Chapter 1

Basic Concepts

This chapter introduces the basic concepts to build discontinuous Galerkin (dG) approximations for the model problems examined in the subsequent chapters. In Sect. 1.1, we present two important results to assert the well-posedness of linear model problems, namely the so-called Banach–Nečas–Babuška (BNB) Theorem, which provides necessary and sufficient conditions for well-posedness, and the Lax–Milgram Lemma, which hinges on coercivity in a Hilbertian setting and provides sufficient conditions for well-posedness. We also state some basic results on Lebesgue and Sobolev spaces. In Sect. 1.2, we describe the main ideas to build finite-dimensional function spaces in the dG setting. The two ingredients are discretizing the domain Ω over which the model problem is posed using a mesh and then choosing a local functional behavior within each mesh element. For simplicity, we focus on a polynomial behavior, thereby leading to so-called broken polynomial spaces. We also introduce important concepts to be used extensively in this book, including mesh faces, jump and average operators, broken Sobolev spaces, and a broken gradient operator. In Sect. 1.3, we outline the key ingredients in the error analysis of nonconforming finite element methods. The error estimates are derived in the spirit of the Second Strang Lemma using discrete stability, (strong) consistency, and boundedness. The advantage of this approach is to deliver error estimates and (quasi-)optimal convergence rates for smooth solutions. This framework for error analysis is frequently used in what follows. Yet, it is not the only tool for analyzing the convergence of dG approximations. In Chaps. 5 and 6, we consider an alternative approach in the context of PDEs with diffusion based on a compactness argument and a different notion of consistency. This approach allows us to prove convergence (without delivering error estimates) with minimal regularity assumptions on the exact solution. Finally, in Sect. 1.4, we present technical, yet important, tools to analyze the convergence of dG methods as the meshsize goes to zero (the so-called *h-convergence*). A crucial issue is then to ensure that some key properties of the mesh hold uniformly in this limit, thereby leading to the important concept of admissible mesh sequences.

D.A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Mathématiques et Applications 69, DOI 10.1007/978-3-642-22980-0_1,

1.1 Well-Posedness for Linear Model Problems

Let X and Y be two Banach spaces equipped with their respective norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ and assume that Y is reflexive. In many applications, X and Y are actually Hilbert spaces. We recall that $\mathcal{L}(X,Y)$ is the vector space spanned by bounded linear operators from X to Y, and that this space is equipped with the usual norm

$$\|A\|_{\mathcal{L}(X,Y)} := \sup_{v\in X\setminus\{0\}} \frac{\|Av\|_Y}{\|v\|_X} \qquad \forall A \in \mathcal{L}(X,Y).$$

We are interested in the abstract linear model problem

$$\text{Find } u \in X \text{ s.t. } a(u,w) = \langle f, w\rangle_{Y',Y} \text{ for all } w \in Y, \tag{1.1}$$

where $a \in \mathcal{L}(X\times Y,\mathbb{R})$ is a bounded bilinear form, $f \in Y' := \mathcal{L}(Y,\mathbb{R})$ is a bounded linear form, and $\langle\cdot,\cdot\rangle_{Y',Y}$ denotes the duality pairing between Y' and Y. Alternatively, it is possible to introduce the bounded linear operator $A \in \mathcal{L}(X,Y')$ such that

$$\langle Av, w\rangle_{Y',Y} := a(v,w) \qquad \forall (v,w) \in X\times Y, \tag{1.2}$$

and to consider the following problem:

$$\text{Find } u \in X \text{ s.t. } Au = f \text{ in } Y'. \tag{1.3}$$

Problems (1.1) and (1.3) are equivalent, that is, u solves (1.1) if and only if u solves (1.3).

Problem (1.1), or equivalently (1.3), is said to be *well-posed* if it admits one and only one solution $u \in X$. The well-posedness of problem (1.3) amounts to A being an isomorphism. In Banach spaces, if $A \in \mathcal{L}(X,Y')$ is an isomorphism, then A^{-1} is bounded, that is, $\|A^{-1}\|_{\mathcal{L}(Y',X)} \le C$ (see, e.g., Ern and Guermond [141, Remark A.37]). As a result, the unique solution $u \in X$ satisfies the a priori estimate

$$\|u\|_X = \|A^{-1}f\|_X \le C\|f\|_{Y'}.$$

1.1.1 The Banach–Nečas–Babuška Theorem

The key result for asserting the well-posedness of (1.1), or equivalently (1.3), is the so-called Banach–Nečas–Babuška (BNB) Theorem. We stress that this result provides *necessary and sufficient* conditions for well-posedness.

Theorem 1.1 (Banach–Nečas–Babuška (BNB))**.** *Let X be a Banach space and let Y be a reflexive Banach space. Let $a \in \mathcal{L}(X\times Y,\mathbb{R})$ and let $f \in Y'$. Then, problem (1.1) is well-posed if and only if:*

(i) *There is $C_{\mathrm{sta}} > 0$ such that*

$$\forall v \in X, \qquad C_{\mathrm{sta}}\|v\|_X \le \sup_{w\in Y\setminus\{0\}} \frac{a(v,w)}{\|w\|_Y}, \tag{1.4}$$

(ii) *For all $w \in Y$,*

$$(\forall v \in X,\ a(v,w) = 0) \Longrightarrow (w = 0). \tag{1.5}$$

Equivalently, the bounded linear operator $A \in \mathcal{L}(X, Y')$ defined by (1.2) is an isomorphism if and only if:

(i) *There is $C_{\mathrm{sta}} > 0$ such that*

$$\forall v \in X, \qquad C_{\mathrm{sta}}\|v\|_X \le \|Av\|_{Y'}, \tag{1.6}$$

(ii) *For all $w \in Y$,*

$$(\forall v \in X,\ \langle Av, w\rangle_{Y',Y} = 0) \Longrightarrow (w = 0). \tag{1.7}$$

Moreover, the following a priori estimate holds true:

$$\|u\|_X \le \frac{1}{C_{\mathrm{sta}}}\|f\|_{Y'}.$$

Condition (1.4) is often called an inf-sup condition since it is equivalent to

$$C_{\mathrm{sta}} \le \inf_{v \in X\setminus\{0\}} \sup_{w \in Y\setminus\{0\}} \frac{a(v,w)}{\|v\|_X\|w\|_Y}.$$

Furthermore, owing to the reflexivity of Y and introducing the adjoint operator $A^t \in \mathcal{L}(Y, X')$ such that, for all $(v,w) \in X \times Y$,

$$\langle A^t w, v\rangle_{X',X} = \langle Av, w\rangle_{Y',Y},$$

condition (1.7) means that, for all $w \in Y$, $A^t w = 0$ in X' implies $w = 0$, or equivalently that A^t is injective. Moreover, a classical result (see, e.g., [141, Lemma A.39]) is that condition (1.6) means that A is injective and that the range of A is closed.

Remark 1.2 (Name of Theorem 1.1). The terminology, proposed in [141], indicates that, from a functional analysis point of view, this theorem hinges on two key results of Banach, the Open Mapping Theorem and the Closed Range Theorem, while the theorem in the form stated above was derived by Nečas [243] and Babuška [21].

1.1.2 The Lax–Milgram Lemma

A simpler, yet less general, condition to assert the well-posedness of (1.1), or equivalently (1.3), is provided by the Lax–Milgram Lemma [222]. In this setting, X is a Hilbert space, $Y = X$, and the following coercivity property is invoked.

Definition 1.3 (Coercivity)**.** Let X be a Hilbert space and let $a \in \mathcal{L}(X \times X, \mathbb{R})$. We say that the bilinear form a is *coercive* on X if there is $C_{\mathrm{sta}} > 0$ such that

$$\forall v \in X, \qquad C_{\mathrm{sta}}\|v\|_X^2 \le a(v,v).$$

Equivalently, we say that the bounded linear operator $A \in \mathcal{L}(X, X')$ defined by (1.2) is *coercive* if there is $C_{\mathrm{sta}} > 0$ such that

$$\forall v \in X, \qquad C_{\mathrm{sta}} \|v\|_X^2 \leq \langle Av, v \rangle_{X',X}.$$

We can now state the Lax–Milgram Lemma. We stress that this result only provides *sufficient* conditions for well-posedness.

Lemma 1.4 (Lax–Milgram)**.** *Let X be a Hilbert space, let $a \in \mathcal{L}(X \times X, \mathbb{R})$, and let $f \in X'$. Then, problem (1.1) is well-posed if the bilinear form a is coercive on X. Equivalently, problem (1.3) is well-posed if the linear operator $A \in \mathcal{L}(X, X')$ is coercive. Moreover, the following a priori estimate holds true:*

$$\|u\|_X \leq \frac{1}{C_{\mathrm{sta}}} \|f\|_{X'}.$$

Proof. Let us verify that if a is coercive, conditions (1.4) and (1.5) hold true. To prove (1.4), we observe that, for all $v \in X \setminus \{0\}$,

$$C_{\mathrm{sta}} \|v\|_X \leq \frac{a(v,v)}{\|v\|_X} \leq \sup_{w \in X \setminus \{0\}} \frac{a(v,w)}{\|w\|_X},$$

and that (1.4) trivially holds true if $v = 0$. To prove (1.5), let $w \in X$ be such that $a(v,w) = 0$, for all $v \in X$. Then, picking $v = w$ yields by coercivity that $\|w\|_X = 0$, i.e., $w = 0$.

□

Remark 1.5 (Lax–Milgram Lemma and Hilbert spaces). Coercivity is essentially a Hilbertian property. Precisely, if X is a Banach space, then X can be equipped with a Hilbert structure with the same topology if and only if there is a coercive operator in $\mathcal{L}(X, X')$; see, e.g., [141, Proposition A.49].

1.1.3 Lebesgue and Sobolev Spaces

In practice, the model problem (1.1) corresponds to a PDE posed over a domain $\Omega \subset \mathbb{R}^d$ with space dimension $d \geq 1$. The domain Ω is a bounded, connected, open subset of $\mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$. The spaces X and Y in (1.1) are then function spaces spanned by functions defined over Ω. For simplicity, we consider scalar-valued functions; the case of vector-valued functions can be treated similarly.

In this section, we briefly present two important classes of function spaces to be used in what follows, namely Lebesgue and Sobolev spaces. We only state the basic properties of such spaces, and we refer the reader to Evans [153, Chap. 5] or Brézis [55, Chaps. 8 and 9] for further background. A thorough presentation can also be found in the textbook of Adams [4].

1.1.3.1 Lebesgue Spaces

We consider functions $v : \Omega \to \mathbb{R}$ that are Lebesgue measurable and we denote by $\int_\Omega v$ the (Lebesgue) integral of v over Ω. Let $1 \le p \le \infty$ be a real number. We set

$$\|v\|_{L^p(\Omega)} := \left(\int_\Omega |v|^p \right)^{1/p} \qquad 1 \le p < \infty,$$

and

$$\begin{aligned} \|v\|_{L^\infty(\Omega)} &:= \operatorname{sup\,ess}\{|v(x)| \text{ for a.e. } x \in \Omega\} \\ &= \inf\{M > 0 \mid |v(x)| \le M \text{ for a.e. } x \in \Omega\}. \end{aligned}$$

In either case, we define the *Lebesgue space*

$$L^p(\Omega) := \{v \text{ Lebesgue measurable} \mid \|v\|_{L^p(\Omega)} < \infty\}.$$

Equipped with the norm $\|\cdot\|_{L^p(\Omega)}$, $L^p(\Omega)$ is a Banach space for all $1 \le p \le \infty$ (see Evans [153, p. 249] or Brézis [55, p. 150]). Moreover, for all $1 \le p < \infty$, the space $C_0^\infty(\Omega)$ spanned by infinitely differentiable functions with compact support in Ω is dense in $L^p(\Omega)$. In the particular case $p = 2$, $L^2(\Omega)$ is a (real) Hilbert space when equipped with the scalar product

$$(v, w)_{L^2(\Omega)} := \int_\Omega vw.$$

A useful tool in Lebesgue spaces is *Hölder's inequality* which states that, for all $1 \le p, q \le \infty$ such that $1/p + 1/q = 1$, all $v \in L^p(\Omega)$, and all $w \in L^q(\Omega)$, there holds $vw \in L^1(\Omega)$ and

$$\int_\Omega vw \le \|v\|_{L^p(\Omega)} \|w\|_{L^q(\Omega)}.$$

The particular case $p = q = 2$ yields the Cauchy–Schwarz inequality, namely, for all $v, w \in L^2(\Omega)$, $vw \in L^1(\Omega)$ and

$$(v, w)_{L^2(\Omega)} \le \|v\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)}.$$

1.1.3.2 Sobolev Spaces

On the Cartesian basis of $\mathbb{R}^d$ with coordinates $(x_1, \ldots, x_d)$, the symbol ∂_i with $i \in \{1, \ldots, d\}$ denotes the distributional partial derivative with respect to x_i. For a d-uple $\alpha \in \mathbb{N}^d$, $\partial^\alpha v$ denotes the distributional derivative $\partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d} v$ of v, with the convention that $\partial^{(0,\ldots,0)} v = v$. For any real number $1 \le p \le \infty$, we define, for all $\xi \in \mathbb{R}^d$ with components $(\xi_1, \ldots, \xi_d)$ in the Cartesian basis of $\mathbb{R}^d$, the norm

$$|\xi|_{\ell^p} := \left(\sum_{i=1}^d |\xi_i|^p \right)^{1/p} \qquad 1 \le p < \infty,$$

and $|\xi|_{\ell^\infty} := \max_{1\le i\le d} |\xi_i|$. The index is dropped for the Euclidean norm obtained with $p = 2$.

Let $m \ge 0$ be an integer and let $1 \le p \le \infty$ be a real number. We define the *Sobolev space*

$$W^{m,p}(\Omega) := \{v \in L^p(\Omega) \mid \forall\alpha \in A_d^m,\ \partial^\alpha v \in L^p(\Omega)\},$$

where

$$A_d^m := \left\{\alpha \in \mathbb{N}^d \mid |\alpha|_{\ell^1} \le m\right\}. \tag{1.8}$$

Thus, $W^{m,p}(\Omega)$ is spanned by functions with derivatives of global order up to m in $L^p(\Omega)$. In particular, $W^{0,p}(\Omega) = L^p(\Omega)$. The Sobolev space $W^{m,p}(\Omega)$ is a Banach space when equipped with the norm

$$\|v\|_{W^{m,p}(\Omega)} := \left(\sum_{\alpha\in A_d^m} \|\partial^\alpha v\|_{L^p(\Omega)}^p\right)^{1/p} \qquad 1 \le p < \infty,$$

and $\|v\|_{W^{m,\infty}(\Omega)} := \max_{\alpha\in A_d^m} \|\partial^\alpha v\|_{L^\infty(\Omega)}$. We also consider the seminorm $|\cdot|_{W^{m,p}(\Omega)}$ by restricting the above definitions to d-uples in the set $\overline{A}_d^m := \left\{\alpha \in \mathbb{N}^d \mid |\alpha|_{\ell^1} = m\right\}$, that is, by keeping only the derivatives of global order equal to m.

For $p = 2$, we use the notation $H^m(\Omega) := W^{m,2}(\Omega)$, so that

$$H^m(\Omega) = \left\{v \in L^2(\Omega) \mid \forall\alpha \in A_d^m,\ \partial^\alpha v \in L^2(\Omega)\right\}.$$

$H^m(\Omega)$ is a Hilbert space when equipped with the scalar product

$$(v, w)_{H^m(\Omega)} := \sum_{\alpha\in A_d^m} (\partial^\alpha v, \partial^\alpha w)_{L^2(\Omega)},$$

leading to the norm and seminorm

$$\|v\|_{H^m(\Omega)} := \left(\sum_{\alpha\in A_d^m} \|\partial^\alpha v\|_{L^2(\Omega)}^2\right)^{1/2}, \qquad |v|_{H^m(\Omega)} := \left(\sum_{\alpha\in \overline{A}_d^m} \|\partial^\alpha v\|_{L^2(\Omega)}^2\right)^{1/2}.$$

To allow for a more compact notation in the case $m = 1$, we consider the gradient $\nabla v = (\partial_1 v, \dots, \partial_d v)^t$ with values in $\mathbb{R}^d$. The norm on $W^{1,p}(\Omega)$ becomes

$$\|v\|_{W^{1,p}(\Omega)} = \left(\|v\|_{L^p(\Omega)}^p + \|\nabla v\|_{[L^p(\Omega)]^d}^p\right)^{1/p} \qquad 1 \le p < \infty,$$

with

$$\|\nabla v\|_{[L^p(\Omega)]^d} := \left(\int_\Omega |\nabla v|_{\ell^p}^p\right)^{1/p} = \left(\int_\Omega \sum_{i=1}^d |\partial_i v|^p\right)^{1/p}.$$

In the case $p = 2$, we obtain

$$(v, w)_{H^1(\Omega)} = (v, w)_{L^2(\Omega)} + (\nabla v, \nabla w)_{[L^2(\Omega)]^d}.$$

Boundary values of functions in the Sobolev space $W^{1,p}(\Omega)$ can be given a meaning (at least) in $L^p(\partial\Omega)$. More precisely (see, e.g., Brenner and Scott [54, Chap. 1]), for all $1 \le p \le \infty$, there is C such that

$$\|v\|_{L^p(\partial\Omega)} \le C\|v\|_{L^p(\Omega)}^{1-1/p}\|v\|_{W^{1,p}(\Omega)}^{1/p} \qquad \forall v \in W^{1,p}(\Omega). \tag{1.9}$$

In particular, for $p = 2$, we obtain

$$\|v\|_{L^2(\partial\Omega)} \le C\|v\|_{L^2(\Omega)}^{1/2}\|v\|_{H^1(\Omega)}^{1/2} \qquad \forall v \in H^1(\Omega). \tag{1.10}$$

The bounds (1.9) and (1.10) are called *continuous trace inequalities*.

Finally, at some instances, we consider Hilbert Sobolev spaces $H^s(\Omega)$ where the exponent s is a *positive real number*. We refer the reader, e.g., to Ern and Guermond [141, p. 484] for the definition of such spaces. In what follows, we use the fact that functions in $H^{1/2+\epsilon}(\Omega)$, $\epsilon > 0$, admit a trace in $L^2(\partial\Omega)$.

1.2 The Discrete Setting

In this section, we present the main ingredients to build finite-dimensional function spaces to approximate the model problem (1.1) using dG methods. The construction of such spaces hinges on discretizing the domain Ω (over which the PDE is posed) using a mesh and choosing a local functional behavior (e.g., polynomial) within each mesh element. This leads to broken polynomial spaces. We also introduce broken Sobolev spaces and a broken gradient operator. Finally, we briefly discuss the function space $H(\mathrm{div};\Omega)$ and its broken version; such spaces are particularly relevant in the context of PDEs with diffusion.

1.2.1 The Domain Ω

To simplify the presentation, we focus, throughout this book, on polyhedra.

Definition 1.6 (Polyhedron in $\mathbb{R}^d$)**.** We say that the set P is a *polyhedron* in $\mathbb{R}^d$ if P is an open, connected, bounded subset of $\mathbb{R}^d$ such that its boundary ∂P is a finite union of parts of hyperplanes, say $\{H_i\}_{1\le i\le n_P}$. Moreover, for all $1 \le i \le n_P$, at each point in the interior of $\partial P \cap H_i$, the set P is assumed to lie on only one side of its boundary.

Assumption 1.7 (Domain Ω)**.** *The domain Ω is a polyhedron in $\mathbb{R}^d$.*

The advantage of Assumption 1.7 is that polyhedra can be exactly covered by a mesh consisting of polyhedral elements. PDEs posed over domains with curved boundary can also be approximated by dG methods using, e.g., isoparametric finite elements to build the mesh near curved boundaries as described, e.g., by Ciarlet [92, p. 224] and Brenner and Scott [54, p. 117].

Definition 1.8 (Boundary and outward normal)**.** The *boundary* of Ω is denoted by $\partial\Omega$ and its (unit) *outward normal*, which is defined a.e. on $\partial\Omega$, by n.

1.2.2 Meshes

The first step is to discretize the domain Ω using a mesh. Various types of meshes can be considered. We examine first the most familiar case, that of simplicial meshes. Such meshes should be familiar to the reader since they are one of the key ingredients to build continuous finite element spaces.

Definition 1.9 (Simplex)**.** Given a family $\{a_0, \ldots, a_d\}$ of $(d+1)$ points in $\mathbb{R}^d$ such that the vectors $\{a_1 - a_0, \ldots, a_d - a_0\}$ are linearly independent, the interior of the convex hull of $\{a_0, \ldots, a_d\}$ is called a non-degenerate *simplex* of $\mathbb{R}^d$, and the points $\{a_0, \ldots, a_d\}$ are called its *vertices*.

By its definition, a non-degenerate simplex is an open subset of $\mathbb{R}^d$. In dimension 1, a non-degenerate simplex is an interval, in dimension 2 a triangle, and in dimension 3 a tetrahedron. The unit simplex of $\mathbb{R}^d$ is the set (cf. Fig. 1.1)

$$S_d := \left\{(x_1, \ldots, x_d) \in \mathbb{R}^d \mid x_i > 0 \ \forall i \in \{1, \ldots, d\} \text{ and } x_1 + \ldots + x_d < 1\right\}.$$

Any non-degenerate simplex of $\mathbb{R}^d$ is the image of the unit simplex by a bijective affine transformation of $\mathbb{R}^d$.

Definition 1.10 (Simplex faces)**.** Let S be a non-degenerate simplex with vertices $\{a_0, \ldots, a_d\}$. For each $i \in \{0, \ldots, d\}$, the convex hull of $\{a_0, \ldots, a_d\} \setminus \{a_i\}$ is called a *face* of the simplex S.

Thus, a non-degenerate simplex has $(d+1)$ faces, and, by construction, a simplex face is a closed subset of $\mathbb{R}^d$. A simplex face has zero d-dimensional Hausdorff measure, but positive $(d-1)$-dimensional Hausdorff measure. In dimension 2, a simplex face is also called an edge, while in dimension 1, a simplex face is a point and its 0-dimensional Hausdorff measure is conventionally set to 1.

Definition 1.11 (Simplicial mesh)**.** A *simplicial mesh* $\mathcal{T}$ of the domain Ω is a finite collection of disjoint non-degenerate simplices $\mathcal{T} = \{T\}$ forming a partition of Ω,

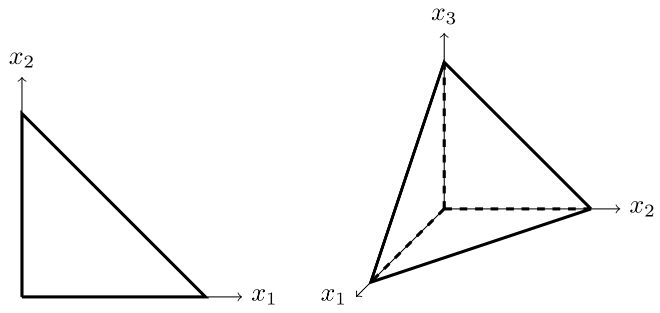

Fig. 1.1: Unit simplex in two (*left*) and three (*right*) space dimensions

$$\overline{\Omega} = \bigcup_{T \in \mathcal{T}} \overline{T}. \tag{1.11}$$

Each $T \in \mathcal{T}$ is called a *mesh element*.

While simplicial meshes are quite convenient in the context of continuous finite elements, dG methods more easily accommodate general meshes.

Definition 1.12 (General mesh)**.** A *general mesh* $\mathcal{T}$ of the domain Ω is a finite collection of disjoint polyhedra $\mathcal{T} = \{T\}$ forming a partition of Ω as in (1.11). Each $T \in \mathcal{T}$ is called a *mesh element*.

Obviously, a simplicial mesh is just a particular case of a general mesh.

Definition 1.13 (Element diameter, meshsize)**.** Let $\mathcal{T}$ be a (general) mesh of the domain Ω. For all $T \in \mathcal{T}$, h_T denotes the *diameter* of T, and the *meshsize* is defined as the real number

$$h := \max_{T \in \mathcal{T}} h_T.$$

We use the notation $\mathcal{T}_h$ for a mesh $\mathcal{T}$ with meshsize h.

Definition 1.14 (Element outward normal)**.** Let $\mathcal{T}_h$ be a mesh of the domain Ω and let $T \in \mathcal{T}_h$. We define n_T a.e. on ∂T as the (unit) *outward normal* to T.

Faces of a *single* polyhedral mesh element are not needed in what follows, and we leave them undefined to avoid confusion with the important concept of mesh faces introduced in Sect. 1.2.3. The key difference is that mesh faces depend on the way neighboring mesh elements come into contact.

Remark 1.15 (General hexahedra). In the present setting, general hexahedra in $\mathbb{R}^3$ cannot be mesh elements since their faces are not parts of (hyper)planes (since four distinct points do not generally belong to the same plane). One possibility is to approximate general hexahedra by subdividing nonplanar faces into two or four triangles.

1.2.3 Mesh Faces, Averages, and Jumps

The concepts of mesh faces, averages, and jumps play a central role in the design and analysis of dG methods.

Definition 1.16 (Mesh faces)**.** Let $\mathcal{T}_h$ be a mesh of the domain Ω. We say that a (closed) subset F of $\overline{\Omega}$ is a *mesh face* if F has positive $(d-1)$-dimensional Hausdorff measure (in dimension 1, this means that F is nonempty) and if either one of the two following conditions is satisfied:

(i) There are distinct mesh elements T_1 and T_2 such that $F = \partial T_1 \cap \partial T_2$; in such a case, F is called an *interface*.

(ii) There is $T \in \mathcal{T}_h$ such that $F = \partial T \cap \partial\Omega$; in such a case, F is called a *boundary face*.

Interfaces are collected in the set $\mathcal{F}_h^i$, and boundary faces are collected in the set $\mathcal{F}_h^b$. Henceforth, we set

$$\mathcal{F}_h := \mathcal{F}_h^i \cup \mathcal{F}_h^b.$$

Moreover, for any mesh element $T \in \mathcal{T}_h$, the set

$$\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$$

collects the mesh faces composing the boundary of T. The maximum number of mesh faces composing the boundary of mesh elements is denoted by

$$N_\partial := \max_{T \in \mathcal{T}_h} \operatorname{card}(\mathcal{F}_T). \tag{1.12}$$

Finally, for any mesh face $F \in \mathcal{F}_h$, we define the set

$$\mathcal{T}_F := \{T \in \mathcal{T}_h \mid F \subset \partial T\}, \tag{1.13}$$

and observe that $\mathcal{T}_F$ consists of two mesh elements if $F \in \mathcal{F}_h^i$ and of one mesh element if $F \in \mathcal{F}_h^b$.

Figure 1.2 depicts an interface between two mesh elements belonging to a simplicial mesh (left) or to a general mesh (right). We observe that in the case of simplicial meshes, interfaces are always parts of hyperplanes, but this is not necessarily the case for general meshes containing nonconvex polyhedra. We now define averages and jumps across interfaces of piecewise smooth functions; cf. Fig. 1.3 for a one-dimensional illustration.

Definition 1.17 (Interface averages and jumps)**.** Let v be a scalar-valued function defined on Ω and assume that v is smooth enough to admit on all $F \in \mathcal{F}_h^i$ a possibly two-valued trace. This means that, for all $T \in \mathcal{T}_h$, the restriction $v|_T$ of v to the open set T can be defined up to the boundary ∂T. Then, for all $F \in \mathcal{F}_h^i$ and a.e. $x \in F$, the *average* of v is defined as

$$\{\!\!\{v\}\!\!\}_F(x) := \frac{1}{2}\Big(v|_{T_1}(x) + v|_{T_2}(x)\Big),$$

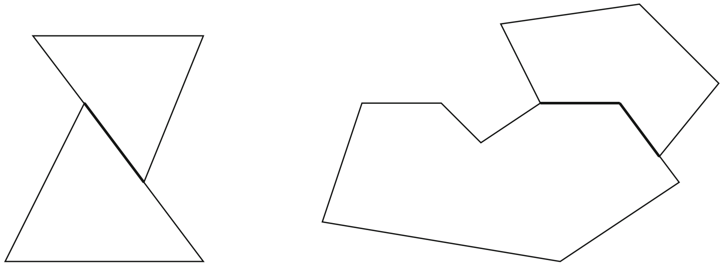

Fig. 1.2: Examples of interface for a simplicial mesh (*left*) and a general mesh (*right*)

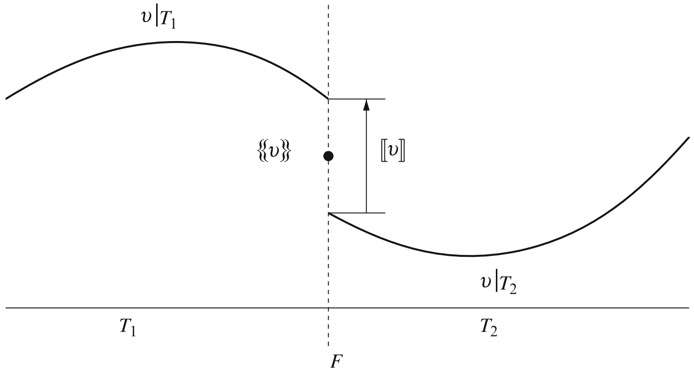

Fig. 1.3: One-dimensional example of average and jump operators; the face reduces to a point separating two adjacent intervals

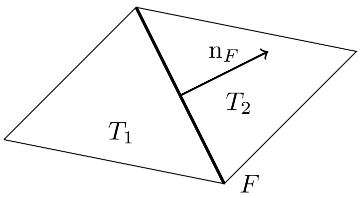

Fig. 1.4: Notation for an interface

and the *jump* of v as

$$[\![v]\!]_F(x) := v|_{T_1}(x) - v|_{T_2}(x).$$

When v is vector-valued, the above average and jump operators act componentwise on the function v. Whenever no confusion can arise, the subscript F and the variable x are omitted, and we simply write $\{\!\{v\}\!\}$ and $[\![v]\!]$.

Definition 1.18 (Face normals)**.** For all $F \in \mathcal{F}_h$ and a.e. $x \in F$, we define the (unit) *normal* n_F to F at x as

(i) n_{T_1}, the unit normal to F at x pointing from T_1 to T_2 if $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$; the orientation of n_F is arbitrary depending on the choice of T_1 and T_2, but kept fixed in what follows. See Fig. 1.4.

(ii) n, the unit outward normal to Ω at x if $F \in \mathcal{F}_h^b$.

Remark 1.19 (Boundary averages and jumps)*.* Averages and jumps can also be defined for boundary faces (this is particularly useful when discretizing PDEs with diffusion as in Chaps. 4–6). One possible definition is to set for a.e. $x \in F$ with $F \in \mathcal{F}_h^b$, $F = \partial T \cap \partial\Omega$, $\{\!\{v\}\!\}_F(x) = [\![v]\!]_F(x) := v|_T(x)$.

Remark 1.20 (Alternative definition of jumps). An alternative definition of jumps consists in setting for all $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$,

$$[\![v]\!]_* := v|_{T_1} \mathrm{n}_{T_1} + v|_{T_2} \mathrm{n}_{T_2},$$

so that $[\![v]\!]_* = \mathrm{n}_F [\![v]\!]$. This alternative definition allows T_1 and T_2 to play symmetric roles; however, the jump of a scalar-valued function is a vector-valued function.

1.2.4 Broken Polynomial Spaces

After having built a mesh of the domain Ω, the second step in the construction of discrete function spaces consists in choosing a certain functional behavior within each mesh element. For the sake of simplicity, we restrict ourselves to polynomial functions; more general cases can also be accommodated (see, e.g., Yuan and Shu [309]). The resulting spaces, consisting of piecewise polynomial functions, are termed *broken polynomial spaces*.

1.2.4.1 The Polynomial Space $\mathbb{P}_d^k$

Let $k \geq 0$ be an integer. Recalling the set A_d^k defined by (1.8), we define the space of polynomials of d variables, of total degree at most k, as

$$\mathbb{P}_d^k := \left\{ p : \mathbb{R}^d \ni x \mapsto p(x) \in \mathbb{R} \mid \exists (\gamma_\alpha)_{\alpha \in A_d^k} \in \mathbb{R}^{\mathrm{card}(A_d^k)} \text{ s.t. } p(x) = \sum_{\alpha \in A_d^k} \gamma_\alpha x^\alpha \right\},$$

with the convention that, for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $x^\alpha := \prod_{i=1}^d x_i^{\alpha_i}$. The dimension of the vector space $\mathbb{P}_d^k$ is

$$\dim(\mathbb{P}_d^k) = \mathrm{card}(A_d^k) = \binom{k+d}{k} = \frac{(k+d)!}{k!d!}. \tag{1.14}$$

The first few values of $\dim(\mathbb{P}_d^k)$ are listed in Table 1.1.

1.2.4.2 The Broken Polynomial Space $\mathbb{P}_d^k(\mathcal{T}_h)$

In what follows, we often consider the broken polynomial space

$$\mathbb{P}_d^k(\mathcal{T}_h) := \left\{ v \in L^2(\Omega) \mid \forall T \in \mathcal{T}_h,\, v|_T \in \mathbb{P}_d^k(T) \right\}, \tag{1.15}$$

Table 1.1: Dimension of the polynomial space $\mathbb{P}_d^k$ for $d \in \{1, 2, 3\}$ and $k \in \{0, 1, 2, 3\}$

k	$d = 1$	$d = 2$	$d = 3$
0	1	1	1
1	2	3	4
2	3	6	10
3	4	10	20

where $\mathbb{P}_d^k(T)$ is spanned by the restriction to T of polynomials in $\mathbb{P}_d^k$. It is clear that

$$\dim(\mathbb{P}_d^k(\mathcal{T}_h)) = \operatorname{card}(\mathcal{T}_h) \times \dim(\mathbb{P}_d^k),$$

since the restriction of a function $v \in \mathbb{P}_d^k(\mathcal{T}_h)$ to each mesh element can be chosen independently of its restriction to other elements.

1.2.4.3 Other Broken Polynomial Spaces

It is possible to consider other broken polynomial spaces. Such spaces are encountered, e.g., when defining local bases using a reference element, e.g., nodal-based local bases associated with quadratures (see Gassner, Lörcher, Munz, and Hesthaven [164]). Further motivations for considering other broken polynomial spaces include, among others, bubble-stabilization techniques (see Burman and Stamm [70]) and inf-sup stable discretizations for incompressible flows (see Toselli [296]).

A relevant example is the space of polynomials of d variables, of degree at most k in each variable, namely

$$\mathbb{Q}_d^k := \left\{ p : \mathbb{R}^d \ni x \mapsto p(x) \in \mathbb{R} \mid \exists (\gamma_\alpha)_{\alpha \in B_d^k} \in \mathbb{R}^{\operatorname{card}(B_d^k)} \text{ s.t. } p(x) = \sum_{\alpha \in B_d^k} \gamma_\alpha x^\alpha \right\},$$

where B_d^k denotes the set of d-uples of ∞-norm smaller than or equal to k,

$$B_d^k := \left\{ \alpha \in \mathbb{N}^d \mid |\alpha|_{\ell^\infty} \leq k \right\}.$$

The dimension of the vector space $\mathbb{Q}_d^k$ is

$$\dim(\mathbb{Q}_d^k) = \operatorname{card}(B_d^k) = (k+1)^d.$$

The first few values of $\dim(\mathbb{Q}_d^k)$ are listed in Table 1.2.

1.2.5 Broken Sobolev Spaces

In this section, we introduce broken Sobolev spaces and a broken gradient operator. We also state the main properties of broken Sobolev spaces to be used in what follows.

Table 1.2: Dimension of the polynomial space $\mathbb{Q}_d^k$ for $d \in \{1,2,3\}$ and $k \in \{0,1,2,3\}$

k	$d=1$	$d=2$	$d=3$
0	1	1	1
1	2	4	8
2	3	9	27
3	4	16	64

Let $\mathcal{T}_h$ be a mesh of the domain Ω. For any mesh element $T \in \mathcal{T}_h$, the Sobolev spaces $H^m(T)$ and $W^{m,p}(T)$ can be defined as above by replacing Ω by T. We then define the *broken Sobolev spaces*

$$H^m(\mathcal{T}_h) := \left\{v \in L^2(\Omega) \mid \forall T \in \mathcal{T}_h,\ v|_T \in H^m(T)\right\}, \tag{1.16}$$

$$W^{m,p}(\mathcal{T}_h) := \{v \in L^p(\Omega) \mid \forall T \in \mathcal{T}_h,\ v|_T \in W^{m,p}(T)\}, \tag{1.17}$$

where $m \geq 0$ is an integer and $1 \leq p \leq \infty$ a real number.

In the context of broken Sobolev spaces, the continuous trace inequality (1.9) can be used to infer that, for all $v \in W^{1,p}(\mathcal{T}_h)$ and all $T \in \mathcal{T}_h$,

$$\|v\|_{L^p(\partial T)} \leq C\|v\|_{L^p(T)}^{1-1/p}\|v\|_{W^{1,p}(T)}^{1/p}, \tag{1.18}$$

while for $p = 2$, we obtain, for all $v \in H^1(\mathcal{T}_h)$ and all $T \in \mathcal{T}_h$,

$$\|v\|_{L^2(\partial T)} \leq C\|v\|_{L^2(T)}^{1/2}\|v\|_{H^1(T)}^{1/2}. \tag{1.19}$$

In what follows, it is implicitly understood that expressions such as $\|v\|_{L^2(\partial T)}$ (or such as $\|v\|_{L^2(F)}$ for a mesh face $F \in \mathcal{F}_T$ of a given mesh element $T \in \mathcal{T}_h$) should be evaluated using the restriction of v to T. A different version of the continuous trace inequality (1.19) is presented in Lemma 1.49 below. Continuous trace inequalities such as (1.18) or (1.19) are important in the context of dG methods to give a meaning to the traces of the exact solution or of its (normal) gradient on mesh faces.

It is natural to define a broken gradient operator acting on the broken Sobolev space $W^{1,p}(\mathcal{T}_h)$. In particular, this operator also acts on broken polynomial spaces.

Definition 1.21 (Broken gradient)**.** The *broken gradient* $\nabla_h : W^{1,p}(\mathcal{T}_h) \to [L^p(\Omega)]^d$ is defined such that, for all $v \in W^{1,p}(\mathcal{T}_h)$,

$$\forall T \in \mathcal{T}_h, \qquad (\nabla_h v)|_T := \nabla(v|_T). \tag{1.20}$$

In what follows, we drop the index h in the broken gradient when this operator appears inside an integral over a fixed mesh element $T \in \mathcal{T}_h$.

It is important to observe that the usual Sobolev spaces are subspaces of the broken Sobolev spaces, and that on the usual Sobolev spaces, the broken gradient coincides with the distributional gradient. For completeness, we detail the proof of this result.

Lemma 1.22 (Broken gradient on usual Sobolev spaces)**.** *Let $m \geq 0$ and let $1 \leq p \leq \infty$. There holds $W^{m,p}(\Omega) \subset W^{m,p}(\mathcal{T}_h)$. Moreover, for all $v \in W^{1,p}(\Omega)$, $\nabla_h v = \nabla v$ in $[L^p(\Omega)]^d$.*

Proof. It is sufficient to prove the inclusion for $m = 1$. Let $v \in W^{1,p}(\Omega)$. We first observe that $\nabla(v|_T) = (\nabla v)|_T$ for all $T \in \mathcal{T}_h$. Indeed, for all $\Phi \in [C_0^\infty(T)]^d$, the extension of Φ by zero to Ω, say $E\Phi$, is in $[C_0^\infty(\Omega)]^d$, so that

$$\begin{aligned}\int_T \nabla(v|_T)\cdot\Phi &= -\int_T v(\nabla\cdot\Phi) = -\int_\Omega v(\nabla\cdot(E\Phi))\\ &= \int_\Omega \nabla v\cdot E\Phi = \int_T (\nabla v)|_T\cdot\Phi.\end{aligned}$$

Since Φ is arbitrary, this implies $\nabla(v|_T) = (\nabla v)|_T$ and, since $T \in \mathcal{T}_h$ is arbitrary, we infer $\nabla_h v = \nabla v$. This equality also shows that $v \in W^{1,p}(\mathcal{T}_h)$.

□

The reverse inclusion of Lemma 1.22 does not hold true in general (except obviously for $m = 0$). The reason is that functions in the broken Sobolev space $W^{1,p}(\mathcal{T}_h)$ can have nonzero jumps across interfaces, while functions in the usual Sobolev space $W^{1,p}(\Omega)$ have zero jumps across interfaces. We now give a precise statement of this important result.

Lemma 1.23 (Characterization of $W^{1,p}(\Omega)$)**.** *Let $1 \le p \le \infty$. A function $v \in W^{1,p}(\mathcal{T}_h)$ belongs to $W^{1,p}(\Omega)$ if and only if*

$$[\![v]\!] = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{1.21}$$

Proof. The proof is again based on a distributional argument. Let $v \in W^{1,p}(\mathcal{T}_h)$. Then, for all $\Phi \in [C_0^\infty(\Omega)]^d$, we observe integrating by parts elementwise that

$$\begin{aligned}\int_\Omega \nabla_h v\cdot\Phi &= \sum_{T\in\mathcal{T}_h}\int_T \nabla(v|_T)\cdot\Phi = -\sum_{T\in\mathcal{T}_h}\int_T v(\nabla\cdot\Phi) + \sum_{T\in\mathcal{T}_h}\int_{\partial T} v|_T(\Phi\cdot\mathrm{n}_T)\\ &= -\int_\Omega v(\nabla\cdot\Phi) + \sum_{F\in\mathcal{F}_h^i}\int_F (\Phi\cdot\mathrm{n}_F)[\![v]\!],\end{aligned} \tag{1.22}$$

since Φ is continuous across interfaces and vanishes on boundary faces. Assume first that (1.21) holds true. Then, (1.22) implies

$$\int_\Omega \nabla_h v\cdot\Phi = -\int_\Omega v(\nabla\cdot\Phi) \qquad \forall \Phi \in [C_0^\infty(\Omega)]^d,$$

meaning that $\nabla v = \nabla_h v$ in $[L^p(\Omega)]^d$. Hence, $v \in W^{1,p}(\Omega)$. Conversely, if $v \in W^{1,p}(\Omega)$, $\nabla v = \nabla_h v$ in $[L^p(\Omega)]^d$ owing to Lemma 1.22, so that (1.22) now implies

$$\begin{aligned}\sum_{F\in\mathcal{F}_h^i}\int_F (\Phi\cdot\mathrm{n}_F)[\![v]\!] &= \int_\Omega \nabla_h v\cdot\Phi + \int_\Omega v(\nabla\cdot\Phi)\\ &= \int_\Omega \nabla_h v\cdot\Phi - \int_\Omega \nabla v\cdot\Phi = 0,\end{aligned}$$

whence we infer (1.21) by choosing the support of Φ intersecting a single interface and since Φ is arbitrary.

□

1.2.6 The Function Space $H(\mathrm{div};\Omega)$ and Its Broken Version

In the context of PDEs with diffusion, the vector-valued field $\sigma = -\nabla u$ can be interpreted as the diffusive flux; here, u solves, e.g., the Poisson problem presented in Sect. 4.1. From a physical viewpoint, it is expected that the normal component of the diffusive flux does not jump across interfaces. From a mathematical viewpoint, the diffusive flux belongs to the function space

$$H(\mathrm{div};\Omega) := \{\tau \in [L^2(\Omega)]^d \mid \nabla\cdot\tau \in L^2(\Omega)\}. \tag{1.23}$$

It is therefore important to specify the meaning of the normal component on mesh faces of functions in $H(\mathrm{div};\Omega)$.

For all $T \in \mathcal{T}_h$, we define the function space $H(\mathrm{div};T)$ by replacing Ω by T in (1.23). We then introduce the broken space

$$H(\mathrm{div};\mathcal{T}_h) := \{\tau \in [L^2(\Omega)]^d \mid \forall T \in \mathcal{T}_h,\ \tau|_T \in H(\mathrm{div};T)\},$$

and the *broken divergence operator* $\nabla_h\cdot : H(\mathrm{div};\mathcal{T}_h) \to L^2(\Omega)$ such that, for all $\tau \in H(\mathrm{div};\mathcal{T}_h)$,

$$\forall T \in \mathcal{T}_h, \qquad (\nabla_h\cdot\tau)|_T := \nabla\cdot(\tau|_T).$$

Proceeding as in the proof of Lemma 1.22, we can verify that, if $\tau \in H(\mathrm{div};\Omega)$, then $\tau \in H(\mathrm{div};\mathcal{T}_h)$ and $\nabla_h\cdot\tau = \nabla\cdot\tau$.

Since the diffusive flux σ belongs to $H(\mathrm{div};\Omega)$, we infer that, for all $T \in \mathcal{T}_h$, $\sigma_T := \sigma|_T$ belongs to $H(\mathrm{div};T)$. This property allows us to give a weak meaning to the normal component $\sigma_{\partial T} := \sigma_T\cdot\mathrm{n}_T$ on ∂T; cf. Remark 1.26 below for further insight. However, in the context of dG methods, we would like to give a meaning to the restriction $\sigma_F := \sigma_{\partial T}|_F$ independently on any face $F \in \mathcal{F}_T$, and additionally, it is convenient to assert $\sigma_F \in L^1(F)$. The fact that $\sigma_T \in H(\mathrm{div};T)$ does not provide enough regularity to assert this, as reflected by the counterexample in Remark 1.25 below. A suitable assumption to achieve $\sigma_F \in L^1(F)$ is

$$\sigma \in [W^{1,1}(\mathcal{T}_h)]^d.$$

Indeed, owing to the trace inequality (1.18), we infer, for all $T \in \mathcal{T}_h$, $\sigma_{\partial T} \in L^1(\partial T)$ so that, for all $F \in \mathcal{F}_T$, $\sigma_F \in L^1(F)$.

We can now characterize functions in $H(\mathrm{div};\Omega) \cap [W^{1,1}(\mathcal{T}_h)]^d$ using the jump of the normal component across interfaces.

Lemma 1.24 (Characterization of $H(\mathrm{div};\Omega)$)**.** *A function $\tau \in H(\mathrm{div};\mathcal{T}_h) \cap [W^{1,1}(\mathcal{T}_h)]^d$ belongs to $H(\mathrm{div};\Omega)$ if and only if*

$$[\![\tau]\!]\cdot\mathrm{n}_F = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{1.24}$$

Proof. The proof is similar to that of Lemma 1.23. Let $\tau \in H(\mathrm{div};\mathcal{T}_h) \cap [W^{1,1}(\mathcal{T}_h)]^d$ and let $\varphi \in C_0^\infty(\Omega)$. Integrating by parts on each mesh element and accounting for the fact that φ is smooth inside Ω and vanishes on $\partial\Omega$, we obtain

$$\begin{aligned}\int_\Omega \tau\cdot\nabla\varphi &= \sum_{T\in\mathcal{T}_h}\int_T \tau\cdot\nabla\varphi = -\sum_{T\in\mathcal{T}_h}\int_T (\nabla\cdot\tau)\varphi + \sum_{T\in\mathcal{T}_h}\int_{\partial T}(\tau\cdot\mathrm{n}_T)\varphi \\ &= -\int_\Omega (\nabla_h\cdot\tau)\varphi + \sum_{F\in\mathcal{F}_h^i}\int_F [\![\tau]\!]\cdot\mathrm{n}_F\varphi,\end{aligned}$$

where we have used the fact that $\tau\cdot\mathrm{n}_T \in L^1(\partial T)$ to write a boundary integral and break it into face integrals. Hence, if (1.24) holds true,

$$\int_\Omega \tau\cdot\nabla\varphi = -\int_\Omega (\nabla_h\cdot\tau)\varphi \qquad \forall\varphi\in C_0^\infty(\Omega),$$

implying that $\nabla\cdot\tau = \nabla_h\cdot\tau \in L^2(\Omega)$, so that $\tau \in H(\mathrm{div};\Omega)$. Conversely, if $\tau\in H(\mathrm{div};\Omega)$, since $\nabla_h\cdot\tau = \nabla\cdot\tau$, the above identity yields

$$\sum_{F\in\mathcal{F}_h^i}\int_F [\![\tau]\!]\cdot\mathrm{n}_F\varphi = 0,$$

whence (1.24) is obtained by choosing the support of φ intersecting a single interface and since φ is arbitrary.

□

Remark 1.25 (Counter-example for $\sigma_F \in L^1(F)$)*.* Following an idea by Carstensen and Peterseim [78], we consider the triangle

$$T = \{(x_1,x_2)\in\mathbb{R}^2 \mid 0 < x_1 < 1,\, 0 < x_2 < x_1\}.$$

For all $\epsilon > 0$, letting $r^2 = x_1^2 + x_2^2$, we define the vector field

$$\sigma_\epsilon = (\epsilon + r^2)^{-1}(x_2, -x_1)^t.$$

A direct calculation shows that σ_ϵ is divergence-free and, since $0 \le r^2 \le 1$,

$$\begin{aligned}\|\sigma_\epsilon\|^2_{[L^2(T)]^d} &= \int_0^{\pi/4}\int_0^{\cos(\theta)^{-1}} \frac{r^3}{(\epsilon+r^2)^2}\,\mathrm{d}r\,\mathrm{d}\theta \\ &\le \int_0^{\pi/4}\int_0^{\sqrt{2}} \frac{r}{(\epsilon+r^2)}\,\mathrm{d}r\,\mathrm{d}\theta = \frac{\pi}{8}\ln(1+2\epsilon^{-1}).\end{aligned}$$

Hence, letting $\tilde\sigma_\epsilon = \ln(1+2\epsilon^{-1})^{-1/2}\sigma_\epsilon$, we infer that $\|\tilde\sigma_\epsilon\|_{[L^2(T)]^d}$ is uniformly bounded in ϵ. Moreover, considering the face $F = \{(x_1,x_2)\in\partial T \mid x_1 = x_2\}$ with normal $\mathrm{n}_F = (-2^{-1/2}, 2^{-1/2})^t$, we obtain

$$\|\tilde\sigma_\epsilon\cdot\mathrm{n}_F\|_{L^1(F)} = \ln(1+2\epsilon^{-1})^{-1/2}\int_0^{\sqrt{2}} \frac{r}{\epsilon+r^2}\,\mathrm{d}r = \frac{1}{2}\ln(1+2\epsilon^{-1})^{1/2},$$

which grows unboundedly as $\epsilon\to 0^+$.

Remark 1.26 (Weak meaning of $\sigma_T{\cdot}n_T$ on ∂T). Let $T \in \mathcal{T}_h$ and let $\sigma_T \in H(\mathrm{div}; T)$. Let $H^{1/2}(\partial T)$ be the vector space spanned by the traces on ∂T of functions in $H^1(T)$. Then, the normal component $\sigma_{\partial T} := \sigma_T{\cdot}n_T$ can be defined in the dual space $H^{-1/2}(\partial T) := (H^{1/2}(\partial T))'$ in such a way that, for all $g \in H^{1/2}(\partial T)$,

$$\langle \sigma_{\partial T}, g \rangle = \int_T \sigma_T{\cdot}\nabla \hat{g} + \int_T (\nabla{\cdot}\sigma_T)\hat{g}, \tag{1.25}$$

where $\hat{g} \in H^1(T)$ is such that its trace on ∂T is equal to g (the right-hand side of (1.25) is independent of the choice of $\hat{g}$). Consider now a face $F \in \mathcal{F}_T$ and let $H^{1/2}_{00}(F)$ be spanned by functions defined on F whose extension by zero to ∂T is in $H^{1/2}(\partial T)$. Then, the restriction $\sigma_F := \sigma_{\partial T}|_F$ can be given a meaning in the dual space $H^{-1/2}(F) := (H^{1/2}_{00}(F))'$ in such a way that, for all $g \in H^{1/2}_{00}(F)$,

$$\langle \sigma_F, g \rangle = \int_T \sigma_T{\cdot}\nabla \hat{g} + \int_T (\nabla{\cdot}\sigma_T)\hat{g},$$

where $\hat{g} \in H^1(T)$ is such that its trace on ∂T is equal to g on F and to zero elsewhere. However, this definition is of little use in the context of dG methods where the normal component σ_F has to act on polynomials on F which do not generally belong to $H^{1/2}_{00}(F)$ (unless they vanish on ∂F).

1.3 Abstract Nonconforming Error Analysis

The goal of this section is to present the key ingredients for the error analysis when approximating the linear model problem (1.1) by dG methods. We assume that (1.1) is well-posed; cf. Sect. 1.1. The error analysis presented in this section is derived in the spirit of Strang's Second Lemma [285] (see also Ern and Guermond [141, Sect. 2.3]). The three ingredients are (1) Discrete stability, (2) (Strong) consistency, and (3) Boundedness.

1.3.1 The Discrete Problem

Let $V_h \subset L^2(\Omega)$ denote a finite-dimensional function space; typically, V_h is a broken polynomial space (cf. Sect. 1.2.4). We are interested in the discrete problem

$$\text{Find } u_h \in V_h \text{ s.t. } a_h(u_h, w_h) = l_h(w_h) \text{ for all } w_h \in V_h, \tag{1.26}$$

with discrete bilinear form a_h defined (so far) only on $V_h \times V_h$ and discrete linear form l_h defined on V_h. We observe that we consider the so-called standard Galerkin approximation where the discrete trial and test spaces coincide. Moreover, since functions in V_h can be discontinuous across mesh elements, $V_h \not\subset X$ and $V_h \not\subset Y$ in general; cf., e.g., Lemma 1.23. In the terminology of finite elements, we say that the approximation is *nonconforming*.

Alternatively, it is possible to introduce the discrete (linear) operator $A_h : V_h \to V_h$ such that, for all $v_h, w_h \in V_h$,

$$(A_h v_h, w_h)_{L^2(\Omega)} := a_h(v_h, w_h), \tag{1.27}$$

and the discrete function $L_h \in V_h$ such that, for all $w_h \in V_h$, $(L_h, w_h)_{L^2(\Omega)} = l_h(w_h)$. This leads to the following problem (obviously equivalent to (1.26)):

$$\text{Find } u_h \in V_h \text{ s.t. } A_h u_h = L_h \text{ in } V_h. \tag{1.28}$$

In what follows, we are often concerned with model problems where $Y \hookrightarrow L^2(\Omega)$ with dense and continuous injection. Identifying $L^2(\Omega)$ with its topological dual space $L^2(\Omega)'$ by means of the Riesz–Fréchet representation theorem, we are thus in the situation where

$$Y \hookrightarrow L^2(\Omega) \equiv L^2(\Omega)' \hookrightarrow Y',$$

with dense and continuous injections. For simplicity (cf. Remark 1.27 for further discussion), we often assume that the datum f is in $L^2(\Omega)$, so that the right-hand side of the model problem (1.1) becomes $(f, w)_{L^2(\Omega)}$, while the right-hand sides of the discrete problems (1.26) and (1.28) become, respectively,

$$l_h(w_h) = (f, w_h)_{L^2(\Omega)}, \qquad L_h = \pi_h f.$$

Here, π_h denotes the $L^2(\Omega)$-*orthogonal projection* onto V_h, that is, $\pi_h : L^2(\Omega) \to V_h$ is defined so that, for all $v \in L^2(\Omega)$, $\pi_h v \in V_h$ with

$$(\pi_h v, y_h)_{L^2(\Omega)} = (v, y_h)_{L^2(\Omega)} \qquad \forall y_h \in V_h. \tag{1.29}$$

We observe that the restriction of $\pi_h v$ to a given mesh element $T \in \mathcal{T}_h$ can be computed independently from other mesh elements. For instance, if $V_h = \mathbb{P}_d^k(\mathcal{T}_h)$, we obtain that, for all $T \in \mathcal{T}_h$, $\pi_h v|_T \in \mathbb{P}_d^k(T)$ is such that

$$(\pi_h v|_T, \xi)_{L^2(T)} = (v, \xi)_{L^2(T)} \qquad \forall \xi \in \mathbb{P}_d^k(T).$$

We refer the reader to Sect. A.1.2 for further insight.

Remark 1.27 (Rough right-hand side). The assumption $f \in L^2(\Omega)$ is convenient so as to define the right-hand side of (1.26) using the $L^2(\Omega)$-scalar product. For rough right-hand sides, i.e., $f \in Y'$ but $f \notin L^2(\Omega)$, the right-hand side of (1.26) needs to be modified since the quantity $\langle f, w_h \rangle_{Y',Y}$ is in general not defined. One possibility is to consider the right-hand side $\langle f, \mathcal{I}_h w_h \rangle_{Y',Y}$ for some smoothing linear operator $\mathcal{I}_h : V_h \to V_h \cap Y$ (cf. Remark 4.9 for an example in the context of diffusive PDEs).

1.3.2 Discrete Stability

To formulate discrete stability, we introduce a norm, say $|\!|\!|\cdot|\!|\!|$, defined (at least) on V_h.

Definition 1.28 (Discrete stability)**.** We say that the discrete bilinear form a_h enjoys *discrete stability* on V_h if there is $C_{\mathrm{sta}} > 0$ such that

$$\forall v_h \in V_h, \qquad C_{\mathrm{sta}} |||v_h||| \leq \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{|||w_h|||}. \tag{1.30}$$

Remark 1.29 (h-dependency). In Definition 1.28, C_{sta} can depend on the meshsize h. In view of convergence analysis, it is important to ensure that C_{sta} be independent of h.

Property (1.30) is referred to as a discrete inf-sup condition since it is equivalent to

$$C_{\mathrm{sta}} \leq \inf_{v_h \in V_h \setminus \{0\}} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{|||v_h||| \, |||w_h|||}.$$

An important fact is that (1.30) is a necessary and sufficient condition for discrete well-posedness.

Lemma 1.30 (Discrete well-posedness)**.** *The discrete problem* (1.26)*, or equivalently* (1.28)*, is well-posed if and only if the discrete inf-sup condition* (1.30) *holds true.*

Proof. Condition (1.30) is the discrete counterpart of condition (1.4) in the BNB Theorem. Hence, owing to this theorem, discrete well-posedness implies (1.30). Conversely, to prove that (1.30) implies discrete well-posedness, we first observe that (1.30) implies that the discrete operator A_h defined by (1.27) is injective. Indeed, $A_h v_h = 0$ yields $a_h(v_h, w_h) = 0$, for all $w_h \in V_h$. Hence, $v_h = 0$ by (1.30). In finite dimension, this implies that A_h is surjective. Hence, A_h is bijective. □

We observe that discrete well-posedness is equivalent to only one condition, namely (1.30), while two conditions appear in the continuous case. This is because injectivity is equivalent to bijectivity when the test and trial spaces have the same finite dimension.

A sufficient, and often easy to verify, condition for discrete stability is coercivity. This property can be stated as follows: There is $C_{\mathrm{sta}} > 0$ such that

$$\forall v_h \in V_h, \qquad C_{\mathrm{sta}} |||v_h|||^2 \leq a_h(v_h, v_h). \tag{1.31}$$

Discrete coercivity implies the discrete inf-sup condition (1.30) since, for all $v_h \in V_h \setminus \{0\}$,

$$C_{\mathrm{sta}} |||v_h||| \leq \frac{a_h(v_h, v_h)}{|||v_h|||} \leq \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{|||w_h|||}.$$

Property (1.31) is the discrete counterpart of that invoked in the Lax–Milgram Lemma.

1.3.3 Consistency

For the time being, we consider a rather strong form of consistency, namely that the exact solution u satisfies the discrete equations in (1.26). To formulate consistency, it is thus necessary to plug the exact solution into the first argument of the discrete bilinear form a_h, and this may not be possible in general since the discrete bilinear form a_h is so far defined on $V_h \times V_h$ only. Therefore, we assume that there is a subspace $X_* \subset X$ such that the exact solution u belongs to X_* and such that the discrete bilinear form a_h can be extended to $X_* \times V_h$ (it is not possible in general to extend a_h to $X \times V_h$). Consistency can now be formulated as follows.

Definition 1.31 (Consistency)**.** We say that the discrete problem (1.26) is *consistent* if for the exact solution $u \in X_*$,

$$a_h(u, w_h) = l_h(w_h) \qquad \forall w_h \in V_h. \tag{1.32}$$

Remark 1.32 (Galerkin orthogonality). Consistency is equivalent to the usual *Galerkin orthogonality property* often considered in the context of finite element methods. Indeed, (1.32) holds true if and only if

$$a_h(u - u_h, w_h) = 0 \qquad \forall w_h \in V_h.$$

1.3.4 Boundedness

The last ingredient in the error analysis is boundedness. We introduce the vector space

$$X_{*h} := X_* + V_h,$$

and observe that the *approximation error* $(u - u_h)$ belongs to this space. We aim at measuring the approximation error using the discrete stability norm $|||\cdot|||$. Therefore, we assume in what follows that this norm can be extended to the space X_{*h}. In the present setting, we want to assert boundedness in the product space $X_{*h} \times V_h$, and not just in $V_h \times V_h$. It turns out that in most situations, it is not possible to assert boundedness using only the discrete stability norm $|||\cdot|||$. This is the reason why we introduce a second norm, say $|||\cdot|||_*$.

Definition 1.33 (Boundedness)**.** We say that the discrete bilinear form a_h is *bounded* in $X_{*h} \times V_h$ if there is C_{bnd} such that

$$\forall (v, w_h) \in X_{*h} \times V_h, \qquad |a_h(v, w_h)| \le C_{\mathrm{bnd}} |||v|||_* |||w_h|||,$$

for a norm $|||\cdot|||_*$ defined on X_{*h} and such that, for all $v \in X_{*h}$, $|||v||| \le |||v|||_*$.

Remark 1.34 (h-dependency). In Definition 1.33, C_{bnd} can depend on the meshsize h. As mentioned above, in view of convergence analysis, it is important to ensure that C_{bnd} be independent of h.

1.3.5 Error Estimate

We can now state the main result of this section.

Theorem 1.35 (Abstract error estimate). *Let u solve (1.1) with $f \in L^2(\Omega)$. Let u_h solve (1.26). Let $X_* \subset X$ and assume that $u \in X_*$. Set $X_{*h} = X_* + V_h$ and assume that the discrete bilinear form a_h can be extended to $X_{*h} \times V_h$. Let $|||\cdot|||$ and $|||\cdot|||_*$ be two norms defined on X_{*h} and such that, for all $v \in X_{*h}$, $|||v||| \le |||v|||_*$. Assume discrete stability, consistency, and boundedness. Then, the following error estimate holds true:*

$$|||u - u_h||| \le C \inf_{y_h \in V_h} |||u - y_h|||_*, \tag{1.33}$$

with $C = 1 + C_{\mathrm{sta}}^{-1} C_{\mathrm{bnd}}$.

Proof. Let $y_h \in V_h$. Owing to discrete stability and consistency,

$$|||u_h - y_h||| \le C_{\mathrm{sta}}^{-1} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(u_h - y_h, w_h)}{|||w_h|||} = C_{\mathrm{sta}}^{-1} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(u - y_h, w_h)}{|||w_h|||}.$$

Hence, owing to boundedness,

$$|||u_h - y_h||| \le C_{\mathrm{sta}}^{-1} C_{\mathrm{bnd}} |||u - y_h|||_*.$$

Estimate (1.33) then results from the triangle inequality, the fact that $|||u - y_h||| \le |||u - y_h|||_*$, and that y_h is arbitrary in V_h.

□

1.4 Admissible Mesh Sequences

The goal of this section is to derive some technical, yet important, tools to analyze the convergence of dG methods as the meshsize goes to zero. We consider a mesh sequence

$$\mathcal{T}_{\mathcal{H}} := (\mathcal{T}_h)_{h \in \mathcal{H}},$$

where $\mathcal{H}$ denotes a countable subset of $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$ having 0 as only accumulation point. Our analysis tools are, on the one hand, inverse and trace inequalities that are instrumental to assert discrete stability and boundedness uniformly in h and, on the other hand, optimal polynomial approximation properties so as to infer from error estimates of the form (1.33) h-convergence rates for the approximation error whenever the exact solution is smooth enough.

In Sects. 1.4.1–1.4.4, we consider the case $d \ge 2$. We first introduce the concept of shape- and contact-regular mesh sequences, which is sufficient to derive inverse and trace inequalities, and then we combine it with an additional requirement on optimal polynomial approximation properties, leading to the concept of admissible mesh sequences. Finally, in Sect. 1.4.5, we deal with the case $d = 1$, where the requirements on admissible mesh sequences are much simpler.

1.4.1 Shape and Contact Regularity

A useful concept encountered in the context of conforming finite element methods is that of matching simplicial meshes.

Definition 1.36 (Matching simplicial mesh)**.** We say that $\mathcal{T}_h$ is a *matching simplicial mesh* if it is a simplicial mesh and if for any $T \in \mathcal{T}_h$ with vertices $\{a_0, \ldots, a_d\}$, the set $\partial T \cap \partial T'$ for any $T' \in \mathcal{T}_h$, $T' \neq T$, is the convex hull of a (possibly empty) subset of $\{a_0, \ldots, a_d\}$.

For instance, in dimension 2, the set $\partial T \cap \partial T'$ for two distinct elements of a matching simplicial mesh is either empty, or a common vertex, or a common edge of the two elements. We now turn to the matching simplicial submesh of a general mesh.

Definition 1.37 (Matching simplicial submesh)**.** Let $\mathcal{T}_h$ be a general mesh. We say that $\mathfrak{S}_h$ is a *matching simplicial submesh* of $\mathcal{T}_h$ if

(i) $\mathfrak{S}_h$ is a matching simplicial mesh,

(ii) For all $T' \in \mathfrak{S}_h$, there is only one $T \in \mathcal{T}_h$ such that $T' \subset T$,

(iii) For all $F' \in \mathfrak{F}_h$, the set collecting the mesh faces of $\mathfrak{S}_h$, there is at most one $F \in \mathcal{F}_h$ such that $F' \subset F$.

The simplices in $\mathfrak{S}_h$ are called *subelements*, and the mesh faces in $\mathfrak{F}_h$ are called *subfaces*. We set, for all $T \in \mathcal{T}_h$,

$$\mathfrak{S}_T := \{T' \in \mathfrak{S}_h \mid T' \subset T\},$$
$$\mathfrak{F}_T := \{F' \in \mathfrak{F}_h \mid F' \subset \partial T\}.$$

We also set, for all $F \in \mathcal{F}_h$,

$$\mathfrak{F}_F := \{F' \in \mathfrak{F}_h \mid F' \subset F\}.$$

Figure 1.5 illustrates the matching simplicial submesh for two polygonal mesh elements, say T_1 and T_2, that come into contact. The triangular subelements composing the sets $\mathfrak{S}_{T_1}$ and $\mathfrak{S}_{T_2}$ are indicated by dashed lines. We observe that the mesh face $F = \partial T_1 \cap \partial T_2$ (highlighted in bold) is not a part of a hyperplane and that the set $\mathfrak{F}_F$ contains two subfaces.

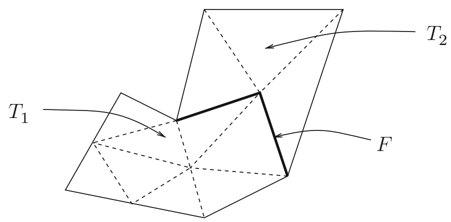

Fig. 1.5: Two polygonal mesh elements that come into contact with corresponding subelements indicated by *dashed lines* and interface indicated in *bold*

Definition 1.38 (Shape and contact regularity)**.** We say that the mesh sequence $\mathcal{T}_{\mathcal{H}}$ is *shape- and contact-regular* if for all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathfrak{S}_h$ such that

(i) The mesh sequence $\mathfrak{S}_{\mathcal{H}}$ is shape-regular in the usual sense of Ciarlet [92], meaning that there is a parameter $\varrho_1 > 0$, independent of h, such that, for all $T' \in \mathfrak{S}_h$,
$$\varrho_1 h_{T'} \le r_{T'},$$
where $h_{T'}$ is the diameter of T' and $r_{T'}$ the radius of the largest ball inscribed in T',

(ii) There is a parameter $\varrho_2 > 0$, independent of h, such that, for all $T \in \mathcal{T}_h$ and for all $T' \in \mathfrak{S}_T$,
$$\varrho_2 h_T \le h_{T'}.$$

Henceforth, the parameters ϱ_1 and ϱ_2 are called the *mesh regularity parameters* and are collectively denoted by the symbol ϱ. Finally, if $\mathcal{T}_h$ is itself matching and simplicial, then $\mathfrak{S}_h = \mathcal{T}_h$ and the only requirement is shape-regularity with parameter $\varrho_1 > 0$ independent of h.

As elaborated in Sect. 1.4.2, the two conditions in Definition 1.38 allow one to control the shape of the elements in $\mathcal{T}_h$ and the way these elements come into contact. The idea of considering a matching simplicial submesh has been proposed, e.g., by Brenner [51] to derive generalized Poincaré–Friedrichs inequalities in broken Sobolev spaces. More recently, in the context of dG methods, a matching simplicial submesh has been considered by Buffa and Ortner [61] for nonlinear minimization problems and by Ern and Vohralík [151] for a posteriori error estimates in the context of PDEs with diffusion.

Remark 1.39 (Anisotropic meshes). Definition 1.38 implies that the mesh is isotropic in the sense that, for all $T \in \mathcal{T}_h$, the d-dimensional measure $|T|_d$ is uniformly equivalent to h_T^d. In applications featuring sharp layers, anisotropic meshes can be advantageous. We refer the reader, e.g., to van der Vegt and van der Ven [297], Sun and Wheeler [287], Georgoulis [166], Georgoulis, Hall, and Houston [167], and Leicht and Hartmann [226] for various aspects of dG methods on anisotropic meshes.

1.4.2 Geometric Properties

This section collects some useful geometric properties of shape- and contact-regular mesh sequences. The first result is a uniform bound on $\mathrm{card}(\mathfrak{S}_T)$.

Lemma 1.40 (Bound on $\mathrm{card}(\mathfrak{S}_T)$)**.** *Let $\mathcal{T}_{\mathcal{H}}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$ and all $T \in \mathcal{T}_h$,* $\mathrm{card}(\mathfrak{S}_T)$ *is bounded uniformly in h.*

Proof. Let $|\cdot|_d$ denote the d-dimensional Hausdorff measure and let $\mathfrak{B}_d$ be the unit ball in $\mathbb{R}^d$. Then,

$$h_T^d \geq |T|_d = \sum_{T' \in \mathfrak{S}_T} |T'|_d \geq \sum_{T' \in \mathfrak{S}_T} |\mathfrak{B}_d|_d r_{T'}^d \geq \sum_{T' \in \mathfrak{S}_T} |\mathfrak{B}_d|_d \varrho_1^d h_{T'}^d$$
$$\geq \sum_{T' \in \mathfrak{S}_T} |\mathfrak{B}_d|_d \varrho_1^d \varrho_2^d h_T^d \geq |\mathfrak{B}_d|_d \varrho_1^d \varrho_2^d \operatorname{card}(\mathfrak{S}_T) h_T^d,$$

yielding the assertion. □

Our next result is a uniform bound on $\operatorname{card}(\mathcal{F}_T)$, $\operatorname{card}(\mathfrak{F}_T)$, the parameter N_∂ defined by (1.12), and $\operatorname{card}(\mathfrak{F}_F)$.

Lemma 1.41 (Bound on $\operatorname{card}(\mathcal{F}_T)$, $\operatorname{card}(\mathfrak{F}_T)$, N_∂, and $\operatorname{card}(\mathfrak{F}_F)$)**.** *Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Then, for all $h \in \mathcal{H}$ and all $T \in \mathcal{T}_h$,* $\operatorname{card}(\mathcal{F}_T)$, $\operatorname{card}(\mathfrak{F}_T)$, *and N_∂ are bounded uniformly in h, while, for all $F \in \mathcal{F}_h$,* $\operatorname{card}(\mathfrak{F}_F)$ *is bounded uniformly in h.*

Proof. We observe that

$$\operatorname{card}(\mathcal{F}_T) \leq \operatorname{card}(\mathfrak{F}_T) \leq (d+1) \operatorname{card}(\mathfrak{S}_T),$$

so that the assertion on $\operatorname{card}(\mathcal{F}_T)$ and $\operatorname{card}(\mathfrak{F}_T)$ follows from Lemma 1.40. The bound on N_∂ results from its definition (1.12) and the bound on $\operatorname{card}(\mathcal{F}_T)$. Finally, to bound $\operatorname{card}(\mathfrak{F}_F)$ for all $F \in \mathcal{F}_h$, we pick $T \in \mathcal{T}_h$ such that $F \in \mathcal{F}_T$ and observe that $\operatorname{card}(\mathfrak{F}_F) \leq (d+1) \operatorname{card}(\mathfrak{S}_T)$, so that the bound on $\operatorname{card}(\mathfrak{F}_F)$ results from the bound on $\operatorname{card}(\mathfrak{S}_T)$. □

Our next result is a lower bound on the diameter of mesh faces.

Lemma 1.42 (Lower bound on face diameters)**.** *Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Then, for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$,*

$$\delta_F \geq \varrho_1 \varrho_2 h_T, \tag{1.34}$$

where δ_F denotes the diameter *of F.*

Proof. Let $T \in \mathcal{T}_h$ and let $F \in \mathcal{F}_T$. Then, we pick $F' \in \mathfrak{F}_F$ and denote by $T' \in \mathfrak{S}_T$ the simplex to which the subface F' belongs. We obtain

$$\delta_F \geq \delta_{F'} \geq r_{T'} \geq \varrho_1 h_{T'} \geq \varrho_1 \varrho_2 h_T,$$

yielding the assertion. □

A direct consequence of Lemma 1.42 is a comparison result on the diameter of neighboring elements.

Lemma 1.43 (Diameter comparison for neighboring elements)**.** *Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Then, for all $h \in \mathcal{H}$ and all $T, T' \in \mathcal{T}_h$ sharing a face F, there holds*

$$\min(h_T, h_{T'}) \geq \varrho_1 \varrho_2 \max(h_T, h_{T'}). \tag{1.35}$$

Proof. Since F is a common face to both T and T', owing to (1.34),

$$\varrho_1 \varrho_2 \max(h_T, h_{T'}) \leq \delta_F \leq \min(h_T, h_{T'}),$$

yielding (1.35).

□

1.4.3 Inverse and Trace Inequalities

Inverse and trace inequalities are useful tools to analyze dG methods. For simplicity, we derive these inequalities on the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined by (1.15); other broken polynomial spaces can be considered.

We begin with the following inverse inequality that delivers a local upper bound on the gradient of discrete functions.

Lemma 1.44 (Inverse inequality)**.** *Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Then, for all $h \in \mathcal{H}$, all $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$, and all $T \in \mathcal{T}_h$,*

$$\|\nabla v_h\|_{[L^2(T)]^d} \leq C_{\mathrm{inv}} h_T^{-1} \|v_h\|_{L^2(T)}, \tag{1.36}$$

where C_{inv} only depends on ϱ, d, and k.

Proof. Let $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$ and let $T \in \mathcal{T}_h$. For all $T' \in \mathfrak{S}_T$, the restriction $v_h|_{T'}$ is in $\mathbb{P}_d^k(T')$. Hence, owing to the usual inverse inequality on simplices (see Brenner and Scott [54, Sect. 4.5] or Ern and Guermond [141, Sect. 1.7]),

$$\|\nabla v_h\|_{[L^2(T')]^d} \leq C_{\mathrm{inv,s}} h_{T'}^{-1} \|v_h\|_{L^2(T')},$$

where $C_{\mathrm{inv,s}}$ only depends on ϱ_1, d, and k. Using point (ii) in Definition 1.38 yields

$$\|\nabla v_h\|_{[L^2(T')]^d} \leq \varrho_2^{-1} C_{\mathrm{inv,s}} h_T^{-1} \|v_h\|_{L^2(T')}.$$

Squaring this inequality and summing over $T' \in \mathfrak{S}_T$ yields (1.36).

□

Remark 1.45 (Nature of (1.36)). The inverse inequality (1.36) is local to mesh elements. As such, it depends on the shape of the mesh elements but not on the way mesh elements come into contact.

We now turn to the following discrete trace inequality that delivers an upper bound on the face values of discrete functions.

Lemma 1.46 (Discrete trace inequality). *Let $\mathcal{T}_{\mathcal{H}}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Then, for all $h \in \mathcal{H}$, all $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$,*

$$h_T^{1/2}\|v_h\|_{L^2(F)} \le C_{\mathrm{tr}}\|v_h\|_{L^2(T)}, \tag{1.37}$$

where C_{tr} only depends on ϱ, d, and k.

Proof. Let $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$, let $T \in \mathcal{T}_h$, and let $F \in \mathcal{F}_T$. We first assume that $\mathcal{T}_h$ is a matching simplicial mesh. Let $\widehat{T}$ be the unit simplex of $\mathbb{R}^d$ and let F_T be the bijective affine map such that $F_T(\widehat{T}) = T$. Let $\widehat{F}$ be any face of $\widehat{T}$. Since the unit sphere in $\mathbb{P}_d^k(\widehat{T})$ for the $L^2(\widehat{T})$-norm is a compact set, there is $\widehat{C}_{d,k}(\widehat{F})$, only depending on d, k, and $\widehat{F}$, such that, for all $\widehat{v} \in \mathbb{P}_d^k(\widehat{T})$,

$$\|\widehat{v}\|_{L^2(\widehat{F})} \le \widehat{C}_{d,k}(\widehat{F})\|\widehat{v}\|_{L^2(\widehat{T})}.$$

Applying the above inequality to the function $\widehat{v} = v_h|_T \circ F_T^{-1}$ which is in $\mathbb{P}_d^k(\widehat{T})$, we infer

$$|F|_{d-1}^{-1/2}\|v_h\|_{L^2(F)} \le \widehat{C}_{d,k}|T|_d^{-1/2}\|v_h\|_{L^2(T)},$$

where $\widehat{C}_{d,k} := \max_{\widehat{F}\in\mathcal{F}_{\widehat{T}}} \widehat{C}_{d,k}(\widehat{F})$. Moreover, we observe that

$$\frac{|T|_d}{|F|_{d-1}} = \frac{1}{d}h_{T,F} \ge \frac{1}{d}r_T \ge \frac{1}{d}\varrho_1 h_T, \tag{1.38}$$

where $h_{T,F}$ denotes the distance of the vertex opposite to F to that face and r_T the radius of the largest ball inscribed in T. As a result,

$$h_T^{1/2}\|v_h\|_{L^2(F)} \le C_{\mathrm{tr,s}}\|v_h\|_{L^2(T)}, \tag{1.39}$$

where $C_{\mathrm{tr,s}} := d^{1/2}\varrho_1^{-1/2}\widehat{C}_{d,k}$ only depends on ϱ_1, d, and k. We now consider the case of general meshes. For each $F' \in \mathfrak{F}_F$, let T' denote the simplex in $\mathfrak{S}_T$ of which F' is a face. Since the restriction $v_h|_{T'}$ is in $\mathbb{P}_d^k(T')$, the discrete trace inequality (1.39) yields

$$h_{T'}^{1/2}\|v_h\|_{L^2(F')} \le C_{\mathrm{tr,s}}\|v_h\|_{L^2(T')} \le C_{\mathrm{tr,s}}\|v_h\|_{L^2(T)}.$$

Squaring this inequality and summing over $F' \in \mathfrak{F}_F$ yields

$$\left(\sum_{F'\in\mathfrak{F}_F} h_{T'}\|v_h\|_{L^2(F')}^2\right)^{1/2} \le C_{\mathrm{tr,s}}\,\mathrm{card}(\mathfrak{F}_F)^{1/2}\|v_h\|_{L^2(T)},$$

whence the assertion follows, since $h_{T'} \ge \varrho_2 h_T$ and $\mathrm{card}(\mathfrak{F}_F)$ is bounded uniformly owing to Lemma 1.41.

□

Remark 1.47 (Variant of (1.37)). Summing over $F \in \mathcal{F}_T$, we infer from (1.37) and the Cauchy–Schwarz inequality that

$$h_T^{1/2}\|v_h\|_{L^2(\partial T)} \le C_{\mathrm{tr}} N_\partial^{1/2} \|v_h\|_{L^2(T)}, \tag{1.40}$$

and we recall from Lemma 1.41 that N_∂ is bounded uniformly in h.

Remark 1.48 (k-dependency). When working with high-degree polynomials, it is important to determine the dependency of C_{inv} and C_{tr} on the polynomial degree k. This turns out to be a delicate question, and precise answers are available only in specific cases. Concerning the discrete trace inequality (1.37), it is proven by Warburton and Hesthaven [304] that on $\mathbb{P}_d^k(\mathcal{T}_h)$, C_{tr} scales as $\sqrt{k(k+d)}$. Moreover, on $\mathbb{Q}_d^k(\mathcal{T}_h)$ with the mesh elements being affine images of the unit hypercube in $\mathbb{R}^d$, one-dimensional results can be exploited using tensor-product polynomials, yielding that C_{tr} scales as $\sqrt{k(k+1)}$; see Canuto and Quarteroni [75], Bernardi and Maday [42], and Schwab [275]. One difficulty concerns the behavior of polynomials near the end points of an interval, and this can be dealt with by considering weighted norms in (1.37); see, e.g., Melenk and Wohlmuth [235]. Concerning the inverse inequality (1.36), C_{inv} scales as k^2 on triangles and parallelograms; see, e.g., Schwab [275].

We also need the following continuous trace inequality, which delivers an upper bound on the face values of functions in the broken Sobolev space $H^1(\mathcal{T}_h)$. We present here a simple proof inspired by Monk and Süli [238] and Carstensen and Funken [77] (see also Stephansen [283, Lemma 3.12]).

Lemma 1.49 (Continuous trace inequality). *Let $\mathcal{T}_{\mathcal{H}}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$, all $v \in H^1(\mathcal{T}_h)$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$,*

$$\|v\|_{L^2(F)}^2 \le C_{\mathrm{cti}}(2\|\nabla v\|_{[L^2(T)]^d} + d h_T^{-1}\|v\|_{L^2(T)})\|v\|_{L^2(T)}, \tag{1.41}$$

with $C_{\mathrm{cti}} := \varrho_1^{-1}$ if $\mathcal{T}_h$ is matching and simplicial, while $C_{\mathrm{cti}} := (1+d)(\varrho_1\varrho_2)^{-1}$ otherwise.

Proof. Let $v \in H^1(\mathcal{T}_h)$, let $T \in \mathcal{T}_h$, and let $F \in \mathcal{F}_T$. Assume first that T is a simplex and consider the $\mathbb{R}^d$-valued function

$$\sigma_F = \frac{|F|_{d-1}}{d|T|_d}(x - a_F),$$

where a_F is the vertex of T opposite to F; cf. Fig. 1.6. The normal component of σ_F is constant and equal to one on F, and it vanishes on all the remaining faces in $\mathcal{F}_T$. (The function σ_F is proportional to the lowest-order Raviart–Thomas–Nédélec shape function in T; see, e.g., Brezzi and Fortin [57, p. 116] or Ern and Guermond [141, Sect. 1.2.7] and also cf. Sect. 5.5.3.) Owing to the divergence theorem,

$$\begin{aligned}\|v\|_{L^2(F)}^2 = \int_F |v|^2 = \int_{\partial T} |v|^2(\sigma_F\cdot \mathrm{n}_T) = \int_T \nabla\cdot(|v|^2\sigma_F)\\ = \int_T 2v\sigma_F\cdot\nabla v + \int_T |v|^2(\nabla\cdot\sigma_F).\end{aligned}$$

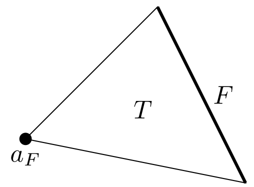

Fig. 1.6: Notation for the proof of Lemma 1.49

Since

$$\|\sigma_F\|_{[L^\infty(T)]^d} \le \frac{|F|_{d-1} h_T}{d|T|_d}, \qquad \nabla{\cdot}\sigma_F = \frac{|F|_{d-1}}{|T|_d},$$

we infer using the Cauchy–Schwarz inequality that

$$\|v\|^2_{L^2(F)} \le \frac{|F|_{d-1} h_T}{d|T|_d} \left(2\|\nabla v\|_{[L^2(T)]^d} + d h_T^{-1} \|v\|_{L^2(T)}\right) \|v\|_{L^2(T)}. \tag{1.42}$$

Using (1.38) yields (1.41) with $C_{\mathrm{cti}} = \varrho_1^{-1}$ on matching simplicial meshes. Consider now the case where T belongs to a general mesh. For each $F' \in \mathfrak{F}_F$, let T' denote the simplex in $\mathfrak{S}_T$ of which F' is a face. Applying the continuous trace inequality using F' and T' yields

$$\|v\|^2_{L^2(F')} \le \varrho_1^{-1} \left(2\|\nabla v\|_{[L^2(T')]^d} + d h_{T'}^{-1} \|v\|_{L^2(T')}\right) \|v\|_{L^2(T')}.$$

Since $h_{T'} \ge \varrho_2 h_T$ and $\varrho_2 \le 1$, we infer

$$\|v\|^2_{L^2(F')} \le (\varrho_1\varrho_2)^{-1} \left(2\|\nabla v\|_{[L^2(T')]^d} + d h_T^{-1} \|v\|_{L^2(T')}\right) \|v\|_{L^2(T')}.$$

Summing over $F' \in \mathfrak{F}_F$ and using the Cauchy–Schwarz inequality yields (1.41) since each $T' \in \mathfrak{S}_T$ appears at most $(d+1)$ times in the summation.

□

We close this section with some useful inverse and trace inequalities in a non-Hilbertian setting. Our first result allows us to compare the $\|\cdot\|_{L^p(T)}$- and $\|\cdot\|_{L^q(T)}$-norms. The proof is skipped since it hinges again on the simplicial submesh and the corresponding inverse inequality on simplices.

Lemma 1.50 (Comparison of $\|\cdot\|_{L^p(T)}$- and $\|\cdot\|_{L^q(T)}$-norms)**.** *Let $\mathcal{T}_{\mathcal{H}}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Let $1 \le p, q \le \infty$ be two real numbers. Then, for all $h \in \mathcal{H}$, all $v_h \in \mathbb{P}^k_d(\mathcal{T}_h)$, and all $T \in \mathcal{T}_h$,*

$$\|v_h\|_{L^p(T)} \le C_{\mathrm{inv},p,q} h_T^{d(1/p-1/q)} \|v_h\|_{L^q(T)}, \tag{1.43}$$

where $C_{\mathrm{inv},p,q}$ only depends on ϱ, d, k, p, and q.

Remark 1.51 (Dependency on p and q). The quantity $C_{\mathrm{inv},p,q}$ in (1.43) can be uniformly bounded in p and q. We first observe that owing to Hölder's inequality, $C_{\mathrm{inv},p,q} = 1$ if $p < q$. Additionally, it is shown by Verfürth [299] that on simplices, for $p > 2$ and $q = 2$, $C_{\mathrm{inv},p,2} \le C_{d,k}^{1-2/p}$ and for $p = 2$ and $q < 2$, $C_{\mathrm{inv},2,q} \le C_{d,k}^{2/q-1}$, where $C_{d,k} := ((2k+2)(4k+2)^{d-1})^{1/2}$. Hence, the largest value for $C_{\mathrm{inv},p,q}$ is obtained in the case where $p > 2 > q$, and observing that $|1 - 2/p| \le 1$, $|1 - 2/q| \le 1$, and $C_{d,k} \ge 1$, we infer that the quantity $C_{\mathrm{inv},p,q}$ can always be bounded by $C_{d,k}^2$. A uniform upper bound in p and q can also be derived on general meshes by summing over the simplicial subelements.

Our second and last result is a non-Hilbertian version of the discrete trace inequality (1.37).

Lemma 1.52 (Discrete trace inequality in $L^p(F)$)**.** *Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence with parameters ϱ. Let $1 \le p \le \infty$ be a real number. Then, for all $h \in \mathcal{H}$, all $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$,*

$$h_T^{1/p}\|v_h\|_{L^p(F)} \le C_{\mathrm{tr},p}\|v_h\|_{L^p(T)}, \tag{1.44}$$

where $C_{\mathrm{tr},p}$ only depends on ϱ, d, k, and p.

Proof. We combine the discrete trace inequality (1.37) with the inverse inequality (1.43) (in F) to infer

$$\begin{aligned} h_T^{1/p}\|v_h\|_{L^p(F)} &\le C_{\mathrm{inv},p,2} h_T^{1/p} \delta_F^{(d-1)(1/p-1/2)}\|v_h\|_{L^2(F)} \\ &\le C_{\mathrm{inv},p,2} C_{\mathrm{tr}} h_T^{1/p-1/2} \delta_F^{(d-1)(1/p-1/2)}\|v_h\|_{L^2(T)} \\ &\le C_{\mathrm{inv},p,2} C_{\mathrm{tr}} C_{\mathrm{inv},2,p} h_T^{1/p-1/2} \delta_F^{(d-1)(1/p-1/2)} h_T^{d(1/2-1/p)}\|v_h\|_{L^p(T)}. \end{aligned}$$

The assertion follows by observing that δ_F in uniformly equivalent to h_T. □

Remark 1.53 (Dependency on p). The quantity $C_{\mathrm{tr},p}$ in (1.44) can be uniformly bounded in p. This is a direct consequence of the above proof and Remark 1.51.

1.4.4 Polynomial Approximation

To infer from estimate (1.33) a convergence rate in h for the *approximation error* $(u - u_h)$ measured in the $|||\cdot|||$-norm when the exact solution u is smooth enough, we need to estimate the right-hand side given by

$$\inf_{y_h \in V_h} |||u - y_h|||_*,$$

when V_h is typically the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined by (1.15); other broken polynomial spaces can be considered. Since $u_h \in V_h$, we infer from (1.33) that

$$\inf_{y_h \in V_h} |||u - y_h||| \le |||u - u_h||| \le C \inf_{y_h \in V_h} |||u - y_h|||_*. \tag{1.45}$$

Definition 1.54 (Optimality, quasi-optimality, and suboptimality of the error estimate)**.** We say that the error estimate (1.45) is

(i) *Optimal* if $|||\cdot||| = |||\cdot|||_*$,

(ii) *Quasi-optimal* if the two norms are different, but the lower and upper bounds in (1.45) converge, for smooth u, at the same convergence rate as $h \to 0$,

(iii) *Suboptimal* if the upper bound converges at a slower rate than the lower bound.

The analysis of the upper bound $\inf_{y_h \in V_h} |||u - y_h|||_*$ depends on the polynomial approximation properties that can be achieved in the broken polynomial space V_h.

Definition 1.55 (Optimal polynomial approximation)**.** We say that the mesh sequence $\mathcal{T}_\mathcal{H}$ has *optimal polynomial approximation* properties if, for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all polynomial degree k, there is a linear interpolation operator $\mathcal{I}_T^k : L^2(T) \to \mathbb{P}_d^k(T)$ such that, for all $s \in \{0, \dots, k+1\}$ and all $v \in H^s(T)$, there holds

$$|v - \mathcal{I}_T^k v|_{H^m(T)} \leq C_{\text{app}} h_T^{s-m} |v|_{H^s(T)} \qquad \forall m \in \{0, \dots, s\}, \tag{1.46}$$

where C_{app} is independent of both T and h.

Remark 1.56 (Nature of (1.46)). As for the inverse inequality (1.36), the optimal polynomial approximation property (1.46) is local to mesh elements. As such, it depends on the shape of the mesh elements, but not on the way mesh elements come into contact.

Definition 1.57 (Admissible mesh sequences)**.** We say that the mesh sequence $\mathcal{T}_\mathcal{H}$ is *admissible* if it is shape- and contact-regular and if it has optimal polynomial approximation properties.

In what follows, we often consider the L^2-orthogonal projection onto the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$. One reason is that its definition (cf. (1.29)) is very simple, even on general meshes.

Lemma 1.58 (Optimality of L^2-orthogonal projection)**.** *Let $\mathcal{T}_\mathcal{H}$ be an admissible mesh sequence. Let π_h be the L^2-orthogonal projection onto $\mathbb{P}_d^k(\mathcal{T}_h)$. Then, for all $s \in \{0, \dots, k+1\}$ and all $v \in H^s(T)$, there holds*

$$|v - \pi_h v|_{H^m(T)} \leq C'_{\text{app}} h_T^{s-m} |v|_{H^s(T)} \qquad \forall m \in \{0, \dots, s\}, \tag{1.47}$$

where C'_{app} is independent of both T and h.

Proof. For $m = 0$, we obtain by definition of the L^2-orthogonal projection,

$$\|v - \pi_h v\|_{L^2(T)} \leq \|v - \mathcal{I}_T^k v\|_{L^2(T)} \leq C_{\text{app}} h_T^s |v|_{H^s(T)}.$$

For $m \geq 1$, we use m times the inverse inequality (1.36) together with the triangle inequality to infer

$$\begin{aligned}|v - \pi_h v|_{H^m(T)} &\leq |v - \mathcal{I}_T^k v|_{H^m(T)} + |\mathcal{I}_T^k v - \pi_h v|_{H^m(T)} \\ &\leq |v - \mathcal{I}_T^k v|_{H^m(T)} + C' h_T^{-m} \|\mathcal{I}_T^k v - \pi_h v\|_{L^2(T)} \\ &\leq |v - \mathcal{I}_T^k v|_{H^m(T)} + 2C' h_T^{-m} \|v - \mathcal{I}_T^k v\|_{L^2(T)},\end{aligned}$$

where C' has the same dependencies as C'_{app}, whence (1.47) owing to (1.46). □

In the analysis of dG methods, we often need to bound polynomial approximation errors on mesh faces. The following result is a direct consequence of (1.47) and of the continuous trace inequality in Lemma 1.49.

Lemma 1.59 (Polynomial approximation on mesh faces)**.** *Under the hypotheses of Lemma 1.58, assume additionally that $s \geq 1$. Then, for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$, there holds*

$$\|v - \pi_h v\|_{L^2(F)} \leq C''_{\mathrm{app}} h_T^{s-1/2} |v|_{H^s(T)},$$

and if $s \geq 2$,

$$\|\nabla(v - \pi_h v)|_T \cdot \mathrm{n}_T\|_{L^2(F)} \leq C'''_{\mathrm{app}} h_T^{s-3/2} |v|_{H^s(T)},$$

where C''_{app} and C'''_{app} are independent of both T and h.

Lemmata 1.58 and 1.59 are instrumental in deriving convergence rates, as the meshsize goes to zero, for the approximation error owing to the error estimate (1.33) which yields

$$|\!|\!| u - u_h |\!|\!| \leq \inf_{y_h \in V_h} |\!|\!| u - y_h |\!|\!|_* \leq |\!|\!| u - \pi_h u |\!|\!|_*.$$

On general meshes, asserting optimal polynomial approximation is a delicate question since this property depends on the shape of mesh elements. In practice, meshes are generated by successive refinements of an initial mesh, and the shape of mesh elements depends on the refinement procedure. It is convenient to identify sufficient conditions on the mesh sequence $\mathcal{T}_{\mathcal{H}}$ to assert optimal polynomial approximation in broken polynomial spaces. One approach is based on the star-shaped property with respect to a ball.

Definition 1.60 (Star-shaped property with respect to a ball)**.** We say that a polyhedron P is *star-shaped with respect to a ball* if there is a ball $\mathfrak{B}_P \subset P$ such that, for all $x \in P$, the convex hull of $\{x\} \cup \mathfrak{B}_P$ is included in $\overline{P}$.

Figure 1.7 displays two polyhedra. The one on the left is star-shaped with respect to the ball indicated in black. Instead, the one on the right is not star-shaped with respect to any ball.

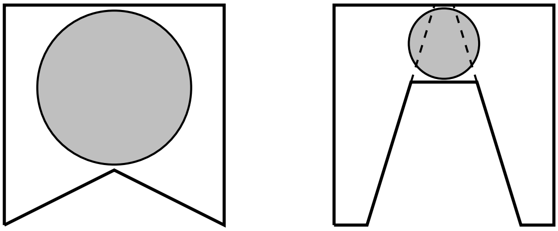

Fig. 1.7: Example (*left*) and counter-example (*right*) of a polyhedron which is star-shaped with respect to a ball

Lemma 1.61 (Mesh sequences with star-shaped property)**.** *Let $\mathcal{T}_{\mathcal{H}}$ be a shape- and contact-regular mesh sequence. Assume that, for all $h \in \mathcal{H}$ and all $T \in \mathcal{T}_h$, the mesh element T is* star-shaped with respect to a ball *with uniformly comparable diameter with respect to h_T. Then, the mesh sequence $\mathcal{T}_{\mathcal{H}}$ is admissible.*

Proof. Optimal polynomial approximation is proven by Brenner and Scott [54, Chap. 4] using averaged Taylor polynomials.

□

Another sufficient condition ensuring optimal polynomial approximation, but somewhat less general than the star-shaped property, is that of finitely shaped mesh sequences. A simple example is that of shape- and contact-regular mesh sequences whose elements are either simplices or parallelotopes in $\mathbb{R}^d$.

Lemma 1.62 (Finitely shaped mesh sequences)**.** *Let $\mathcal{T}_{\mathcal{H}}$ be a shape- and contact-regular mesh sequence. Assume that $\mathcal{T}_{\mathcal{H}}$ is* finitely shaped *in the sense that there is a finite set $\widehat{\mathcal{R}} = \{\widehat{T}\}$ whose elements are reference polyhedra in $\mathbb{R}^d$ and such that, for all $h \in \mathcal{H}$, each $T \in \mathcal{T}_h$ is the image of a reference polyhedron in $\widehat{\mathcal{R}}$ by an affine bijective map F_T. Then, the mesh sequence $\mathcal{T}_{\mathcal{H}}$ is admissible.*

Proof. The proof is sketched since it uses classical finite element techniques. Let $h \in \mathcal{H}$ and let $T \in \mathcal{T}_h$. Let $v \in H^s(T)$. Since T is such that $T = F_T(\widehat{T})$ for some $\widehat{T} \in \widehat{\mathcal{R}}$, we set $\widehat{v} = v \circ F_T$ and observe that $\widehat{v} \in H^s(\widehat{T})$. Let $k \geq 0$ and let $s \in \{0, \ldots, k+1\}$. Owing to the Deny–Lions Lemma (see Deny and Lions [123] or Ern and Guermond [141, Lemma B.67]), we infer

$$|\widehat{v} - \pi^k_{\widehat{T}} \widehat{v}|_{H^m(\widehat{T})} \leq C_{\widehat{T}} |\widehat{v}|_{H^s(\widehat{T})} \qquad \forall m \in \{0, \ldots, s\},$$

where $\pi^k_{\widehat{T}}$ is the L^2-orthogonal projection onto $\mathbb{P}^k_d(\widehat{T})$. Since T contains a ball with diameter comparable to h_T, transforming back to T yields

$$|v - \mathcal{I}^k_T v|_{H^m(T)} \leq C'_{\widehat{T}} h_T^{s-m} |v|_{H^s(T)},$$

where $\mathcal{I}^k_T v = (\pi^k_{\widehat{T}} \widehat{v}) \circ F_T^{-1}$. The assertion follows by taking the maximal value of $C'_{\widehat{T}}$ for all $\widehat{T} \in \widehat{\mathcal{R}}$, and this yields a bounded quantity since the set $\widehat{\mathcal{R}}$ is finite.

□

Remark 1.63 (Role of finitely-shaped assumption). As reflected in the above proof, the assumption that the mesh sequence is finitely shaped allows us to derive a uniform upper bound on the quantities $C'_{\widehat{T}}$ resulting from the application of the Deny–Lions Lemma in each mesh element.

Remark 1.64 (Quadrangular meshes). In dimension 2, it is possible to consider more general quadrangular meshes, where the elements are generated using bilinear mappings from the reference unit square in $\mathbb{R}^2$. Then, under the regularity conditions derived by Girault and Raviart [170, Sect. A.2], optimal polynomial approximation is achieved using polynomials in $\mathbb{Q}_d^k$.

1.4.5 The One-Dimensional Case

The situation is much simpler in dimension 1 where all the mesh elements are intervals which can come into contact only through their endpoints. Thus, shape and contact regularity are void concepts. Moreover, optimal polynomial approximation properties can be classically asserted. However, we need the counterpart of Lemma 1.43 to compare the diameter of neighboring mesh intervals. This is the only requirement for admissibility of mesh sequences.

Definition 1.65 (Admissible mesh sequences, $d = 1$)**.** We say that the mesh sequence $\mathcal{T}_{\mathcal{H}}$ is *admissible* if there is $\varrho > 0$ such that, for all $h \in \mathcal{H}$ and all $T, T' \in \mathcal{T}_h$ with $\overline{T} \cap \overline{T}'$ nonempty,

$$\min(h_T, h_{T'}) \geq \varrho \max(h_T, h_{T'}).$$

The important properties derived in Sects. 1.4.3 and 1.4.4 are available on admissible mesh sequences in dimension 1, namely:

(a) The inverse inequality (1.36) holds true.

(b) The discrete trace inequality (1.37) holds true (recalling that, in dimension 1, face integrals reduce to pointwise evaluation) and the continuous trace inequality (1.41) also holds true (recalling that the 0-dimensional Hausdorff measure of a face is conventionally set to 1).

(c) The optimal polynomial approximation property (1.46) holds true, together with the optimal bounds on the L^2-projection (cf. Lemmata 1.58 and 1.59).