualitatively and quantitatively, we proceed in several steps. Implementation issues are addressed in Section 5.7 followed by a unique experimental analysis investigating the characteristics of information transfer between non-nested meshes in Section 5.8. The practical application to the semi-geometric multilevel methods is studied by numerical experiments in Chapter 6.

5.7 Implementation aspects

We have already pointed out some essential guidelines concerning the implementation of the respective operators with the introduction of each single transfer concept (Section 5.1 to Section 5.5). We have also described how non-nested transfer routines are incorporated into a multilevel framework in Section 3.6. In this section, we focus on the realization of the specific transfer concepts in practical finite element codes. This complements the remarks in Section 3.6 about the implementation aspects of the semi-geometric framework itself. All described methods are implemented in the developed module `nnmglib` of `obslib++`. A part of our implementation at an earlier stage has also been used in [83] to formulate weak coupling conditions between simulations from continuum mechanics and molecular dynamics. More precisely, between the corresponding function spaces, transfer concepts similar to the ones investigated here play a crucial role.

For the computation of a matrix representation of a linear operator from $X_{\ell-1}$ to X_ℓ, e. g., in setupSGMG (Algorithm 3.3), one needs to deal with quantities associated with different meshes without any usable a priori relation. Recall that we may yet assume, by a quadtree or octree structure mentioned in Section 3.6.2, suitable neighborhood relations to be given. In particular, for each node $p \in \mathcal{N}_\ell$, a set $\mathcal{N}_{\ell-1}^p \subset \mathcal{N}_{\ell-1}$ containing a sufficiently small number of nodes has been determined such that

$$q \in \mathcal{N}_{\ell-1},\, \mathrm{int}(\omega_q) \cap \mathrm{int}(\omega_p) \neq \emptyset \implies q \in \mathcal{N}_{\ell-1}^p.$$

Then, all terms which appear in the presented discrete operators may evidently be computed only based on these local subsets. We elaborate on this in the following.

Note that the assembly process is in general not symmetric. This is easy to understand in the context of the nodal interpolation where coarse level basis functions need to be evaluated at fine level nodes. Besides, the only part which is symmetric in the sense that coarse and fine level quantities are treated equally is the rectangular matrix $\boldsymbol{B}_\ell$. It appears in the representations of $\mathcal{Q}_{\ell-1}^\ell$, $\widetilde{\mathcal{Q}}_{\ell-1}^\ell$, $\widehat{\mathcal{Q}}_{\ell-1}^\ell$ and also $\widetilde{\mathcal{R}}_{\ell-1}^{\ell\prime}$ and merely contains the L^2-inner products of coarse and fine level basis functions.

Numerical integration

For all transfer concepts, with the exception of the nodal interpolation $\mathcal{I}$, inner products of functions associated with different meshes need to be computed. This is obvious for the above indicated operators involving the sparse but global coupling matrix $\boldsymbol{B}_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ defined in (5.13). The analogon for the pseudo-L^2-projection $\mathcal{P}$ requires the entries

$$(\psi_p^\ell, \lambda_q^{\ell-1})_{L^2(\omega_p \cap \omega_q)}, \quad \forall\, p \in \mathcal{N}_\ell,\ q \in \mathcal{N}_{\ell-1}. \tag{5.30}$$

We turn to the other mappings which employ local orthogonal projections below.

To evaluate (5.13) or (5.30) in the setup phase (Algorithm 3.3) exactly, one has to compute the intersections of the elements in the consecutive meshes. As we have previously done in [69, 71] for the intersection of locally projected non-matching interface meshes, we employ the quickhull algorithm in an implementation by [16] for this purpose. After a suitable remeshing of the computed intersection polytopes, one achieves an exact integration, up to roundoff errors, by the application of low order quadrature rules*. We have implemented the methods concerning element intersections in a module `cutlib`.

In practice, good results may be obtained by an approximate numerical integration via a quadrature rule solely based on the finer mesh. The order of the employed quadrature rules should be adequate such that they are exact at least in case of nested meshes. This requires order two for the above operators and order $r+1$ for the Clément quasi-interpolations. We are aware of the fact that such an approach might fail to retain optimal (discretization) error estimates, for instance, in the mortar finite element setting; see [85, 141]. However, the situation is different here and we do not experience any problems. In addition, let us refer to our numerical studies in Section 5.8.3, where we show that the error in the operator itself due to approximate integration is small if the quadrature rule is chosen adequately.

Computation of orthogonal projections

The implementation issues concerning the operator $\mathcal{Q}_{\ell-1}^\ell$, which is the orthogonal projection to the space X_ℓ with respect to the L^2-inner product, are discussed in detail in Section 5.3. For the stated reasons, it is not suitable for an application in the multilevel algorithms. But, for the experiments in Section 5.8, we employ the direct sparse solver `pardiso` [165, 166] to decompose the appearing mass matrices. This is more efficient than an iterative solver in this special case as the respective inverse needs to be applied to a large number of vectors.

The pseudo-L^2-projection is defined via a global variational formulation, too. This has direct consequences for the implementation as no subproblems need to be solved in case of unstructured simplicial meshes. Moreover, the system of biorthogonal test functions is designed to obviate the inversion of a global matrix. Note that the quasi-projections from Section 5.4 also yield simple formulations; they are not projections, though. The same holds true for the quasi-interpolation concept in case the trial and test spaces are one-dimensional, e. g., for the Clément operator with $r = 0$.

In contrast, we have seen that general transfer concepts may require the evaluation of local orthogonal projections. In the following, we sketch the implementation of the operators

*We always refer to numerical integration as quadrature, no matter what the dimension d is.

$\mathcal{R}_{\ell-1}^{\ell}$ with $r > 0$ and $\widetilde{\mathcal{R}}_{\ell-1}^{\ell}$. To solve the corresponding local variational equations (5.5) and (5.9) for the right hand sides given by the coarse level basis functions, one needs to compute coarse-to-fine coupling matrices and mass matrices similar to the ones in (5.12) but associated with the local spaces. Let $(\varphi_i^p)_{i=1,\dots,n_p}$ be a basis of the considered trial space at $p \in \mathcal{N}_\ell$. Then,

$$(\boldsymbol{M}_p)_{ij} = (\varphi_i^p, \varphi_j^p)_{L^2(\omega_p)}, \quad \forall\, 1 \le i, j \le n_p, \tag{5.31}$$

and

$$(\boldsymbol{B}_p)_{iq} = (\varphi_i^p, \lambda_q^{\ell-1})_{L^2(\omega_p)}, \quad \forall\, 1 \le i \le n_p,\ q \in \mathcal{N}_{\ell-1}, \tag{5.32}$$

are the respective local matrices. We omit the level index ℓ as it is clear from the choice of p.

For the Clément quasi-interpolation operators, the trial and test spaces are obtained by restrictions of global polynomial spaces to the patches. Therefore, one may choose a universal basis for the implementation; for instance, $(\varphi_i)_{i=1\dots,d+1}$ with $\varphi_i(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{e}_i$ for $i \le d$ and $\varphi_{d+1} \equiv 1$ is a convenient choice in case $r = 1$. For the alternative operator $\widetilde{\mathcal{R}}_{\ell-1}^{\ell}$, the dimension of the local finite element spaces depends on the complexity of the patches. An appropriate basis of X_ℓ^p from (5.10), $p \in \mathcal{N}_\ell$, is immediately obtained by restricting the nodal basis functions to ω_p. The issues concerning the numerical integration of (5.32) are solved as before for the global coupling matrices.

As usual in finite element assembly algorithms, a single loop over all elements in $\mathcal{T}_\ell$ makes sure that no redundant computations are carried out; each integral is only computed once. However, to store the local information rather than merely summing up the element contributions, some additional structure is needed as the essential objects at this point, namely the patches, are usually not included in customary finite element codes. For a flexible handling of all above cases, we introduced a class `nodepatch` whose instances manage the local data, i. e., essentially the matrices from (5.31) and (5.32). The objects also require access to the information which nodes belong to the respective patch in case of $\widetilde{\mathcal{R}}_{\ell-1}^{\ell}$. This is crucial for the local numberings of the sets $(\varphi_i^p)_{i=1,\dots,n_p}$, $p \in \mathcal{N}_\ell$; we assume p is always associated with $i = 1$. For $\mathcal{R}_{\ell-1}^{\ell}$ this is not necessary if the bases are chosen as described above, namely as globally defined polynomials. In any case, for each fine level node, the coupling coarse level nodes are determined during the assembly routine. This is also true for the assembly of the global matrices in case of the other operators discussed before. The number of candidates is evidently limited by the sets $\mathcal{N}_{\ell-1}^p$, $p \in \mathcal{N}_\ell$, generated by the quadtree or octree structure.

Finally, the entries of the global matrix $\widetilde{\boldsymbol{\mathcal{R}}}_{\ell-1}^{\ell} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ are obtained by the scalar products

$$(\widetilde{\boldsymbol{\mathcal{R}}}_{\ell-1}^{\ell})_{pq} = \sum_{i=1}^{n_p} (\boldsymbol{M}_p^{-1})_{i1} (\boldsymbol{B}_p)_{iq}, \quad \forall\, p \in \mathcal{N}_\ell,\ q \in \mathcal{N}_{\ell-1}.$$

Similarly, the matrix representations of the Clément operators read as

$$(\boldsymbol{\mathcal{R}}_{\ell-1}^{\ell})_{pq} = \sum_{i=1}^{n_p} \sum_{j=1}^{n_p} (\boldsymbol{M}_p^{-1})_{ij} (\boldsymbol{B}_p)_{jq} \varphi_i(p), \quad \forall\, p \in \mathcal{N}_\ell,\ q \in \mathcal{N}_{\ell-1}.$$

These formulas are immediately derived by solving the variational equations (5.5) and (5.9), respectively, for the basis functions $(\lambda_q^{\ell-1})_{q\in\mathcal{N}_\ell}$ and evaluating the result at the node p. Both approaches require the inversion of n_ℓ local $(n_p \times n_p)$-matrices. In the latter, the local dimension is constant; we have $n_p = \dim(\mathbb{P}_r(\omega_p)) = \frac{(d+r)!}{d!\,r!}$.

5.8 Numerical results

Let us now focus on the practical properties of the operators generated by the previously described transfer concepts. In this section, we report on various numerical experiments which are performed to assess the stability properties of the single operators and the interconnections between them. Subjecting these mappings to a close examination, we aspire to make out the fundamental characteristics of the information transfer between non-nested finite element spaces as such. We consider this approach very helpful to get the feel of the diverse operators and approximate variants when it comes to their properties in practice. A detailed performance analysis of the application of the transfer concepts to the prolongation in the semi-geometric multilevel methods is studied afterwards in Section 6.

To investigate the behaviors of the mappings between two non-nested finite element spaces, we introduce suitable operator norms with respect to the L^2-norm and the H^1-semi-norm associated with the appropriate domains. By a sampling procedure explained in Section 5.8.2, we obtain results on the accuracy of the approximate numerical integration. Moreover, the numerical experiments indicate that mostly the constants in the H^1-stability estimates required for the analysis in Section 3.4 are bounded by one; the dependence of the operator norm on the mesh size is studied in Section 5.8.4. Finally, we examine quantitative differences of the transfer concepts by measuring certain distances between the generated operators in Section 5.8.5. Let us mention, as early as now, that we confirmed that the operators $\widehat{\mathcal{Q}}$ and $\widetilde{\mathcal{Q}}$ coincide for simplicial meshes. We also learned that the rather expensive operator $\widetilde{\mathcal{R}}$, which works with local projections onto restricted finite element spaces, is identical to $\mathcal{P}$ in this particular case. The investigation of other element types and higher order trial functions needs to be done somewhere else.

The development of the rather exceptional system for the assessment outlined above is motivated by the desire to become a little more familiar with the application of (to a greater or lesser extent) sophisticated (quasi-)interpolation and (quasi-)projection operators in practical computations in the overall context of this thesis. It is a quite remarkable method which, in the end, allows for "drawing a map" arranging the operators in a suitable sense. To our knowledge, the evaluation of operators for the information transfer between finite element spaces associated with non-nested meshes has never been studied in such a manner so far. The experiments designed and conducted here are completely new.

5.8.1 Setup of the experiments

For the experiments to be carried out in this section, we consider a number of independently generated meshes of the unit ball. This is an appropriate geometric setting as one can easily obtain completely independent unstructured volume meshes for a large variety of different mesh sizes by standard tetrahedral mesh generation tools, e. g., from CUBIT [61].

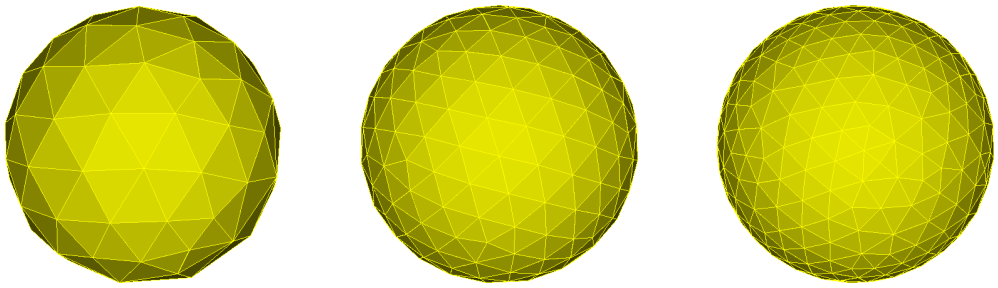


Figure 5.6. Three of the fifteen unrelated meshes ($\mathcal{B}_1$, $\mathcal{B}_3$, $\mathcal{B}_4$, from left to right) for the numerical experiments of this section.

In addition, it yields very good reproducibility. Note that the setting is also sufficiently general. On the one hand, this can be seen in an illustrative example in Remark 5.21. On the other hand, we have tested other geometries with essentially the same results.

We make use of a set of meshes $(\mathcal{B}_i)_{i=1,\dots,15}$ of the unit ball with their characteristics given in Table 5.2 ordered by the number of elements. Let the respective domains be denoted by $(\Omega_i)_{i=1,\dots,15}$. The meshes $\mathcal{B}_1$, $\mathcal{B}_3$ and $\mathcal{B}_4$ are illustrated in Figure 5.6. Note that the situation between the single meshes is sufficiently general in the sense that there are no mutual relations other than that they approximate the same domain. In particular, none of the meshes stems from a refinement routine; they are all imported separately.

To avoid conflicts with the meaning of meshes in the other parts of this thesis, we use a different notation here, namely the letter $\mathcal{B}$ with some index unequal ℓ. This is in contrast to the meshes $\mathcal{T}_\ell$; the latter are always associated with a specific level in a finite element space hierarchy. Here, we consider mappings between all these different meshes and introduce the notation, again, by using the generic operator symbol Π with i and j as indices and exponents. Let $(X_i)_{i=1,\dots,15}$ be the standard finite element spaces associated with the meshes $(\mathcal{B}_i)_{i=1,\dots,15}$ without any boundary modifications. Then, we denote the connecting operators, e. g., by $\Pi_i^j : X_i \to X_j$ for $1 \leq i, j \leq 15$. The studied concrete versions are generated by the transfer concepts presented before.

Let us clarify that we examine the information transfer between the given non-nested finite element spaces rather than the nested ones which have been constructed in Section 3.2.1. It is these original operators which appear in the analysis of Section 3.4. Recall that the corresponding matrices represent the natural embeddings of the newly constructed spaces in the semi-geometric setting as discussed previously.

In the following, several quantities of different natures are studied. Usually, operator norms play a central role; for $\Pi_i^j, \widetilde{\Pi}_i^j \in \mathrm{Lin}(X_i, X_j)$, we estimate terms of the form

$$\sup_{v \in X_i \cap H_0^1(\Omega_i), \|v\|_i \neq 0} \frac{\|\Pi_i^j v\|_j}{\|v\|_i} \quad \text{or} \quad \sup_{v \in X_i \cap H_0^1(\Omega_i), \|v\|_i \neq 0} \frac{\|\Pi_i^j v - \widetilde{\Pi}_i^j v\|_j}{\|v\|_i} \tag{5.33}$$

where $\|\cdot\|_i$ and $\|\cdot\|_j$ are suitably chosen (semi-)norms in X_i and X_j, respectively.

All mentioned transfer concepts may be employed to construct operators mapping an infinite-dimensional function space to a finite element space which is normally a subspace. However, we emphasize that we do *not* consider a general Hilbert space setting but restrict

	#elements	#nodes in $\overline{\Omega}$	#nodes on $\partial\Omega$
$\mathcal{B}_1$	566	151	94
$\mathcal{B}_2$	1,380	325	165
$\mathcal{B}_3$	2,314	520	241
$\mathcal{B}_4$	4,111	882	363
$\mathcal{B}_5$	9,307	1,896	659
$\mathcal{B}_6$	13,383	2,678	868
$\mathcal{B}_7$	20,077	3,970	1,225
$\mathcal{B}_8$	31,728	6,162	1,733
$\mathcal{B}_9$	48,320	9,228	2,364
$\mathcal{B}_{10}$	75,884	14,388	3,539
$\mathcal{B}_{11}$	93,620	17,647	4,179
$\mathcal{B}_{12}$	122,578	23,046	5,388
$\mathcal{B}_{13}$	181,789	33,067	5,643
$\mathcal{B}_{14}$	203,618	36,868	6,099
$\mathcal{B}_{15}$	262,365	47,348	7,539

Table 5.2. Characteristics of the independently generated meshes $(\mathcal{B}_i)_{i=1,\dots,15}$ of the unit ball. The meshes do not stem from a refinement routine; they cover a broad range of sizes.

the attention to the case of two finite element spaces which is relevant for this thesis. Therefore, the suprema in (5.33) and below are taken over finite element functions in X_i only. This is in perfect agreement with the analysis of Section 3.4 and with the comments about the fully discrete case made throughout the first part of this chapter. Still, it has major consequences for the expected results of the numerical studies as each considered operator generally does *not* map a Hilbert space to some subspace. Simple and well-known statements such as "a projection operator has norm greater or equal one" do *not* hold true in our case. Finally, we require the test functions to be in $H^1_0(\Omega_i)$ such that their extensions by zero to the possibly larger domain Ω_j are continuous and, thus, weakly differentiable.

5.8.2 A sampling procedure

For the numerical evaluation or rather estimation of quantities of the form (5.33), we introduce a sampling technique. In the fully discrete setting, a variety of trial functions is employed to scan the behaviors of the considered operators Π_i^j between the respective finite element spaces X_i and X_j.

For this purpose, choose a set of functions $\{v_k\}_{k=1,\dots,N} \subset X_i \cap H^1_0(\Omega_i)$ for an integer $N \gg 1$. Assume that v_k is not identically zero for all $k = 1, \dots, N$; thus, $\|v_k\|_{L^2(\Omega_i)} > 0$ and $|v_k|_{H^1(\Omega_i)} > 0$. Then, the norms of an operator $\Pi_i^j \in \mathrm{Lin}(X_i, X_j)$ with respect to the L^2-norm and the H^1-semi-norm are approximated by

$$\|\Pi_i^j\|_{L^2} := \max_{k=1,\dots,N} \frac{\|\Pi_i^j v_k\|_{L^2(\Omega_j)}}{\|v_k\|_{L^2(\Omega_i)}} \quad \text{and} \quad |\Pi_i^j|_{H^1} := \max_{k=1,\dots,N} \frac{|\Pi_i^j v_k|_{H^1(\Omega_j)}}{|v_k|_{H^1(\Omega_i)}}, \tag{5.34}$$

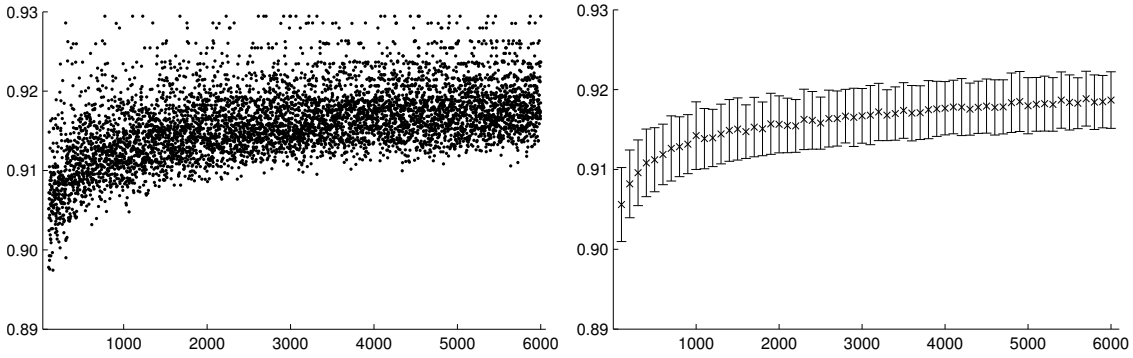


Figure 5.7. Sampling procedure. Exemplary illustration of the improving accuracy with increasing sample size. Values computed by one independent sampling according to (5.34) for each $N \geq 100$ (left); approximation of mean and standard deviation of the sampling procedure depending on N (right).

respectively. For brevity, in the notations of the approximate operator norms, we omit the two different spaces with the two different domains and the dependence on the sample. Naturally, $|\cdot|_{H^1(\Omega_i)}$ is a norm in $H_0^1(\Omega_i)$. Distances between operators may be measured by $\|\cdot\|_{L^2}$ or $|\cdot|_{H^1}$ introduced in (5.34). We also consider relative quantities, namely terms of the form

$$\frac{\|\Pi_i^j - \widetilde{\Pi}_i^j\|_{L^2}}{\|\Pi_i^j\|_{L^2}} \quad \text{and} \quad \frac{|\Pi_i^j - \widetilde{\Pi}_i^j|_{H^1}}{|\Pi_i^j|_{H^1}}$$

for some $\Pi_i^j, \widetilde{\Pi}_i^j \in \mathrm{Lin}(X_i, X_j)$.

For the computation of a quantity involving Π_i^j, we proceed as outlined in Section 5.7 to obtain a numerical representation of the operator. Then, N matrix-vector multiplications with a prolongation matrix in $\mathbb{R}^{n_j \times n_i}$ are required to map the sample from X_i to X_j. To evaluate the L^2-projection, an additional forward-backward substitution per function is necessary. The finite element input data $\{v_k\}_{k=1,\dots,N}$ is gathered as vectors in $\mathbb{R}^{n_i}$ generated by standard pseudo-random numbers (almost) uniformly distributed in $[-1, +1]$.

We do not aspire to investigate the behavior of the sampling procedure itself too closely. Let us just illustrate that the results we achieve seem absolutely reasonable and deduce a presumable accuracy of the estimation by Figure 5.7. In the left part, we see that the variation of the computed maximum (here, the approximation $|\mathcal{P}_1^4|_{H^1}$) is quite small. For each size $100 \leq N \leq 6000$, one independent random sample has been chosen; the values of the corresponding estimates according to (5.34) are marked by dots. To confirm this, we estimate the standard deviation of the sampling procedure depending on the number of samples by performing a large number of independent samplings for each size $N = 100, 200, \dots, 6000$. The computed mean and standard deviation are depicted in the right part of the figure. In addition, we experience that the quality only weakly depends on n_i. Note that it is also robust with respect to the choice of the specific type of the operator. On the whole, our tests indicate that the error for any of the norms is in the order of $0.01 - 0.02$ for adequate N. Consequently, the collected data we present in the following is sufficiently reliable.

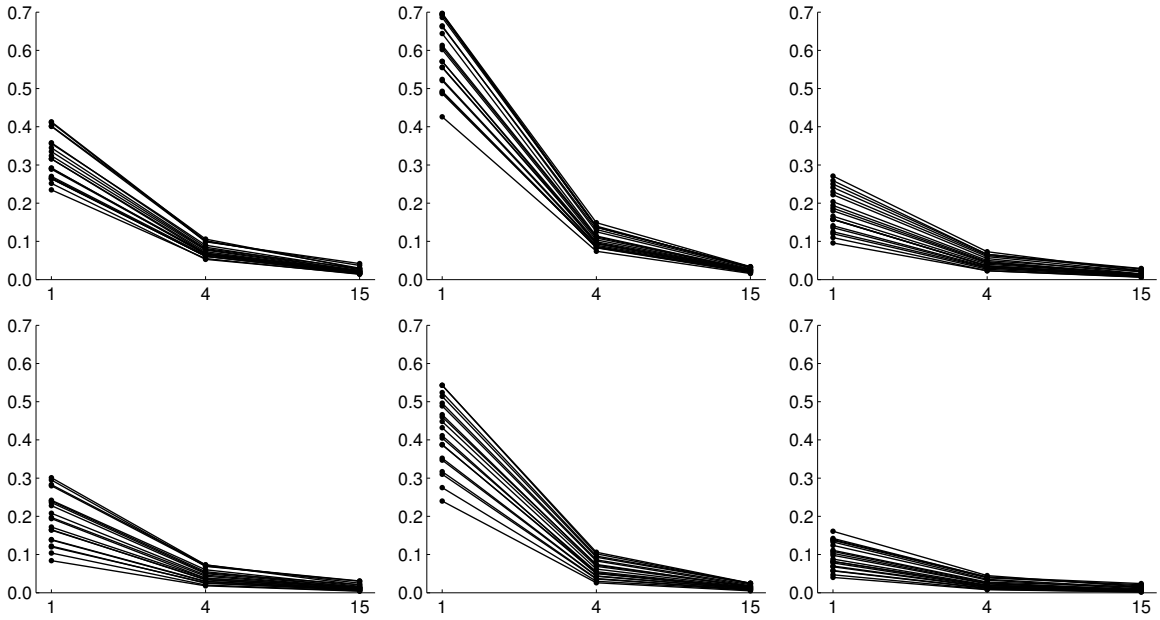


Figure 5.8. The differences between an operator computed by exact integration on intersections and the approximate evaluation based on inexact numerical integration becomes small for higher order quadrature rules. For the operators $\mathcal{Q}$, $\mathcal{P}$, $\widehat{\mathcal{Q}}$ (from left to right), the diagrams show the relative error with respect to $|\cdot|_{H^1}$ (first row) and $\|\cdot\|_{L^2}$ (second row) depending on the number of integration points per element. Each line represents one of the combinations $(\mathcal{B}_i, \mathcal{B}_j)$ with $i < j \leq 7$.

5.8.3 Influence of numerical integration

In this paragraph, we consider the inexact integration of the coupling terms between the basis functions of X_i and X_j by means of a quadrature rule solely associated with $\mathcal{T}_j$, as described in Section 5.7. We verify that this approximation is very accurate in case of sufficiently many function evaluations per element. Therefore, one can avoid the rather expensive computation of the intersections of the elements in the involved meshes in practice, even if one puts emphasis on the application of a particular transfer operator.

To quantify the effect not on the integrals as such but on the actual mappings, we estimate the relative differences between the operators generated by the transfer concepts $\mathcal{Q}$, $\mathcal{P}$ and $\widehat{\mathcal{Q}}$ on the one hand and approximate versions on the other hand. Inner products of finite element functions associated with the same mesh are always evaluated exactly except for roundoff errors. Moreover, neglecting roundoff errors also in the intersection procedure (and, for $\mathcal{Q}$, in the application of the inverse mass matrix), we may indeed assume that it is possible to evaluate the operators exactly. In contrast, the inexact evaluation is based on quadrature merely associated with the mesh of the target finite element space.

In Figure 5.8, we show the results for three standard quadrature rules with one, four and fifteen points, which are exact for polynomials of order one, two and five, respectively. In our experiments, a rule with eight points (order three) only performed comparable to the one with four points. The results are given for the combinations of the first seven meshes

$\mathcal{Q}$: H^1-error	$\mathcal{Q}$: L^2-error	$\mathcal{P}$: H^1-error	$\mathcal{P}$: L^2-error	$\widehat{\mathcal{Q}}$: H^1-error	$\widehat{\mathcal{Q}}$: L^2-error
2.3%	1.4%	2.5%	1.5%	1.5%	0.9%

Table 5.3. Average relative errors of the approximate operators in case the higher order quadrature rule is applied.

$(\mathcal{B}_i)_{i=1,\dots,7}$, namely we investigate $\Pi_i^j : X_i \to X_j$ for $i < j \leq 7$ and $\Pi \in \{\mathcal{Q}, \mathcal{P}, \widehat{\mathcal{Q}}\}$.

As expected, the quality of the approximation improves considerably as the number of integration points is increased. We also note that, for fixed coarse mesh $\mathcal{B}_i$, the error becomes smaller with increasing index j. This is obvious but cannot be seen in the figure as we do not intend to label all the single lines. Other than that, we do not experience any dependence on the mesh size. In particular, for the most critical combinations $(\mathcal{B}_i, \mathcal{B}_{i+1})_{i=1,\dots,6}$, the errors depicted in Figure 5.8 do not grow with increasing i. We may also allow operators $\Pi_i^j : X_i \to X_j$ with $i > j$ provided h_j is not a lot larger than h_i. The experiments indicate that the error is under control for the studied examples in case $i \leq j + 3$.

Further, let us comment on the slightly different error decay for the three transfer concepts in Figure 5.8. The coarse-to-fine integrand is the same for the orthogonal projection $\mathcal{Q}$ and the lumped version $\widehat{\mathcal{Q}}$; see Section 5.7. But the error transport via the inverse mass matrix is different from the one via scaling with its row totals, apparently yielding larger error for $\mathcal{Q}$. For the pseudo-L^2-projection $\mathcal{P}$, the respective integrand (5.30) has larger derivatives because the biorthogonal test functions are employed instead of the nodal basis, which leads to a larger quadrature error. The scaling of the rectangular matrix is the same as for $\widehat{\mathcal{Q}}$.

Finally, an important conclusion from our studies is that the higher order quadrature rule always produces a very accurate approximation of the operators. For the concepts other than the ones investigated in detail here, no new issues arise. We summarize the average relative errors in Table 5.3.

5.8.4 Stability of the operators

As discussed in detail in Section 3.4.4, the analysis of the semi-geometric multigrid method involves the product of the stability constants of the employed prolongation operators. This is because we exploit the recursive structure of the semi-geometric space hierarchy in the proofs. As we are now able to study norms of operators between finite element spaces in a practical way, let us examine this point more carefully.

In this paragraph, we provide experimental evidence that the H^1-stability constants of the operators generated by most of the transfer concepts are bounded by one. This cannot hold true in a general Hilbert space setting as, e. g., a projection operator to a subspace, which is not identically zero, obviously has operator norm greater or equal one. However, the first hint for the case of two finite element spaces is given in Section 5.1 about the standard nodal interpolation, where this assertion is proved for $\Pi = \mathcal{I}$ and $d = 1$ in Lemma 5.2.

The numerical results can be found in Appendix A starting on page 165. Table A.1 to Table A.8 present the H^1-stability constants, i. e., the norms $|\cdot|_{H^1}$ of the operators $\mathcal{I}$,

	0°	1°	2°	3°	4°	5°	8°	16°	32°	64°
$\mathcal{B}_7 \to \mathcal{B}_8$	0.736	0.734	0.735	0.736	0.734	0.733	0.732	0.736	0.735	0.736
$\mathcal{B}_7 \to \mathcal{B}_9$	0.791	0.790	0.789	0.790	0.790	0.790	0.789	0.787	0.791	0.794
$\mathcal{B}_7 \to \mathcal{B}_{12}$	0.871	0.871	0.871	0.869	0.869	0.869	0.867	0.872	0.869	0.871
$\mathcal{B}_7 \to \mathcal{B}_{15}$	0.923	0.923	0.923	0.923	0.924	0.924	0.922	0.923	0.922	0.922

Table 5.4. The setting is sufficiently general. The computed estimates of the operator norms are independent of rotations of the meshes. We show exemplarily the H^1-stability of $\mathcal{P}$ between meshes of different sizes.

$\mathcal{Q}$, $\mathcal{P}$, $\mathcal{R}$ (for $r = 0, 1, 2$), $\widehat{\mathcal{Q}}$ and $\widetilde{\mathcal{R}}'$, each time considered between the spaces associated with $(\mathcal{B}_i, \mathcal{B}_j)$ for $i < j \leq 15$. Here, i and j are the row and column indices, respectively. The numbers are measured quite accurately in the previously described fashion. Apart from the comments at the end of this chapter (Section 5.9), we do not study the presented operators for an application to fine-to-coarse transfer. However, for completeness, we give the respective stability estimates in the tables for $j < i \leq j + 3$ and also include $i = j$. As we are less interested in these cases here, they appear shaded gray.

First, one notes that the estimated quantities are less or equal one. With the exception of the operator $\widetilde{\mathcal{R}}'$, this holds true over the whole range of depicted problem settings with the numbers of elements varying from 566 to 262, 365 (a factor of a little more than 460). The corresponding ratios of the typical mesh sizes of $\mathcal{B}_j$ to $\mathcal{B}_i$ for the considered combinations roughly vary from $\frac{1}{8}$ to $\frac{2}{1}$; the range in between is well covered. Therefore, our extensive numerical experiments make us confident that this observation about the boundedness of the operators with respect to the H^1-semi-norm is valid for the "well-behaved" transfer concepts in general situations. As a consequence, let us make perfectly clear that this would rule out the possible dependence of the constants in the H^1-stability and in the L^2-approximation properties (3.12) and (3.13) of the constructed fine-to-coarse operators $\mathcal{Q}_\ell^V$ on the number of levels in the analysis of Section 3.4.

Second, we confirm that the mappings generated by Clément's quasi-interpolation, namely $\mathcal{R}$ with $r = 0, 1, 2$, and the quasi-projection $\widehat{\mathcal{Q}}$ are not even close to projections; the norms on the diagonal, which describes the cases of identical meshes, are considerably less than one. The former are also denoted by $\mathcal{R}_{r=0}$, $\mathcal{R}_{r=1}$, $\mathcal{R}_{r=2}$ in the following.

Third, concerning the dependence on the mesh size, one notes two aspects in addition to the mentioned boundedness. Again, let us exclude $\widetilde{\mathcal{R}}'$ from the considerations. Then, on the one hand, the norm estimate increases up to one with increasing j for fixed i. This is due to the fact that the operators reproduce constant functions and even an oscillating function associated with a very coarse mesh appears locally constant on a very fine mesh. On the other hand, along the off-diagonal lines, which roughly contain cases with comparable ratios between the respective element numbers, the estimates vary only moderately.

Now, for $\widetilde{\mathcal{R}}'$ the previous considerations are not true. The fact that this naive transfer concept does not locally reproduce constant functions affects the numerical results, which can be seen in Table A.8. The qualitative and quantitative disagreement between $\widetilde{\mathcal{R}}'$ and the other operators is explained in more detail in the next paragraph.

Remark 5.21. *To further illustrate that the relations between the employed meshes of the unit ball are sufficiently general, we present the following experiment. Consider a couple of rotations of the mesh $\mathcal{B}_7$ about the axis spanned by the sum of the standard basis vectors $(\boldsymbol{e}_i)_{1\leq i\leq 3}$ by different angles. Then, Table 5.4 states the estimates $|\cdot|_{H^1}$ for the operator $\mathcal{P}$ between these rotated meshes (for the angles specified at the head) and the meshes $\mathcal{B}_8$, $\mathcal{B}_9$, $\mathcal{B}_{12}$ and $\mathcal{B}_{15}$. Note that the target meshes are of very different sizes.*

If there were distinguished relations between the unrotated mesh and (some of) the other meshes, one would expect some differences in the stability estimates. This is not the case in present and other studies we performed. Although the sampling does not yield this accuracy, we state the numbers up to the third decimal place to show that the differences we obtain are yet marginal.

5.8.5 Quantitative analysis of the relations between the transfer concepts

In this paragraph, we present a quantitative study of the diversity of the operators generated by the relevant geometric transfer concepts. This eventually allows for arranging them in a map-like sketch illustrating similarities and differences. Here, we explain some interesting relations achieved by measuring mutual distances. In the charts designed for this purpose, we mark the operators by the symbols specified in Table 5.5 where they appear. As indicated before, the mapping $\widetilde{\mathcal{R}}'$ is treated separately afterwards.

$\mathcal{I}$	$\mathcal{Q}$	$\mathcal{P}$	$\widehat{\mathcal{Q}}$	$\mathcal{R}_{r=0}$	$\mathcal{R}_{r=1}$	$\mathcal{R}_{r=2}$
circle $\circ$	star $*$	dot $\bullet$	triangle $\triangle$	plus $+$	square $\square$	crossing $\times$

Table 5.5. Symbols of the operators in the charts.

To start with, for each practicable choice of two finite element meshes from $(\mathcal{B}_i)_{i=1,\dots,15}$, we consider the distance of the generated operators to the L^2-projection $\mathcal{Q}$. Figure 5.9 shows the relative difference with respect to $|\cdot|_{H^1}$. The diagram consists of two rows and is arranged such that a section marked by $\mathcal{B}_i$ below (for some index i) comprises the results for the situations $(\mathcal{B}_i, \mathcal{B}_j)$, $i < j$, each time ordered by increasing j from left to right.

We point out two distinct facts established by the performed experiments and readily understood by the figures. First, with decreasing ratio between fine and coarse mesh size, all depicted operators approximate $\mathcal{Q}$ more accurately. This is because they have the common property to preserve the constant functions, which has been mentioned before. In a certain sense, a very fine mesh is "almost nested" in a very coarse mesh and the coarse function is "almost constant" in the patches of the fine mesh; thus, the operators asymptotically become more and more like the identity if the coarse mesh is fixed.

The second, even more important result is the following. We see that, for all experiments, the pseudo-L^2-projection $\mathcal{P}$ is clearly the closest to the actual L^2-projection. In fact, it is remarkable how much closer the operators generated by this transfer concept are to the orthogonal projection compared to all other approaches. The standard interpolation and the Clément-type interpolation with local polynomial degree $r = 2$, although being only

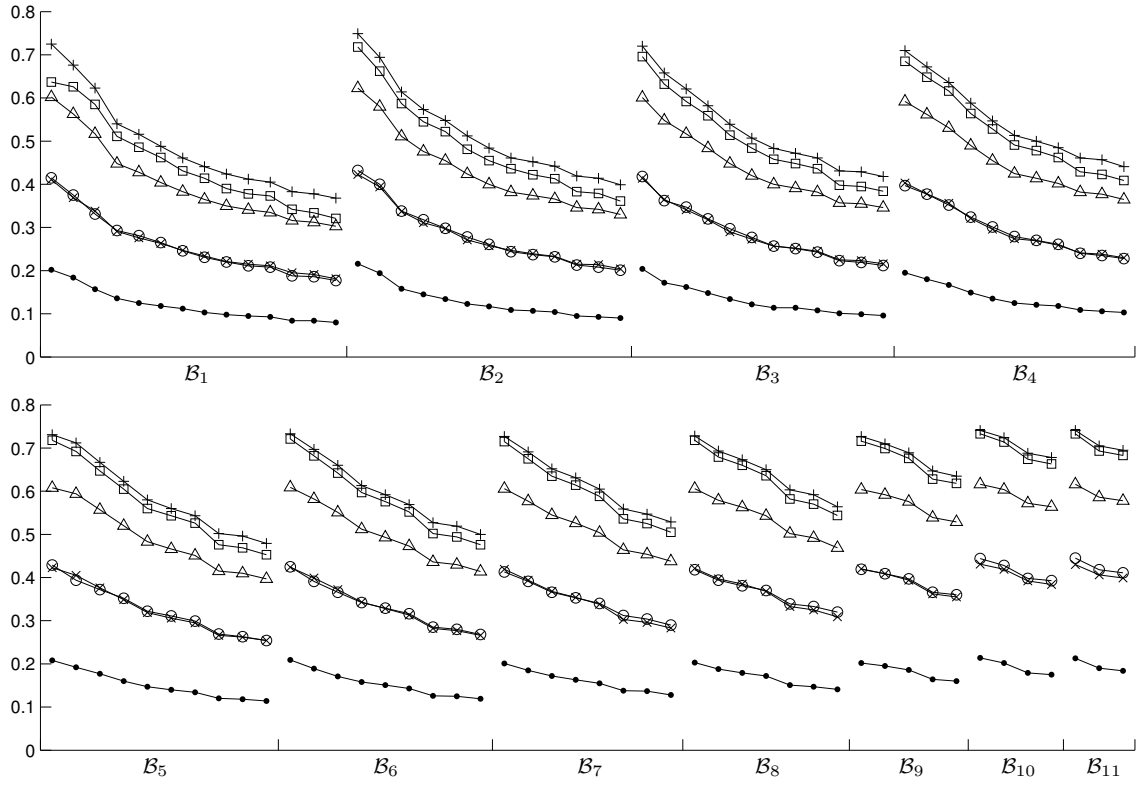


Figure 5.9. Relative distances to the L^2-projection $\mathcal{Q}$ with respect to $|\cdot|_{H^1}$. The labels $\mathcal{B}_1, \ldots, \mathcal{B}_{11}$ indicate the particular space X_i. In each of these sections, the results are given for increasing j from left to right.

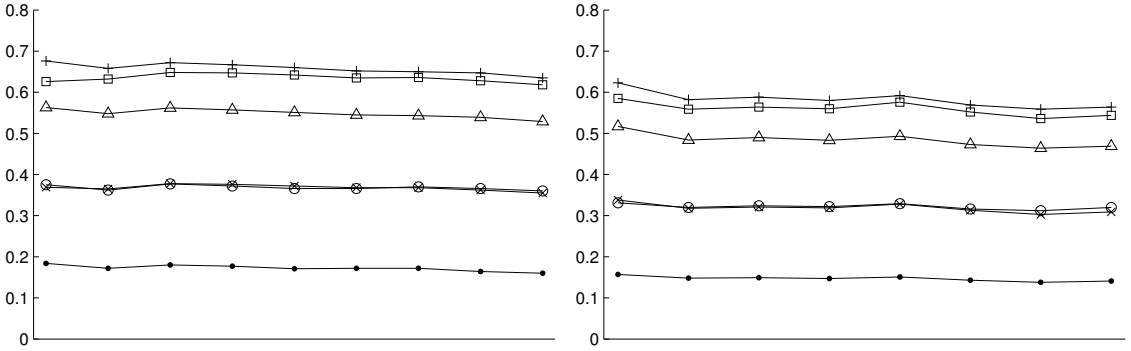


Figure 5.10. The distances primarily depend on the ratio of the number of elements. Here, we state the relative H^1-difference to $\mathcal{Q}$ in selected situations. The estimates are shown for all combinations of meshes from $(\mathcal{B}_i)_{i=1,\dots,15}$ where this number varies between 3.3 and 4.2 (left) and 7.0 and 9.2 (right).

	$\mathcal{Q}$	$\mathcal{P}$	$\mathcal{I}$	$\mathcal{R}_{r=2}$	$\widehat{\mathcal{Q}}$	$\mathcal{R}_{r=1}$	$\mathcal{R}_{r=0}$	$\widetilde{\mathcal{R}}'$
$\mathcal{Q}$		0.15	0.32	0.32	0.48	0.56	0.58	0.67
$\mathcal{P}$	0.16		0.23	0.25	0.43	0.51	0.54	0.76
$\mathcal{I}$	0.40	0.26		0.24	0.41	0.46	0.52	1.03
$\mathcal{R}_{r=2}$	0.43	0.31	0.26		0.28	0.35	0.40	1.10
$\widehat{\mathcal{Q}}$	0.79	0.66	0.54	0.34		0.27	0.16	1.50
$\mathcal{R}_{r=1}$	1.00	0.85	0.67	0.47	0.30		0.31	1.80
$\mathcal{R}_{r=0}$	1.06	0.91	0.76	0.55	0.18	0.31		1.80
$\widetilde{\mathcal{R}}'$	0.44	0.46	0.54	0.54	0.60	0.66	0.65	

Table 5.6. Relative distances with respect to $|\cdot|_{H^1}$ in the situation $(\mathcal{B}_5, \mathcal{B}_{10})$. The value in a cell is relative to the operator specified by the row.

moderately close to each other as we show shortly, have a very similar (almost identical) distance to $\mathcal{Q}$. These two concepts are the next closest to the orthogonal projection; they are roughly twice as far away from $\mathcal{Q}$ as the pseudo-L^2-projection is. The other operators are considerably further away.

Another important point is that the ratio between fine and coarse mesh size is most relevant for the considered distances but not the mesh size itself. This is illustrated in Figure 5.10. Here, we have collected all cases with a roughly comparable ratio of the numbers of elements, namely between 3.3 and 4.2 on the left and between 7.0 and 9.2 on the right. This classification is somewhat arbitrary; however, the figure clearly shows that the approximate differences to $\mathcal{Q}$ are almost constant over all the considered situations. Note that the charts have the same scale on the vertical axes as before.

To highlight the interconnections, let us give the complete data, namely the mutual relative distances between the operators with respect to $|\cdot|_{H^1}$ for one typical setting. The results for the mappings generated between the spaces associated with $\mathcal{B}_5$ and $\mathcal{B}_{10}$ are given in Table 5.6, ordered by their proximity to the orthogonal projection. In each cell, we state

	$\mathcal{Q}$	$\mathcal{P}$	$\mathcal{I}$	$\mathcal{R}_{r=2}$	$\widehat{\mathcal{Q}}$	$\mathcal{R}_{r=1}$	$\mathcal{R}_{r=0}$	$\widetilde{\mathcal{R}}'$
$\mathcal{Q}$		0.08	0.22	0.22	0.34	0.41	0.41	0.98
$\mathcal{P}$	0.09		0.16	0.17	0.29	0.36	0.37	1.03
$\mathcal{I}$	0.25	0.18		0.13	0.23	0.28	0.30	1.24
$\mathcal{R}_{r=2}$	0.26	0.19	0.13		0.17	0.23	0.25	1.28
$\widehat{\mathcal{Q}}$	0.44	0.37	0.26	0.18		0.15	0.09	1.50
$\mathcal{R}_{r=1}$	0.57	0.49	0.34	0.27	0.16		0.14	1.69
$\mathcal{R}_{r=0}$	0.56	0.49	0.36	0.29	0.10	0.14		1.65
$\widetilde{\mathcal{R}}'$	0.51	0.52	0.56	0.56	0.60	0.63	0.62	

Table 5.7. Relative distances with respect to $\|\cdot\|_{L^2}$ in the situation $(\mathcal{B}_5, \mathcal{B}_{10})$.

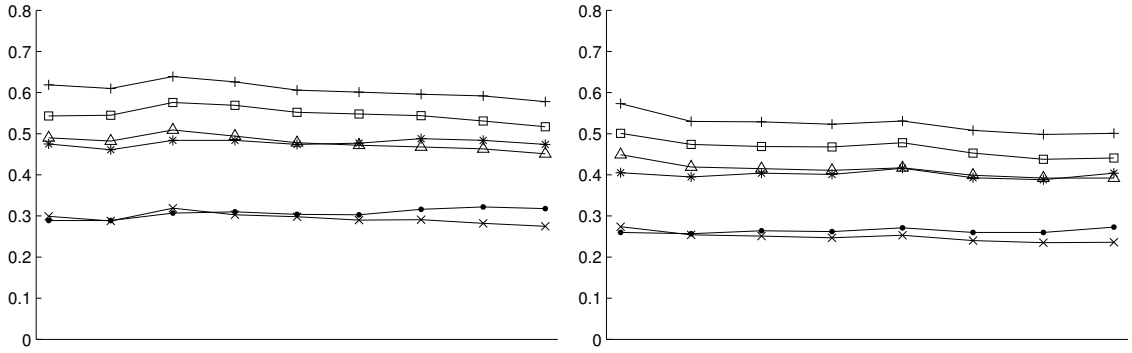


Figure 5.11. Relative H^1-distances to $\mathcal{I}$ in the situations selected in Figure 5.10.

the relative difference of the two specified operators with respect to the one in the current row. Therefore, the upper right half is the more interesting part. For completeness, the corresponding quantities with respect to $\|\cdot\|_{L^2}$ are given in Table 5.7. The tables belong to the setting which is fourth in the "columns" of the right part of Figure 5.10. Please be assured that this is an absolutely typical example and indeed representative.

Concerning the relative differences with respect to the single operators, we do not provide a complex chart like the one in Figure 5.9 for every type of transfer. Let us instead offer a condensed view and comment on the distances to the operators we consider next important, namely the nodal interpolation $\mathcal{I}$ and the pseudo-L^2-projection $\mathcal{P}$. In the fashion of Figure 5.10, we choose all cases with relatively similar element ratios. Again, this results in only slightly varying distances in the single charts; see Figure 5.11 and Figure 5.12. We recognize that the operators $\mathcal{R}_{r=2}$ and $\mathcal{P}$ are closest to $\mathcal{I}$. For the projection $\mathcal{P}$, the closest operator is $\mathcal{I}$ followed by $\mathcal{Q}$ and $\mathcal{R}_{r=2}$. We summarize the overall state at the end of this section.

The operator $\widetilde{\mathcal{R}}'$ has not appeared in the figures so far. As this transfer concept does not preserve the constant functions, it is different in some respects. For instance, the distance to the L^2-projection increases with decreasing mesh size ratio, which is contrary to the other operators. To further illustrate this disagreement, we display a diagram, which would be a detail of Figure 5.9, but now with the absolute values of the distances. Figure 5.13

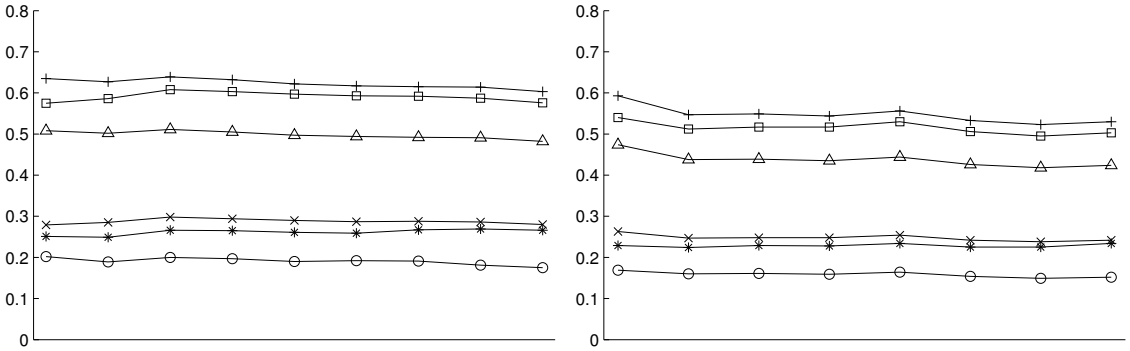


Figure 5.12. Relative H^1-distances to $\mathcal{P}$ in the situations selected in Figure 5.10.

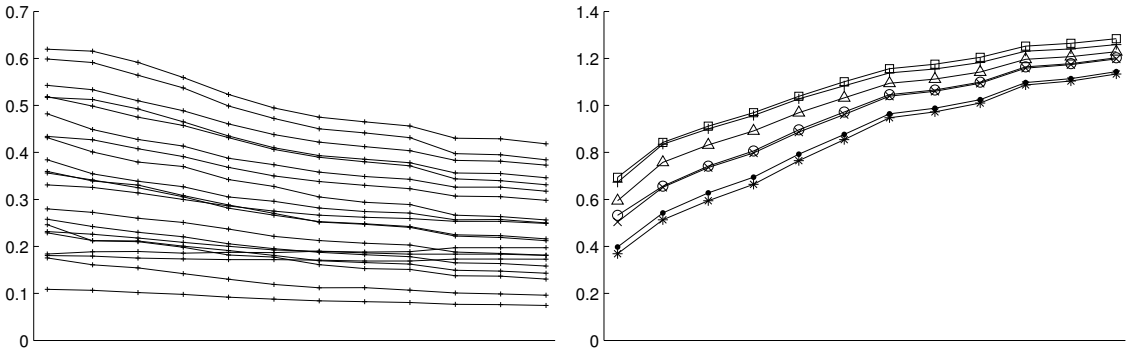


Figure 5.13. Absolute distances in the situations $(\mathcal{B}_3, \mathcal{B}_4), \ldots, (\mathcal{B}_3, \mathcal{B}_{15})$, in both parts ordered from left to right. The left chart shows the mutual differences between all operators but $\widetilde{\mathcal{R}}'$. The remaining differences, namely the ones between $\widetilde{\mathcal{R}}'$ and all other operators, are depicted in the right chart. Note the different scales on the vertical axes.

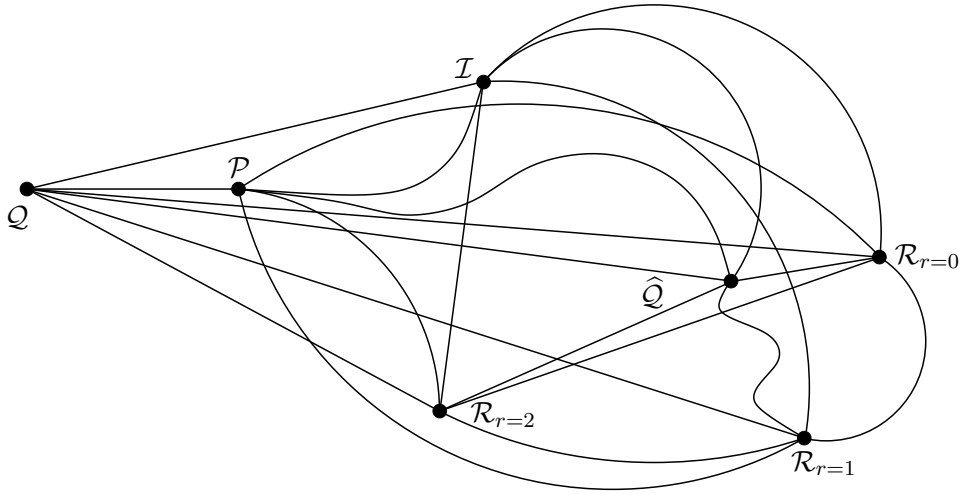


Figure 5.14. The mutual relations of the single operators visualized as a map-like graph. The length of each connecting line represent the absolute H^1-distance between the respective operators.

(left) shows the mutual differences in the situations involving $\mathcal{B}_3$ for the "well-behaved" mappings, i. e., all but $\widetilde{\mathcal{R}}'$. For the present purpose, we do not need to specify which of the 21 lines represents the difference between which two operators; these relations are visualized in a more convenient way at the end of this section. Let us just mention that the line at the bottom is the difference of $\widehat{\mathcal{Q}}$ and $\mathcal{R}_{r=0}$; the line at the top is the distance between $\mathcal{Q}$ and $\mathcal{R}_{r=0}$. Moreover, note that all differences are significant, namely clearly greater than zero, also for very small ratios $\frac{h_j}{h_i}$. This also holds for the corresponding relative quantities as the norms appear to be bounded; see Section 5.8.4. Now, Figure 5.13 (right) depicts the distances between $\widetilde{\mathcal{R}}'$ and the other operators. One notices that these quantities increase relatively fast with increasing j. In contrast, the differences in the left part decrease or increase only moderately.

Remark 5.22. *This illustrates that one needs to be careful in dealing with a seemingly natural transfer concept which the operator $\widetilde{\mathcal{R}}'$ certainly is. Both the derivation in Section 5.2.4 and especially the formula (5.11) look promising to approximate the L^2-orthogonal projection. However, in the context of the quasi-interpolation concepts reviewed in Section 5.2, the replacement of each polynomial trial space by the span of the associated single basis function is apparently not practicable. This is because the constant functions are left out of account. Likewise, in the derivation of transfer formulas of the type (5.11), (5.16), (5.18) and (5.25) or in the choice of discrete inner products such as (5.14) and (5.17), one cannot choose the scaling freely. For instance, the scaling by the diagonal entries of the mass matrix in (5.11), which seems reasonable at first glance, is not appropriate, whereas the scaling by its row totals in (5.18) is.*

Let us conclude this section with the promised overview. We visualize the interconnections between the transfer concepts by Figure 5.14, which, admittedly, a little more colorful

but without the annotations, could be mistaken for a piece of abstract art. In this map-like graph, the lengths of the lines represent the absolute distances of the connected operators with respect to $|\cdot|_{H^1}$. We pick a typical situation; here, the operators generated from $\mathcal{B}_3$ to $\mathcal{B}_7$ are considered. By the studies throughout this paragraph, we see that other situations or some averages yield essentially the same result as the sizes of the mutual distances are reasonably stable. The operator $\widetilde{\mathcal{R}}'$ is not in the picture as it is almost equally far away from all the depicted "well-behaved" ones.

5.9 Fine-to-coarse transfer of primal information

Throughout this chapter, we have investigated the information transfer between non-nested finite element spaces. So far, the aim of the considerations was to find stable and efficient approximation operators for the coarse-to-fine transfer. In the context of the semi-geometric multigrid methods developed in Chapter 3, the respective operators are employed to prolongate coarse level corrections to finer levels. More precisely, they serve as a means to create a nested space hierarchy with sufficient approximation properties. Let us consider a different point of view in this section.

A very special case of information transfer between non-nested finite element spaces appears in the following. Let

$$X := X_L \supset \ldots \supset X_\ell \supset \ldots \supset X_0$$

be a sequence of nested finite element spaces as in Section 2.2 about standard geometric multigrid methods. In non-linear multigrid methods such as the full approximation scheme or full approximation storage algorithm, see [37], one needs an approximation of the current iterate at the coarser levels. The approximation may be realized by suitable fine to-coarse mappings

$$\Pi_\ell^{\ell-1} : X_\ell \to X_{\ell-1}.$$

This is in addition to finding a coarse level model or objective function; see [147] for an example in discrete non-linear optimization. We do not go into detail here.

Note that a coarse level approximation of poor quality does not only slow down the "convergence" at the coarser levels but also might degenerate the overall convergence. This is because the coarse level correction depends crucially on the initial coarse level iterate. In fact, the correction which is added at the finer level is some interpolated difference between the initial coarse level iterate and the final coarse level iterate.

The fine-to-coarse operator which is already available is the standard restriction operator which acts on the residuals by multiplication of the transposed interpolation matrix. Admittedly, this mapping, commonly called full weighting operator, is generally not suitable for this transfer of primal information. Certain scaled variants are reported to work well, though; see, e. g., [94]. However, the operator used in the overwhelming majority of papers is the standard nodal finite element interpolation. Again, this seems the most attractive choice from a computational point of view. We remark that the linear interpolation from fine to coarse between node-nested meshes is usually called injection as only the weights zero and one appear. Colleagues [101] have been advertising the usage of the L^2-projection or the lumped variant in this context.

The extensive numerical studies presented in Section 5.8 suggest that it would be interesting to investigate the application of some of the considered transfer concepts in this context. This goes beyond the scope of this thesis, though. Let us just comment on a practical aspect. For the computation of approximation operators from X_ℓ to $X_{\ell-1}$, one may exploit the special structure of the problem, namely the relation between the nested meshes and spaces. In common finite element codes, it is usually much more convenient to deal with quantities at the same level than at different levels.

To illustrate this, let $\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ be the matrix representation of the linear interpolation from $X_{\ell-1}$ to X_ℓ with respect to the standard nodal bases. As the spaces are nested, the entries of this matrix are entirely determined by the logical connections. Then, an integral of $v \in L^1(\Omega)$ with test function $\lambda_p^{\ell-1}$, $p \in \mathcal{N}_{\ell-1}$, may be computed as a weighted sum of the entries of the vector $\boldsymbol{v} := \big((v, \lambda_q^\ell)_{L^2(\Omega)}\big)_{q \in \mathcal{N}_\ell}$, i. e.,

$$(v, \lambda_p^{\ell-1})_{L^2(\Omega)} = \sum_{q \in \mathcal{N}_\ell} (\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell})_{qp}\, v_q.$$

For example, the fine-to-coarse coupling matrix $\boldsymbol{B}_\ell^{\ell-1} \in \mathbb{R}^{n_{\ell-1} \times n_\ell}$, which is the analogon of the coarse-to-fine matrix in (5.13) appearing in the formulas of the L^2-projection and the L^2-quasi-projections, reads as

$$\boldsymbol{B}_\ell^{\ell-1} = (\boldsymbol{\mathcal{I}}_{\ell-1}^{\ell})^T \boldsymbol{M}_\ell.$$

In fact, this is equivalent to an exact composite quadrature formula. The same applies to the computation of the local projections for the evaluation of the quasi-interpolation operators of Section 5.2.

6 Numerical results

In this chapter, we focus on the practical properties of the devised semi-geometric multilevel methods. We report on broad experiences with this general class of multigrid methods and preconditioners based on non-nested meshes. Because of the geometric nature of the construction of the coarse level spaces we consider it especially important to analyze its dependence on the interaction of the meshes. In particular, the robustness of the convergence behavior with respect to the choice of the coarse meshes is demonstrated by numerical experiments. We verify the quality of the constructed coarse level hierarchies and compare the convergence rates to the ones obtained with a geometric multigrid method in case of a simple geometry. Here, only the practically most relevant case $d = 3$ is considered.

In Section 6.1, the flexible applicability of the developed semi-geometric framework and the performance of the multilevel algorithms are demonstrated. In Section 6.2, we introduce a semi-geometric monotone multigrid method for the solution of variational inequalities. To this end, we explain the required extensions to the linear method and present numerical examples illustrating the performance in case of Signorini's problem.

6.1 Semi-geometric multilevel methods

This section provides evidence that the semi-geometric concept is a flexible technique for the solution of problems with unstructured meshes. We report on a number of experiments carried out to assess the performance of the multilevel methods in a variety of respects. In Section 6.1.1, fundamental experiments on the asymptotic convergence behavior of the presented multigrid methods are described. Section 6.1.2 illustrates the flexibility of the approach by studying the dependence on the choice of the coarse meshes. In Section 6.1.3, we present experiments on the additive algorithms including an analysis of the two different setup variants. Section 6.1.4 is devoted to studying the almost nested case, which allows for a comparison of the convergence rates with the ones obtained by geometric multigrid methods. In Section 6.1.5, we apply different transfer concepts in the semi-geometric framework. Finally, Section 6.1.6 prepares the discussion of the non-linear methods to be considered the next section by a short description of an example from linear elasticity.

6.1.1 Asymptotic semi-geometric multigrid convergence

We start with an example to demonstrate the asymptotic convergence behavior of the semi-geometric multigrid method. As usual for the investigation of linear solvers, we first measure the effective convergence rates numerically on the basis of a setting with trivial solution $u_L = 0$. For this purpose, we consider the Poisson problem with zero Dirichlet boundary values on $\partial\Omega$ and zero right hand side. In this case, we can compute the energy norm of the algebraic error $\|e_L^k\|_A = |e_L^k|_{H^1(\Omega)} = |u_L^k|_{H^1(\Omega)}$ and the corresponding contraction rates ρ^k, $k \geq 1$. Finite element functions generated by pseudo-random data in the interval $[-1, +1]$ are employed as initial guesses u_L^0. This is a reasonable setting to estimate the asymptotic convergence behavior if the assumption holds that these initial iterates are

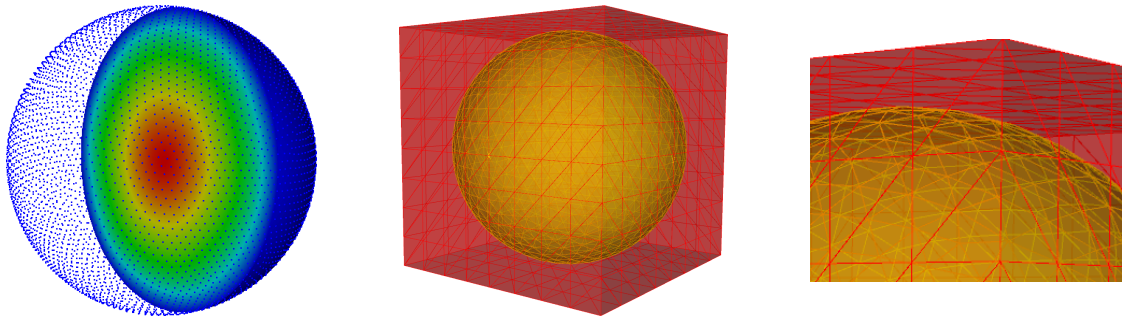


Figure 6.1. Sketch of the solution of the model problem (left); one of the coarse meshes used later, depicted in relation to the fine mesh (center) with zoom in on a corner of the cube (right). The meshes are completely unrelated.

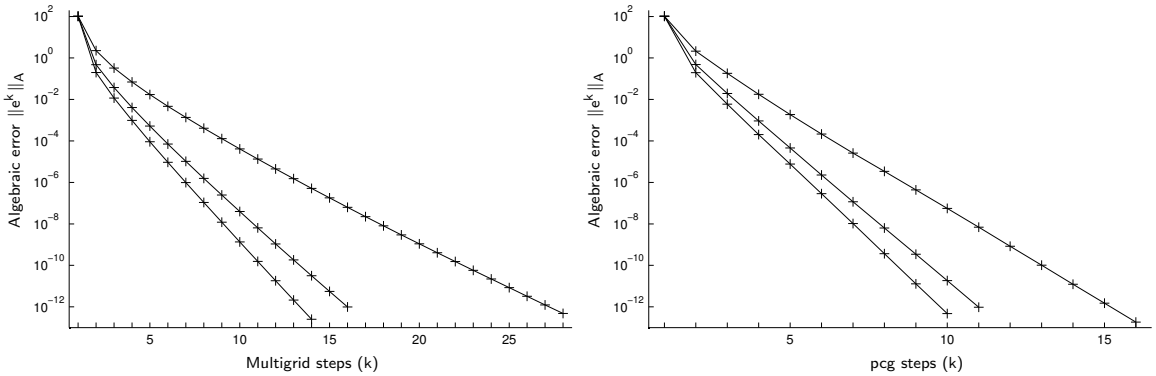


Figure 6.2. Error reduction during the semi-geometric multigrid iteration (left) and the pcg iteration (right) for a problem with 244,932 degrees of freedom. In both charts, the three lines represent the $\mathcal{V}(1,1)$-, the $\mathcal{V}(2,2)$- and the $\mathcal{V}(3,3)$-cycle (top down). Note the different scales on the horizontal axes.

sufficiently general in the sense that they contain contributions from the different parts of the spectrum of the operator A_L.

Recall from Section 3.6.1 that, in case of non-nested meshes with geometrically induced prolongation operators, one always needs to specify a parameter $\varepsilon_{\text{tr}} > 0$. The experiments in this paragraph employ $\varepsilon_{\text{tr}} = 0.20$; we discuss this parameter in more detail in Section 6.1.2, Section 6.1.4 and Section 6.1.5. Moreover, we choose $\Pi = \mathcal{I}$ for now.

For the first series of experiments, the domain $\Omega = \Omega_L$ may be the unit ball. As before for the studies in Section 5.8, this geometry is chosen because, for one thing, it yields good reproducibility and, for another thing, it allows for the generation of completely independent meshes for a large variety of different mesh sizes. Unstructured meshes approximating the unit ball have already been depicted in Figure 5.6. The solution of the Poisson problem with constant right hand side not equal to zero is illustrated in Figure 6.1 (left).

Let us now consider the convergence behavior of the semi-geometric multigrid method (Algorithm 3.4) with appropriately chosen coarse meshes. The typical error reduction

#elements	#dof	$\mathcal{C}_{\text{gr}}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(1,1)}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}_{\mathcal{V}(3,3)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(1,1)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(3,3)}$
1,380	325	1.17	1.26	0.047	0.012	0.005	0.016	0.004	0.001
4,111	882	1.17	1.31	0.072	0.018	0.008	0.024	0.007	0.003
13,383	2,678	1.18	1.37	0.092	0.028	0.017	0.030	0.010	0.006
48,320	9,228	1.15	1.35	0.108	0.037	0.025	0.038	0.014	0.008
93,620	17,647	1.15	1.34	0.137	0.048	0.030	0.050	0.018	0.011
181,789	33,067	1.15	1.36	0.162	0.052	0.031	0.057	0.021	0.012
361,907	64,833	1.15	1.38	0.202	0.060	0.035	0.071	0.024	0.014
719,951	127,787	1.14	1.35	0.189	0.071	0.046	0.068	0.028	0.017
1,017,911	179,831	1.11	1.27	0.229	0.091	0.062	0.087	0.033	0.021
1,390,421	244,932	1.10	1.24	0.272	0.094	0.061	0.096	0.036	0.021

Table 6.1. Convergence rates of the semi-geometric multigrid method and the corresponding pcg for a broad range of problem sizes.

during a multigrid or pcg iteration is illustrated in Figure 6.2. However, in the following, we are more interested in the average or the asymptotic convergence rate than in the actual convergence history.

Important characteristic quantities of a constructed multilevel hierarchy are the complexities $\mathcal{C}_{\text{gr}}$ and $\mathcal{C}_{\text{op}}$ defined in (3.24) as they reflect both the memory consumption of the structure and the execution cost of each cycle. Moreover, the average convergence rate of the multigrid iteration with a semi-geometric $\mathcal{V}(\nu_1, \nu_2)$-cycle is denoted by $\bar{\rho}_{\mathcal{V}(\nu_1,\nu_2)}$, the one of the respective pcg iteration by $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(\nu_1,\nu_2)}$. Here and in the following, rates and complexities are rounded to three and two digits, respectively.

Table 6.1 demonstrates the characteristics of the semi-geometric method for several different problem sizes. Note that the table contains more data, namely more problems, than usually considered in studies about geometric multigrid methods. This is to make perfectly clear that we do not just examine refinements of certain meshes but rather consider problems which are associated with independently generated unstructured meshes. The factor of the number of elements from the largest to the smallest problem is around $8^{\frac{10}{3}}$; thus, the range of problem sizes corresponds to more than three times uniform refinement ($d = 3$).

For this study, we pursue the following strategy to generate the coarse mesh hierarchy. First of all, a simple enclosing geometry is used. In the present case, we are free to choose the unit ball itself or a scaled version. Then, for the problem sizes listed in the table, the maximum level L and the respective coarsest meshes $\mathcal{T}_0$ are chosen such that the number of elements $|\mathcal{T}_0|$ is in the order of 10^{-L} to 8^{-L} times the number of elements of the given fine mesh. Finally, the meshes $(\mathcal{T}_\ell)_{1,\dots,L-1}$ are generated by regular refinement from $\mathcal{T}_0$. This means that the coarse meshes $(\mathcal{T}_\ell)_{\ell<L}$ are nested, but $\mathcal{T}_L$ is not. Consequently, we have $X_0 \subset \ldots \subset X_{L-1} \not\subset X_L$. The more general case is considered later.

For example, for the problem with 1,017,911 elements we have $L = 4$, $|\mathcal{T}_0| = 196$ and $n_0 = 64$; three uniform refinement steps have been carried out to obtain $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$ from $\mathcal{T}_0$. The given unstructured fine mesh is referred to as $\mathcal{T}_L = \mathcal{T}_4$. The other problems are treated similarly with slightly varying coarse mesh sizes; the number of coarse levels L increases

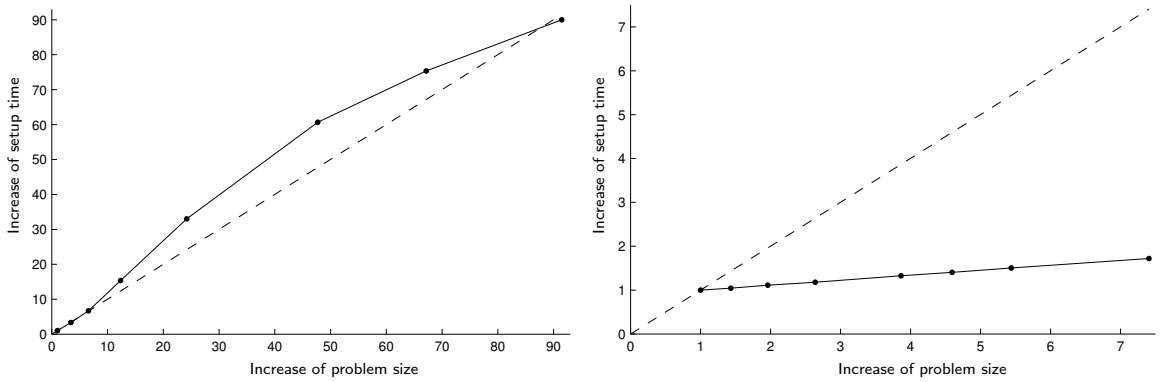


Figure 6.3. Increase of the total setup time of the semi-geometric multilevel hierarchy plotted versus the increase of the problem size (degrees of freedom). The dashed lines represent the linear functions with slope one. The data corresponds to the problems of Table 6.1 (left) and of Table 6.5 A (right), respectively.

from one to four. In all cases, the coarse level problems associated with $\ell = 0$ can be solved efficiently by a direct method. As we usually achieve $n_0 < 100$, both the computation of an LU decomposition and the execution of the corresponding forward-backward substitution schemes are inexpensive.

The data in Table 6.1 indicates that we have indeed found a means to construct multilevel hierarchies such that, on the one hand, $\mathcal{C}_{\mathrm{gr}}$ and $\mathcal{C}_{\mathrm{op}}$ stay in a reasonable range and, on the other hand, the convergence rates increase only moderately. For comparison, in the fully nested geometric case, the complexity measures are around 1.15 as already estimated in Section 3.6.1. The obtained rates are in fact not far off the ones usually observed in geometric multigrid methods. For instance, to reduce the error by a factor of at least 10^{-10}, in case of the largest problem with 244,932 degrees of freedom one needs ten $\mathcal{V}(2,2)$-cycles as compared with six $\mathcal{V}(2,2)$-cycles in case of the smallest problem with 325 degrees of freedom. We study the relations between geometric and semi-geometric methods more closely in Section 6.1.4. For comparison, the convergence rate of the symmetric Gauß–Seidel method degenerates to 0.984 for the largest problem. A conjugate gradient iteration converges with a rate of 0.907. If preconditioned with the symmetric Gauß–Seidel method, the rate can only be improved to 0.758.

It is important to note that the setup time increases roughly linearly with the number of degrees of freedom, which is illustrated in Figure 6.3 (left). This is a consequence of the fact that all operations for the assembly of the matrices $(\mathbf{\Pi}_{\ell-1}^{\ell})_{\ell=1,\dots,L}$ are local and the matrices themselves have limited bandwidth, as discussed in detail in our remarks on implementation aspects in Section 3.6 and Section 5.7. We comment on the right part of Figure 6.3 in Section 6.1.2.

Next, we consider an example where also the coarse meshes are non-nested. Let us point out that this is a variant which is absolutely covered by the developed theory. The coarse mesh hierarchies are constructed following a similar heuristic as before; we choose $L > 0$ appropriately and use independently generated meshes of the unit ball such that the ratio $\frac{|\mathcal{T}_{\ell-1}|}{|\mathcal{T}_{\ell}|}$, $\ell \in \{1, \dots, L\}$, is in the order of eight to ten. In Table 6.2, one sees that the results are

#elements	#dof	$\mathcal{C}_{\mathrm{gr}}$	$\mathcal{C}_{\mathrm{op}}$	$\bar{\rho}_{\mathcal{V}(1,1)}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}_{\mathcal{V}(3,3)}$	$\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(1,1)}$	$\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(3,3)}$
719,951	127,787	1.12	1.32	0.264	0.093	0.049	0.085	0.033	0.017
1,017,911	179,831	1.10	1.26	0.287	0.110	0.070	0.108	0.042	0.024
1,390,421	244,932	1.10	1.24	0.366	0.143	0.098	0.126	0.052	0.036

Table 6.2. Convergence rates of the semi-geometric multigrid method and the corresponding pcg in case the coarse meshes are also non-nested.

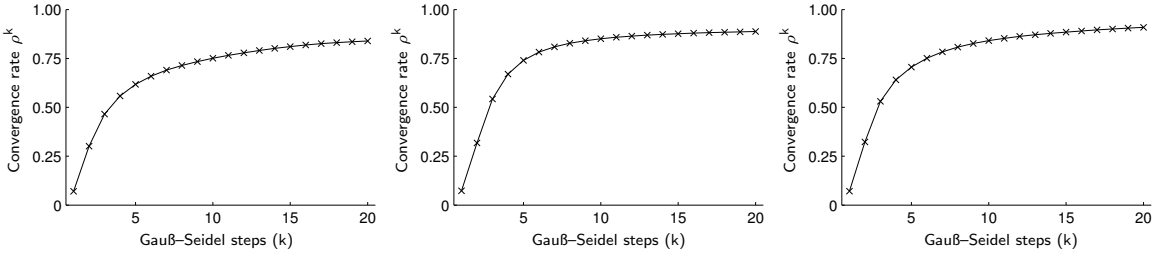


Figure 6.4. Convergence behavior of the symmetric Gauß–Seidel method in the constructed coarse spaces V_{L-1} for the homogeneous coarse problems associated with the fifth, seventh and ninth problem from Table 6.1 (from left to right). The reduction rates are rather small in the first few iteration steps but increase rapidly.

not as good as in the previous study although the convergence behavior is still acceptable for practical usage. Note that we obtain grid and operator complexities comparable to the ones before. We return to this issue when considering the additive variants in Section 6.1.3.

As a result, these first expriments show that the idea to employ finite element spaces associated with non-nested meshes to construct multigrid methods by the framework developed in Chapter 3 works very well. However, we also see that it does not seem to be beneficial to employ coarse level meshes which are completely independent of each other. The approach with $X_0 \not\subset \ldots \not\subset X_L$ may still yield an optimal method, which is also demonstrated in Section 3.4. But, in practice, better convergence rates with comparable or smaller complexities are obtained by choosing $X_0 \subset \ldots \subset X_{L-1} \not\subset X_L$.

Finally, we briefly focus on the smoothing properties of the standard linear iterative methods in the new spaces $(V_\ell)_{\ell<L}$. Above all, the excellent convergence of the presented multilevel methods is evidence for the adequacy of the employed smoothers. However, it is difficult to illustrate their effect similarly to the standard geometric case as in Section 2.1. The least we can do is examine the convergence behavior of the common smoothing iterations in the constructed coarse spaces. For this purpose, consider the homogeneous problem associated with the coarse level operator A_{L-1} in the non-standard space V_{L-1}. Then, for three different problems, Figure 6.4 clearly shows that some components of the error are reduced very fast in the first few steps, presumably the high-frequent contributions in the geometric sense. This is the exact same behavior as is well-known from the classical iterative methods applied to problems set in standard finite element spaces.

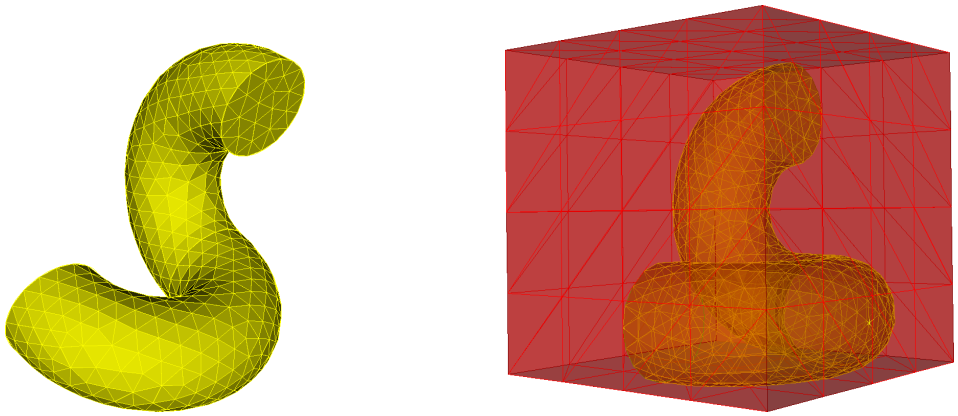


Figure 6.5. Snake-like geometry. The semi-geometric space hierarchy is generated by structured coarse meshes associated with an enclosing cube.

#elements	#dof	$\mathcal{C}_{\text{gr}}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$
44,780	8,893	1.11	1.19	0.053	0.018	0.069	0.026	0.068	0.025
126,224	24,102	1.12	1.24	0.063	0.024	0.067	0.025	0.066	0.023
263,122	49,182	1.12	1.27	0.062	0.021	0.066	0.024	0.068	0.025
405,195	75,035	1.13	1.30	0.068	0.027	0.089	0.029	0.085	0.030
904,880	165,351	1.10	1.28	0.158	0.056	0.285	0.087	0.270	0.087
1,397,664	254,069	1.07	1.19	0.230	0.075	0.360	0.109	0.327	0.096
				$\Gamma_D = \partial\Omega_L$		mixed (i)		mixed (ii)	

Table 6.3. Convergence rates of the semi-geometric multigrid methods for the snake-like geometry with different boundary conditions.

Complicated domain with mixed boundary conditions

Here, we consider the convergence of the semi-geometric multigrid methods for an example with a more complicated domain and address the dependence on the type of the boundary conditions. This also illustrates that the same or similar sets of coarse meshes, here the ones associated with a cube, may be employed to construct a suitable multilevel hierarchy for very different problems. For this purpose, let the computational domain $\Omega = \Omega_L$ be given by the snake-like geometry depicted in Figure 6.5. As before, we measure the mean convergence rates and the complexities for different problem sizes; each treated fine mesh has been generated independently.

Table 6.3 contains the results of the convergence study. Regarding the influence of the boundary conditions, we consider three configurations: pure Dirichlet boundary, i. e., the case $\Gamma_D = \partial\Omega$, and two different settings with mixed boundary conditions, namely first (i), homogeneous Neumann boundary conditions at the circular tails and Dirichlet conditions elsewhere and second (ii), vice versa.

The convergence rates are a little larger than for the example with the unit ball. It is fair to say that, in practice, similar increases in the convergence rates are generally observed for most iterative methods when turning from a rather simple geometry to a more complicated one, although it is usually difficult to quantify this effect, also for the analysis. However, the characteristic behavior of the multilevel methods, namely moderately increasing rates with increasing problem sizes is achieved for the more complicated problems, too. Still, for the largest problem with 254,069 degrees of freedom, one needs nine to eleven pcg steps preconditioned by the semi-geometric $\mathcal{V}(2,2)$-cycle to reduce the error by a factor of at least 10^{-10}, depending on the boundary conditions.

6.1.2 Flexible choice of coarse meshes

In this section, the influence of the choice of the coarse meshes is studied in more detail. First, we make sure that the setting of the previous paragraph is sufficiently general by showing that the results do not depend on the exact positions of the employed meshes. Second, we examine the behavior of the semi-geometric method in case the same set of coarse meshes is employed to construct multilevel spaces for different problems. This study shows that, in practice, it is not necessary to put too much effort into finding the most appropriate set of coarse meshes; very good convergence may be obtained even if the coarse meshes "fit only roughly".

As indicated before, we feel obliged to demonstrate that all above described qualitative and quantitative behaviors are not due to any special relations between the meshes. For this purpose, we consider several rotated settings. Then, Figure 6.6 depicts the measured convergence rates $\bar{\rho}$ for three of the problems from Table 6.1 in case the respective fine mesh is rotated about the axis spanned by the sum of the standard basis vectors $(\boldsymbol{e}_i)_{1\leq i\leq 3}$ by different angles. The obtained rates are almost constant over the entire range of studied configurations; the complexity measures (not illustrated here) do not change either. The influence of small changes to the coarse meshes on the multilevel hierarchy is examined in more detail in Section 6.1.4.

Let us now turn to the issue of robustness of the convergence behavior with respect to the

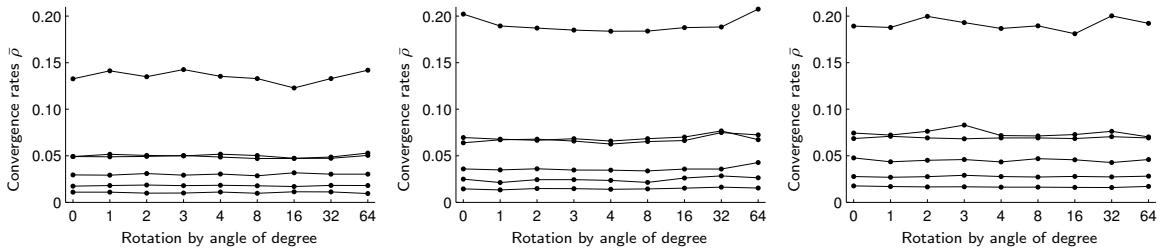


Figure 6.6. The measured convergence rates are virtually independent of rotations of the meshes. Each line represents one of the quantities $\bar{\rho}$ for the $\mathcal{V}$-cycles and the pcg iterations, respectively, plotted versus the angle of rotation. The three diagrams are associated with the fifth, seventh and eighth problem from Table 6.1 (from left to right).

		#elements			#nodes		
A	ball	744	5,952	47,616	184	1,210	8,767
B	cube	768	6,144	49,152	189	1,241	9,009
C	cube	1,729	13,832	110,656	421	2,798	20,339

Table 6.4. Overview of the characteristics of the meshes employed in the convergence study with fixed coarse meshes reported on in Table 6.5.

choice of the coarse meshes. In Section 6.1.1, we have seen that the multigrid convergence rates increase only moderately in case the coarse meshes are chosen by an appropriate heuristic. The power of the semi-geometric idea results from the advantage that the same set of coarse meshes $(\mathcal{T}_\ell)_{\ell<L}$ may be applied successfully to solve very different problems. We demonstrate this by the following study.

To investigate the influence of the choice of the coarse meshes, three different sets $(\mathcal{T}_\ell)_{\ell=0,\dots,L-1}$ are considered; the examples are labeled A, B and C. For this purpose, we choose an enclosing ball or cube and each time generate a simple coarsest mesh $\mathcal{T}_0$. Then, two uniform refinement steps are performed resulting in a total of four levels ($L = 3$) for each of the given fine problems to be considered. The characteristic data of the coarse meshes are given in Table 6.4. Note that the numbers are very similar for the examples A and B. However, the first one is an unstructured mesh of a ball whereas the second one is a structured mesh of a cube; cf. Figure 6.1. The meshes of example C are also associated with a cube but a little finer and unstructured.

We report on the convergence of the multigrid method and the respective complexity measures for problems of different sizes ranging from 33,067 to 244,932 degrees of freedom in Table 6.5. For each of the coarse mesh hierarchies specified before (A, B and C), the table contains the measured convergence rates $\bar{\rho}_{\mathcal{V}(2,2)}$ and $\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(2,2)}$, now for $\varepsilon_{\mathrm{tr}} = 0.20$ and $\varepsilon_{\mathrm{tr}} = 0.05$. Note that the behavior for these two choices of the parameter $\varepsilon_{\mathrm{tr}}$ is essentially the same although mostly the rates are significantly smaller for the more accurate but more expensive information transfer with $\varepsilon_{\mathrm{tr}} = 0.05$, as expected.

First, within each example, i. e., in each part labeled A, B and C, it is evident that

	#elements	#dof	$\mathcal{C}_{\text{gr}}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\mathcal{C}_{\text{gr}}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$
A	181,789	33,067	1.28	1.74	0.031	0.014	1.29	1.95	0.020	0.008
	262,365	47,348	1.20	1.52	0.047	0.019	1.21	1.68	0.032	0.012
	361,907	64,833	1.15	1.38	0.066	0.024	1.15	1.50	0.047	0.018
	490,617	87,244	1.11	1.28	0.086	0.031	1.11	1.37	0.066	0.023
	719,951	127,787	1.08	1.19	0.131	0.046	1.08	1.26	0.098	0.036
	858,429	151,930	1.07	1.16	0.154	0.050	1.07	1.22	0.121	0.045
	1,017,911	179,831	1.06	1.13	0.180	0.060	1.06	1.18	0.144	0.050
	1,390,421	244,932	1.04	1.09	0.222	0.075	1.04	1.13	0.181	0.068
B	181,789	33,067	1.20	1.53	0.046	0.016	1.21	1.68	0.028	0.010
	262,365	47,348	1.14	1.36	0.064	0.021	1.15	1.47	0.042	0.015
	361,907	64,833	1.11	1.26	0.087	0.030	1.11	1.35	0.064	0.023
	490,617	87,244	1.08	1.19	0.109	0.037	1.08	1.26	0.087	0.030
	719,951	127,787	1.05	1.13	0.155	0.052	1.06	1.18	0.131	0.044
	858,429	151,930	1.05	1.11	0.182	0.061	1.05	1.15	0.156	0.051
	1,017,911	179,831	1.04	1.09	0.205	0.069	1.04	1.13	0.180	0.062
	1,390,421	244,932	1.03	1.06	0.265	0.091	1.03	1.09	0.228	0.082
C	181,789	33,067	1.35	1.94	0.026	0.011	1.36	2.19	0.013	0.006
	262,365	47,348	1.25	1.66	0.043	0.014	1.26	1.84	0.025	0.009
	361,907	64,833	1.19	1.48	0.050	0.019	1.19	1.62	0.037	0.015
	490,617	87,244	1.14	1.35	0.088	0.031	1.14	1.46	0.074	0.019
	719,951	127,787	1.10	1.24	0.092	0.035	1.10	1.32	0.075	0.028
	858,429	151,930	1.08	1.20	0.113	0.043	1.09	1.27	0.088	0.033
	1,017,911	179,831	1.07	1.17	0.146	0.048	1.07	1.23	0.112	0.041
	1,390,421	244,932	1.05	1.12	0.162	0.066	1.05	1.17	0.186	0.059
			$\varepsilon_{\text{tr}} = 0.20$				$\varepsilon_{\text{tr}} = 0.05$			

Table 6.5. Convergence rates of the semi-geometric multigrid method and the corresponding pcg in case the coarse meshes remain fixed within each of the parts A, B and C. The results are given for two choices of the parameter ε_{tr}.

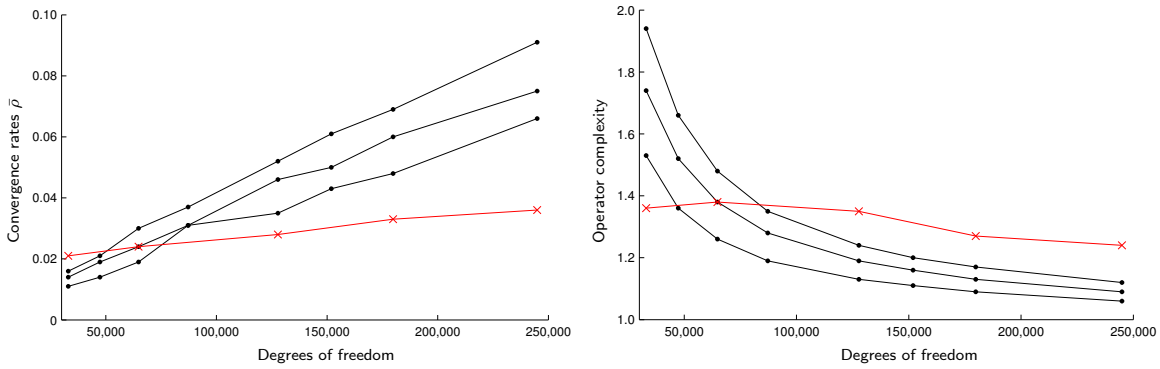


Figure 6.7. Comparison of the convergence behaviors: rate $\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(2,2)}$ (left) and complexity $\mathcal{C}_{\mathrm{op}}$ (right) plotted versus the number of degrees of freedom. The rather flat red lines marked with $\times$ represent the values from Table 6.1 where appropriate coarse meshes have been chosen for each single problem size. The other three lines illustrate the increase of the convergence rate and the decrease of the operator complexity in case the coarse meshes remain unchanged. The latter data is from the examples A, B and C from Table 6.5; the paramter $\varepsilon_{\mathrm{tr}}$ is always chosen as 0.20.

the convergence rates increase and the complexity measures decrease with increasing fine problem sizes. Second, one notes that the coarse meshes of example B yield space hierarchies with slightly worse convergence rates but at the same time better complexities. This is due to the fact that a number of nodes or elements of the coarse meshes lies far outside the actual computational domain Ω_L. As explained earlier, such nodes or elements do not contribute to the constructed coarse spaces $(V_\ell)_{\ell<L}$. The meshes of example C are finer, which results in smaller convergence rates over the entire range of considered fine problems with slighly larger complexities.

Note that the rates for the largest problems can obviously not compete with the ones in Table 6.1. Starting from very small values, they grow substantially faster because the coarse level hierarchy is not capable of representing the defect problem sufficiently accurately for increasing problem sizes. This is also reflected in the considerably smaller complexities $\mathcal{C}_{\mathrm{op}}$ for the larger problems. Both effects are illustrated in Figure 6.7. Moreover, the increase of the setup time is significantly slower than linear in case the coarse meshes remain unchanged; see Figure 6.3 (right). This is as expected and has two reasons. First, clearly less effort needs to be spent computing the coupling terms compared with the case of appropriately chosen (much finer) coarse meshes. Second, the coarse level spaces have relatively few degrees of freedom, which results in a faster assembly of the coarse level matrices.

All in all, the experiments show that a set of chosen coarse meshes is applicable for a respectable range of fine problem sizes. Consequently, one may choose the coarse mesh hierarchies in a rather flexible way. The fact that similar coarse meshes, occasionally even the same sets of meshes, are favorable for many fine meshes of different natures has further been illustrated by the more complicated geometry at the end of Section 6.1.1.

#elements	#dof	$\mathcal{C}_{\mathrm{op}}$	$\mathcal{C}_{\Pi}$	$\bar{\rho}_{\mathrm{std}}^{\mathrm{pcg}}$	$\bar{\rho}_{\mathrm{mod}}^{\mathrm{pcg}}$	$\mathcal{C}_{\mathrm{op}}$	$\mathcal{C}_{\Pi}$	$\bar{\rho}_{\mathrm{std}}^{\mathrm{pcg}}$	$\bar{\rho}_{\mathrm{mod}}^{\mathrm{pcg}}$
13,383	2,678	1.37	0.16	0.266	0.090	1.35	0.28	0.216	0.049
48,320	9,228	1.35	0.17	0.288	0.099	1.33	0.30	0.279	0.107
93,620	17,647	1.34	0.17	0.347	0.184	1.31	0.45	0.331	0.158
181,789	33,067	1.36	0.18	0.359	0.198	1.32	0.48	0.403	0.220
361,907	64,833	1.38	0.19	0.367	0.206	1.34	0.48	0.459	0.294
719,951	127,787	1.35	0.19	0.412	0.258	1.30	0.64	0.515	0.352
1,017,911	179,831	1.27	0.18	0.423	0.275	1.24	0.65	0.536	0.378
1,390,421	244,932	1.24	0.18	0.436	0.292	1.21	0.65	0.564	0.393
		setupSGMG				setupSGMG$_{\mathrm{imm}}$			

Table 6.6. Convergence rates of the pcg iteration with the additive semi-geometric multilevel preconditioners.

6.1.3 Additive semi-geometric preconditioners

This paragraph is concerned with the additive semi-geometric preconditioners introduced in Section 3.2.3. We study the relevant variants in the setting described at the beginning of this chapter (Section 6.1.1). Note that, as is often observed for this type of method, the quality of the additive preconditioner (Algorithm 3.5) does not depend significantly on whether the coarse problem at level $\ell = 0$ is solved exactly. It turns out that an inexact coarse level solve by the standard smoothing operator is sufficient.

We have seen that there are two equally straightforward ways for the construction of the coarse level hierarchy, namely the recursive procedure setupSGMG (Algorithm 3.3), which is also employed for the multigrid method, and the "immediate" variant setupSGMG$_{\mathrm{imm}}$ as described in Section 3.2.3. These two methods result in considerably different structures of the prolongation and restriction routines. We also study this in the experiments.

For this purpose, let us consider another complexity measure which is less common. Recall that n_ℓ is the number of degrees of freedom at level ℓ and n_ℓ^A is the number of non-zero entries in the matrix $\boldsymbol{A}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$, $\ell \in \{0, \dots, L\}$. In addition, we denote by n_ℓ^Π the number of non-zero entries in $\boldsymbol{\Pi}_{\ell-1}^{\ell}$ or in $\boldsymbol{\Pi}_{\ell-1}^{L}$, $\ell \in \{1, \dots, L\}$, depending on which matrices have been computed. Then, to quantify the memory requirements and also the cost of prolongation and restriction more accurately, we introduce the ratio

$$\mathcal{C}_\Pi := \frac{\sum_{\ell=1}^{L} n_\ell^\Pi}{n_L^A}.$$

For comparison, $\mathcal{C}_\Pi$ is around 0.04 for the medium-sized problem with completely nested meshes which is studied in the next paragraph (Section 6.1.4).

In Table 6.6, we consider the performance of the pcg iteration preconditioned by the (standard) additive semi-geometric method given by Algorithm 3.5 in case $X_0 \subset \ldots \subset X_{L-1} \not\subset X_L$. The results are given for both proposed variants for the setup of the operators $(\Pi_\ell^L)_{\ell=0,\dots,L}$. To ensure comparability of the data with a modified version, we report on the convergence rates $\bar{\rho}_{\mathrm{std}}^{\mathrm{pcg}}$ in case $\nu = 2$ smoothing steps are performed at each level.

#elements	#dof	$\mathcal{C}_{\text{op}}$	$\mathcal{C}_{\Pi}$	$\bar{\rho}_{\text{std}}^{\text{pcg}}$	$\bar{\rho}_{\text{mod}}^{\text{pcg}}$	$\mathcal{C}_{\text{op}}$	$\mathcal{C}_{\Pi}$	$\bar{\rho}_{\text{std}}^{\text{pcg}}$	$\bar{\rho}_{\text{mod}}^{\text{pcg}}$
719,951	127,787	1.32	0.19	0.354	0.180	1.28	0.33	0.487	0.321
1,017,911	179,831	1.26	0.19	0.400	0.252	1.22	0.48	0.518	0.344
1,390,421	244,932	1.24	0.19	0.416	0.265	1.21	0.48	0.538	0.369
		setupSGMG				setupSGMG$_{\text{imm}}$			

Table 6.7. Convergence rates of the pcg iteration with the additive semi-geometric multilevel preconditioners in case the coarse meshes are also non-nested.

The modified algorithm already mentioned in Section 3.2.3 is obtained by splitting the smoothing steps at level L and performing one before the restriction and one after the prolongation such that the numbers of operations are the same in the two versions. We denote the respective convergence rates by $\bar{\rho}_{\text{mod}}^{\text{pcg}}$. Table 6.7 contains the results for the examples with $X_0 \not\subset \ldots \not\subset X_L$.

The experiments show that the convergence rates of the pcg with the standard additive algorithm are significantly larger than the ones of the multiplicative versions. They may indeed be improved considerably by modifying the order of the operations at the finest level as described above. It is interesting to note that the rates in Table 6.7 with the completely non-nested coarse meshes are slightly better. In the analogous experiment for the multiplicative algorithms, this is not the case.

Let us now focus on the complexity measures of the constructed multilevel hierarchies. Evidently, if setupSGMG is employed, all complexities are the same as for the multiplicative methods. In this case, we see that $\mathcal{C}_{\Pi}$ is almost constant over the entire range of problem sizes. This is different for the hierarchies produced by setupSGMG$_{\text{imm}}$. Although $\mathcal{C}_{\text{op}}$ stays in a quite reasonable range (so does $\mathcal{C}_{\text{gr}}$ which is not given in the table), the quantity $\mathcal{C}_{\Pi}$ is considerably larger. In contrast to before, it grows with increasing problem size. This effect seems to be less severe for the second variant (Table 6.7) where also the coarse meshes are non-nested.

The examples of the behavior of the additive multilevel methods in this paragraph employ $\varepsilon_{\text{tr}} = 0.20$. Note that the observations only weakly depend on the choice of this parameter.

6.1.4 Studying the almost nested case

In this paragraph, we investigate the connection between the semi-geometric multigrid methods and the truly geometric variant more closely. This is done by considering a sufficiently simple geometry and treating the geometric multigrid method as a special case in a family of almost nested settings. By this experiment, we aspire to give a more precise idea of what kind of results can be achieved in the general unstructured case.

For this purpose, consider a hierarchy of four nested meshes of the unit cube where the coarsest mesh consists of 768 elements with 189 nodes. Throughout the study, we keep the finest mesh $\mathcal{T}_L = \mathcal{T}_3$ with 393,216 elements and 68,705 nodes fixed. In contrast, the coarse domains and the corresponding coarse meshes are scaled with a different factor

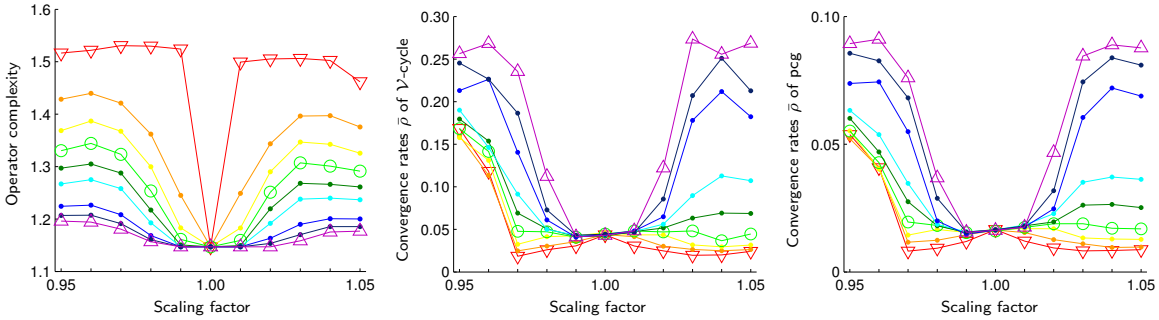


Figure 6.8. The complexity measure $\mathcal{C}_{\text{op}}$ (left) and the convergence rates $\bar{\rho}_{\mathcal{V}(2,2)}$ (center) and $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$ (right) of a semi-geometric multigrid method, plotted versus the scale of the coarse meshes. Each line represents a different parameter $\varepsilon_{\text{tr}} \in [0.01, 0.49]$. The marked lines correspond to the values 0.01 ($\triangledown$), 0.20 ($\circ$) and 0.49 ($\triangle$), respectively.

between 0.95 and 1.05 for each set of tests. We study the complexity of the constructed space hierarchy and the convergence of the semi-geometric multigrid method for a variety of values for the parameter ε_{tr} in $[0.01, 0.49]$. Note that, for linear finite elements associated with simplicial meshes, it does generally not make sense to choose ε_{tr} greater than or equal 0.5. This is because such a choice would result in deleting entries even in case of perfectly nested meshes, leaving nodes without direct coupling to the next coarser level.

The results are illustrated in Figure 6.8. Note that each single line represents either the complexity $\mathcal{C}_{\text{op}}$ or one of the convergence rates $\bar{\rho}_{\mathcal{V}(2,2)}$ and $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$ for a fixed parameter ε_{tr} plotted versus the scale of the coarse meshes. The lines corresponding to the extreme ε_{tr}-values 0.01 and 0.49 are marked by a downward and upward triangle, respectively; the intermediate value of 0.20 is marked by a circle. Table 6.8 contains the numbers for these three values. We stop with the scales 0.95 and 1.05, respectively. For smaller factors, the convergence rates further increase quite fast as less and less of the computational domain $\Omega = \Omega_L$ is covered by the coarse meshes; the complexity measures do not change much in this case. For larger factors, the convergence rates slowly increase whereas the complexity measures decrease. This is due to the fact that more and more elements of the coarse meshes lie completely outside the computational domain.

This series of experiments is well suited to explain several phenomena. As expected and observed in the vast majority of experiments, the convergence rates principally increase with increasing truncation parameter. Note that the deterioration of the convergence behavior is usually rather slow, though. It is evident that the semi-geometric methods, which leave the coarse meshes flexible, coincide with the standard geometric variants in the special case of nested meshes. In addition, an important observation is that both the complexities and the convergence rates of the geometric multigrid methods are retained in case the meshes are almost nested if a suitable parameter ε_{tr} is applied; see the discussion below. This also indicates that our construction is robust in the sense that the coarse level hierarchy (and with it the multigrid convergence) only varies slightly if the coarse meshes themselves change slightly. Perturbations of the meshes are irrelevant for the efficiency of the methods.

scale	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$
0.95	1.52	0.169	0.054	1.33	0.168	0.055	1.20	0.256	0.089
0.96	1.52	0.118	0.041	1.34	0.142	0.043	1.19	0.268	0.091
0.97	1.53	0.018	0.008	1.32	0.048	0.020	1.18	0.235	0.076
0.98	1.53	0.026	0.009	1.25	0.047	0.018	1.16	0.112	0.037
0.99	1.52	0.031	0.012	1.16	0.041	0.015	1.15	0.041	0.016
1.00	1.15	0.044	0.016	1.15	0.044	0.016	1.15	0.044	0.016
1.01	1.50	0.031	0.012	1.16	0.048	0.017	1.15	0.048	0.018
1.02	1.51	0.025	0.009	1.25	0.047	0.019	1.15	0.122	0.047
1.03	1.51	0.020	0.008	1.31	0.048	0.019	1.16	0.273	0.085
1.04	1.50	0.020	0.008	1.30	0.037	0.017	1.18	0.256	0.089
1.05	1.46	0.024	0.009	1.29	0.045	0.017	1.18	0.269	0.088
	$\varepsilon_{\text{tr}} = 0.01$			$\varepsilon_{\text{tr}} = 0.20$			$\varepsilon_{\text{tr}} = 0.49$		

Table 6.8. Studying the convergence behavior for a family of almost nested meshes associated with the unit cube. The middle row (scale 1.00) corresponds to the completely nested case in which the approach coincides with the standard geometric multigrid method.

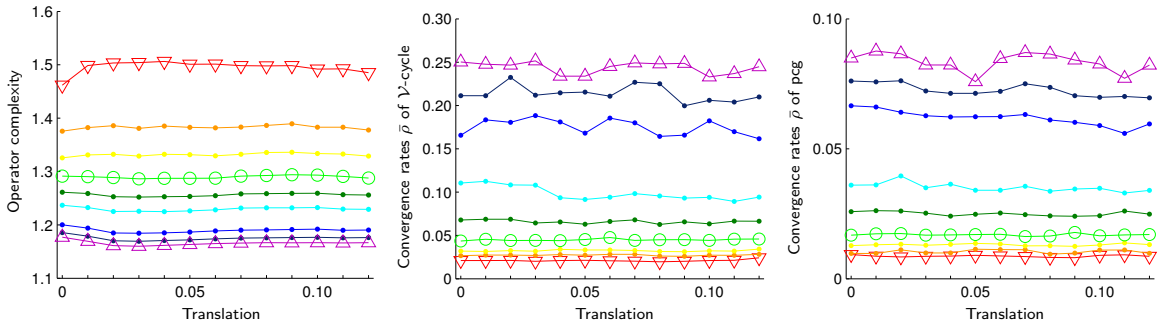


Figure 6.9. The complexity measure $\mathcal{C}_{\text{op}}$ (left) and the convergence rates $\bar{\rho}_{\mathcal{V}(2,2)}$ (center) and $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$ (right) of a semi-geometric multigrid method, plotted versus the size of the translation of the coarse meshes for different parameters ε_{tr}.

This can also be seen in another experiment illustrated in Figure 6.9. Here, we consider different translations of the coarse meshes associated with the cube of scale 1.05 in direction of the unit vector $(\frac{2}{3}, \frac{2}{3}, \frac{1}{3})^T \in \mathbb{R}^3$ by sizes up to 0.12. In this case, the computational domain is covered for almost the entire range of translations.

As a general rule, we observe the following effects. The larger the parameter ε_{tr} the less sensitive is the complexity $\mathcal{C}_{\text{op}}$ to changes of the coarse meshes. The smaller ε_{tr} the less sensitive are the convergence rates to changes of the coarse meshes. In our scaling examples, the convergence actually improves in case of small perturbations for sufficiently small ε_{tr}. This is of course accompanied by a rapid increase of $\mathcal{C}_{\text{op}}$. The choice $\varepsilon_{\text{tr}} = 0.20$ is a reasonable attempt to achieve the two competing goals. It manages to keep the convergence rates almost constant for a rather broad range of different problem sizes while leading to a moderate increase of $\mathcal{C}_{\text{op}}$; see Figure 6.8.

Note that the size of the problem studied here (68,705 degrees of freedom on a structured mesh) is roughly comparable to the seventh setting (64,833 degrees of freedom on an unstructured mesh) of the very first experiment in Section 6.1.1; see Table 6.1. This highlights the quality of the developed multigrid methods and preconditioners for unstructured meshes, respectively, as the measured semi-geometric convergence rates, $\bar{\rho}_{\mathcal{V}(2,2)} = 0.060$ and $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)} = 0.024$, are not much worse than the ones produced by the geometric method on the cube, $\bar{\rho}_{\mathcal{V}(2,2)} = 0.044$ and $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)} = 0.016$. However, for unstructured meshes without natural coarse level hierarchy, it seems impossible to achieve this convergence with an operator complexity as small as 1.15 which is easily obtained in the structured case.

6.1.5 Application of other transfer concepts

In this paragraph, we study several interesting phenomena regarding the application of different transfer concepts in the semi-geometric methods. So far, we have presented our convergence results employing the standard finite element interpolation $\mathcal{I}$ to construct the semi-geometric multilevel hierarchy from the non-nested sequence of meshes. The experiments in the following justify this choice as it turns out that the operator $\mathcal{I}$ is most reasonable in a rather general sense. Here, we compare the nodal interpolation and the other transfer concepts investigated in Chapter 5 apart from the L^2-projection, which is not practicable due to its global character, as discussed before. However, we have seen in Section 5.8.5 that some of the local operators approximate the L^2-orthogonal projection between finite element spaces associated with non-nested meshes quite well.

For the purposes of the present analysis, it is useful to introduce an over-relaxation parameter $\alpha \geq 1$. In the context of multilevel iterative methods, the term over-relaxation stands for a multiplication of the coarse level correction $\mathbf{\Pi}^{\ell}_{\ell-1}\boldsymbol{x}'$ in the resepective cycles at level ℓ (see Algorithm 3.4) by a number greater than one. On the one hand, it is generally reasonable to examine whether an over-relaxation of the coarse level correction can improve the convergence, especially in view of the affinity of our method to aggregation-based algebraic multigrid; see [20, 176]. On the other hand, the study of this effect allows us to identify some differences between the transfer concepts. The choice of α influences the convergence behaviors of the studied multigrid methods to different degrees. In particular, we consider the dependence on the truncation procedure, namely on the size of ε_{tr}.

#elements	ε_{tr}	$\mathcal{I}$	$\widehat{\mathcal{Q}}$	$\widetilde{\mathcal{R}}'$	$\mathcal{R}_{r=0}$	$\mathcal{R}_{r=1}$	$\mathcal{R}_{r=2}$	$\mathcal{P}$
93,620	0.01	1.48	1.83	1.83	1.98	1.97	1.90	2.01
	0.05	1.45	1.61	1.61	1.70	1.70	1.54	1.56
	0.10	1.41	1.51	1.51	1.57	1.58	1.45	1.41
	0.20	1.34	1.39	1.39	1.43	1.44	1.35	1.32
719,951	0.01	1.49	1.91	1.91	2.08	2.07	1.99	2.11
	0.05	1.46	1.66	1.66	1.77	1.76	1.57	1.60
	0.10	1.42	1.54	1.54	1.62	1.61	1.47	1.42
	0.20	1.35	1.40	1.40	1.45	1.45	1.35	1.32

Table 6.9. For two selected problems of different size from Figure 6.10, we give the measured operator complexities $\mathcal{C}_{\text{op}}$ depending on the parameter ε_{tr} and the choice of the transfer Π.

Recall that it has been demonstrated in Section 5.8.5 that the pseudo-L^2-projection is the closest to the L^2-projection. We are now interested in determining whether this is beneficial for the application in the semi-geometric methods. Note that the theoretical investigations in Chapter 5 still imply that most of the geometrically inspired transfer concepts, although in part very different from each other, should yield optimal multilevel methods.

We have conducted a large number of experiments; the most relevant results are summarized in two pages of diagrams (Figure 6.10 and Figure 6.11) which are explained in the following. The arrangement of the diagrams is designed to give an insight into the dependence of the convergence rates on the following variables:

- the type of the employed transfer operators, $\Pi \in \{\mathcal{I}, \widehat{\mathrm{Q}}, \widetilde{\mathcal{R}}', \mathcal{R}_{r=0}, \mathcal{R}_{r=1}, \mathcal{R}_{r=2}, \mathcal{P}\}$,
- the truncation parameter, $\varepsilon_{\text{tr}} \in \{0.01, 0.05, 0.10, 0.20\}$,
- the over-relaxation parameter, $\alpha \in [1.0, 1.4]$,
- the problem size, meshes from Table 6.1 and Table 6.5 (A), respectively.

The operator type is fixed in each of the seven rows whereas the size of the parameter ε_{tr} is fixed in each of the four columns. The charts show the convergence rates $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$ plotted versus the parameter α. Each line represents a different problem size; we have ten lines per diagram in Figure 6.10 and eight lines per diagram in Figure 6.11. The nine measuring points for the parameter α, which constitute the respective lines, are marked by dots. Note that all axes have the same scales.

We present our results for both variants of the convergence studies performed at the beginning of this section. More precisely, Figure 6.10 depicts the convergence rates in case appropriate coarse meshes are chosen as in Section 6.1.1. In Figure 6.11, we study the behaviors for the fixed coarse meshes from example A of Section 6.1.2. The case $\varepsilon_{\text{tr}} = 0.01$ is not of practical relevance as the constructed coarse level problems are usually too dense; see, e. g., Section 6.1.4 and the complexities in Table 6.9. We still give the measured data as it corresponds to the most accurate usage of the respective transfer concepts.

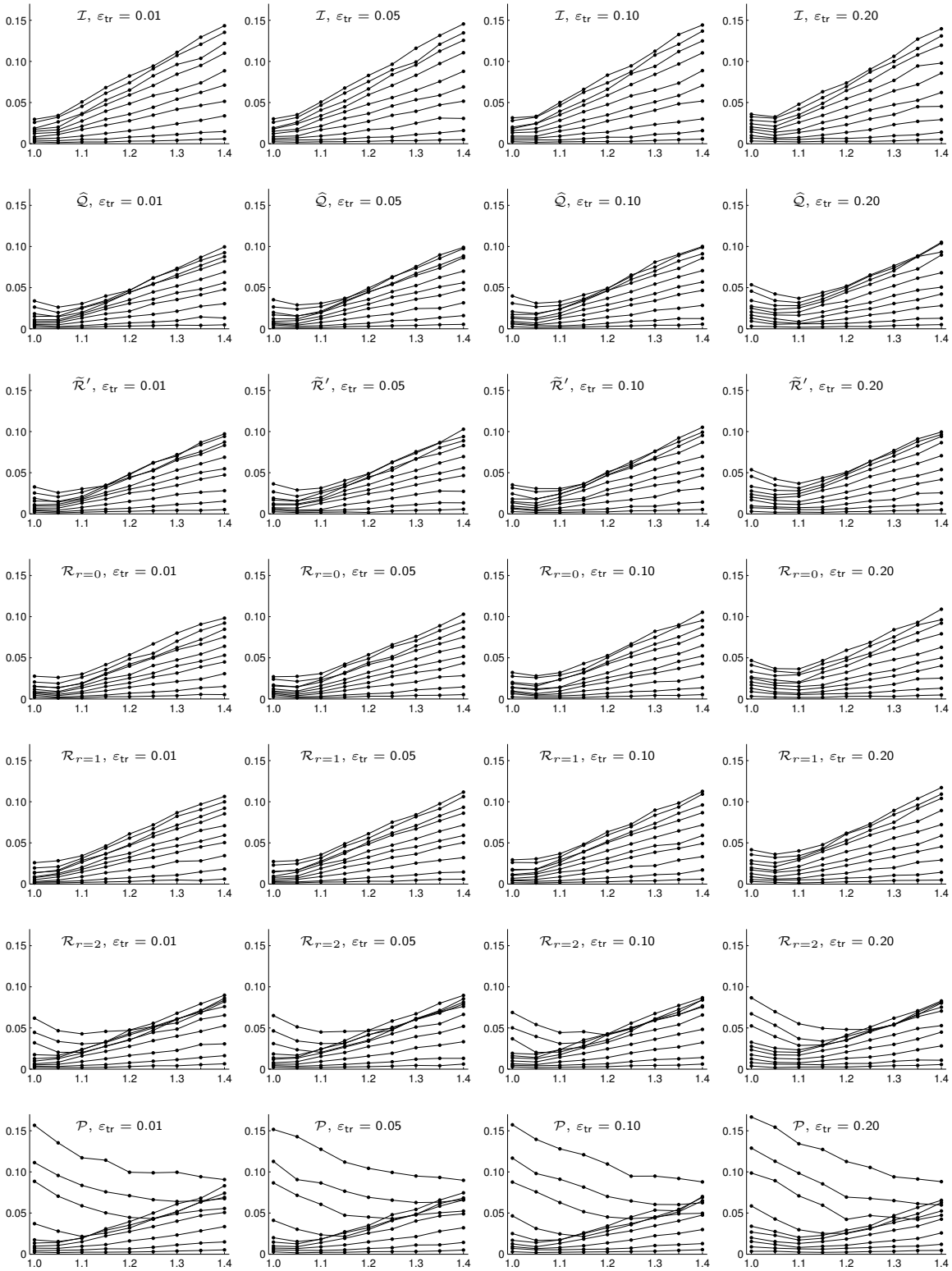


Figure 6.10. Convergence rates $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$ plotted versus the over-relaxation parameter $\alpha \in [1.0, 1.4]$. The operator type is fixed in each row; the level of truncation by the parameter ε_{tr} is fixed in each column. Each line represents a different problem size.

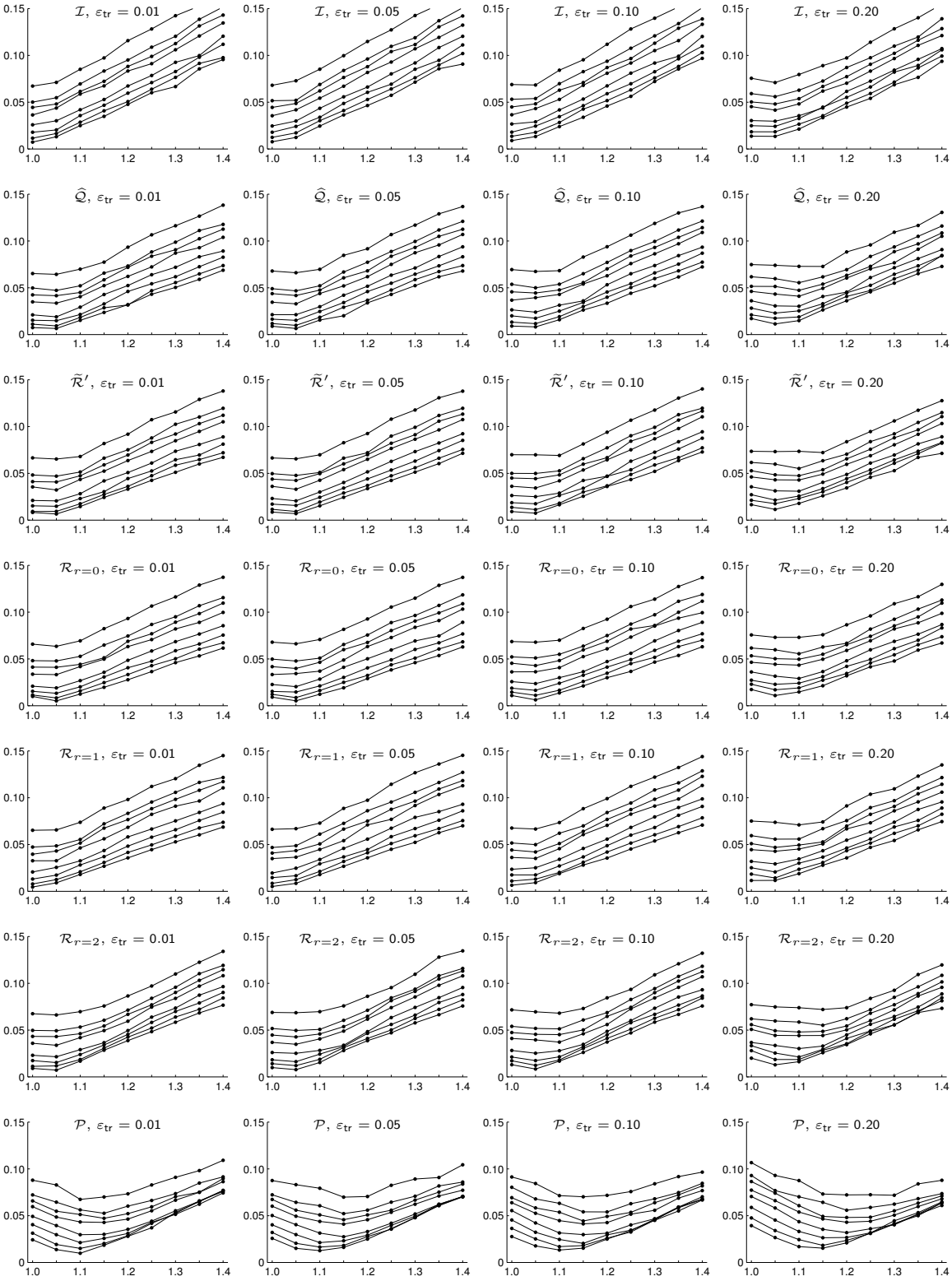


Figure 6.11. Convergence rates $\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$ plotted versus the over-relaxation parameter $\alpha \in [1.0, 1.4]$. This addresses the case of a fixed set of coarse meshes, which has initially been studied in Section 6.1.2. Each line represents a different problem size.

First, note that the convergence rates slighly worsen with increasing $\varepsilon_{\mathrm{tr}}$. This holds true for every operator; it can be seen in each row of both Figure 6.10 and Figure 6.11 as the curves are shifted upward a bit from left to right. We have studied this effect in detail for the nodal interpolation $\mathcal{I}$ in the previous examples. This confirms that a geometrically inspired transfer concept shows the following reasonable behavior: Neglecting (small) parts of the available information (slightly) affects the inherent approximation power and thus (slightly) worsens the convergence.

Second, recall that we have seen in Section 5.8.5 that the studied operators exhibit significant quantitative differences. However, the results when applied in the construction of the multigrid method are not very different, except for the operator $\mathcal{P}$ in the last row of Figure 6.10. For $\mathcal{P}$, the optimal relaxation parameter seems to depend highly on the problem size, whereas for the other operators it does not change much if at all. Beside $\mathcal{P}$ only $\mathcal{R}_{r=2}$ shows this tendency. The effect is less prominent for the example with fixed coarse meshes in Figure 6.11.

Third, we observe that over-relaxation is not necessary to obtain fast convergence in the semi-geometric approach; in all cases, the convergence rates increase moderately with the problem size and can be improved at most slightly, again with the exception of $\mathcal{P}$. The magnitude of the convergence rates confirms the applicability of the studied transfer concepts. This includes the operator $\widetilde{\mathcal{R}}'$ despite the findings of Section 5.8.5. The fact that over-relaxation is not suitable is in contrast to algebraic multigrid based on aggregation, where the idea to take advantage of this technique was motivated by the fact that the constructed coarse level (basis) functions exhibit regions in which they are very flat. In this case, a dramatic improvement could be achieved by over-relaxation with α as large as 1.8; see [24]. This observation suggests that the shapes of our coarse level basis functions are more similar to the standard ones. This conclusion is supported by the fact that the optimal over-relaxation parameter tends to depend on the level of truncation, which is certainly a procedure "increasing flatness".

We have no conclusive explanation yet why $\mathcal{P}$ behaves differently in this context. The reason might lie in the fact that this operator reflects a concentration of the local weighting towards the centers of the patches, compared to the other operators comprising a weighting. However, the nodal interpolation $\mathcal{I}$, which is in a sense the "most local operator possible", behaves similarly to all other transfer concepts.

6.1.6 Linear elastic problems

Before proceeding to the non-linear problems in the next section, let us briefly address the performance of the (linear) semi-geometric multigrid methods for systems of partial differential equations. Here, we perform two of the previous experiments once again in the different setting, namely for a linear elastic problem. We do this to keep the effects originating from the system of partial differential equations (compared with the scalar equation) and from the treatment of the contact conditions separate.

Let the constants in the linear constitutive law (1.9) be $E = 200\,\mathrm{GPa}$ and $\nu = 0.3$, i. e., $\lambda \approx 115.38\,\mathrm{GPa}$ and $\mu \approx 76.92\,\mathrm{GPa}$. The material described by these parameters may be considered a selected steel grade. Evidently, iterative solvers are robust with respect to

#elements	#dof	$\mathcal{C}_{\text{gr}}$	$\mathcal{C}_{\text{op}}$	$\bar{\rho}_{\mathcal{V}(1,1)}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}_{\mathcal{V}(3,3)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(1,1)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\text{pcg}}_{\mathcal{V}(3,3)}$
1,380	975	1.17	1.26	0.162	0.048	0.019	0.053	0.014	0.006
4,111	2,646	1.17	1.31	0.178	0.064	0.035	0.063	0.022	0.011
13,383	8,034	1.18	1.37	0.224	0.086	0.045	0.077	0.029	0.016
48,320	27,684	1.15	1.35	0.257	0.109	0.065	0.091	0.039	0.024
93,620	52,941	1.15	1.34	0.310	0.140	0.092	0.115	0.051	0.030
181,789	99,201	1.15	1.36	0.351	0.164	0.098	0.133	0.059	0.035
361,907	194,499	1.15	1.38	0.392	0.188	0.115	0.152	0.072	0.043
719,951	383,361	1.14	1.35	0.394	0.205	0.139	0.154	0.076	0.049
1,017,911	539,493	1.11	1.27	0.451	0.250	0.161	0.178	0.088	0.055
1,390,421	734,796	1.10	1.24	0.480	0.255	0.178	0.188	0.095	0.063

Table 6.10. Convergence rates of the semi-geometric multigrid method and the corresponding pcg applied to a linear elastic problem.

Young's modulus E. We do not study the behavior for the critical case $\nu \to 0.5$ here.

It is important to note that the cost of the setup relative to the number of degrees of freedom is much less here. This is because we treat the different physical unknowns separately, i. e., the scalar displacements in direction of the basis vectors $(\boldsymbol{e}_i)_{1\le i\le 3}$, such that the coarse level hierarchy is the same in each component. Recall the basis $\boldsymbol{\Lambda}_\ell = (\lambda_p^\ell \boldsymbol{e}_i)_{p\in\mathcal{N}_\ell, 1\le i\le 3}$ of the vector-valued space $\boldsymbol{X}_\ell := (X_\ell)^3$ defined in Section 1.3. Then, the (3×3)-block between the nodes $p \in \mathcal{N}_\ell$ and $q \in \mathcal{N}_{\ell-1}$ is

$$(\bar{\boldsymbol{\Pi}}^\ell_{\ell-1})_{pq} := \begin{pmatrix} (\boldsymbol{\Pi}^\ell_{\ell-1})_{pq} & 0 & 0 \\ 0 & (\boldsymbol{\Pi}^\ell_{\ell-1})_{pq} & 0 \\ 0 & 0 & (\boldsymbol{\Pi}^\ell_{\ell-1})_{pq} \end{pmatrix}, \tag{6.1}$$

for $\ell \in \{1,\dots,L\}$, where $(\boldsymbol{\Pi}^\ell_{\ell-1})_{pq} \in \mathbb{R}$ is the entry of the respective prolongation matrix in the scalar case. These blocks constitute the matrix representations $\bar{\boldsymbol{\Pi}}^\ell_{\ell-1} \in \mathbb{R}^{3n_\ell \times 3n_{\ell-1}}$, $\ell \in \{1,\dots,L\}$, of the prolongation operators. If the diagonal structure of the blocks in (6.1) is not changed, one may obviously save memory by storing only a single scalar quantity per pair of nodes. In addition, the bases of the semi-geometric spaces $(\boldsymbol{V}_\ell)_{\ell=0,\dots,L-1}$ read as $\widetilde{\boldsymbol{\Lambda}}_\ell = (\widetilde{\lambda}^\ell_p \boldsymbol{e}_i)_{p\in\mathcal{N}_\ell, 1\le i\le 3}$ for $\ell \in \{0,\dots,L-1\}$ in this case. Here, the functions $(\widetilde{\lambda}^\ell_p)_{p\in\mathcal{N}_\ell}$ are defined by the exact same formulas as in the scalar case, namely by (3.4) and (3.5). Therefore, the complexity measures are the same as for the scalar experiments; we state them again in the tables for clarity. We reconsider the block structure of the prolongation matrices in Section 6.2.1 in the context of the monotone multigrid methods.

First, we report on an example which uses the same meshes as described in Section 6.1.1. In the smoothing iterations, collective relaxation of the degrees of freedom associated with the same node is employed as indicated in Section 2.4. The results of the experiments stated in Table 6.10 show a moderate increase of the convergence rates with increasing problem size, comparable to the behavior of the method in the scalar case. Note that the largest problem has 734,796 degrees of freedom here.

To compare the performance with geometric multigrid methods, let us also consider

scale	$\mathcal{C}_{\mathrm{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(2,2)}$	$\mathcal{C}_{\mathrm{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(2,2)}$	$\mathcal{C}_{\mathrm{op}}$	$\bar{\rho}_{\mathcal{V}(2,2)}$	$\bar{\rho}^{\mathrm{pcg}}_{\mathcal{V}(2,2)}$
0.95	1.52	0.253	0.091	1.33	0.268	0.097	1.20	0.388	0.154
0.96	1.52	0.230	0.073	1.34	0.232	0.082	1.19	0.394	0.147
0.97	1.53	0.095	0.034	1.32	0.136	0.049	1.18	0.350	0.127
0.98	1.53	0.098	0.037	1.25	0.140	0.052	1.16	0.208	0.072
0.99	1.52	0.118	0.044	1.16	0.138	0.052	1.15	0.139	0.051
1.00	1.15	0.144	0.053	1.15	0.144	0.053	1.15	0.144	0.053
1.01	1.50	0.121	0.045	1.16	0.146	0.054	1.15	0.159	0.056
1.02	1.51	0.110	0.040	1.25	0.146	0.055	1.15	0.231	0.086
1.03	1.51	0.109	0.041	1.31	0.139	0.051	1.16	0.373	0.137
1.04	1.50	0.115	0.042	1.30	0.140	0.050	1.18	0.397	0.145
1.05	1.46	0.121	0.043	1.29	0.154	0.053	1.18	0.404	0.151
	$\varepsilon_{\mathrm{tr}} = 0.01$			$\varepsilon_{\mathrm{tr}} = 0.20$			$\varepsilon_{\mathrm{tr}} = 0.49$		

Table 6.11. Studying the almost nested case for the linear elastic setting by the experiment from Section 6.1.4. The qualitative behavior is the same as for the scalar problems. The convergence of the geometric multigrid method, namely the case of completely nested meshes, is presented in the middle row.

the almost nested case which has been studied for scalar problems in Section 6.1.4. The problem associated with the unit cube now has 206,115 degrees of freedom. The results of the analogous experiments are given in Table 6.11. Again, note that the complexities are the same as before. The corresponding sketch, namely the analogon of Figure 6.8 is very similar to before but with the curves shifted upward; we omit it here.

6.2 Semi-geometric monotone multigrid methods

It is worthwile to consider an extension of the semi-geometric multigrid methods to a nonlinear iteration for the solution of variational inequalities. In this section, we briefly discuss how to convert the linear method to a monotone multigrid method [121, 122, 124] and present numerical examples for Signorini's problem. For an overview of other methods for the solution of elastic contact problems, let us refer to [125] and the references therein as well as to the monographs [135, 192].

We emphasize that the multilevel finite element spaces employed in the semi-geometric monotone multigrid methods are non-standard in a double sense. The finite element functions of the given non-nested coarse spaces $(\boldsymbol{X}_\ell)_{\ell=0,\dots,L-1}$ are altered twice. First, the semi-geometric hierarchy is constructed by means of the framework developed in Chapter 3. Second, the coarse spaces are modified depending on the current fine level iterate by truncating certain basis functions. This is explained in Section 6.2.1. It is interesting that this double modification does not really seem to impact the asymptotic convergence behavior.

Although, in contrast to methods employing outer iterations, the treatment of the con-

tact constraints is well incorporated into the multigrid method, the modifications of our particular implementation are relatively small. This is because we have a flexible toolbox `obslib++` for the extended problem class ready to hand; see [124] and also the work [70, 71, 72, 73, 100, 125, 126, 127, 128, 129] based thereon.

The contact constraints are handled by a projected block Gauß–Seidel smoother in the finest space $\boldsymbol{X}_L$. Note that the monotone multigrid methods do not require coarse level approximations of the current iterates. This is in contrast to other non-linear multigrid algorithms; see the discussion in Section 5.9. By a suitable truncation of the coarse level bases, every action at the coarser levels is completely defined by the current fine level iterate. This means that the non-linearity is treated, on the one hand, by employing a non-linear smoother at the finest level and, on the other hand, by modifying the space hierarchy appropriately.

6.2.1 Conversion into a monotone multigrid method

In this paragraph, we outline the additional ingredients required to change the linear semi-geometric multigrid method into a monotone method for contact problems. The details about the characteristics of `obslib++` can be found in [124]. We focus on the issues which are special to the present purpose.

Let a sequence of non-nested vector-valued finite element spaces $(\boldsymbol{X}_\ell)_{\ell=0,\dots,L}$ be given. Also assume that the prolongation matrices $(\bar{\boldsymbol{\Pi}}^\ell_{\ell-1})_{\ell=1,\dots,L}$ with the block structure (6.1) have been computed according to a suitable transfer concept. To solve the Signorini problem, namely the variational inequality (1.19) in the finite element space $\boldsymbol{X}_L$, one needs to distinguish between the components of the displacements in the normal direction and the tangential directions at the possible contact boundary Γ_C. This may be done by an orthogonal transformation, e. g., realized as local Householder reflections. At each node $p \in \mathcal{N}_L \cap \Gamma_C$, the basis $(\boldsymbol{e}_i)_{1\le i\le 3}$ is rotated to a new system $(\boldsymbol{e}^p_i)_{1\le i\le 3}$ such that $\boldsymbol{e}^p_1 = \boldsymbol{n}(p)$; see [129]. The Euclidean basis vectors remain unchanged at the majority of nodes, i. e., we have $\boldsymbol{e}^p_i = \boldsymbol{e}_i$ for $p \in \mathcal{N}_L \setminus (\mathcal{N}_L \cap \Gamma_C)$, $1 \le i \le 3$. This yields a locally modified basis $\boldsymbol{\Lambda}'_L = (\lambda^L_p \boldsymbol{e}^p_i)_{p\in\mathcal{N}_L, 1\le i\le 3}$ of $\boldsymbol{X}_L$. In particular, the non-penetration condition becomes

$$\boldsymbol{\mathcal{K}}_L = \{\boldsymbol{v} \in \boldsymbol{X}_L \mid \boldsymbol{v}(p)\cdot\boldsymbol{e}_1 \le g(p),\ \forall\, p \in \mathcal{N}_L \cap \Gamma_C\}$$

if $\boldsymbol{v} \in \boldsymbol{X}_L$ is written with respect to $\boldsymbol{\Lambda}'_L$ as we do in the following.

The blocks from (6.1) of the prolongation matrix $\bar{\boldsymbol{\Pi}}^L_{L-1} \in \mathbb{R}^{3n_L\times 3n_{L-1}}$ which are associated with possible contact nodes are adjusted accordingly. We denote the resulting matrix, which is the representation of the transfer operator from $\boldsymbol{X}_{L-1}$ to $\boldsymbol{X}_L$ with respect to the standard basis $\boldsymbol{\Lambda}_{L-1}$ and the modified ("rotated") basis $\boldsymbol{\Lambda}'_L$, by $\bar{\boldsymbol{\Pi}}'^L_{L-1}$. For the nodes $p \in \mathcal{N}_L$ and $q \in \mathcal{N}_{L-1}$, let the entries of the block $(\bar{\boldsymbol{\Pi}}'^L_{L-1})_{pq} \in \mathbb{R}^{3\times 3}$ be denoted by $(\bar{\boldsymbol{\Pi}}^L_{L-1})^{ij}_{pq}$, $1 \le i, j \le 3$. We have the analogous notation for the blocks $(\bar{\boldsymbol{\Pi}}^\ell_{\ell-1})_{pq}$ from (6.1) associated with nodes $p \in \mathcal{N}_\ell$ and $q \in \mathcal{N}_{\ell-1}$.

Consequently, the definition of the coarse level bases is slightly more involved than before in Section 6.1.6, where the block entries of the prolongation matrices were diagonal

at all levels. Here, the basis $\widetilde{\boldsymbol{\Lambda}}_{L-1} = (\widetilde{\boldsymbol{\lambda}}_{q,j}^{L-1})_{q\in\mathcal{N}_{L-1},1\le j\le 3}$ of $\boldsymbol{V}_{L-1}$ reads as

$$\widetilde{\boldsymbol{\lambda}}_{q,j}^{L-1} := \sum_{p\in\mathcal{N}_L}\sum_{i=1}^{3} (\bar{\boldsymbol{\Pi}}'^{L}_{L-1})^{ij}_{pq}\,\lambda_p^L \boldsymbol{e}_i^p, \quad \forall\, q\in\mathcal{N}_{L-1},\, 1\le j\le 3.$$

Further, note that the bases of the spaces $\boldsymbol{X}_\ell$ have not been modified at the coarse levels, i. e., for $\ell < L$. Therefore, to construct the spaces $(\boldsymbol{V}_\ell)_{\ell=0,\dots,L-2}$, we have the recursive relation

$$\widetilde{\boldsymbol{\lambda}}_{q,j}^{\ell} := \sum_{p\in\mathcal{N}_{\ell+1}} (\boldsymbol{\Pi}_\ell^{\ell+1})_{pq}\widetilde{\boldsymbol{\lambda}}_{p,j}^{\ell+1}, \quad \forall\, q\in\mathcal{N}_\ell,\, 1\le j\le 3, \tag{6.2}$$

for the bases $\widetilde{\boldsymbol{\Lambda}}_\ell = (\widetilde{\boldsymbol{\lambda}}_{q,j}^{\ell})_{q\in\mathcal{N}_\ell,1\le j\le 3}$ with $\ell\in\{0,\dots,L-2\}$.

As indicated before, we employ a non-linear block Gauß–Seidel method, which is described in detail in [124]. The non-penetration conditions are treated only on the possible contact boundary represented at the finest level, i. e., at the nodes in $\mathcal{N}_L\cap\Gamma_C$. Let $\boldsymbol{u}_L^k\in\boldsymbol{V}_L=\boldsymbol{X}_L$ be some intermediate iterate. For this approximate solution, we denote the set of active nodes where the constraints are binding by

$$\mathcal{A}_L^k := \{p\in\mathcal{N}_L\cap\Gamma_C \mid \boldsymbol{u}_L^k(p)\cdot\boldsymbol{e}_1 = g(p)\}.$$

The paradigm of monotone multigrid methods is that the coarse level correction must not change the active constraints. Therefore, a linear multilevel preconditioner depending on the current iterate is employed which acts only on the subspace

$$\boldsymbol{V}_L^k := \{\boldsymbol{v}\in\boldsymbol{V}_L \mid \boldsymbol{v}(p)\cdot\boldsymbol{e}_1 = 0,\ \forall\, p\in\mathcal{A}_L^k\}\subset\boldsymbol{V}_L.$$

For this purpose, the coarse spaces need to be constructed in a non-trivial way as, in general, the constraints in $\mathcal{A}_L^k$ cannot be represented at the coarser levels $\ell < L$ in the standard multilevel basis. This difficulty is overcome by using so-called truncated basis functions. As the approach developed in [121, 124] is of purely algebraic character, this idea can be applied to the semi-geometric framework quite straightforwardly.

To derive a multilevel hierarchy of subspaces of $\boldsymbol{V}_L^k$ from the semi-geometric spaces $(\boldsymbol{V}_\ell)_{\ell=0,\dots,L-1}$, consider the sets

$$\mathcal{A}_\ell^k := \{p\in\mathcal{N}_\ell \mid \exists\, q\in\mathcal{A}_{\ell+1}^k,\ (\boldsymbol{\Pi}_\ell^{\ell+1})_{pq} > 0\}$$

recursively defined for $\ell\in\{0,\dots,L-1\}$. We remark that the scalar quantities $(\boldsymbol{\Pi}_\ell^{\ell+1})_{pq}$ are indeed appropriate in this definition. Moreover, note that these sets are used for a recursive modification of the spaces exclusively. No special treatment of the nodes in $\mathcal{A}_\ell^k$ by the smoothing operators at the coarse levels is necessary; the standard smoothers are adequate.

We obtain the truncated coarse level spaces by local modifications. The respective bases $(\widetilde{\boldsymbol{\Lambda}}_\ell)_{\ell=0,\dots,L-1}$ are changed to $\widetilde{\boldsymbol{\Lambda}}_\ell^k = (\widetilde{\boldsymbol{\lambda}}_{q,j}^{\ell,k})_{q\in\mathcal{N}_\ell,1\le j\le 3}$ with

$$\widetilde{\boldsymbol{\lambda}}_{q,j}^{L-1,k} := \widetilde{\boldsymbol{\lambda}}_q^{L-1} - \sum_{p\in\mathcal{A}_L^k} (\bar{\boldsymbol{\Pi}}'^{L}_{L-1})^{1j}_{pq}\,\boldsymbol{\lambda}_{p,1}^{L,k}, \quad \forall\, q\in\mathcal{N}_{L-1},\, 1\le j\le 3, \tag{6.3}$$

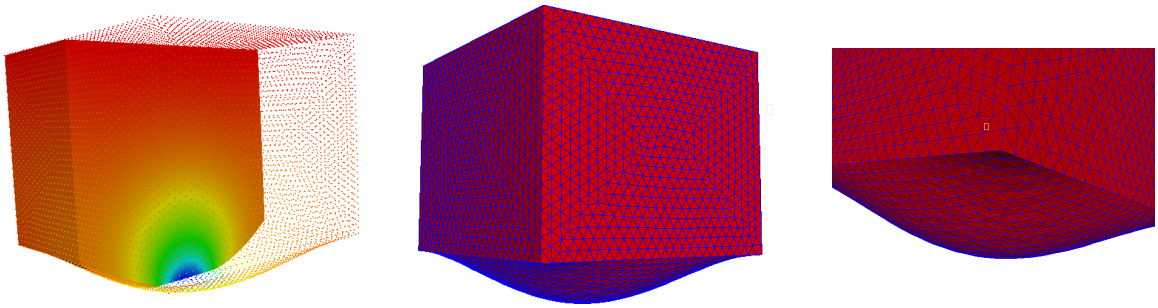


Figure 6.12. Illustration of the considered contact problem: sketch of the solution (left); an unstructured mesh of the warped geometry (center) with zoom in on a corner (right). The coloring of the left image reflects the component of the displacement field in direction of the obstacle, i. e., downward.

and then, recursively for $\ell \in \{0, \dots, L-2\}$,

$$\widetilde{\boldsymbol{\lambda}}_{q,j}^{\ell,k} := \sum_{p \in \mathcal{N}_{\ell+1}} (\boldsymbol{\Pi}_\ell^{\ell+1})_{pq} \widetilde{\boldsymbol{\lambda}}_{p,j}^{\ell+1,k}, \quad \forall\, q \in \mathcal{N}_\ell,\ 1 \le j \le 3, \tag{6.4}$$

which corresponds to (6.2). This yields coarse spaces $(\boldsymbol{V}_\ell^k)_{\ell=0,\dots,L-1}$ contained in $\boldsymbol{V}_L^k$. Note that generally $\boldsymbol{V}_\ell^k \not\subset \boldsymbol{V}_\ell$.

The outlined truncation procedure is efficiently implemented by local algebraic modifications. Concerning the prolongation matrices, we see by (6.3) and (6.4) that only entries between the levels $L-1$ and L need to be modified, namely set to zero, in blocks associated with nodes in $\mathcal{A}_L^k$. In contrast, the stiffness matrices at the levels $\ell < L$ change at the nodes in $(\mathcal{A}_\ell^k)_{\ell=0,\dots,L-1}$. This holds true in the standard (not semi-geometric) case of [124], too.

All in all, this approach leads to a convergent non-linear iteration process which reduces to a linear subspace correction method as soon as the actual contact boundary is identified; see [122, 124]. Note that the truncated basis functions are of crucial importance for the convergence and the efficiency of the method, particularly for complicated geometries in case $d = 3$. As a matter of fact, monotone multigrid methods for contact show to be of optimal complexity, even for frictional problems; see also [125, 126].

6.2.2 Numerical results

In the following experiment, we study the asymptotic convergence behavior of the semi-geometric monotone multigrid method for Signorini's problem. For this purpose, we consider the cube-like geometry depicted in Figure 6.12 where one of the six faces, namely the possible contact boundary, is warped. Let the rigid obstacle be the half space given by the tangential plane of the center point of the warped surface. This problem remotely reminds of the one studied analytically by Hertz [110] as early as in the year 1881. For the assessment of the performance of solvers for contact problems, it is crucial to choose a setting where at least the geometry or the obstacle are not flat. Otherwise, the correct discrete contact zone is likely to be identified in the very first step of the iteration.

#elements	#dof	$\lvert\mathcal{A}_L\rvert$	$\mathcal{C}_{\text{gr}}$	$\mathcal{C}_{\text{op}}$	$K_{\mathcal{V}(2,2)}$	$\tilde{\rho}_{\mathcal{V}(2,2)}$	$K_{\mathcal{V}(3,3)}$	$\tilde{\rho}_{\mathcal{V}(3,3)}$	$K_{\mathcal{V}(4,4)}$	$\tilde{\rho}_{\mathcal{V}(4,4)}$
6,568	4,197	9	1.15	1.34	19 (4)	0.238	16 (4)	0.166	14 (3)	0.133
63,645	36,102	44	1.13	1.33	21 (7)	0.295	18 (7)	0.202	16 (7)	0.162
310,198	168,978	155	1.11	1.33	20 (7)	0.325	17 (6)	0.226	15 (6)	0.158
543,408	293,346	240	1.11	1.32	23 (8)	0.357	19 (8)	0.237	18 (7)	0.196
1,037,557	555,198	418	1.12	1.36	25 (8)	0.422	20 (8)	0.309	18 (8)	0.231
1,206,114	643,704	478	1.10	1.31	27 (9)	0.449	21 (9)	0.340	19 (8)	0.269

Table 6.12. Convergence of the semi-geometric monotone multigrid method.

The setting of this study needs to differ from the ones before as the problem is non-linear. In particular, it is not sufficient to study the convergence behavior to the trivial solution only. We prescribe non-zero Dirichlet boundary conditions for the displacements at the top quadratic surface of the domain, pointing towards the obstacle. The influence of the size of the boundary values is discussed below; the data does not need to be physically reasonable for our purposes. May the material parameters of the linear elastic body be chosen as in Section 6.1.6.

A usual estimate for the algebraic error in the context of iterative methods for variational inequalities is the energy norm of the computed correction as indicated in (1.26); see [121]. We denote by $K_{\mathcal{V}(\nu_1,\nu_2)}$ the number of monotone multigrid steps, i. e., of non-linear $\mathcal{V}(\nu_1,\nu_2)$-cycles specified by the construction in Section 6.2.1, to reach an estimated algebraic error less than 10^{-10} starting from the initial iterate $\boldsymbol{u}_L^0 = \mathbf{0}$. Let $\tilde{\rho}_{\mathcal{V}(\nu_1,\nu_2)}$ be the corresponding approximate asymptotic convergence rate given by (1.26) with $k = K_{\mathcal{V}(\nu_1,\nu_2)} - 1$.

Table 6.12 contains the results of the convergence study in case of Dirichlet values of size 0.08. For comparison, the size of the domain is two in all space dimensions. Moreover, the initial distance to the obstacle at the center point of the warped surface is zero. Similar to before in Section 6.1.1, we have generated several completely independent meshes of different sizes, which are then treated as given fine level problems. Then, the coarse meshes are chosen appropriately by means of the outlined strategy to construct a suitable coarse level hierarchy. In addition to the usual problem data, namely the number of elements and the number of degrees of freedom, we state the quantity $\lvert\mathcal{A}_L\rvert := \lvert\mathcal{A}_L^{K_{\mathcal{V}(\nu_1,\nu_2)}}\rvert$. To describe the convergence behavior over this range of problems, we consider the mentioned rates $\tilde{\rho}$, the total number of multigrid steps and also the number of included non-linear steps constituting the transient phase at the beginning of the iteration. The latter are given in brackets. Both has not been necessary for the linear problems presented in the previous section; there, the convergence behavior was completely characterized by the quantities $\bar{\rho}$.

The presented numbers show a moderate increase of the convergence rates and iteration counts with increasing problem size. Operator and grid complexities stay in a rather small range. The monotone multigrid cycles tend to need one or two more smoothing iterations to achieve results comparable to the linear method. This is an observation which has in principle been made for the case where the truncated spaces are constructed from standard multilevel finite elements in [124], too. However, comparing the results in Table 6.12 with the ones in Table 6.10, recall that both the problem settings and the error measurements

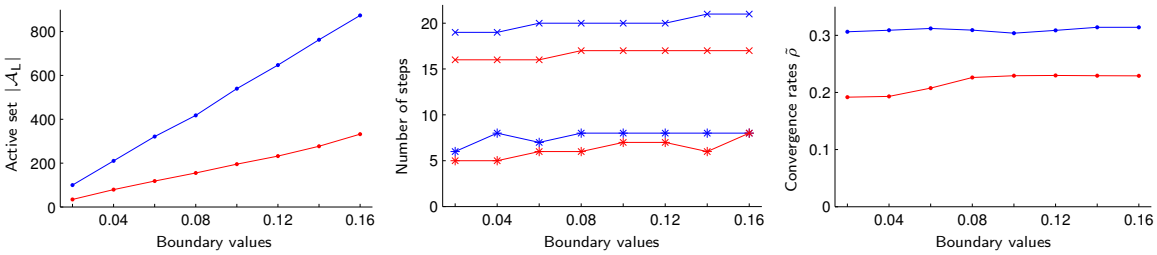


Figure 6.13. Studying the dependence of the convergence on the magnitude of the prescribed boundary data: number of nodes in the actual contact zone (left); total number of semi-geometric monotone multigrid steps $K_{\mathcal{V}(3,3)}$ marked by $\times$ and included non-linear steps marked by $*$ (center); convergence rates $\tilde{\rho}_{\mathcal{V}(3,3)}$ (right). The values are given for the third and fifth problem of Table 6.12 in red and blue, respectively.

are different. A detailed investigation of other practical issues in the context of monotone multigrid methods for Signorini's problem can be found in [124]. For instance, we do not study the effect of highly varying normals here.

Let us now examine the dependence of the convergence on the magnitude of the prescribed boundary data. For this purpose, we consider two of the above problems of different sizes and apply Dirichlet boundary values at the top surface in direction of the obstacle between 0.02 and 0.16 in eight independent runs each. Before, in Table 6.12, a medium size of 0.08 was studied. It turns out that the convergence is only weakly affected. The result of this study is illustrated in Figure 6.13 for the monotone multigrid iteration with $\mathcal{V}(3,3)$-cycles. As an illustration, we give the increasing number of nodes which are in contact, namely $|\mathcal{A}_L|$, in the left diagram.

7 Multigrid methods based on parametric finite elements

In this chapter, we turn our attention to a selected technique for the application of elementary multilevel ideas to problems with complicated boundaries. This is done in the context of the numerical simulation of elastic contact problems. Combining the general multilevel setting with a different perspective, namely an advanced modeling point of view, we present a (monotone) multigrid method based on a hierarchy of parametric finite element spaces. For this purpose, a full-dimensional parameterization is employed which accurately represents the computational domain.

Although the developed concept is related to several considerations made throughout this thesis, especially to the discussion in Section 4.1, we prefer to organize this as a supplement at the end. Indeed, the development of this particular focus does not need to be linked too closely to the previous chapters. However, we mainly do this because the purpose of the parametric finite element discretization put forward in this chapter is two-fold. On the one hand, it allows for an elegant multilevel hierarchy to be used in the mentioned multigrid algorithms. But, on the other hand, it comes with particular advantages for the modeling of contact problems. As a matter of fact, a combination of the parametric concept with other ideas, which can take advantage of an enhanced representation of the computational domain to improve some modeling aspects, is certainly advisable. This is elaborated in more detail in Section 7.1. After all, the long-term objective lies in an increased flexibility of *hp*-adaptive methods for contact problems.

7.1 Introduction

In the numerical simulation of elastic contact problems, the treatment of the non-penetration conditions at the potential contact boundary is of particular importance for both the quality of a finite element approximation and the overall efficiency of the algorithms. A vital challenge is to achieve an accurate description of geometric features, e. g., of warped surfaces, often incorporated in three-dimensional models from computer-aided design (CAD). Here, we investigate a new connection of different numerical methods, namely modern discretization techniques for partial differential equations on complex geometries on the one side and fast multilevel solvers for constrained minimization problems on the other side.

It is fair to say that the development of *hp*-adaptive methods for contact problems has not yet reached a mature state; see, e. g., [54] and the references therein. Partly, this is due to the difficulties concerning the geometric representation of the computational domain. A generally accepted paradigm is, though, that high order (finite element or boundary element) methods need high order meshes [114, 140]. This is especially difficult for three-dimensional multi-body contact problems. In this case, the application of non-conforming domain decomposition techniques [173] to realize the information transfer across geometrically non-matching warped contact interfaces is a highly demanding task. For low order finite elements, this has been achieved, among others, by the author; see [71].

The perspective we offer here is a parametric finite element method. For *hp*-adaptive methods it is convenient to have a parameterization describing the geometry accurately ready to hand. This is because a change of the computational domain due to locally altered polynomial degree is not desirable. Therefore, it is reasonable to uncouple the representation of the geometry on the one hand and of a scale of approximation spaces for the discrete solution on the other hand. These two purposes are usually not separated properly. But of course, one can find curved elements of other than isoparametric structure in some form or another in the literature; see, e. g., [93, 205] or the monograph [56] and the references therein. Note that, for similar reasons, an "isogeometric" concept, which uses NURBS bases for both the description of the geometry and the discrete solution of the differential equation, has been introduced in [114].

For practical computations, the development of fast and robust solvers is equally important. As this issue has not yet been in the main focus of, e. g., the isogeometric analysis [114], we would like to contribute ideas from the field of multilevel methods for variational inequalities. More precisely, as indicated before, we show how to use a monotone multigrid method to efficiently solve the non-linear contact problem discretized with low order parametric finite elements. Note that the actual treatment of higher order elements is beyond the scope of the present discussion.

To obtain multilevel parametric finite element spaces in case $d = 3$, we use a full-dimensional parameterization, constructed by tetrahedral transfinite interpolation [158] of CAD data, to lift standard Lagrange elements to the computational domain. Note that, similarly, a surface parameterization has been used in a wavelet Galerkin scheme for boundary integral equations, see [108]. Such a procedure may serve as an essential prerequisite to tackle the problems mentioned above. In particular, many of the issues arising in the generation of *p*-version meshes for curved boundaries [140] can be avoided in a quite elegant way. In this sense, although rather expensive, the use of a high order parameterization permits maximal freedom in an *hp*-adaptive discretization scheme. We presume that the present concept can also be combined with the ideas in [71].

All in all, the results presented in this excursus constitute at least a little progress on the way to an efficient *hp*-adaptive numerical simulation of contact problems in case of complex three-dimensional geometries.

7.2 Parametric finite elements

In this section, we introduce a parametric finite element discretization of the contact problem stated in Section 1.2.2. On the one hand, this method uses much more geometric information from a CAD model than standard finite elements; on the other hand, we do not use the same functions for the discrete approximation of the displacement field as for the representation of the geometry, which is done in the so-called "isogeometric analysis" introduced in [114]. This allows for a reasonable multilevel hierarchy in case of low order trial functions to be discussed in the next section.

In the following, the symbols $\boldsymbol{\varphi}$ with some indices stand for certain parameterizations or transformations; this must not be confused with the notation of the deformation in the continuum mechanical model. We denote the (closed) d-simplex by Δ^d and its faces

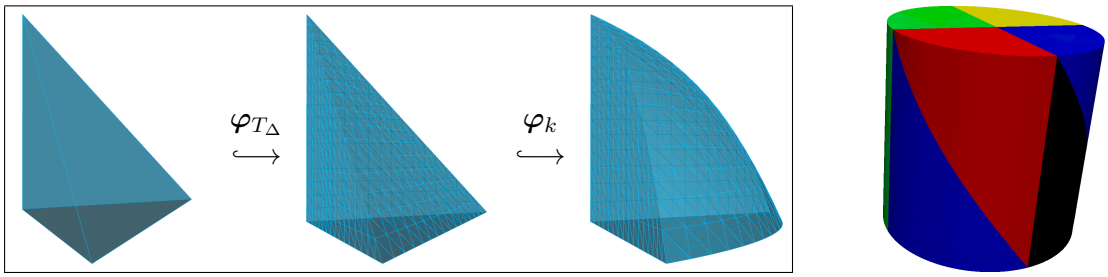


Figure 7.1. From left to right: the reference element $\widehat{T} = \Delta^3$; a mesh of the simplex Δ^3; a parametric mesh (here, $K = 1$) where each element is an image of an affine element; a tetrahedral decomposition of a cylinder with $K = 8$.

by Δ^d_j, $j \in \{1, \dots, d+1\}$. To describe the elastic body (here, $d = 3$) by a practicable parameterization, we consider a non-overlapping simplicial decomposition of $\Omega \subset \mathbb{R}^d$ into a fixed number of $K \geq 1$ subdomains. Formally this reads as

$$\overline{\Omega} = \bigcup_{k=1}^{K} \overline{\Omega}_k = \bigcup_{k=1}^{K} \varphi_k(\Delta^d),$$

where the notation already indicates that the subdomains $(\Omega_k)_{k=1,\dots,K}$ appear as particular images of suitable parameterizations $(\varphi_k)_{k=1,\dots,K}$. This is illustrated in Figure 7.1 (right).

Let us assume that the faces of the simplicial cells Ω_k, namely the surfaces $\varphi_k(\Delta^d_j)$, $k \in \{1, \dots, K\}$, $j \in \{1, \dots, d+1\}$, are given as B-patches. This way to represent polynomial surfaces is analyzed in [66]. In this case, the author of [158] proposes to construct the full-dimensional mappings $\varphi_k : \Delta^d \to \mathbb{R}^d$, $k \in \{1, \dots, K\}$, as transfinite interpolations of the surface values from the CAD model using certain blending functions. Particularly, the single parameterizations are smooth and they match across these B-patch surfaces if the surfaces themselves match. This gives rise to a consistent global parameterization which we do not write down explicitly. We note that this global mapping is continuous but not necessarily differentiable across the interior interfaces. In addition, one can guarantee that each parameterization φ_k satisfies the regularity assumption

$$\det(\boldsymbol{\nabla} \varphi_k) > 0 \quad \text{in } \Delta^d. \tag{7.1}$$

In fact, this is one of the main results of [158].

In the following, we define the parametric finite element spaces in a rather straightforward way via a lift of standard Lagrange finite elements. For this purpose, let $(\mathcal{T}^k_\ell)_{\ell \in \mathbb{N}}$ be a family of nested simplicial meshes of Δ^d for each $k \in \{1, \dots, K\}$. To keep the global finite element spaces conforming, we assume that the meshes meeting at the faces of the simplicial subdomains Ω_k of Ω match at each level $\ell \in \mathbb{N}$. Let $\widehat{T}$ be the reference element; here, $\widehat{T} = \Delta^d$. Then, for each $T_\Delta \in \mathcal{T}^k_\ell$, there is an affine mapping φ_{T_Δ} such that $\varphi_{T_\Delta}(\widehat{T}) = T_\Delta$.

Now, we give a concise description of the parametric elements in Ω by employing the special finite element transformations

$$\varphi_T := \varphi_k \circ \varphi_{T_\Delta} : \widehat{T} \to \mathbb{R}^d, \tag{7.2}$$

which are diffeomorphisms between the reference element $\widehat{T}$ and the actual elements. That way, the parametric elements at level $\ell \in \mathbb{N}$ are identified as the images of the elements of the meshes $(\mathcal{T}_\ell^k)_{k=1,\dots,K}$; see Figure 7.1. More precisely, a family of parametric meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ of Ω can be defined by

$$\mathcal{T}_\ell := \left\{ T = \boldsymbol{\varphi}_T(\widehat{T}) = \boldsymbol{\varphi}_k(\boldsymbol{\varphi}_{T_\Delta}(\widehat{T})) \mid 1 \le k \le K,\ T_\Delta \in \mathcal{T}_\ell^k \right\}, \quad \forall\, \ell \in \mathbb{N}.$$

Assume that this family of global meshes is shape regular and quasi-uniform according to the equations (1.16) and (1.17). Note that assumption (7.1), combined with the continuous differentiability of the mappings $(\boldsymbol{\varphi}_k)_{k=1,\dots,K}$, in the compactum Δ^d, implies that it is sufficient to ensure these conditions for each sequence $(\mathcal{T}_\ell^k)_{\ell\in\mathbb{N}}$ separately as far as we keep K fixed.

Finally, let $\mathbb{P} := \mathbb{P}_r(\widehat{T})$ be the space of polynomials of degree r in $\widehat{T}$. Then, for $\ell \in \mathbb{N}$, the parametric finite element space associated with the parametric mesh $\mathcal{T}_\ell$ is

$$\begin{aligned} X_\ell &:= \left\{ v \in \mathcal{C}^0(\Omega) \mid \forall\, T \in \mathcal{T}_\ell\ \exists\, w \in \mathbb{P} :\ v(\boldsymbol{x}) = w(\boldsymbol{\varphi}_T^{-1}(\boldsymbol{x})),\ \forall\, \boldsymbol{x} \in T \right\} \\ &= \left\{ v \in \mathcal{C}^0(\Omega) \mid v \circ \boldsymbol{\varphi}_T \in \mathbb{P},\ \forall\, T \in \mathcal{T}_\ell \right\}. \end{aligned} \tag{7.3}$$

Note that, in principle, the above definition makes sense for any reasonable set of finite element transformations $(\boldsymbol{\varphi}_T)_{T\in\mathcal{T}_\ell}$. In case the mappings are constructed as in (7.2) via the high order parameterization from [158], this is a "superparametric" concept if the degree r is small. This is in contrast to the subparametric or isoparametric finite elements which are usually considered in the literature; see [56].

From a practical point of view, virtually every kind of parameterization can be employed with the following qualification. For an efficient assembly of the stiffness matrix and the right hand side via sufficiently accurate numerical quadrature, the derivatives of the resulting finite element transformations (7.2) and the mappings themselves must be easy to evaluate; see, e. g., [9].

Let us now apply the above concept. We suppose that the Dirichlet data has been treated appropriately. Then, a discretization of Signorini's problem described in Section 1.2.2 is obtained by specifying a suitable set of admissible displacements $\boldsymbol{\mathcal{K}}_\ell$ using the vector-valued parametric finite element space $\boldsymbol{X}_\ell := (X_\ell)^3$ defined by (7.3) with $r = 1$. The discrete variational problem reads exactly as before in the standard case of Section 1.3 if, as usual, the non-penetration conditions on the possible contact boundary Γ_C are merely enforced at the nodes $\mathcal{N}_\ell \cap \Gamma_C$. For clarity, we recall the variational inequality: find $\boldsymbol{u}_\ell \in \boldsymbol{\mathcal{K}}_\ell$ such that

$$a(\boldsymbol{u}_\ell, \boldsymbol{v} - \boldsymbol{u}_\ell) \ge \mathcal{F}(\boldsymbol{v} - \boldsymbol{u}_\ell), \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{K}}_\ell := \{ \boldsymbol{v} \in \boldsymbol{X}_\ell \mid \boldsymbol{v}(p) \cdot \boldsymbol{n}(p) \le g(p),\ \forall\, p \in \mathcal{N}_\ell \cap \Gamma_C \}.$$

Remark 7.1. *Although, from a modeling point of view, as much geometric information as possible should be used for an accurate description of contact phenomena, we remark that a strong pointwise non-penetration condition everywhere on Γ_C is usually not suitable for the variational formulation the (parametric) finite element method relies on. Besides, a decoupled set of constraints is preferable for a variety of reasons. The common remedy is to prescribe the contact constraints with respect to a suitable cone of Lagrange multipliers.*

This requires the introduction of appropriate sets of functionals in $(H^{\frac{1}{2}}(\Gamma_C))'$. To retain inequality constraints which can be enforced merely by looking at the nodes, one can employ discontinuous test spaces similar to the ones we used for the pseudo-L^2-projection as pointed out in Section 5.5.2. This connection is elaborated in more detail, e. g. in our preprint [73].

The quality of a priori error estimates for the above discretization certainly depends on a number of aspects which have to be examined more closely. Beside regularity assumptions for the solution of (1.15), the balance of the primal degrees of freedom and the constraints by means of an inf-sup condition and certain properties of the parameterization, e. g., the regularity (7.1), influence the error analysis.

7.3 Monotone multigrid for parametric elements

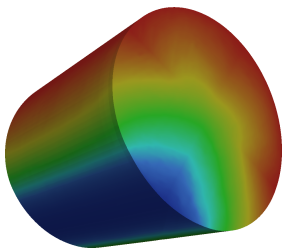
For the solution of the arising discrete variational inequalities, we propose a monotone multigrid method; see Section 6.2. Similar to some of the approaches reviewed in Section 4.1, the scale of parametric finite element spaces constitutes an adjusted discretization technique which allows for an almost straightforward application of multilevel ideas. For this purpose, we examine the constructed space hierarchy, which we presume to possess the required approximation properties, and the corresponding natural transfer operators in a little more detail.

By construction, the spaces defined by (7.3) are nested. This is an immediate consequence of the fact that the parameterization is fixed and does not change with the index ℓ. Still, let us formulate this statement in the following lemma and give an elementary proof of the assertion.

Lemma 7.2. *The parametric finite element spaces $(X_\ell)_{\ell\in\mathbb{N}}$ are nested.*

Proof. For $\ell \geq 1$, let $v \in X_{\ell-1}$ be arbitrary. Then, for $T \in \mathcal{T}_{\ell-1}$ there is a unique element $T_\Delta \in \mathcal{T}_{\ell-1}^k$ for some $k \in \{1,\dots,K\}$ such that $\boldsymbol{\varphi}_k(T_\Delta) = T$. Let $(T_\Delta^i)_{i=1,\dots,N}$ be the children of T_Δ in $\mathcal{T}_\ell^k$. In general, $1 \leq N \leq 2^d$; in case of standard uniform refinement of the simplices, it is $N = 2^d$. We have the corresponding elements $(T^i)_{i=1,\dots,N}$ in $\mathcal{T}_\ell$ with $T^i = \boldsymbol{\varphi}_k(T_\Delta^i)$ for $i \in \{1,\dots,N\}$. By assumption, $v \circ \boldsymbol{\varphi}_T = v \circ \boldsymbol{\varphi}_k \circ \boldsymbol{\varphi}_{T_\Delta} \in \mathbb{P}$. Therefore, $v \circ \boldsymbol{\varphi}_{T^i} = v \circ \boldsymbol{\varphi}_k \circ \boldsymbol{\varphi}_{T_\Delta^i} \in \mathbb{P}$ because $T_\Delta^i \subset T_\Delta$ and the finite element transformations are affine. As each element of $\mathcal{T}_\ell$ appears as the child of an element in $\mathcal{T}_{\ell-1}$ in the above fashion, we obtain $v \in X_\ell$. Consequently, $X_{\ell-1} \subset X_\ell$ for all $\ell \geq 1$. □

Therefore, no advanced transfer concepts need to be studied here as the canonical inclusion $\mathcal{I}_{\ell-1}^\ell : X_{\ell-1} \to X^\ell$ is the most natural operator to be used as prolongation. Note that these operators only depend on the logical structure; as in the standard nested case of Section 2.2.1, the representing matrices contain the entries 0, 0.5 and 1 and may be computed from the neighborhood relations in and between the simplicial meshes $(\mathcal{T}_{\ell-1}^k)_{k=1,\dots,K}$ and $(\mathcal{T}_\ell^k)_{k=1,\dots,K}$. This is because the respective multilevel basis is defined via a lift by proceeding as in (7.3). As a result, for a fixed finest level L, the computation of the matrices $(\boldsymbol{\mathcal{I}}_{\ell-1}^\ell)_{\ell=1,\dots,L}$ between the nested spaces $(X_\ell)_{\ell=0,\dots,L}$ does not need the parameterization. However, the computation of the outer normals $(\boldsymbol{n}(p))_{p\in\mathcal{N}_L\cap\Gamma_C}$ and also of the



#elements	#dof	#steps	$\tilde{\rho}$	$\lvert\mathcal{A}_L\rvert$
96	107	8 (2)	0.032	3
768	615	10 (3)	0.031	15
6,144	3,915	11 (4)	0.065	58
49,152	27,795	13 (6)	0.091	199
393,216	209,187	14 (6)	0.102	753
3,145,728	1,622,595	15 (8)	0.114	2,984

Figure 7.2. Contact problem of a parameterized cylinder with a rigid obstacle shaped like a broad channel. The colors indicate the displacement in $\boldsymbol{e}_3$-direction.

values $(g(p))_{p\in\mathcal{N}_L\cap\Gamma_C}$ for the prescription of the contact constraints may require access to the mappings $(\boldsymbol{\varphi}_k)_{k=1,\dots,K}$.

We anticipate that the constructed coarse spaces have the desired multilevel approximation properties. More precisely, under mild assumptions on the employed parameterization mappings $(\boldsymbol{\varphi}_k)_{k=1,\dots,K}$, the relevant Jackson- and Bernstein-type inequalities mentioned in Section 2.3.2 transfer from the standard finite element spaces to the parametric spaces; see also [107].

Finally, we would like to point out that no modifications are necessary in the code of the solver provided that the local normal/tangential coordinate systems can be computed from the parameterization. Consequently, a monotone multigrid method can be employed for contact problems discretized with parametric finite elements in the quite straightforward way outlined above. See Figure 7.2 for a numerical example illustrating the performance of the method. We report on the asymptotic convergence rate of a conjugate gradient method with the monotone multigrid ($\mathcal{V}(3,3)$-cycle) as preconditioner. Starting with the initial iterate zero at each refinement level (i. e., no nested iteration), we list the number of total steps needed to reduce the residual norm to less than 10^{-10}. The count of included non-linear steps is given in brackets.

Concluding remarks

The results briefly described in this excursus certainly have preliminary character; the performance of the presented algorithms needs to be studied in more detail. This is work in progress*. However, the experiments so far show that (monotone) multigrid methods based on parametric finite elements work as expected; see Figure 7.2. Still, the effort of constructing a (high order) parameterization by the methodology developed in [158] especially pays if there is also a considerable gain on the modeling side. Here, the effect of this special resolution of the boundary on the discrete approximation of contact phenomena or general boundary effects needs to be investigated more closely.

*We would like to thank Helmut Harbrecht and Maharavo Randrianarivony for bringing this topic to our attention. Moreover, we acknowledge the latter for providing his code for the tetrahedral transfinite interpolation described in [158]. The valuable assistance of Lukas Döring in the implementation of a flexible interface of the parameterization concept to our finite element code is highly appreciated.

Conclusion

We have presented a class of multilevel methods based on non-nested meshes. The devised algorithms can indeed be used to efficiently solve problems associated with unstructured meshes, which do not exhibit any natural multilevel structures. Our numerical experiments provide convincing arguments for both the applicability and the flexibility of the semi-geometric framework.

Relevant connections to other geometry-based multilevel techniques have been explained at different points throughout the presentation. Let us emphasize that the assumptions of our convergence analysis are particularly weak. We have been able to prove quasi-optimal convergence and preconditioning results without any of the surely tighter relations between the hierarchies of non-nested meshes or spaces that come into play with other geometric multilevel algorithms.

In the thorough investigation of the information transfer between finite element spaces associated with non-nested meshes, we have identified several local operators, which are suitable for an application in the semi-geometric framework. All studied transfer concepts between non-nested finite element spaces have been geometrically motivated and thus involved the corresponding non-nested meshes.

In particular, we have seen in a series of experiments that the introduced pseudo-L^2-projection operator is by far the closest to the actual L^2-projection when applied in the present context. However, the numerical experiments concerning the multilevel iterations based on the semi-geometric approach have shown that the standard nodal interpolation is the most appropriate choice in practice. It turns out that this quite simple transfer concept yields the most efficient multilevel hierarchies. Besides, the fact that most of the studied operators are well-suited provides more proof for the practical applicability of the proposed multilevel methods.

Another finding of the extensive numerical studies is the robustness of the semi-geometric approach with respect to the choice of the coarse meshes and the required truncation of the prolongation matrices. It also seems that it is generally appropriate to choose coarse meshes which are nested such that $X_0 \subset \ldots \subset X_{L-1} \not\subset X_L$. This is supported by the experiments. However, the case of completely non-nested hierarchies is covered by the analysis we have carried out in Chapter 3 and also works in practice.

An issue which needs to be investigated more closely is the possible local adaptation of the coarse meshes to a given fine mesh. This may be necessary for highly non-uniform meshes to more accurately retain the approximation power of the coarse level spaces. In addition, the described methods could be analyzed for finite element spaces of higher order. Moreover, the brief discussion of the parametric multigrid idea in Chapter 7 demonstrates that the first steps have been taken towards a flexible combination of the parametric concept and the indicated multilevel approach. We will study this in more detail in future work.

Appendix

A Tables

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	1.00	0.76	0.75	0.80	0.83	0.85	0.87	0.88	0.90	0.91	0.91	0.92	0.93	0.93	0.94
$\mathcal{B}_2$	0.52	1.00	0.68	0.72	0.77	0.79	0.82	0.83	0.85	0.87	0.88	0.88	0.90	0.91	0.92
$\mathcal{B}_3$	0.45	0.57	1.00	0.67	0.74	0.76	0.78	0.81	0.83	0.85	0.86	0.86	0.89	0.89	0.90
$\mathcal{B}_4$	0.37	0.46	0.54	1.00	0.68	0.71	0.74	0.77	0.79	0.81	0.83	0.84	0.87	0.87	0.88
$\mathcal{B}_5$		0.36	0.42	0.49	1.00	0.64	0.67	0.70	0.74	0.77	0.78	0.79	0.83	0.84	0.85
$\mathcal{B}_6$			0.37	0.44	0.55	1.00	0.63	0.67	0.71	0.74	0.76	0.77	0.81	0.82	0.83
$\mathcal{B}_7$				0.39	0.49	0.54	1.00	0.64	0.67	0.71	0.73	0.75	0.78	0.79	0.81
$\mathcal{B}_8$					0.43	0.47	0.53	1.00	0.63	0.67	0.69	0.71	0.75	0.76	0.77
$\mathcal{B}_9$						0.42	0.47	0.53	1.00	0.63	0.64	0.67	0.72	0.73	0.75
$\mathcal{B}_{10}$							0.41	0.46	0.53	1.00	0.60	0.63	0.69	0.69	0.72
$\mathcal{B}_{11}$								0.44	0.49	0.55	1.00	0.61	0.66	0.67	0.70
$\mathcal{B}_{12}$									0.46	0.52	0.55	1.00	0.63	0.64	0.67
$\mathcal{B}_{13}$										0.46	0.48	0.51	1.00	0.63	0.62
$\mathcal{B}_{14}$											0.46	0.50	0.60	1.00	0.60
$\mathcal{B}_{15}$												0.46	0.52	0.54	1.00

Table A.1. H^1-stability of $\mathcal{I}$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	1.00	0.93	0.96	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$\mathcal{B}_2$	0.61	1.00	0.86	0.92	0.96	0.97	0.97	0.98	0.98	0.99	0.99	0.99	1.00	1.00	1.00
$\mathcal{B}_3$	0.49	0.68	1.00	0.86	0.94	0.95	0.96	0.97	0.97	0.98	0.99	0.99	1.00	1.00	1.00
$\mathcal{B}_4$	0.37	0.52	0.64	1.00	0.88	0.92	0.94	0.95	0.96	0.97	0.98	0.98	0.99	0.99	1.00
$\mathcal{B}_5$		0.35	0.44	0.55	1.00	0.80	0.87	0.91	0.94	0.96	0.96	0.97	0.98	0.98	0.99
$\mathcal{B}_6$			0.37	0.48	0.65	1.00	0.81	0.88	0.92	0.95	0.95	0.96	0.98	0.98	0.98
$\mathcal{B}_7$				0.39	0.56	0.64	1.00	0.82	0.89	0.93	0.94	0.95	0.97	0.97	0.98
$\mathcal{B}_8$					0.46	0.54	0.63	1.00	0.82	0.89	0.91	0.93	0.97	0.97	0.98
$\mathcal{B}_9$						0.44	0.53	0.64	1.00	0.83	0.86	0.90	0.95	0.96	0.97
$\mathcal{B}_{10}$							0.43	0.54	0.65	1.00	0.80	0.85	0.92	0.94	0.97
$\mathcal{B}_{11}$								0.49	0.60	0.71	1.00	0.82	0.90	0.92	0.96
$\mathcal{B}_{12}$									0.54	0.66	0.71	1.00	0.87	0.90	0.94
$\mathcal{B}_{13}$										0.54	0.60	0.67	1.00	0.87	0.89
$\mathcal{B}_{14}$											0.58	0.65	0.82	1.00	0.88
$\mathcal{B}_{15}$												0.60	0.74	0.78	1.00

Table A.2. H^1-stability of $\mathcal{Q}$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	1.00	0.87	0.89	0.92	0.93	0.94	0.95	0.96	0.97	0.97	0.97	0.98	0.98	0.98	0.98
$\mathcal{B}_2$	0.57	1.00	0.76	0.81	0.89	0.90	0.91	0.93	0.94	0.95	0.96	0.96	0.97	0.97	0.98
$\mathcal{B}_3$	0.45	0.61	1.00	0.77	0.85	0.87	0.89	0.91	0.93	0.94	0.95	0.95	0.97	0.97	0.97
$\mathcal{B}_4$	0.35	0.48	0.57	1.00	0.78	0.82	0.85	0.88	0.90	0.92	0.93	0.93	0.95	0.96	0.96
$\mathcal{B}_5$		0.33	0.43	0.51	1.00	0.72	0.78	0.82	0.85	0.88	0.90	0.91	0.93	0.94	0.94
$\mathcal{B}_6$			0.35	0.44	0.60	1.00	0.73	0.78	0.83	0.86	0.88	0.89	0.92	0.93	0.93
$\mathcal{B}_7$				0.37	0.52	0.59	1.00	0.74	0.79	0.83	0.85	0.87	0.90	0.91	0.92
$\mathcal{B}_8$					0.43	0.49	0.57	1.00	0.73	0.79	0.81	0.84	0.88	0.89	0.91
$\mathcal{B}_9$						0.42	0.50	0.58	1.00	0.74	0.77	0.80	0.86	0.88	0.89
$\mathcal{B}_{10}$							0.41	0.50	0.59	1.00	0.71	0.76	0.83	0.85	0.88
$\mathcal{B}_{11}$								0.47	0.55	0.65	1.00	0.74	0.81	0.83	0.87
$\mathcal{B}_{12}$									0.51	0.60	0.65	1.00	0.78	0.81	0.85
$\mathcal{B}_{13}$										0.51	0.56	0.61	1.00	0.81	0.80
$\mathcal{B}_{14}$											0.54	0.60	0.77	1.00	0.79
$\mathcal{B}_{15}$												0.57	0.68	0.72	1.00

Table A.3. H^1-stability of $\mathcal{P}$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	0.43	0.47	0.55	0.62	0.68	0.72	0.77	0.79	0.82	0.84	0.85	0.86	0.89	0.89	0.90
$\mathcal{B}_2$	0.25	0.40	0.41	0.47	0.60	0.62	0.65	0.70	0.73	0.77	0.78	0.80	0.83	0.84	0.85
$\mathcal{B}_3$	0.20	0.27	0.40	0.40	0.51	0.55	0.60	0.65	0.69	0.73	0.74	0.77	0.80	0.81	0.82
$\mathcal{B}_4$	0.13	0.20	0.26	0.39	0.42	0.48	0.53	0.58	0.63	0.67	0.69	0.71	0.76	0.77	0.78
$\mathcal{B}_5$		0.13	0.16	0.22	0.38	0.36	0.41	0.47	0.53	0.58	0.61	0.63	0.69	0.70	0.72
$\mathcal{B}_6$			0.13	0.18	0.27	0.38	0.37	0.43	0.48	0.54	0.56	0.59	0.65	0.66	0.69
$\mathcal{B}_7$				0.14	0.22	0.26	0.38	0.37	0.43	0.49	0.51	0.55	0.61	0.62	0.65
$\mathcal{B}_8$					0.17	0.21	0.25	0.38	0.37	0.43	0.46	0.49	0.56	0.57	0.60
$\mathcal{B}_9$						0.12	0.15	0.20	0.25	0.30	0.33	0.37	0.44	0.45	0.49
$\mathcal{B}_{10}$							0.12	0.16	0.20	0.25	0.27	0.31	0.38	0.40	0.43
$\mathcal{B}_{11}$								0.19	0.24	0.29	0.37	0.36	0.43	0.44	0.48
$\mathcal{B}_{12}$									0.21	0.26	0.29	0.37	0.40	0.41	0.45
$\mathcal{B}_{13}$										0.21	0.23	0.26	0.37	0.36	0.39
$\mathcal{B}_{14}$											0.22	0.25	0.33	0.37	0.38
$\mathcal{B}_{15}$												0.23	0.30	0.32	

Table A.4. H^1-stability of $\mathcal{R}$, $r = 0$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	0.39	0.47	0.54	0.61	0.68	0.71	0.74	0.77	0.79	0.81	0.83	0.83	0.86	0.86	0.87
$\mathcal{B}_2$	0.20	0.31	0.37	0.46	0.54	0.58	0.63	0.67	0.71	0.74	0.75	0.77	0.80	0.81	0.82
$\mathcal{B}_3$	0.15	0.23	0.30	0.36	0.46	0.51	0.56	0.61	0.66	0.69	0.71	0.73	0.77	0.78	0.79
$\mathcal{B}_4$	0.12	0.16	0.20	0.28	0.38	0.43	0.47	0.54	0.59	0.63	0.66	0.68	0.73	0.74	0.75
$\mathcal{B}_5$		0.11	0.13	0.17	0.27	0.30	0.36	0.42	0.48	0.53	0.56	0.58	0.65	0.66	0.68
$\mathcal{B}_6$			0.11	0.14	0.22	0.26	0.31	0.36	0.43	0.48	0.51	0.54	0.61	0.62	0.64
$\mathcal{B}_7$				0.12	0.17	0.21	0.26	0.31	0.37	0.43	0.46	0.49	0.56	0.57	0.60
$\mathcal{B}_8$					0.13	0.16	0.20	0.26	0.31	0.37	0.39	0.43	0.50	0.52	0.55
$\mathcal{B}_9$						0.13	0.16	0.20	0.25	0.31	0.34	0.37	0.45	0.46	0.50
$\mathcal{B}_{10}$							0.12	0.16	0.20	0.25	0.28	0.32	0.39	0.40	0.44
$\mathcal{B}_{11}$								0.14	0.18	0.23	0.25	0.29	0.36	0.38	0.42
$\mathcal{B}_{12}$									0.16	0.20	0.23	0.25	0.33	0.34	0.38
$\mathcal{B}_{13}$										0.16	0.18	0.21	0.25	0.29	0.32
$\mathcal{B}_{14}$											0.17	0.20	0.26	0.25	0.31
$\mathcal{B}_{15}$												0.18	0.23	0.25	

Table A.5. H^1-stability of $\mathcal{R}$, $r = 1$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	0.63	0.65	0.71	0.75	0.80	0.82	0.83	0.86	0.87	0.89	0.89	0.90	0.92	0.92	0.93
$\mathcal{B}_2$	0.37	0.59	0.57	0.63	0.71	0.74	0.77	0.79	0.82	0.84	0.85	0.86	0.88	0.89	0.90
$\mathcal{B}_3$	0.28	0.40	0.59	0.57	0.66	0.70	0.73	0.77	0.79	0.83	0.83	0.84	0.87	0.87	0.88
$\mathcal{B}_4$	0.21	0.31	0.38	0.58	0.58	0.62	0.67	0.71	0.75	0.78	0.79	0.81	0.84	0.85	0.86
$\mathcal{B}_5$		0.20	0.25	0.33	0.57	0.51	0.56	0.62	0.67	0.71	0.73	0.75	0.79	0.80	0.81
$\mathcal{B}_6$			0.20	0.27	0.39	0.57	0.51	0.57	0.63	0.68	0.70	0.72	0.77	0.78	0.79
$\mathcal{B}_7$				0.23	0.33	0.39	0.56	0.52	0.58	0.63	0.66	0.68	0.73	0.75	0.77
$\mathcal{B}_8$					0.26	0.31	0.37	0.56	0.51	0.58	0.61	0.64	0.70	0.71	0.73
$\mathcal{B}_9$						0.25	0.31	0.38	0.56	0.52	0.55	0.59	0.65	0.67	0.70
$\mathcal{B}_{10}$							0.25	0.31	0.38	0.56	0.49	0.53	0.61	0.62	0.66
$\mathcal{B}_{11}$								0.28	0.35	0.43	0.56	0.50	0.58	0.60	0.64
$\mathcal{B}_{12}$									0.31	0.39	0.42	0.55	0.55	0.57	0.61
$\mathcal{B}_{13}$										0.32	0.35	0.39	0.55	0.53	0.55
$\mathcal{B}_{14}$											0.34	0.38	0.49	0.55	0.54
$\mathcal{B}_{15}$												0.35	0.44	0.46	

Table A.6. H^1-stability of $\mathcal{R}$, $r = 2$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	0.43	0.47	0.55	0.62	0.68	0.72	0.77	0.79	0.82	0.84	0.85	0.86	0.89	0.89	0.90
$\mathcal{B}_2$	0.25	0.40	0.41	0.47	0.60	0.62	0.65	0.70	0.73	0.77	0.78	0.80	0.83	0.84	0.85
$\mathcal{B}_3$	0.20	0.27	0.40	0.40	0.51	0.55	0.60	0.65	0.69	0.73	0.74	0.77	0.80	0.81	0.82
$\mathcal{B}_4$	0.13	0.20	0.26	0.39	0.42	0.48	0.53	0.58	0.63	0.67	0.69	0.71	0.76	0.77	0.78
$\mathcal{B}_5$		0.13	0.16	0.22	0.38	0.36	0.41	0.47	0.53	0.58	0.61	0.63	0.69	0.70	0.72
$\mathcal{B}_6$			0.13	0.18	0.27	0.38	0.37	0.43	0.48	0.54	0.56	0.59	0.65	0.66	0.69
$\mathcal{B}_7$				0.14	0.22	0.26	0.38	0.37	0.43	0.49	0.51	0.55	0.61	0.62	0.65
$\mathcal{B}_8$					0.17	0.21	0.25	0.38	0.37	0.43	0.46	0.49	0.56	0.57	0.60
$\mathcal{B}_9$						0.17	0.20	0.26	0.38	0.37	0.40	0.44	0.51	0.52	0.56
$\mathcal{B}_{10}$							0.16	0.21	0.26	0.37	0.35	0.38	0.45	0.47	0.51
$\mathcal{B}_{11}$								0.19	0.24	0.29	0.37	0.36	0.43	0.44	0.48
$\mathcal{B}_{12}$									0.21	0.26	0.29	0.37	0.40	0.41	0.45
$\mathcal{B}_{13}$										0.21	0.23	0.26	0.27	0.36	0.39
$\mathcal{B}_{14}$											0.22	0.25	0.33	0.37	0.38
$\mathcal{B}_{15}$												0.23	0.30	0.32	

Table A.7. H^1-stability of $\widehat{\mathcal{Q}}$.

from\to	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$	$\mathcal{B}_4$	$\mathcal{B}_5$	$\mathcal{B}_6$	$\mathcal{B}_7$	$\mathcal{B}_8$	$\mathcal{B}_9$	$\mathcal{B}_{10}$	$\mathcal{B}_{11}$	$\mathcal{B}_{12}$	$\mathcal{B}_{13}$	$\mathcal{B}_{14}$	$\mathcal{B}_{15}$
$\mathcal{B}_1$	1.04	1.19	1.37	1.50	1.72	1.81	1.88	1.96	2.03	2.10	2.12	2.14	2.21	2.23	2.25
$\mathcal{B}_2$	0.59	1.01	0.99	1.17	1.43	1.54	1.63	1.74	1.84	1.91	1.95	2.00	2.08	2.09	2.13
$\mathcal{B}_3$	0.46	0.67	0.98	0.99	1.27	1.39	1.49	1.62	1.73	1.82	1.86	1.90	1.99	2.02	2.06
$\mathcal{B}_4$	0.32	0.51	0.64	0.98	1.07	1.20	1.31	1.47	1.59	1.68	1.73	1.78	1.89	1.91	1.96
$\mathcal{B}_5$		0.31	0.40	0.53	0.96	0.90	1.04	1.18	1.32	1.46	1.51	1.59	1.72	1.74	1.80
$\mathcal{B}_6$			0.33	0.44	0.67	0.95	0.91	1.06	1.21	1.35	1.41	1.49	1.63	1.66	1.72
$\mathcal{B}_7$				0.36	0.55	0.66	0.95	0.94	1.08	1.23	1.29	1.37	1.52	1.56	1.63
$\mathcal{B}_8$					0.43	0.52	0.63	0.94	0.92	1.07	1.14	1.23	1.40	1.44	1.51
$\mathcal{B}_9$						0.41	0.51	0.65	0.94	0.93	1.00	1.09	1.27	1.31	1.39
$\mathcal{B}_{10}$							0.41	0.52	0.65	0.94	0.87	0.96	1.13	1.17	1.27
$\mathcal{B}_{11}$								0.47	0.59	0.73	0.93	0.89	1.07	1.11	1.21
$\mathcal{B}_{12}$									0.52	0.66	0.73	0.93	0.99	1.03	1.13
$\mathcal{B}_{13}$										0.52	0.58	0.66	0.93	0.91	0.98
$\mathcal{B}_{14}$											0.55	0.63	0.83	0.93	0.95
$\mathcal{B}_{15}$												0.58	0.74	0.79	

Table A.8. H^1-stability of $\widetilde{\mathcal{R}}'_\ell$.

References

[1] R. A. Adams. *Sobolev Spaces.* Academic Press: New York, 1975.

[2] L. M. Adams, T. P. Chartier. New geometric immersed interface multigrid solvers. *SIAM J. Sci. Comput.* 2004; **25**(5):1516–1533.

[3] L. M. Adams, T. P. Chartier. A comparison of algebraic multigrid and geometric immersed interface multigrid methods for interface problems. *SIAM J. Sci. Comput.* 2005; **26**(3):762–784.

[4] L. M. Adams, Z. Li. The immersed interface/multigrid methods for interface problems. *SIAM J. Sci. Comput.* 2002; **24**(2):463–479.

[5] H. Ainsworth. *Octree C++ General Component.* 2005.

[6] B. Aksoylu, A. Khodakovsky, P. Schröder. Multilevel solvers for unstructured surface meshes. *SIAM J. Sci. Comput.* 2005; **26**(4):1146–1165.

[7] R. E. Alcouffe, A. Brandt, J. E. Dendy jun., J. W. Painter. The multi-grid method for the diffusion equation with strongly discontinuous coefficients. *SIAM J. Sci. Stat. Comput.* 1981; **2**(4):430–454.

[8] T. Apel. Interpolation in h-version finite element spaces. In *Encyclopedia of Computational Mechanics. Vol. 1. Fundamentals,* E. Stein, R. de Borst, T. J. R. Hughes (eds). Wiley: Chichester, 2004; 55–72.

[9] S. Bartels, C. Carstensen, A. Hecht. P2Q2Iso2D = 2D isoparametric FEM in Matlab. *J. Comput. Appl. Math.* 2006; **192**(2):219–250.

[10] I. Babuška, J. M. Melenk. The partition of unity method. *Int. J. Numer. Methods Engrg.* 1997; **40**(4):727–758.

[11] I. Babuška, U. Banerjee, J. E. Osborn. Generalized finite element methods – main ideas, results and perspective. *Int. J. Comput. Methods* 2004; **1**(1):67–103.

[12] N. S. Bakhvalov. On the convergence of a relaxation method with natural constraints on the elliptic operator. *U.S.S.R. Comput. Math. Math. Phys.* 1966; **6**(5):101–135.

[13] R. E. Bank, T. F. Dupont, H. Yserentant. The hierarchical basis multigrid method. *Numer. Math.* 1988; **52**(4):427–458.

[14] R. E. Bank, J. Xu. The hierarchical basis multigrid method and incomplete LU decomposition. In *Domain Decomposition Methods in Scientific and Engineering Computing,* D. E. Keyes, J. Xu (eds). Contemporary Mathematics 180. AMS: Providence, RI, USA, 1994; 163–173.

[15] R. E. Bank, J. Xu. An algorithm for coarsening unstructured meshes. *Numer. Math.* 1996; **73**(1):1–36.

[16] C. M. Barber, D. P. Dobkin, H. Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* 1996; **22**(4):469–483.

[17] P. Bastian, K. Birken, K. Johannsen, S. Lang, N. Neuß, H. Rentz-Reichert, C. Wieners. UG – a flexible software toolbox for solving partial differential equations. *Comput. Vis. Sci.* 1997; **1**(1):27–40.

[18] F. Ben Belgacem. The mortar finite element method with Lagrange multipliers. *Numer. Math.* 1999; **84**(2):173–197.

[19] C. Bernardi, Y. Maday, A. T. Patera. A new nonconforming approach to domain decomposition: the mortar element method. In *Nonlinear Partial Differential Equations and Their Applications,* H. Brezis, J. L. Lions (eds). Pitman Research Notes in Mathematics 299. Harlow: Longman Scientific & Technical: New York, 1994; 13–51.

[20] R. Blaheta. Algebraic multilevel methods with aggregations: An overview. In *Large-Scale Scientific Computing,* I. Lirkov, S. Margenov, J. Waśniewski (eds). Lecture Notes in Computer Science 3743. Springer: Berlin, 2006; 3–14.

[21] F. A. Bornemann, P. Deuflhard. The cascadic multigrid method for elliptic problems. *Numer. Math.* 1996; **75**(2):135–152.

[22] F. A. Bornemann, R. Krause. Classical and cascadic multigrid – a methodological comparison. In *Ninth International Conference on Domain Decomposition Methods,* P. E. Bjørstad, M. S. Espedal, D. E. Keyes (eds). DDM.org, 1998; 64–71.

[23] F. A. Bornemann, H. Yserentant. A basis norm equivalence for the theory of multilevel methods. *Numer. Math.* 1993; **64**(4):455–476.

[24] D. Braess. Towards algebraic multigrid for elliptic problems of second order. *Computing* 1995; **55**(4):379–393.

[25] D. Braess. *Finite Elemente.* Springer: Berlin, 2007.

[26] D. Braess, R. Verfürth. Multigrid methods for nonconforming finite element methods. *SIAM J. Numer. Anal.* 1990; **27**(4):979–986.

[27] J. H. Bramble, J. E. Pasciak. New convergence estimates for multigrid algorithms. *Math. Comp.* 1987; **49**(180):311–329.

[28] J. H. Bramble, J. E. Pasciak. The analysis of smoothers for multigrid algorithms. *Math. Comp.* 1992; **58**(198):467–488.

[29] J. H. Bramble, J. E. Pasciak. New estimates for multilevel algorithms including the V-cycle. *Math. Comp.* 1993; **60**(202):447–471.

[30] J. H. Bramble, J. E. Pasciak, O. Steinbach. On the stability of the L^2-projection in $H^1(\Omega)$. *Math. Comp.* 2002; **71**(237):147–156.

[31] J. H. Bramble, J. E. Pasciak, P. S. Vassilevski. Computational scales of Sobolev norms with applications to preconditioning. *Math. Comp.* 2000; **69**(230):463–480.

[32] J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.* 1991; **57**(195):1–21.

[33] J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu. Convergence estimates for multigrid algorithms without regularity assumptions. *Math. Comp.* 1991; **57**(195):23–45.

[34] J. H. Bramble, J. E. Pasciak, J. Xu. Parallel multilevel preconditioners. *Math. Comp.* 1990; **55**(191):1–22.

[35] J. H. Bramble, J. E. Pasciak, J. Xu. The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms. *Math. Comp.* 1991; **56**(193):1–34.

[36] J. H. Bramble, J. Xu. Some estimates for a weighted L^2 projection. *Math. Comp.* 1991; **56**(194):463–476.

[37] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.* 1977; **31**(138):333–390.

[38] S. C. Brenner. Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems. *Math. Comp.* 2004; **73**(247):1041–1066.

[39] S. C. Brenner, L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Texts in Applied Mathematics 15. Springer: Berlin, 2002.

[40] M. Brezina, A. J. Cleary, R. D. Falgout, V. E. Henson, J. E. Jones, T. A. Manteuffel, S. F. McCormick, J. W. Ruge. Algebraic multigrid based on element interpolation (AMGe). *SIAM J. Sci. Comput.* 2000; **22**(5):1570–1592.

[41] F. Brezzi, W. W. Hager, P. A. Raviart. Error estimates for the finite element solution of variational inequalities. Part I. Primal theory. *Numer. Math.* 1977; **28**(4):431–443.

[42] C. Burstedde, L. C. Wilcox, O. Ghattas. p4est: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees. 2010. (unpublished)

[43] X.-C. Cai. The use of pointwise interpolation in domain decomposition methods with non-nested meshes. *SIAM J. Sci. Comput.* 1995; **16**(1):250–256.

[44] C. Carstensen. Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for H^1-stability of the L^2-projection onto finite element spaces. *Math. Comp.* 2002; **71**(237):157–163.

[45] C. Carstensen. Clément interpolation and its role in adaptive finite element error control. In *Partial Differential Equations and Functional Analysis – The Philippe Clément Festschrift,* E. Koelink, J. van Neerven, B. de Pagter, G. Sweers (eds). Operator Theory: Advances and Applications 168. Birkhäuser: Basel, 2006; 27–43.

[46] J. C. Cavendish. Automatic triangulation of arbitrary planar domains for the finite element method. *Int. J. Numer. Methods Engrg.* 1974; **8**(4):679–696.

[47] T. F. Chan, S. Go, J. Zou. Boundary treatments for multilevel methods on unstructured meshes. *SIAM J. Sci. Comput.* 1999; **21**(1):46–66.

[48] T. F. Chan, B. F. Smith. Domain decomposition and multigrid algorithms for elliptic problems on unstructured meshes. *Electron. Trans. Numer. Anal.* 1994; **2**:171–182.

[49] T. F. Chan, B. F. Smith, J. Zou. Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids. *Numer. Math.* 1996; **73**(2):149–167.

[50] T. F. Chan, J. Xu, L. T. Zikatanov. An agglomeration multigrid method for unstructured grids. In *Domain Decomposition Methods 10,* J. Mandel, C. Farhat, X.-C. Cai (eds). Contemporary Mathematics 218. AMS: Providence, RI, USA, 1998; 67–81.

[51] T. F. Chan, J. Zou. Additive Schwarz domain decomposition methods for elliptic problems on unstructured meshes. *Numer. Algorithms* 1994; **13**(2–4):329–346.

[52] T. F. Chan, J. Zou. A convergence theory of multilevel additive Schwarz methods on unstructured meshes. *Numer. Algorithms* 1996; **13**(3–4):365–398.

[53] T. Chartier, R. D. Falgout, V. E. Henson, J. E. Jones, T. A. Manteuffel, S. F. McCormick, J. W. Ruge, P. S. Vassilevski. Spectral AMGe (ρAMGe). *SIAM J. Sci. Comput.* 2003; **25**(1):1–26.

[54] A. Chernov, M. Maischak, E. P. Stephan. A priori estimates for *hp* penalty BEM for contact problems in elasticity. *Comput. Methods Appl. Mech. Eng.* 2007; **196**(37–40):3871–3880.

[55] D. Cho, L. T. Zikatanov. A multilevel preconditioning for generalized finite element method problems on unstructured simplicial meshes. *J. Numer. Math.* 2007; **15**(3):163–180.

[56] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* Studies in Mathematics and its Applications 4. North-Holland: Amsterdam, 1978.

[57] P. G. Ciarlet. *Mathematical Elasticity, Vol. I: Three-Dimensional Elasticity.* Studies in Mathematics and its Applications 20. North-Holland: Amsterdam, 1988.

[58] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* 1975; **9**(R-2):77–84.

[59] A. Cohen, I. Daubechies, J. C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* 1992; **45**(5):485–560.

[60] M. Crouzeix, V. Thomée. The stability in L^p and W_p^1 of the L^2-projection onto finite element function spaces. *Math. Comp.* 1987; **48**(178):521–532.

[61] CUBIT. Sandia National Laboratories. `http://cubit.sandia.gov/`.

[62] W. Dahmen, S. Dekel, P. Petrushev. Multilevel preconditioning for partition of unity methods: some analytic concepts. *Numer. Math.* 2007; **107**(3):503–532.

[63] W. Dahmen, B. Faermann, I. G. Graham, W. Hackbusch, S. A. Sauter. Inverse inequalities on non-quasi-uniform meshes and application to the mortar element method. *Math. Comp.* 2004; **73**(247):1107–1138.

[64] W. Dahmen, A. Kunoth. Multilevel preconditioning. *Numer. Math.* 1992; **63**(3):315–344.

[65] W. Dahmen, A. Kunoth, K. Urban. Biorthogonal spline wavelets on the interval – stability and moment conditions. *Appl. Comput. Harmon. Anal.* 1999; **6**(2):132–196.

[66] W. Dahmen, C. A. Micchelli, H. P. Seidel. Blossoming begets B-spline bases built better by B-patches. *Math. Comp.* 1992; **59**(199):97–115.

[67] P. M. de Zeeuw. Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *J. Comput. Appl. Math.* 1990; **33**(1):1–27.

[68] P. Deuflhard. Cascadic conjugate gradient methods for elliptic partial differential equations: algorithm and numerical results. In *Domain Decomposition Methods in Scientific and Engineering Computing,* D. E. Keyes, J. Xu (eds). Contemporary Mathematics 180. AMS: Providence, RI, USA, 1994; 29–42.

[69] T. Dickopf. *Konstruktion stabiler Kopplungsoperatoren für Mehrkörperkontaktprobleme und Anwendungen in der Biomechanik.* Diploma thesis, Universität Bonn, 2007.

[70] T. Dickopf, C. Groß, R. Krause, C. Mohr, J. Steiner. Efficient simulation techniques for heterogeneous models in biomechanics. In *Proceedings of the 8th International Symposium on Computer Methods in Biomechanics and Biomedical Engineering, Porto, Portugal,* Arup: Solihull, 2009; ISBN: 978-0-9562121-0-8.

[71] T. Dickopf, R. Krause. Efficient simulation of multi-body contact problems on complex geometries: A flexible decomposition approach using constrained minimization. *Int. J. Numer. Methods Engrg.* 2009; **77**(13):1834–1862.

[72] T. Dickopf, R. Krause. Weak information transfer between non-matching warped interfaces. In *Domain Decomposition Methods in Science and Engineering XVIII*, M. Bercovier, M. J. Gander, R. Kornhuber, O. B. Widlund (eds). Lecture Notes in Computational Science and Engineering 70. Springer: Berlin, 2009; 283–290.

[73] T. Dickopf, R. Krause. A contact method for parametric finite elements. INS Preprint No. 0910, Universität Bonn, September 2009.

[74] T. Dickopf, R. Krause. A study of prolongation operators between non-nested meshes. INS Preprint No. 0913, Universität Bonn, November 2009. (accepted for publication)

[75] Z. Ding. A proof of the trace theorem of Sobolev spaces on Lipschitz domains. *Proc. Am. Math. Soc.* 1996; **124**(2):591–600.

[76] A. Dirks. *Non-Standard Coarse Spaces for Domain Decomposition with Complicated Subdomains.* Diploma thesis, Universität Bonn, 2009.

[77] J. Douglas jun., T. F. Dupont, L. Wahlbin. The stability in L^q of the L^2-projection into finite element function spaces. *Numer. Math.* 1975; **23**(3):193–197.

[78] M. Dryja, M. V. Sarkis, O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.* 1996; **72**(3):313–348.

[79] H. Y. Duan, S. Q. Gao, R. C. E. Tan, S. Zhang. A generalized BPX multigrid framework covering nonnested V-cycle methods. *Math. Comp.* 2007; **76**(257):137–152.

[80] C. Eck, H. Garcke, P. Knabner. *Mathematische Modellierung.* Springer: Berlin, 2008.

[81] C. Engwer. *An Unfitted Discontinuous Galerkin Scheme for Micro-scale Simulations and Numerical Upscaling.* PhD thesis, Universität Heidelberg, 2009.

[82] L. C. Evans. *Partial Differential Equations.* Graduate Studies in Mathematics 19. AMS: Providence, RI, USA, 1998.

[83] K. Fackeldey. *The Weak Coupling Method for Coupling Continuum Mechanics with Molecular Dynamics.* PhD thesis, Universität Bonn, 2009.

[84] R. S. Falk. Error estimates for the approximation of a class of variational inequalities. *Math. Comp.* 1974; **28**(128):963–971.

[85] S. Falletta. The approximate integration in the mortar method constraint. In *Domain Decomposition Methods in Science and Engineering XVI,* O. B. Widlund, D. E. Keyes (eds). Lecture Notes in Computational Science and Engineering 55. Springer: Berlin, 2007; 555–563.

[86] R. P. Fedorenko. A relaxation method for solving elliptic difference equations. *U.S.S.R. Comput. Math. Math. Phys.* 1962; **1**(4):1092–1096.

[87] R. P. Fedorenko. The speed of convergence of one iterative process. *U.S.S.R. Comput. Math. Math. Phys.* 1964; **4**(3):227–235.

[88] B. Flemisch, B. Wohlmuth. Stable Lagrange multipliers for quadrilateral meshes of curved interfaces in 3D. *Comput. Methods Appl. Mech. Eng.* 2007; **196**(8):1589–1602.

[89] M. J. Gander, C. Japhet. An algorithm for non-matching grid projections with linear complexity. In *Domain Decomposition Methods in Science and Engineering XVIII*, M. Bercovier, M. J. Gander, R. Kornhuber, O. B. Widlund (eds). Lecture Notes in Computational Science and Engineering 70. Springer: Berlin, 2009; 185–192.

[90] B. Gmeiner, T. Gradl, F. Gaspar, U. Rüde. Validation and optimization of the convergence rate on semi-structured meshes using the LFA. Talk at 11th Copper Mountain Conference on Iterative Methods, 2010.

[91] S. Go. *Multilevel Methods on Unstructured Grids.* PhD thesis, University of California, Los Angeles, 1999.

[92] G. H. Golub, C. F. Van Loan. *Matrix computations.* The Johns Hopkins University Press: Baltimore, MD, USA, 1983.

[93] W. J. Gordon, C. A. Hall. Transfinite element methods: blending-function interpolation over arbitrary curved element domains. *Numer. Math.* 1973; **21**(2):109–129.

[94] S. Gratton, A. Sartenaer, P. L. Toint. Recursive trust-region methods for multiscale nonlinear optimization. *SIAM J. Optim.* 2008; **19**(1):414–444.

[95] M. Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen.* Teubner Skripten zur Numerik. Teubner: Stuttgart, 1994.

[96] M. Griebel. Multilevel algorithms considered as iterative methods on semidefinite systems. *SIAM J. Sci. Comput.* 1994; **15**(3):547–565.

[97] M. Griebel, P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.* 1995; **70**(2):163–180.

[98] M. Griebel, K. Scherer, M. A. Schweitzer. Robust norm equivalencies for diffusion problems. *Math. Comp.* 2007; **76**(259):1141–1161.

[99] M. Griebel, M. A. Schweitzer. A particle-partition of unity method. Part III: A multilevel solver. *SIAM J. Sci. Comput.* 2002; **24**(2):377–409.

[100] C. Groß, R. Krause. Import of geometries and extended information into Obslib++ using the Exodus II and Exodus parameter file formats. INS Preprint No. 0712, Universität Bonn, January 2008.

[101] C. Groß, R. Krause. On the convergence of recursive trust-region methods for multiscale non-linear optimization and applications to non-linear mechanics. *SIAM J. Numer. Anal.* 2009; **47**(4):3044–3069.

[102] H. Guillard. Node-nested multigrid method with Delaunay coarsening. Rapports de Recherche No. 1898, Institut National de Recherche en Informatique et en Automatique, March 1993.

[103] W. Hackbusch. *Multi-Grid Methods and Applications.* Springer Series in Computational Mathematics 4. Springer: Berlin, 1985.

[104] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations.* Applied Mathematical Sciences 95. Springer: Berlin, 1994.

[105] W. Hackbusch, S. A. Sauter. Composite finite elements for the approximation of PDEs on domains with complicated micro-structures. *Numer. Math.* 1997; **75**(4):447–472.

[106] W. Hackbusch, S. A. Sauter. Composite finite elements for problems containing small geometric details. II: Implementation and numerical results. *Comput. Vis. Sci.* 1997; **1**(1):15–25.

[107] H. Harbrecht. A finite element method for elliptic problems with stochastic input data. *Appl. Numer. Math.* 2010; **60**(3):227–244.

[108] H. Harbrecht, M. Randrianarivony. From computer aided design to wavelet BEM. *Comput. Visual. Sci.* 2010; **13**(2):69–82.

[109] V. E. Henson, P. S. Vassilevski. Element-free AMGe: General algorithms for computing interpolation weights in AMG. *SIAM J. Sci. Comput.* 2001; **23**(2):629–650.

[110] H. Hertz. Über die Berührung fester elastischer Körper. *Journal für die reine und angewandte Mathematik* 1881; **92**:156–171.

[111] M. R. Hestenes, E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand.* 1952; **49**(6):409–436.

[112] R. Hiptmair, G. Widmer, J. Zou. Auxiliary space preconditioning in $H_0(\mathrm{curl};\Omega)$. *Numer. Math.* 2006; **103**(3):435–459.

[113] R. Hiptmair, J. Xu. Nodal auxiliary space preconditioning in $H(\mathrm{curl})$ and $H(\mathrm{div})$ spaces. *SIAM J. Numer. Anal.* 2007; **45**(6):2483–2509.

[114] T. J. R. Hughes, J. A. Cottrell, Y. Bazilevs. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput. Methods Appl. Mech. Eng.* 2005; **194**(39–41):4135–4195.

[115] J. E. Jones, P. S. Vassilevski. AMGe based on element agglomeration. *SIAM J. Sci. Comput.* 2001; **23**(1):109–133.

[116] B. N. Khoromskij, J. M. Melenk. Boundary concentrated finite element methods. *SIAM J. Numer. Anal.* 2003; **41**(1):1–36.

[117] N. Kikuchi, J. T. Oden. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods.* Studies in Applied Mathematics 8. SIAM: Philadelphia, PA, USA, 1988.

[118] C. Kim, R. D. Lazarov, J. E. Pasciak, P. S. Vassilevski. Multiplier spaces for the mortar finite element method in three dimensions. *SIAM J. Numer. Anal.* 2001; **39**(2):519–538.

[119] D. Kinderlehrer, G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications.* Pure and Applied Mathematics 88. Academic Press: New York, 1980.

[120] B. Koobus, M.-H. Lallemand, A. Dervieux. Unstructured volume-agglomeration MG: Solution of the Poisson equation. *Int. J. Numer. Methods Fluids* 1994; **18**(1):27–42.

[121] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems.* Teubner: Stuttgart, 1997.

[122] R. Kornhuber, R. Krause. Adaptive multigrid methods for Signorini's problem in linear elasticity. *Comput. Visual. Sci.* 2001; **4**(1):9–20.

[123] R. Kornhuber, H. Yserentant. Multilevel methods for elliptic problems on domains not resolved by the coarse grid. In *Domain Decomposition Methods in Scientific and Engineering Computing,* D. E. Keyes, J. Xu (eds). Contemporary Mathematics 180. AMS: Providence, RI, USA, 1994; 49–60.

[124] R. Krause. *Monotone Multigrid Methods for Signorini's Problem with Friction.* PhD thesis, Freie Universität Berlin, 2001.

[125] R. Krause. From inexact active set strategies to nonlinear multigrid methods. In *Analysis and Simulation of Contact Problems,* P. Wriggers, U. Nackenhorst (eds). Lecture Notes in Applied and Computational Mechanics 27. Springer: Berlin, 2006; 13–21.

[126] R. Krause. A nonsmooth multiscale method for solving frictional two-body contact problems in 2d and 3d with multigrid efficiency. *SIAM J. Sci. Comput.* 2009; **31**(2):1399–1423.

[127] R. Krause, C. Mohr. Level set based multi-scale methods for large deformation contact problems. ICS Preprint No. 2009-01, University of Lugano, April 2009.

[128] R. Krause, M. Walloth. A time discretization scheme based on Rothe's method for dynamical contact problems with friction. *Comput. Methods Appl. Mech. Eng.* 2009; **199**(1–4):1–19.

[129] R. Krause, B. Wohlmuth. Nonconforming domain decomposition techniques for linear elasticity. *East-West J. Numer. Math.* 2000; **8**(3):177–206.

[130] M.-H. Lallemand, H. Steve, A. Dervieux. Unstructured multigridding by volume agglomeration: Current status. *Comput. Fluids* 1992; **21**(3):397–433.

[131] B. P. Lamichhane. *Higher Order Mortar Finite Elements with Dual Lagrange Multiplier Spaces and Applications.* PhD thesis, Universität Stuttgart, 2006.

[132] B. P. Lamichhane, B. Wohlmuth. Biorthogonal bases with local support and approximation properties. *Math. Comp.* 2007; **76**(257):233–249.

[133] B. P. Lamichhane. A mixed finite element method for non-linear and nearly incompressible elasticity based on biorthogonal systems. *Int. J. Numer. Methods Engrg.* 2009; **79**(7):870–886.

[134] C. Lasser, A. Toselli. Convergence of some two-level overlapping domain decomposition preconditioners with smoothed aggregation coarse spaces. In *Recent Developments in Domain Decomposition Methods,* L. F. Pavarino, A. Toselli (eds). Lecture Notes in Computational Science and Engineering 23. Springer: Berlin, 2002; 95–117.

[135] T. A. Laursen. *Computational Contact and Impact Mechanics.* Springer: Berlin, 2002.

[136] R. J. LeVeque, Z. Li. The immersed interface method for elliptic equations with discontinuous coefficients and singular sources. *SIAM J. Numer. Anal.* 1994; **31**(4):1019–1044.

[137] F. Liehr, T. Preusser, M. Rumpf, S. A. Sauter, L. O. Schwen. Composite finite elements for 3D image based computing. *Comput. Vis. Sci.* 2009; **12**(4):171–188.

[138] J. L. Lions, G. Stampacchia. Variational inequalities. *Commun. Pure Appl. Math.* 1967; **20**(3):493–519.

[139] R. Löhner, K. Morgan. An unstructured multigrid method for elliptic problems. *Int. J. Numer. Methods Engrg.* 1987; **24**(1):101–115.

[140] X. J. Luo, M. S. Shephard, J. F. Remacle, R. M. O'Bara, M. W. Beall, B. Szabó, R. Actis. *p*-version mesh generation issues. In *Proceedings of the 11th International Meshing Roundtable,* 2002; 343–354.

[141] Y. Maday, F. Rapetti, B. Wohlmuth. The influence of quadrature formulas in 2d and 3d mortar element methods. In *Recent Developments in Domain Decomposition Methods,* L. F. Pavarino, A. Toselli (eds). Lecture Notes in Computational Science and Engineering 23. Springer: Berlin, 2002; 203–221.

[142] J. Mandel, S. F. McCormick, R. E. Bank. Variational multigrid methods. In *Multigrid Methods,* S. F. McCormick (ed). Frontiers in Applied Mathematics 3. SIAM: Philadelphia, PA, USA, 1987; 131–177.

[143] D. J. Mavriplis, A. Jameson. Multigrid solution of the two-dimensional Euler equations on unstructured triangular meshes. Preprint AIAA-87-0353, American Institute of Aeronautics and Astronautics, January 1987.

[144] D. J. Mavriplis, A. Jameson. Multigrid solution of the Euler equations on unstructured and adaptive meshes. In *Multigrid methods,* S. F. McCormick (ed). Lecture Notes in Pure and Applied Mathematics 110. Marcel Dekker: New York, 1988; 413–429.

[145] J. M. Melenk, I. Babuška. The partition of unity finite element method: Basic theory and applications. *Comput. Methods Appl. Mech. Eng.* 1996; **139**(1–4):289–314.

[146] G. L. Miller, D. Talmor, S.-H. Teng. Optimal coarsening of unstructured meshes. *J. Algorithms* 1999; **31**(1):29–65.

[147] S. G. Nash. A multigrid approach to discretized optimization problems. *Optim. Methods Softw.* 2000; **14**(1–2):99–116.

[148] Y. Notay. An aggregation-based algebraic multigrid method. *Electron. Trans. Numer. Anal.* 2010; **37**:123–146.

[149] C. Ollivier-Gooch. Coarsening unstructured meshes by edge contraction. *Int. J. Numer. Methods Engrg.* 2003; **57**(3):391–414.

[150] P. Oswald. On function spaces related to finite element approximation theory. *Z. Anal. Anwend.* 1990; **9**(1):43–64.

[151] P. Oswald. *Multilevel Finite Element Approximation. Theory and Applications.* Teubner Skripten zur Numerik. Teubner: Stuttgart, 1994.

[152] P. Oswald. Preconditioners for nonconforming discretizations. *Math. Comp.* 1996; **65**(215):923–941.

[153] P. Oswald. Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations. *Appl. Numer. Math.* 1997; **23**(1):139–158.

[154] P. Oswald. On the robustness of the BPX-preconditioner with respect to jumps in the coefficients. *Math. Comp.* 1999; **68**(226):633–650.

[155] P. Oswald. Subspace correction methods and multigrid theory. In *Multigrid,* U. Trottenberg, C. W. Oosterlee, A. Schüller. Academic Press: Orlando, FL, USA, 2001; 533–572.

[156] P. Oswald, B. Wohlmuth. On polynomial reproduction of dual FE bases. In *Thirteenth International Conference on Domain Decomposition Methods,* N. Debit, M. Garbey, R. Hoppe, J. Périaux, D. E. Keyes, Y. Kuznetsov (eds). CIMNE: Barcelona, 2002; 85–96.

[157] T. Preusser, M. Rumpf, L. O. Schwen. Finite element simulation of bone microstructures. In *Proceedings of the 14th Workshop on the Finite Element Method in Biomedical Engineering, Biomechanics and Related Fields.* University of Ulm, 2007; 52–66.

[158] M. Randrianarivony. Tetrahedral transfinite interpolation with B-patch faces: construction and regularity. INS Preprint No. 0803, Universität Bonn, June 2008.

[159] J. W. Ruge, K. Stüben. Algebraic multigrid. In *Multigrid Methods,* S. F. McCormick (ed). Frontiers in Applied Mathematics 3. SIAM: Philadelphia, PA, USA, 1987; 73–130.

[160] M. V. Sarkis. Partition of unity coarse spaces. In *Fluid Flow and Transport in Porous Media: Mathematical and Numerical Treatment,* Z. Chen, R. Ewing (eds). Contemporary Mathematics 295. AMS: Providence, RI, USA, 2002; 445–456.

[161] M. V. Sarkis. Partition of unity coarse spaces and Schwarz methods with harmonic overlap. In *Recent Developments in Domain Decomposition Methods,* L. F. Pavarino, A. Toselli (eds). Lecture Notes in Computational Science and Engineering 23. Springer: Berlin, 2002; 77–94.

[162] M. V. Sarkis. Partition of unity coarse spaces: enhanced versions, discontinuous coefficients and applications to elasticity. In *Domain Decomposition Methods in Science and Engineering,* I. Herrera, D. E. Keyes, O. B. Widlund, R. Yates (eds). UNAM: Mexico City, 2003; 149–158.

[163] S. A. Sauter. Composite finite elements for problems with complicated boundary. Part III: Essential boundary conditions. Technical Report 97-16, Universität Kiel, 1997.

[164] R. Scheichl, E. Vainikko. Additive Schwarz with aggregation-based coarsening for elliptic problems with highly variable coefficients. *Computing* 2007; **80**(4):319–343.

[165] O. Schenk, K. Gärtner. Solving unsymmetric sparse systems of linear equations with PARDISO. *Future Generation Computer Systems* 2004; **20**(3):475–487.

[166] O. Schenk, K. Gärtner. On fast factorization pivoting methods for sparse symmetric indefinite systems. *Electron. Trans. Numer. Anal.* 2006; **23**:158–179.

[167] K. Scherer. Weighted norm-equivalences for preconditioning. In *Domain Decomposition Methods in Science and Engineering,* R. Kornhuber, R. Hoppe, J. Périaux, O. Pironneau, O. B. Widlund, J. Xu (eds). Lecture Notes in Computational Science and Engineering 40. Springer: Berlin, 2005; 405–413.

[168] K. Scherer. Robust norm equivalencies and preconditioning. In *Domain Decomposition Methods in Science and Engineering XVII,* U. Langer, M. Discacciati, D. E. Keyes, O. B. Widlund, W. Zulehner (eds). Lecture Notes in Computational Science and Engineering 60. Springer: Berlin, 2008; 373–380.

[169] H. A. Schwarz. Über einen Grenzübergang durch alternierendes Verfahren. *Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich* 1870; **15**:272–286.

[170] M. A. Schweitzer. *A Parallel Multilevel Partition of Unity Method for Elliptic Partial Differential Equations.* Lecture Notes in Computational Science and Engineering 29. Springer: Berlin, 2003.

[171] L. R. Scott, S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* 1990; **54**(190):483–493.

[172] L. R. Scott, S. Zhang. Higher-dimensional nonnested multigrid methods. *Math. Comp.* 1992; **58**(198):457–466.

[173] P. Seshaiyer, M. Suri. Uniform *hp* convergence results for the mortar finite element method. *Math. Comp.* 2000; **69**(230):521–546.

[174] O. Steinbach. On a generalized L^2-projection and some related stability estimates in Sobolev spaces. *Numer. Math.* 2002; **90**(4):775–786.

[175] O. Steinbach. *Stability Estimates for Hybrid Coupled Domain Decomposition Methods.* Lecture Notes in Mathematics 1809. Springer: Berlin, 2003.

[176] K. Stüben. An introduction to algebraic multigrid. In *Multigrid,* U. Trottenberg, C. W. Oosterlee, A. Schüller. Academic Press: Orlando, FL, USA, 2001; 413–532.

[177] A. Toselli, O. B. Widlund. *Domain Decomposition Methods – Algorithms and Theory.* Springer Series in Computational Mathematics 34. Springer: Berlin, 2005.

[178] U. Trottenberg, C. W. Oosterlee, A. Schüller. *Multigrid.* Academic Press: Orlando, FL, USA, 2001.

[179] R. S. Tuminaro, J. Xu, Y. Zhu. Auxiliary space preconditioners for mixed finite element methods. In *Domain Decomposition Methods in Science and Engineering XVIII*, M. Bercovier, M. J. Gander, R. Kornhuber, O. B. Widlund (eds). Lecture Notes in Computational Science and Engineering 70. Springer: Berlin, 2009; 99–109.

[180] P. Vaněk. Acceleration of convergence of a two-level algorithm by smoothing transfer operators. *Appl. Math., Praha* 1992; **37**(4):265–274.

[181] P. Vaněk, M. Brezina, J. Mandel. Convergence of algebraic multigrid based on smoothed aggregation. *Numer. Math.* 2001; **88**(3):559–579.

[182] P. Vaněk, J. Mandel, M. Brezina. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing* 1996; **56**(3):179–196.

[183] V. Venkatakrishnan, D. J. Mavriplis. Agglomeration multigrid for the three-dimensional Euler equations. *AIAA J.* 1995; **33**(4):633–640.

[184] R. Verfürth. Error estimates for some quasi-interpolation operators. *Math. Model. Numer. Anal.* 1999; **33**(4):695–713.

[185] J. W. L. Wan, X.-D. Liu. A boundary condition-capturing multigrid approach to irregular boundary problems. *SIAM J. Sci. Comput.* 2004; **25**(6):1982–2003.

[186] H. F. Wang, M. P. Anderson. *Introduction to Groundwater Modeling.* Freeman: San Francisco, CA, USA, 1982.

[187] O. B. Widlund. Some Schwarz methods for symmetric and nonsymmetric elliptic problems. In *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations,* D. E. Keyes, T. F. Chan, G. Meurant, J. S. Scroggs, R. G. Voigt (eds). SIAM: Philadelphia, PA, USA, 1992; 19–36.

[188] O. B. Widlund. Accomodating irregular subdomains in domain decomposition theory. In *Domain Decomposition Methods in Science and Engineering XVIII*, M. Bercovier, M. J. Gander, R. Kornhuber, O. B. Widlund (eds). Lecture Notes in Computational Science and Engineering 70. Springer: Berlin, 2009; 87–98.

[189] O. B. Widlund. The development of coarse spaces for domain decomposition algorithms. In *Domain Decomposition Methods in Science and Engineering XVIII*, M. Bercovier, M. J. Gander, R. Kornhuber, O. B. Widlund (eds). Lecture Notes in Computational Science and Engineering 70. Springer: Berlin, 2009; 241–248.

[190] B. Wohlmuth. A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.* 2000; **38**(3)989–1012.

[191] B. Wohlmuth. *Discretization Methods and Iterative Solvers Based on Domain Decomposition.* Lecture Notes in Computational Science and Engineering 17. Springer: Berlin, 2001.

[192] P. Wriggers. *Computational Contact Mechanics.* Wiley: Chichester, 2002.

[193] J. Xu. *Theory of Multilevel Methods.* PhD thesis, Cornell University, 1989.

[194] J. Xu. Iterative methods by subspace decomposition and subspace correction. *SIAM Rev.* 1992; **34**(4):581–613.

[195] J. Xu. The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids. *Computing* 1996; **56**(3):215–235.

[196] D. M. Young. *Iterative solution of large linear systems.* Academic Press: New York, 1971.

[197] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.* 1986; **49**(4):379–412.

[198] H. Yserentant. Two preconditioners based on the multi-level splitting of finite element spaces. *Numer. Math.* 1990; **58**(2):163–184.

[199] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica* 1993; **2**:285–326.

[200] H. Yserentant. Coarse grid spaces for domains with a complicated boundary. *Numer. Algorithms* 1999; **21**(1–4):387–392.

[201] S. Zhang. Optimal-order nonnested multigrid methods for solving finite element equations I: On quasi-uniform meshes. *Math. Comp.* 1990; **55**(191):23–36.

[202] S. Zhang. Optimal-order nonnested multigrid methods for solving finite element equations II: On non-quasi-uniform meshes. *Math. Comp.* 1990; **55**(192):439–450.

[203] S. Zhang. Optimal-order nonnested multigrid methods for solving finite element equations III: On degenerate meshes. *Math. Comp.* 1995; **64**(209):23–49.

[204] X. Zhang. Multilevel Schwarz methods. *Numer. Math.* 1992; **63**(4):521–539.

[205] M. Zlámal. The finite element method in domains with curved boundaries. *Int. J. Numer. Methods Engrg.* 1973; **5**(3):367–373.