On Discontinuous- and Continuous-In-Time Unfitted Space-Time Methods for PDEs on Moving Domains

Master's thesis

Institute for Numerical and Applied Mathematics
University of Göttingen

October 2020

Supervisor: Prof. Dr. Christoph Lehrenfeld
Second Assessor: Prof. Dr. Gert Lube
Author: Fabian Heimann

Contents

1 Introduction

Partial differential equations can be used to describe a broad variety of problems in physics and engineering. Often, interesting geometries of the physical setting are so complicated that analytical methods cannot deliver solutions to the differential equation, so that numerical methods are needed. One class of such methods are Finite Element methods, which subdivide the domain into a mesh and turn the differential equation problem into a (linear) algebra problem by mapping the mesh elements to a reference configuration where the so called Finite Element is defined. While the Finite Element method is typically introduced first in application to an elliptic problem such as the Poisson equation, it also can be applied to a broader class of problems.

In this thesis, we want to focus on the convection-diffusion equation as a parabolic model problem. Imagine we are given a domain Ω, typically $\Omega \subseteq \mathbb{R}^2$ or $\Omega \subseteq \mathbb{R}^3$, and a quantity (such as a liquid) inside this domain. The concentration of the quantity can then by modelled by a function $u\colon \Omega \to \mathbb{R}$ at each time. Further assume that there is a transport field $\mathbf{w}$ acting on the quantity, because, for instance, the liquid is mixed. Then the convection process will be governed by a generalisation of the simple one dimensional transport equation, which reads

$$\partial_t u(x,t) + c\partial_x u(x,t) = 0 \quad \text{for all } (x,t) \in \Omega \times [0,T], c \in \mathbb{R}.$$

The n-dimensional generalisation of this equation would then be

$$\partial_t u(x,t) + \operatorname{div}(\mathbf{w}u(x,t)) = 0 \quad \text{for all } (x,t) \in \Omega \times [0,T].$$

Now, for the convection-diffusion problem, we add diffusion as a physical process. This process is known from the elliptic Poisson case or the heat equation and causes a high concentration at some point to smoothen out in the vicinity. A typical real-world example for this would be a drop of ink poured into water. In addition, we allow for a source of concentration by means of a right-hand side f in order to arrive at the differential equation

$$\partial_t u(x,t) + \operatorname{div}(\mathbf{w}u(x,t)) - \alpha\Delta u(x,t) = f(x,t). \tag{1}$$

This is the *convection-diffusion equation*. It needs to be supplied with suitable initial and boundary condition, which we will do later.

So far, we assumed the domain Ω to be fixed over the total time of interest. This is a reasonable assumption covering many relevant examples, but the generalisation of a time dependent domain $\Omega(t)$ is of interest, too. One example for this general framework is given

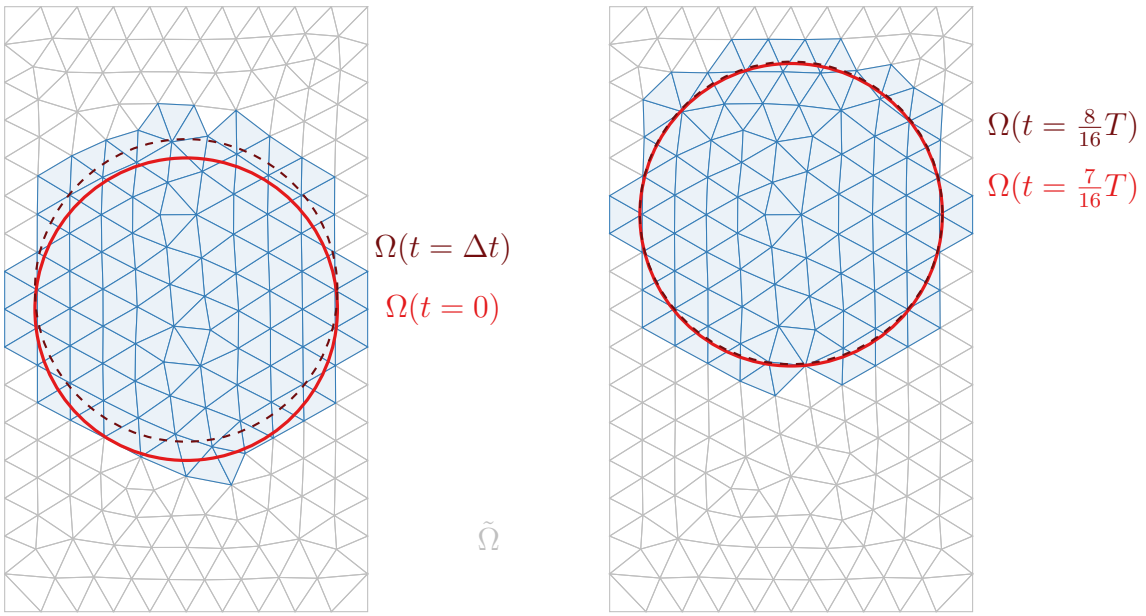


Figure 1: Example of a domain $\Omega(t)$ moving with time. Some aspects of the unfitted DG method introduced later are already hinted at. The convection field points merely in y-direction and changes with time.

by a domain which is shifted itself by the convection field $\mathbf{w}$. In addition, the setting of the convection-diffusion equation posed on a varying domain is important as a preparation for more involved problems such as two-phase flow problems, where the same problem (but vector-valued) is essentially posed on two domains of different species. This is why we want to study the convection-diffusion equation posed on a time-dependent domain in this thesis. In Fig. 1, a first illustration of a moving domain problem from this thesis is given.

The domain Ω—time-dependent or not—might have a computationally simple or complex structure. Let us for a moment consider the time-independent case. There, simple examples of domains $\Omega \subseteq \mathbb{R}^2$ would be polygonals such as the unit square $\Omega = [0, 1]^2$. In that case, the Finite Element method can take into account the domain directly in the triangulation $\mathcal{T}_h$; the triangles $T \in \mathcal{T}_h$ could be chosen such that $\partial\Omega$ would be subdivided into edges of triangles. But not all domains of interest are polygonal. For a more complicated example consider the circle

$$\Omega = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \leq 1\}.$$

Obviously, the strategy of representing $\partial\Omega$ by edges of the mesh triangles does not work exactly in this case. If one is interested in a method of second order accuracy in space, one can approximate Ω by a suiting polygonal. The accuracy of such an approach can be improved by applying a mapping Ψ to the effect that the edges of $\Psi(T)$ approximate the boundary $\partial\Omega$

with higher order accuracy (see [6, Subsection 3.7] for more details). In general, we call such methods which construct the mesh in accordance with the boundary of the domain *fitted discretisations*.

Apart from fitted discretisations, there exists an important class of methods called *unfitted discretisations*, where the geometry of the domain is not taken into consideration when the mesh is generated. More specifically, a background domain $\tilde{\Omega} \supseteq \Omega$ is meshed. In our example of the circle, we could for instance choose $\tilde{\Omega} = [-2, 2]^2$. Then, the Finite Element space is modified with the cut elements in mind in order to account for the fact that potentially only a part of the triangle represents the domain. That poses new challenges such as the generation of quadrature rules for the cut elements, but allows for a higher flexibility with more complicated geometries. (See e.g. [5, Section 2] for more details, or [2], or [10].)

The flexibility of unfitted methods is particularly relevant for time-dependent domains. There, a fitted method naively would need to generate a new mesh for each discrete time $t_i \in [0, T]$, which would be computationally expensive. There are strategies to deal with this issue, such as Arbitrary-Lagrangian-Eulerian formulations, but, for instance, topology changes in time still pose a challenge for such approaches. [5, p. 69] Unfitted methods can easily accomodate for a time-dependent domain. This is why we will consider unfitted methods for our model problem from now on in this thesis.

This summarises two relevant classes of options concerning the discretisation in space. In addition, our partial differential equation contains a time variable. This means that we also need to think about a discretisation in time. Basically, one can distinguish between *time stepping methods* and *space-time methods*. While the former—roughly speaking—extend the spatial method by a time stepping loop, often inspired by the according methods for ordinary differential equations, such as the implicit or explicit Euler scheme, the latter fully discretise the problem on a space-time domain.

In recent years, progress has been made with both options for time discretisation for the convection-diffusion problem on a varying domain, within the realm of unfitted methods. Concerning time stepping methods, Lehrenfeld and Olshanskii [7] recently suggested an implicit Euler-type based discretisation in time that makes use of extensions of functions from the physical domain to small neighbourhoods and performed a careful analysis. Generalisations of this methods to discretisations of higher order in space and time are subject of current research.

Concerning space-time methods, one such method was suggested and analysed by Preuß in [9]. This method and its analysis represent the starting point for this thesis. We would like to call the method a *discontinuous* (in time) *Galerkin* method because it allows for jumps in

the numerical solution across time discretisation points. Preuß introduced the method and performed an error analysis under the assumption of an exact handling of the geometries. This should mean the following: In the continuous problem, the space-time domain has the following structure[1]

$$Q = \bigcup_t \Omega(t) \times \{t\}.$$

This construction is of theoretical importance, but might in general lead to a complicated structure. Hence, in the numerical simulation we replace this integration region with a computationally feasible approximation Q^h and perform the integration over that region instead. The regions Q and Q^h are only allowed to differ in an controlled manner. This allows one to accomodate the "variational crime" of replacing the integration region by a Strang Lemma. The roadmap for one important line of thought of this thesis will be to transfer this general strategy to the specific case.

Another way to extend the results about the Discontinuous Galerkin method is to consider a corresponding *Continuous Galerkin* method. There, the numerical solution is constructed in a continuous-in-time manner. This reduces the computational effort because on each time slice there is one degree of freedom less in the time direction. However, new challenges for the analysis appear. This is why we will only introduce the CG method rigorously and just give numerical examples to illustrate how the method performes in practice.

1.1 Outline of this thesis

In line with this general plan, the remainder of this thesis is organised as follows. In the next section, Section 2, several unfitted space-time methods for our model problem will be introduced. For the reader's convenience, we will start by introducing the method of [9] as a starting point for our generalisation and its Continuous Galerkin counterpart. Afterwards, we will describe the isoparametric mapping which is used to represent the geometry with high order accuracy. That allows us then to present the final Discontinuous Galerkin method we will analyse in this thesis. This analysis is presented in the following Section 3. In Section 4, we illustrate our theoretical finding with numerical experiments. Finally, Section 5 will give a summary and contains suggestions for directions for further research.

[1] We omit a specification of the interval of t to just sketch the idea here. Concrete definition will be given later.

2 Unfitted Space-Time methods

In this section, we want to introduce several space-time-methods in detail. Before we go through those, we start with a complete introduction of the model problem. Afterwards, the DG method of [9] assuming an exact handling of geometries will be introduced, followed by its CG counterpart. Then, we will turn to our variant of the method with a weaker assumption concerning the handling of geometries.

2.1 The model problem

We already gave the strong form of the convection-diffusion equation in (1). This is the condition which has to hold everywhere in $\Omega(t)$, for all $t \in [0,T]$. Now, this equation is supplemented with appropriate initial and boundary conditions. First, at $t = 0$, the solution should be given as u_0. Then, we assume that there is no transport of the species along the (possibly moving) boundary, $\nabla u \cdot \mathbf{n}_{\partial\Omega} = 0$. Finally, we assume the convection field $\mathbf{w}$ to be divergence-free. Then, we arrive at the following full strong formulation of the problem: (c.f. also [9, Eq. 2.1])

$$\begin{aligned} \partial_t u + \mathbf{w} \cdot \nabla u - \Delta u &= f \quad \text{in } \Omega(t) \text{ for } t \in [0,T] \\ \operatorname{div}(\mathbf{w}) &= 0 \quad \text{in } \Omega(t) \text{ for } t \in [0,T] \\ \nabla u \cdot \mathbf{n}_{\partial\Omega} &= 0 \quad \text{on } \partial\Omega(t) \text{ for } t \in [0,T] \\ u(\ldots,0) &= u_0 \quad \text{in } \Omega(0). \end{aligned} \tag{2}$$

In the first line, we applied the product rule for the divergence. Further, we fixed the diffusion coefficient to $\alpha = 1$. This is the strong form of our model problem. As a technical detail, we will assume that the right-hand side function f and the convection field $\mathbf{w}$ are not only defined on $\Omega(t)$ for all $t \in [0,T]$, but also in a neighbourhood which is sufficiently large to contain the discrete versions of the space-time domain to be introduced later.

2.2 Space-Time DG assuming an exact handling of the geometry

In this subsection, we will introduce the space-time DG method of [9]. It assumes an exact handling of the geometry. This method will be the base case for the further methods to be introduced later.

2.2.1 Discrete regions

We start with the introduction of certain regions which will be relevant to the respective function spaces.

First, we follow Preuß by assuming that we are given a decomposition of the time interval $[0, T]$ into intervals $I_n = (t_{n-1}, t_n]$ $(n = 1, \dots, N)$, where each pair of adjacent time points has the same distance, $\Delta t = t_n - t_{n-1}$ for all $n = 1, \dots, N$. This is a simplification that we make especially for ease of presentation of the analysis section. Let us mention that it is however not a necessary restriction for the method itself. Exploiting the decomposition of the time interval, we abbreviate $\Omega^n := \Omega(t_n)$. Next, we can subdivide the whole space-time region into slices as follows:

$$Q^n = \bigcup_{t \in I_n} \Omega(t) \times \{t\}, \quad Q = \bigcup_{n=1}^{N} Q^n.$$

The aim of this division of the space-time domain into slices is the reduction of numerical effort. By this technique, it is not necessary to solve the problem for the full time interval $[0, T]$ in one big matrix, but to solve the problem for each smaller slice, step by step.

For the space-time boundary, we want introduce the following notation:

$$\Gamma_* = \bigcup_{t \in (0,T]} \partial\Omega(t) \times \{t\}, \quad \Gamma_*^n = \bigcup_{t \in I_n} \partial\Omega(t) \times \{t\}.$$

We assume that the domain $\Omega(t)$ is contained in a background domain $\tilde{\Omega}$, where a shape regular triangulation $\tilde{\mathcal{T}_h}$ is given. We can define space-time slices on the background domain as well:

$$\tilde{Q}^n = \bigcup_{t \in I_n} \tilde{\Omega} \times \{t\}, \quad \tilde{Q} = \bigcup_{n=1}^{N} \tilde{Q}^n = \tilde{\Omega} \times [0, T].$$

For each time slice n, we define an extension of the exact domain Q^n which has tensor product structure:

$$\mathcal{E}(\Omega^n) := \{x \in T \text{ for some } T \in \mathcal{T}_h \text{ such that } (T \times I_n) \cap Q^n \neq \varnothing\}, \quad \mathcal{E}(Q^n) = \mathcal{E}(\Omega^n) \times I_n.$$

In addition, we define an interior domain with tensor product structure as

$$\mathcal{I}(\Omega^n) := \{x \in T \text{ for some } T \in \mathcal{T}_h \text{ such that } (T \times I_n) \subseteq Q^n\}, \quad \mathcal{I}(Q^n) = \mathcal{I}(\Omega^n) \times I_n.$$

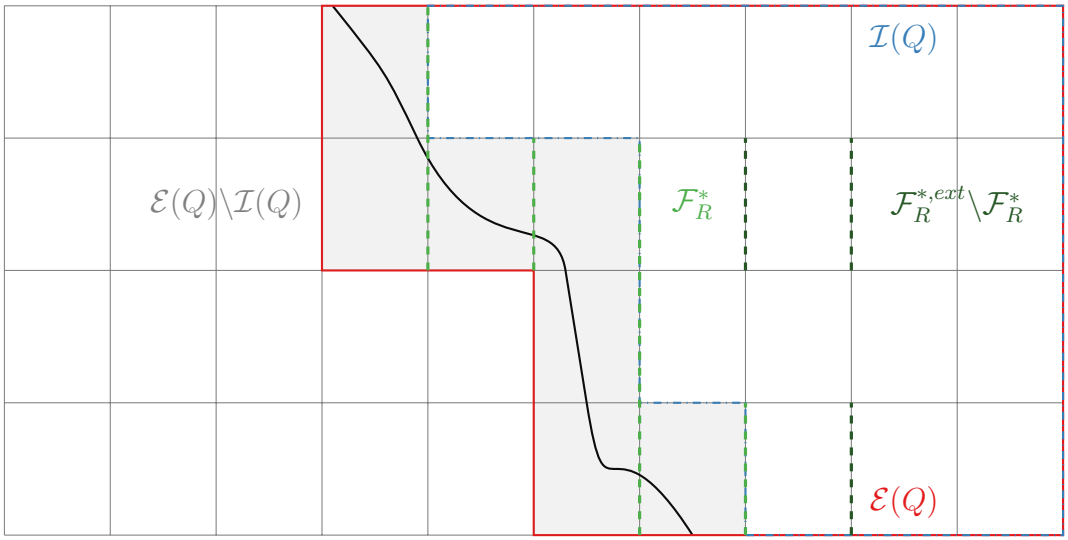


Figure 2: Discrete regions in the definition of the ghost-penalty of the plain DG method. By $\mathcal{F}_R^*$ we mean the union of all $\mathcal{F}_R^{*,n}$ and respectively for $\mathcal{F}_R^{*,ext}$.

Moreover, we collect those regions

$$\mathcal{E}(Q) = \bigcup_{n=1}^{N} \mathcal{E}(Q^n), \quad \mathcal{I}(Q) = \bigcup_{n=1}^{N} \mathcal{I}(Q^n).$$

As we are dealing with an unfitted discretisation, ill-posed cut configurations pose a challenge to the numerical method which is solved by a Ghost penalty stabilisation. See e.g. [1] for a general introduction of the Ghost penalty technique.

In our setting, the space of facets which will determine the classical ghost penalty is then given in two steps: First, we define $\mathcal{F}_R^{*,n}$:

$$\mathcal{F}_R^{*,n} = \{F \in \mathcal{F} \text{ such that } \exists T_1 \neq T_2, T_1 \in \mathcal{E}(\Omega^n)\backslash\mathcal{I}(\Omega^n), T_2 \in \mathcal{E}(\Omega^n) \text{ with } F = T_1 \cap T_2\}.$$

Second, this set of facets will be extended to a set $\mathcal{F}_R^{*,n,ext} \supseteq \mathcal{F}_R^{*,n}$ in line with the Assumption 3.1 stated explicitly below.

All those regions and facet sets are depicted in Fig. 2 for the reader's convenience.

2.2.2 Function spaces

After we have defined the relevant geometrical regions, we continue with introducing the relevant function spaces:

We consider a tensor product function space in space and time. For those components, we

define

$$W_{h,space}^{k_s} = \{v \in H^1(\tilde{\Omega}) \,|\, v|_T \in \mathcal{P}^{k_s}(T) \;\; \forall T \in \tilde{\mathcal{T}}_h\},$$
$$W_{h,time,n}^{k_t} = \mathcal{P}^{k_t}([t_{n-1}, t_n]).$$

For the spatial part, we just consider a standard continuous discrete function space. In time, we basically take standard polynomials on each time slice.

If we move forward in time, our interface might change. Hence, it seems reasonable to define the discrete function on each slice only relative to the necessary finite elements of our triangulation, which are given by the extension $\mathcal{E}(\Omega^n)$. Therefore, we define

$$W_{h,space,n}^{k_s} = \{v|_{\mathcal{E}(\Omega^n)} \,|\, v \in W_{h,space}^{k_s}\}.$$

Then on each of the time slices n, those spaces are taken together with the time spaces to yield

$$W_n^{k_s,k_t} = W_{h,space,n}^{k_s} \otimes W_{h,time,n}^{k_t} =: \{\sum_{i,j} c_{ij}\phi_i \cdot \psi_j \,|\, \phi_i \in W_{h,space,n}^{k_s}, \psi_j \in W_{h,time,n}^{k_t}\}.$$

Finally, those slices are collected for the different times, and we require the function to vanish outside of the extended domain $\mathcal{E}(Q)$:

$$W_h = \{v \in L^2(\mathcal{E}(Q)) \,|\, v|_{\mathcal{E}(Q^n)} \in W_n^{k_s,k_t} \, \forall n = 1, \dots, N\}.$$

2.2.3 The discrete problem

With the help of the discrete regions and function spaces introduced above, we are now able to formulate the variational formulation of the DG method under the assumption of an exact handling of geometries. We define

$$\begin{aligned} B(u,v) = &\sum_{n=1}^{N} (\partial_t u + \mathbf{w} \cdot \nabla u, v)_{Q^n} \\ &+ (\nabla u, \nabla v)_{Q^n} + \sum_{n=1}^{N-1} ([u]^n, v_+^n)_{\Omega(t_n)} + (u_+^0, v_+^0)_{\Omega(0)}, \end{aligned}$$

where $(\dots, \dots)$ is the L^2 scalar product, $[u]^n$ describes the jump of the discrete function u along the time-slice boundary at t_n. Accordingly, $u_\pm^n = \lim_{s \to 0, s>0} u(\dots, t_n \pm s)$.

The method is fully described by introducing a ghost penalty term in addition. To this end,

we define

$$j_h^n(u,v) = \int_{t_{n-1}}^{t_n} \tilde{\gamma}_J \sum_{F \in \mathcal{F}_R^{*,n,ext}} \int_{\omega_F} \frac{1}{h^2} [u]_{\omega_F} [v]_{\omega_F} \mathrm{d}x\mathrm{d}t$$

$$J(u,v) = \sum_{n=1}^{N} j_h^n(u,v)$$

In the sum over all $F \in \mathcal{F}_R^{*,n,ext}$, we introduced the following notation: For a facet $F \in \mathcal{F}$, the facet patch ω_F is defined as $\omega_F = T_1 \cup T_2$, where T_1 and T_2 are the elements adjacent to F. The facet patch jump $[\ldots]_{\omega_F}$ is defined as the difference between the function as defined on one of the elements T_i and the canonical polynomial extension of the function defined on the other element T_{3-i}.

Finally, we also introduce a right hand side:[2]

$$f(v) := \sum_{n=1}^{N} (f,v)_{Q^n} + (u_0, v_+^0)_{\Omega(0)}.$$

Then, the discrete problem reads: Find $u \in W_h$ such that

$$B(u,v) + J(u,v) = f(v) \quad \forall v \in W_h.$$

Remark. This is one of two forms to introduce the bilinear form. Because it will be important for our analysis later, we would like to remark here that it is possible to derive an mass conserving form of the bilinear form also and refer to [9, Lemma 2.1] for further details.

2.3 Space-Time CG

Next, we want to introduce a variant of the discontinuous Galerkin method presented in the previous subsection. Roughly, the idea is to enforce continuity along time slice boundaries by setting the value for the first discrete time point of the slice fix to the last value from the former slice. By means of that, the dimension of the problem matrix is reduced by the same amount as choosing one polynomial order less in time would lead to. This makes the CG variant interesting from a computational point of view, in particular for low and moderate polynomial orders in time.

[2]Note that we slightly abuse notation here by using f both as a symbol for the functional and the right-hand-side function in Eq. 2. However, the context should make clear whether a functional or a space-time function is meant.

2.3.1 Discrete regions

In detail, we start to introduce the method by defining discrete regions: Fix a time slice $I_n = (t_{n-1}, t_n]$. In order to define the discrete regions for the ghost penalty, we will only take into consideration the interface configuration, i.e. $\Omega(t_n)$ or $\Gamma(t_n) := \partial\Omega(t_n)$. We assume that there exists a (purely spatial) extension function $\mathcal{D}$ which satisfies at least the following assumptions:[3]

(1) For each n, $\Gamma(t) \subseteq \mathcal{D}(\Gamma(t_n))$ for all $t \in I_n$.

(2) $\mathcal{D}(\Gamma(t_n))$ contains all elements with cuts on the space-time slice and an α/ β strip in- or outside from the boundary:

$$\mathcal{D}(\Gamma(t_n)) \supseteq \{x \in T \text{ for some } T \in \mathcal{T}_h \text{ s.t. } (T \times I_n) \cap \Gamma_*^n \neq \varnothing \text{ or } \operatorname{dist}((T \times I_n), \Gamma_*^n) \leq \begin{cases} \alpha, \text{ if } (T \times I_n) \subseteq Q^n \\ \beta, \text{ if } (T \times I_n) \subseteq \tilde{Q}^n \backslash Q^n. \end{cases} \}$$

(3) For each n and $T \in \mathcal{T}_h$ such that $(T \times I_{n+1}) \cap Q^{n+1} \neq \varnothing$, $T \subseteq \mathcal{D}(\Gamma(t_n))$.

Now we define the extended domain in terms of this operator:

$$\mathcal{E}^{PG}(\Omega^n) := \{x \in T \text{ for some } T \in \mathcal{T}_h \text{ s.t. } (T \times I_n) \subseteq Q^n \text{ or } T \cap \mathcal{D}(\Gamma(t_n)) \neq \varnothing\},$$
$$\mathcal{E}^{PG}(Q^n) = \mathcal{E}^{PG}(\Omega^n) \times I_n.$$

In addition we define an operator for the outer and extended interface part:

$$\mathcal{O}^{PG}(\Omega^n) := \{x \in T \text{ for some } T \in \mathcal{T}_h \text{ s.t. } (T \times I_n) \subseteq \tilde{Q}^n \backslash Q^n \text{ or } T \cap \mathcal{D}(\Gamma(t_n)) \neq \varnothing\},$$
$$\mathcal{O}^{PG}(Q^n) = \mathcal{O}^{PG}(\Omega^n) \times I_n.$$

Those operators are now used to define the relevant facet set for the new ghost penalty:

$$\mathcal{F}_R^{*,n,PG} = \{F \in \mathcal{F} \text{ such that } \exists T_1 \neq T_2, T_1 \in \mathcal{O}^{PG}(\Omega^n) \cap \mathcal{E}^{PG}(\Omega^n), T_2 \in \mathcal{E}^{PG}(\Omega^n) \text{ with } F = T_1 \cap T_2\}.$$

[3] "at least" should mean in this context that we would like to describe some properties of the extension $\mathcal{D}$ which illustrate its basic functionality. In the course of a detailed analysis, further assumptions on the mapping and other aspects of the discretisation would be necessary (compare with Section 3 of this thesis). But we leave the task of analysing the CG method in detail for further investigations at this point.

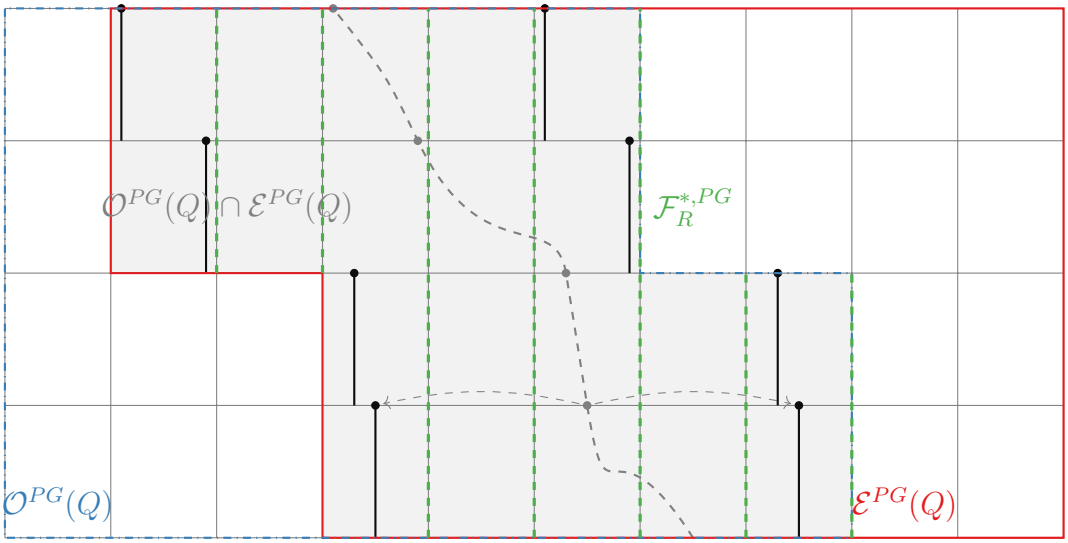


Figure 3: Discrete regions in the definition of the ghost-penalty of the CG method. Again, we denote the union over all $\mathcal{F}_R^{*,n,PG}$ by $\mathcal{F}_R^{*,PG}$.

As a last bit of notation, we introduce a symbol for the union over all extended/outer regions on each time slice:

$$\mathcal{E}^{PG}(Q) = \bigcup_{n=1}^{N} \mathcal{E}^{PG}(Q^n), \quad \mathcal{O}^{PG}(Q) = \bigcup_{n=1}^{N} \mathcal{O}^{PG}(Q^n).$$

Again, we give a sketch of this construction in Fig. 3.

2.3.2 Function spaces

After having defined the relavant geometrical regions, we continue with introducing the relevant function spaces:

For the discrete variational problem, we consider a tensor product function space in space and time. For those components, we define

$$W_{h,space}^{k_s} = \{v \in H^1(\tilde{\Omega}) \,|\, v|_T \in \mathcal{P}^{k_s}(T) \forall T \in \tilde{\mathcal{T}}_h\}$$
$$W_{h,time,n}^{k_t} = \mathcal{P}^{k_t}([t_{n-1}, t_n]).$$

For the spatial part, we just consider a standard continuous discrete function space. In time, we basically take standard polynomials on each time slice.

If we move forward in time, our interface might change. Hence, it seems reasonable to define the discrete function on each slice only relative to the necessary finite elements of our

triangulation, which are given by the extension $\mathcal{E}^{PG}(\Omega^n)$. Therefore, we define

$$W^{k_s}_{h,space,n} = \{v|_{\mathcal{E}^{PG}(\Omega^n)} \,|\, v \in W^{k_s}_{h,space}\}.$$

Then on each of the time slices n, those spaces are taken together with the time spaces to yield

$$W^{k_s,k_t}_n = W^{k_s}_{h,space,n} \otimes W^{k_t}_{h,time,n} =: \{\sum_{i,j} c_{ij}\phi_i \cdot \psi_j \,|\, \phi_i \in W^{k_s}_{h,space,n}, \psi_j \in W^{k_t}_{h,time,n}\}.$$

Finally, those slices are collected for the different times, and we require the function to vanish outside of the extended domain $\mathcal{E}^{PG}(Q)$:

$$W^{k_s,k_t}_h = \{v \in H^1(\tilde{\Omega} \times [0,T]) \,|\, v|_{\mathcal{E}^{PG}(Q^n)} \in W^{k_s,k_t}_n \,\forall n = 1,\dots,N \\ \text{and } v \text{ vanishes outside of } \mathcal{E}^{PG}(Q)\}.$$

Note that we imposed continuity along the time slice boundaries by the restriction posed on v. This function space will be the space of our ansatz functions, so we will write

$$U_h := W^{k_s,k_t}_h.$$

For the space of test function, the continuity condition is relaxed:

$$W^{k_s,k_t}_{h,dc} = \{v \in L^2(\tilde{\Omega} \times [0,T]) \,|\, v|_{\mathcal{E}^{PG}(Q^n)} \in W^{k_s,k_t}_n \,\forall n = 1,\dots,N \text{and } v|_{\tilde{\Omega}\times[0,T]\setminus\mathcal{E}^{PG}(Q)} = 0 \\ \text{and } v|_{\tilde{\Omega}\times I_n} \in H^1(\tilde{\Omega} \times I_n) \forall n\}.$$

Furthermore, we define the abbreviation

$$V_h := W^{k_s,k_t-1}_{h,dc}.$$

Note that the polynomial order in time is one less as usual with Petrov Galerkin methods.

2.3.3 Projection operators

We are now going to exploit these function space definitions in order to define projections Π and Π^*. We start by presenting these operators for polynomials in 1D and extend these definitions towards finite element functions in the paragraph afterwards. Note that the operator Π will be necessary to state the CG version of the method later. Strictly speaking, the

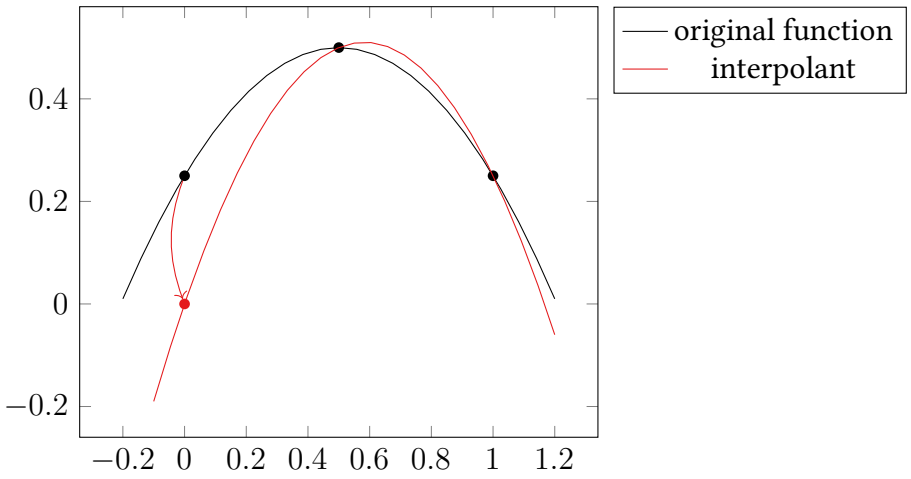


Figure 4: The projection of type $V_k \to V_{k,0}$.

operator Π^* will not be needed for the modest purposes of this thesis. We will present both operators anyhow, because they are easily introduced together and we believe that Π is of importance for the highly interesting task of analysing the CG method (c.f. also Section 5).

Projection operators in 1D In the following, we are going to introduce two types of interpolation for polynomials. Both will rely on a Lagrange basis for a concrete definition and neglect the information from the first integration point in one or the other sense. In particular, these two senses are

(1) Instead of the value of the function at the first integration point, fix this value to zero. $V_k \to V_{k,0}$

(2) Forget about the integration point and consider the polynomial space of one order less spanned by the remaining integration points. $V_k \to V_{k-1}$

For the case of a unit interval and a parabola, those two options are illustrated in Fig. 4 and 5.

The undivided case We are considering the space of 1D polynomials defined on the interval $[0, 1]$ of the order k:

$$V_k = \mathcal{P}^k([0, 1]), \quad \dim(V_k) = k + 1$$

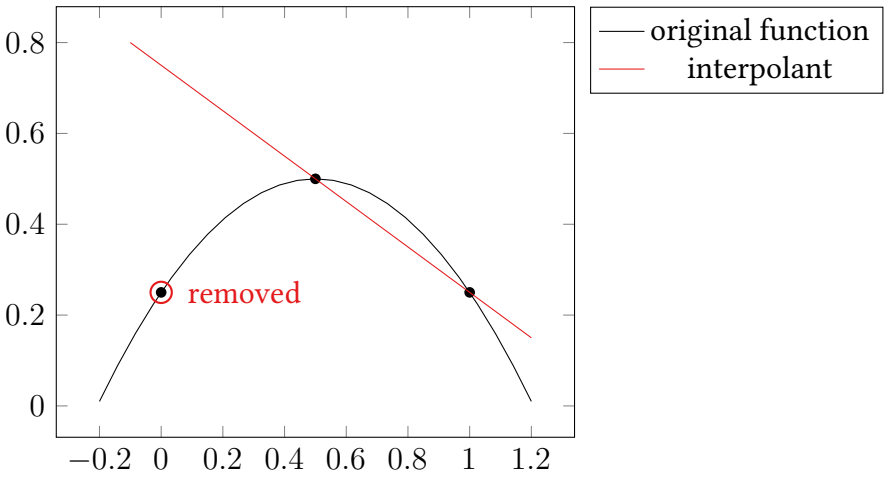


Figure 5: The projection of type $V_k \to V_{k-1}$.

and define the subspace

$$V_{k,0} = \{v \in V_k \mid v(0) = 0\}, \quad \dim(V_{k,0}) = k.$$

In order to arrive at a Langrage basis for this space of functions, we need to pick $k+1$ reference points. Let those be $0 = \hat{x}_0 < \hat{x}_1 < \hat{x}_2 < \ldots < \hat{x}_k = 1$ equidistant

$$\hat{x}_i = \frac{i}{k} \quad \text{for } i = 0, \ldots, k.$$

Then we can obtain, from the usual definition, the Lagrange basis functions $(l_i)_{i=0,\ldots,k}$. Hence, we can split any function in this basis,

$$v(x) = \sum_{i=0}^{k} l_i(x) v(\hat{x}_i) \quad \text{for } v \in V_k.$$

We define the projection $\Pi_{[0,1]} : V_k \to V_{k,0}$ of a function v as

$$\Pi_{[0,1]} v = \sum_{i=1}^{k} l_i(x) v(\hat{x}_i),$$

and in addition the projection $\Pi^*_{[0,1]} : V_k \to V_{k-1}$ of a function v as

$$\Pi^*_{[0,1]} v = \sum_{i=0}^{k-1} l^*_i(x) v(\hat{x}_{i+1}),$$

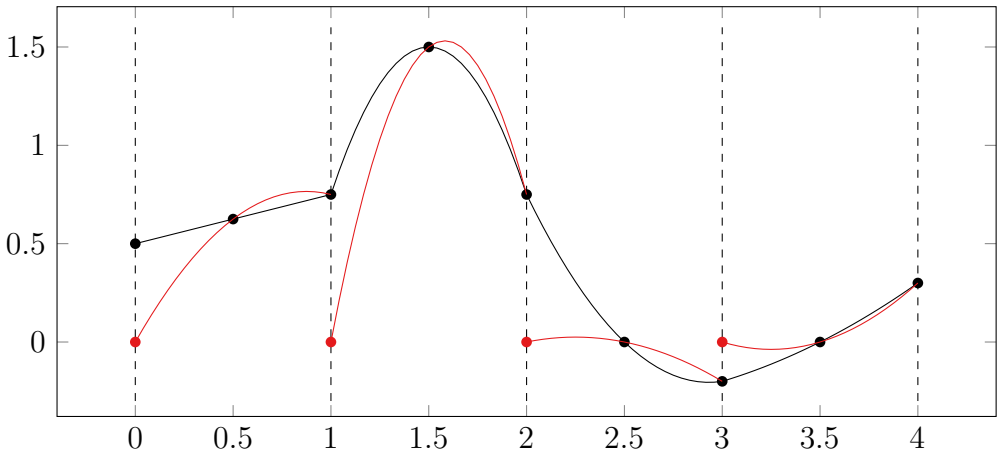


Figure 6: The projection of type $W_k \to W_{k,0}$.

where l_i^* should be the Lagrange interpolation functions stemming from the points $\hat{x}_1, \ldots, \hat{x}_k$. Note that—crucially—$\hat{x}_0$ is "missing".

The divided case Now we consider the divided case: The interval $[0, T]$ is subdivided into N intervals I_j, $I_0 = [0, \underbrace{\frac{T}{N}}_{=:t^1}]$, $I_1 = [\underbrace{\frac{T}{N}}_{=:t^1}, \underbrace{\frac{2T}{N}}_{=:t^2}]$, ... Let $\Phi_i : [0,1] \to I_i,\ \Phi_i(\hat{t}) = t^i + \hat{t} \cdot (t^{i+1} - t^i),\ i = 0, .., N-1$. We define the following space of piecewise polynomials:

$$W_k = \{v \in C^0([0,T]) \,|\, v|_{I_j} \in \mathcal{P}^k(I_j) \text{ for all } j\}$$

where $\mathcal{P}^k(I_j)$ is the space of polynomials up to degree k. We also define the discontinuous spaces

$$W_k^{\text{disc}} = \{v \in L^2([0,T]) \,|\, v|_{I_j} \in \mathcal{P}^k(I_j) \text{ for all } j\}$$
$$W_{k,0} = \{v \in L^2([0,T]) \,|\, v|_{I_j} \in \mathcal{P}_0^k(I_j) \text{ for all } j\}$$

with $\mathcal{P}_0^k(I_j) = \{v \in \mathcal{P}^k(I_j) | v(t^j) = 0\}$. Note that $W_{k,0} \not\subset W_k$. By application of the previously introduced projection on $[0,1]$ to the respective intervals one defines

$$\Pi : W_k \to W_{k,0}, \quad \Pi v|_{I_i} := \Pi_{[0,1]}(v \circ \Phi_i) \circ \Phi_i^{-1},\ i = 0, .., N-1,$$
$$\Pi^* : W_k \to W_{k-1}^{\text{disc}}, \quad \Pi^* v|_{I_i} := \Pi^*_{[0,1]}(v \circ \Phi_i) \circ \Phi_i^{-1},\ i = 0, .., N-1.$$

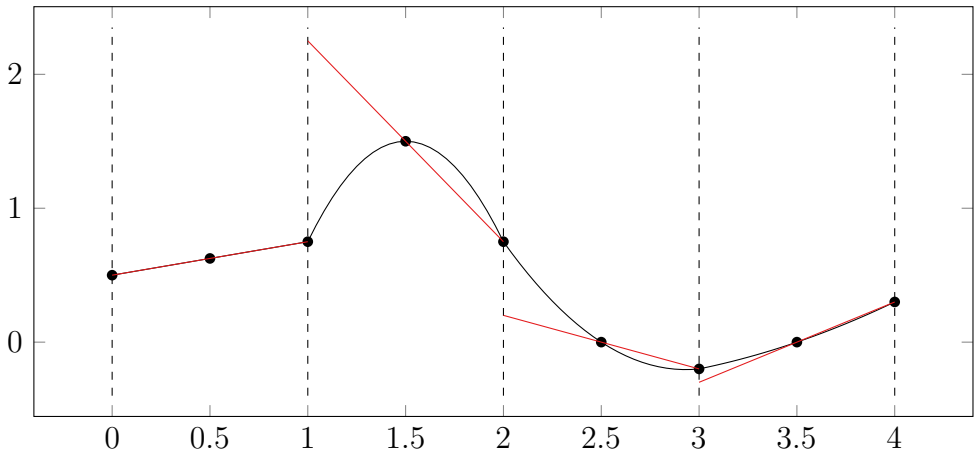


Figure 7: The projection of type $W_k \to W_{k-1}^{\text{disc}}$.

Projection operators for discrete functions In the next step, we will now extend the definition of the previous paragraph towards our discrete functions: The type of the projection Π^* is easily specified as $\Pi^* \colon U_h = W_h^{k_s,k_t} \to W_{h,dc}^{k_s,k_t-1} = V_h$. For Π, the according statement is given in terms of a modified version of the discontinuous space:

$$\begin{aligned} W_{h,dc,0}^{k_s,k_t} = \{&v \in L^2(\tilde{\Omega} \times [0,T]) \,|\, v|_{\mathcal{E}^{PG}(Q^n)} \in W_{n,0}^{k_s,k_t} \; \forall n = 1,\dots,N \\ &\text{and } v|_{\tilde{\Omega}\times[0,T]\setminus\mathcal{E}^{PG}(Q)} = 0 \\ &\text{and } v|_{\tilde{\Omega}\times I_n} \in H^1(\tilde{\Omega} \times I_n) \forall n\}, \\ W_{n,0}^{k_s,k_t} &= W_{h,space,n}^{k_s} \otimes \mathcal{P}_0^{k_t}([t_{n-1},t_n]), \\ \mathcal{P}_0^{k_t}([t_{n-1},t_n]) &= \{p \in \mathcal{P}^{k_t}([t_{n-1},t_n]) \,|\, p(t_{n-1}) = 0\}. \end{aligned}$$

Note that $\mathcal{P}_0^{k_t}$ has the same dimension as $\mathcal{P}^{k_t-1}$, and hence $V_h = W_{h,dc}^{k_s,k_t-1}$ and $W_{h,dc,0}^{k_s,k_t}$ also have the same dimension. Then for Π we have $\Pi \colon U_h = W_h^{k_s,k_t} \to W_{h,dc,0}^{k_s,k_t}$. Now let us fix some $u \in U_h$ and specify the image under Π, Πu. For each point $x \in \tilde{\Omega}$, we have

$$\Pi u(x,\dots) = \Pi(u(x,\dots)),$$

where the Π on the right hand side is the projection on 1D piecewise polynomials introduced above.

The same procedure of definition applies to Π^*, where we define

$$\Pi^* u(x,\dots) = \Pi^*(u(d,\dots)).$$

Remark. As an alternative to this pointwise definition, we could also fix an element T and project on each of its Lagrange points. Then, we would get the final function by picking the polynomial which has the projected values on the Lagrange points at each t.

2.3.4 The discrete problem

Now we have all tools in place to state the variational formulation of the CG method under the assumption of an exact handling of the geometries. It reads: Find $u \in U_h$ s.t.

$$B(u,v) + J(u,v) = f(v) \quad \forall v \in V_h,$$

where

$$B(u,v) = \sum_{n=1}^{N} (\partial_t u + w\nabla u, v)_{Q^n} + \sum_{n=1}^{N} (\nabla u, \nabla v)_{Q^n} + (u_+^0, v_+^0)_{\Omega^0},$$

$$f(v) := \sum_{n=1}^{N} (f,v)_{Q^n} + (u_0, v_+^0)_{\Omega(0)}$$

and

$$j_h^n(u,v) = \int_{t_{n-1}}^{t_n} \gamma \sum_{F \in \mathcal{F}_R^{*,n,PG}} \int_{\omega_F} \frac{1}{h^2} [\Pi u]_{\omega_F} [v]_{\omega_F}, \quad J(u,v) = \sum_{n=1}^{N} j_h^n(u,v).$$

In this equation, the set $\mathcal{F}_R^{*,n,PG}$ denotes the facet set for the ghost penalty introduced above.

The projection operator Π sets all contributions of the discrete function u from the exact discretisation points in time, t_i, to zero, as described above.

2.4 Geometry handling

In this subsection, we will explain two ways how to achieve higher order accuracy with respect to the geometry handling involved in the methods of the previous subsections. Both suggestions are generalisations of the isoparametric mapping of [4] and [8] in a space-time setting. Hence, we will start this subsection with a recapitulation of that base case. Afterwards, we will present the first way of generalising the construction in a space-time setting. This generalisation will resemble the method actually implemented in the upcoming numerical examples. Third, we will also introduce another space-time generalisation of the mapping which will be counterpart of the construction actually implemented for the purposes of our analysis.

2.4.1 The merely spatial base case

Let us start to illustrate the mechanism of the isoparametric mapping with a simple example: Imagine our time-independent domain should have the structure $\Omega = \{x \in \mathbb{R}^{+2} \,|\, \|x\|_2 \leq \frac{1}{2}\}$, which is embedded in the background domain $\tilde{\Omega} = [0,1]^2$. Then, Ω can be represented by the levelset function $\phi(x,y) = \sqrt{x^2+y^2} - \frac{1}{2}$. We now want to perform a numerical integration on the domain Ω. We assume our usual background mesh $\mathcal{T}_h$ to be given, as well as an according node-wise interpolation operator yielding the elementwise linear approximation ϕ^{lin}. Because the piecewise linear function ϕ^{lin} leads to straight cuts on each element, numerical integration can be performed there by a simple tesselation of the region. The domain stemming from the function ϕ^{lin} should be called Ω^{lin}.

Next, we want to improve the accuracy of this second-order approximation of Ω. To this end, a mesh deformation $\Theta_h \colon \tilde{\Omega} \to \mathbb{R}^d$ is applied. To each point in the domain $x \in \tilde{\Omega}$, we assign a slightly shifted point $\Theta_h(x)$, such that $\Theta_h(\Omega^{lin})$ is an approximation of Ω of high order accuracy. The explicit and detailed construction of this mesh deformation is involved and we refer the reader to [4] for a detailed account. Here, we only provide a short summary in terms of a recipe which is an adaption of the material of [9]:

(1) We start from an exact levelset function ϕ and derive two interpolations, $\phi^h = I_{k_s}\phi$ and $\phi^{lin} = I_1\phi^h$. Here, ϕ^h should be an elementwise polynomial of degree k_s, and ϕ^{lin} should be of first order.

(2) Exploiting the functions ϕ^{lin} and ϕ^h from the previous step, we search for a mapping Θ_h with the property

$$\phi^{lin} \approx \phi^h \circ \Theta_h.$$

This will lead to $\Theta_h(\Omega^{lin}) \approx \Omega^h$ as desired, where Ω^h denotes the region induced by ϕ^h. In detail, the mapping Θ_h is constructed along the following steps:

0. First, we need to fix some notation. We define $\Gamma^{lin} = \{x \in \hat{\Omega} \,|\, \phi^{lin}(x) = 0\}$, the approximation of the boundary induced by ϕ^{lin}. In terms of our triangulation $\mathcal{T}_h$ of $\tilde{\Omega}$, we collect certain elements in certain sets related to the interface:

$$\begin{aligned}
\mathcal{T}^\Gamma &= \{T \in \mathcal{T}_h \,|\, \overline{T} \cap \overline{\Gamma^{lin}} \neq \varnothing\} \\
\Omega^\Gamma &= \bigcup \mathcal{T}^\Gamma \\
\mathcal{T}^\Gamma_+ &= \{T \in \mathcal{T}_h \,|\, \overline{T} \cap \overline{\Omega^\Gamma} \neq \varnothing\} \\
\Omega^\Gamma_+ &= \bigcup \mathcal{T}^\Gamma_+
\end{aligned}$$

1. A candidate for the mapping is constructed element-wise:
   Fix some $T \in \mathcal{T}^\Gamma$. Our aim is to construct a preliminary version of the mapping Θ_h just on this element. To this end, we consider the polynomial extension of $(\phi^h)|_T$, and call it $\mathcal{E}^p\phi_h$. Next, we fix a search direction $G_h(x)$. Typical choices would be $G_h(x) = \nabla\phi^h$, or $G_h(x) = P_h^\Gamma(\nabla\phi_h)$, where P_h^Γ is a projection operator such as an Oswald-type projection (see [4] for more details). The idea is then that we start from a point $x \in T$ and search for the point y where $\mathcal{E}^p\phi_h(y)$ can serve as a proxy for $\phi^{lin}(x)$, i.e. their values conincide. Formally, this is defined as follows: For δ sufficiently small, define $d_h \colon T \to [-\delta, \delta]$ to be the function such that $d_h(x)$ is the smallest number such that

$$\mathcal{E}^p\phi_h(x + d_h(x)G_h(x)) = \phi^{lin}(x) \quad \text{for } x \in T.$$

   In practice, this search procedure can be done by a Newton search.

   In terms of this function d_h, we can now define our first candidate for the final mapping, $\Theta^\Gamma_{h,dc}(x)$:

$$\Theta^\Gamma_{h,dc}(x) = x + d_h(x)G_h(x)$$

   Here, the subindex dc should indicate that the function so far is discontinuous. Note that we deviate from the "standard" notation here, where this mapping would be called Ψ_h^Γ.

2. Translating the element-wise construction to a function on Ω^Γ:
   As noted before, our mapping so far is discontinuous along elements. This is undesirable and can be fixed by applying the projection operator P_h^Γ known from the previous step. In terms of it, we write

$$\Theta_h^\Gamma := P_h^\Gamma(\Theta^\Gamma_{h,dc}) = \mathrm{id} + P_h^\Gamma(d_h G_h).$$

   This transformation turns out to have the property that $\Theta_h^\Gamma(\Gamma^{lin}) = \Gamma^h$, where Γ^h should be the interface stemming from the high order approximation ϕ^h of the levelset function. This leads in turn to a high order approximation property of $\Theta_h^\Gamma(\Gamma^{lin})$ as we aimed for.

3. Extending the mapping on all of $\tilde{\Omega}$:
   In the last step, we start with the function Θ_h^Γ from the previous step and want to extend it to all of $\tilde{\Omega}$. Here, we use the strip $\Omega_+^\Gamma \backslash \Omega^\Gamma$ as a "transition region", i.e. the final mapping should be identical to Θ_h^Γ on all of Ω^Γ and the identity outside

of Ω_+^Γ,

$$\Theta_h(x) = \begin{cases} \Theta_h^\Gamma(x), & \text{if } x \in \Omega^\Gamma, \\ x + \mathcal{E}^{\partial\Omega^\Gamma}(\Theta_h^\Gamma - \mathrm{id})(x) & \text{if } x \in \Omega_+^\Gamma \backslash \Omega^\Gamma, \\ x & \text{else.} \end{cases}$$

Here, the extension operator $\mathcal{E}^{\partial\Omega^\Gamma}$ is a standard extension operator for a piecewise smooth function defined on $\partial\Omega^\Gamma$. Such a function defined on the boundary is extended to a piecewise smooth function defined on the volume, which vanishes on all of $\tilde{\Omega}\backslash\Omega_+^\Gamma$.

(3) After we found Θ_h, we use it to improve the geometry approximation of Ω^{lin}. Correspondingly, we use the new geometry approximation also in our weak formulation. The regions of integration are adapted accordingly, i.e. instead of integrating some function f on Ω^{lin}, we integrate on $\Theta_h(\Omega^{lin})$, leading to

$$\int_{\Theta_h(\Omega^{lin})} f \mathrm{d}x = \int_{\Omega^{lin}} (f \circ \Theta_h) |\det(D\Theta_h)|$$

The integrals of the later kind are again easy to compute in practice.

(4) Last, the function space needs to take care of the modified domain of definition. If we e.g. started with a discrete function u on Ω^{lin} we now want to introduce the discrete space on the curved elements as the result of first going back to Ω^{lin} and then evaluating u. So we define

$$\mathcal{V}_h := W_h \circ \Theta_h^{-1} := \{v \circ \Theta_h^{-1} \,|\, v \in W_h\},$$

where W_h is the respective discrete function space of the merely spatial problem on Ω^{lin}.

2.4.2 Geometry handling of the numerical space-time DG method

In this subsection we will extend the construction of the isoparametric mapping for merely spatial problems to the setting of space-time problem. We will first develop a description which resembles the implementation accurately.

space-time level sets As the isoparametric method for stationary problems starts with a merely spatial levelset function, in our setting we first assume to be given a space-time levelset function, $\phi = \phi(x,t)$ for $x \in \tilde{\Omega}$ and $t \in [0,T]$. We follow the discretisation in time

and hence construct the mapping for a time interval $I_n = [t_{n-1}, t_n]$ now, or in other words for the slice $\tilde{Q}^n = \tilde{\Omega} \times I_n$.

(1) The discretisation in time depends on a parameter k_t resembling the polynomial order. Our goal is it to approximate the exact levelset ϕ by discrete functions of different polynomial degree in space and degree k_t in time. For matters of illustration, we want to present the polynomials in time in terms of a Lagrange basis. To this end, we assume that the time slab I_n is in turn subdivided as follows:

$$t_{n-1} = t_0^{n,*}, t_1^{n,*}, \ldots, t_{k_t}^{n,*} = t_n, \quad t_{j+1}^{n,*} - t_j^{n,*} = \text{const.}$$

Then a basis for the polynomials $\mathcal{P}^{k_t}$ of order k_t on I_n will be given as the Lagrange polynomials l_i defined relative to the evaluation points $(t_j^{n,*})_j$. So we assume that we are given a high order interpolation of ϕ in the space $V_h^{k_s} \otimes \mathcal{P}^{k_t}$ for some spatial discrete space $V_h^{k_s}$ of order k_s. This can be written as

$$\phi^h(x,t) = \sum_{i=0}^{k_t} l_i(t) \cdot \phi_i^h(x) \quad \phi_i^h \in V_h^{k_s}.$$

One way to obtain such an approximation consists of projecting the levelset function onto the intermediate time points $(t_j^{n,*})_j$ and applying the spatial interpolation known from the merely spatial construction then. However, for the purpose of the analysis it is not relevant how the high-order interpolation is obtained in detail.

In addition to this high-order approximation of the domain in space, we introduce a low-order variant by considering $\phi_i^{lin} := I_1 \phi_i^h$ for each i. Then we define

$$\phi^{lin}(x,t) = \sum_{i=0}^{k_t} l_i(t) \cdot \phi_i^{lin}(x).$$

The spatially polygonal approximation of $\Omega(t)$ stemming from this function will serve the job of the base case of which the mapping is applied, as before.

space-time deformation In the second step, we want to derive a deformation function in space and time. The construction will in the end mostly be an application of the merely spatial construction at all intermediate time points. One slight deviation is the definition of new discrete regions for the whole time slice:

(2) We define

$$\begin{aligned}
\mathcal{T}^{\Gamma_n} &= \{T \in \mathcal{T}_h \,|\, \overline{T} \cap \overline{\Gamma^{lin}(t)} \neq \varnothing \quad \text{for some } t \in I_n\} \\
\Omega^{\Gamma_n} &= \bigcup \mathcal{T}^{\Gamma_n} \\
\mathcal{T}_+^{\Gamma_n} &= \{T \in \mathcal{T}_h \,|\, \overline{T} \cap \overline{\Omega^{\Gamma,n}} \neq \varnothing\} \\
\Omega_+^{\Gamma_n} &= \bigcup \mathcal{T}_+^{\Gamma_n}
\end{aligned}$$

Fix some $t_j^{n,*}$. Then let $\Theta_{h,j}$ be the isoparametric mapping derived from the merely spatial functions ϕ_j^h and ϕ_h^{lin} in terms of the "new" regions $\mathcal{T}^{\Gamma_n}$ etc. as defined above. We apply this procedure for all intermediate time points and collect the results as follows:

$$\Theta_h(x,t) = \sum_{i=0}^{k_t} l_i(t) \cdot \Theta_{h,i}(x)$$

This function then has the property that $\Theta_h(\ldots,t)(\Omega^{lin}(t)) = \Omega^h(t)$, so that it approximates $\Omega(t)$, for each $t \in I_n$.

2.4.3 Geometry handling for the analysis under a weakened assumption of an exact handling

Lastly, we want to describe a variant of the construction of the previous paragraphs which will be used in the analysis. Roughly, we will assume there that we are able to perform the integration in time exactly, but the construction in space remains as it is.

We fix a time slice I_n and some time point $t \in I_n$. Then, our goal is to construct the mapping $\Theta_h(x,t)$ for that time. To this end, we proceed as follows:

(1) We fix the time t in the levelset function and derive ϕ^h and ϕ^{lin} in the merely spatial manner: $\phi^h = I_{k_s}\phi(\ldots,t)$, $\phi^{lin} = I_1\phi^h$.

(2) We generate the mapping Θ^h at t by performing the construction of the merely spatial case in terms of the regions $\mathcal{T}^{\Gamma_n}$ etc. as defined above.

Next, we want to extend this function towards a full spacetime mapping, where the time coordinate remains fixed. To this end, we define

Definition 2.1. The space-time version of the isoparametric mapping Θ_h^{st} is defined as

$$\Theta_h^{st}(x,t) := (\Theta_h(x,t),t)^T,$$

where $\Theta_h(x,t)$ refers to the above construction for a time point t.

The next lemma collects some properties of the function so defined. For each time $t \in [0,T]$, the properties follow from the according properties of the merely spatial mapping, as the extension of the region of definition does not affect them:

Lemma 2.2. *For each time $t \in [0,T]$, $t \in I_n$, the following holds with $\Omega_+^{\Gamma_n}$ denoting the extended domain of cut elements as defined above*

$$\Theta_h^{st}(x,t) = x \quad \text{for all } x \text{ vertices of } \mathcal{T}_h \text{ or } x \in \tilde{\Omega} \backslash \Omega_+^{\Gamma_n},$$
$$\|\Theta_h^{st}(x,t) - x\|_{\infty,\tilde{\Omega}} \lesssim h^2 \quad \|D\Theta_h^{st}(x,t) - I\|_{\infty,\tilde{\Omega}} \lesssim h$$

In terms of Θ_h^{st}, we can define our discrete approximation of the space-time domain

Definition 2.3. We define the approximated space-time domain Q^h and the slices thereof as

$$Q^h := \Theta_h^{st}(Q^{lin}), \quad Q^{h,n} := \Theta_h^{st}(Q^{lin,n}).$$

2.4.4 Mapping to the exact geometry and properties of Θ

The mapping Θ_h^{st} allows us to map the linear approximation of the space-time region Q^{lin} to a high-order approximation $Q^h = \Theta_h^{st}$. In order to describe this high-order approximation property exactly (and for other purposes), we assume in addition the existence of a mapping Ψ^{st} between Q^{lin} and the exact domain Q:

Assumption 2.4. *There exists a smooth bijective function $\Psi^{st} \colon Q^{lin} \to Q$ such that*

$$\|u\|_{L^2(Q)} \simeq \|u \circ \Psi^{st}\|_{L^2(Q^{lin})} \quad \text{for all } u \in H^1(Q).$$

In terms of this mapping, we can extend Lemma 2.2 as follows:

Lemma 2.5. *For each time $t \in [0,T]$, $t \in I_n$*

$$\|\Theta_h^{st}(\ldots,t) - \Psi^{st}(\ldots,t)\|_{\infty,\tilde{\Omega}} + h\|D(\Theta_h^{st}(\ldots,t) - \Psi^{st}(\ldots,t))\|_{\infty,\tilde{\Omega}} \lesssim h^{k_s+1}.$$

Proof. Transfer the spatial proof for each time. □

Assumption 2.6. *In addition, we assume for each $t \in [0,T]$, $t \in I_n$*

$$\|\frac{\partial}{\partial t}(\Theta_h^{st} - \Psi^{st})(\ldots,t)\|_{\infty,\tilde{\Omega}} \lesssim h^{k_s+1}.$$

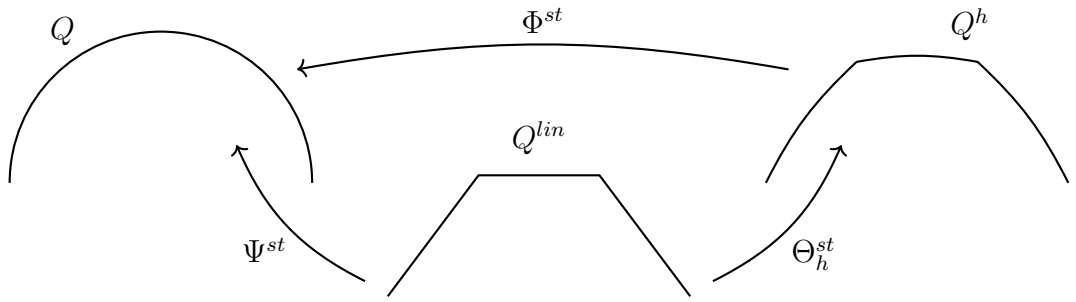


Figure 8: Schematic illustration of the mappings Ψ^{st}, Θ_h^{st}, and Φ^{st}. Note that the situation is depicted just for one time, while the mappings are actually full space-time mappings.

Remark. If we would consider the construction of 2.7.2 instead of that of 2.7.3, we would expect

$$\begin{aligned}&\|\Theta_h^{st}(\dots) - \Psi^{st}(\dots)\|_{\infty,\tilde{\Omega}\times[0,T]} + h\|D_x(\Theta_h^{st} - \Psi^{st})\|_{\infty,\tilde{\Omega}\times[0,T]} \\ &+ \Delta t\|\frac{\partial}{\partial t}(\Theta_h^{st} - \Psi^{st})\|_{\infty,\tilde{\Omega}\times[0,T]} \lesssim h^{k_s+1} + \Delta t^{k_t+1}.\end{aligned}$$

Introducing Φ In the previous paragraphs, we introduced a mapping $\Theta_h^{st}\colon Q^{lin} \to Q^h$ and a mapping $\Psi^{st}\colon Q^{lin} \to Q$. Naturally, one might also want to switch back and forth between Q^h and Q directly. For this purpose, we want to introduce a mapping Φ^{st} as follows

Definition 2.7. The mapping $\Phi^{st}\colon Q^h \to Q$ is defined as

$$\Phi^{st} := \Psi^{st} \circ (\Theta_h^{st})^{-1}.$$

Corollary 2.8. *It holds*

$$\Psi^{st} = \Phi^{st} \circ \Theta_h^{st}.$$

For the reader's convenience, we give an overview over all three mappings in Fig. 8.

2.5 Space-Time DG under a weakened assumption on the handling of geometries

In the previous subsection, we defined a high order accurate approximation of the geometry Q^n of each time-slab, $Q^{h,n} = \Theta_h^{st}(Q^{lin,n})$. In terms of this new region, we now want to define the discrete method of high order.

A first step towards this goal consists of introducing the relevant function spaces. The introduction of the function spaces for the high order method follows the known recipe:

The function space is first defined for the linear approximation, and those discrete functions are then lifted to the curved geometry. So we start the definition of the function spaces by treating the piecewise linear case.

From the piecewise linear in space function ϕ^{lin}, we can obtain the space-time regions

$$\begin{aligned}\Omega^{lin}(t) &= \{x \in \tilde{\Omega} \,|\, \phi^{lin}(x,t) = 0\}\\ Q^{lin,n} &= \bigcup_{t\in I_n} \Omega^{lin}(t) \times \{t\}, \quad Q^{lin} = \bigcup_{n=1}^{N} Q^{lin,n}.\end{aligned}$$

This change in the computational domains also affects the extended and interior space-time regions. Hence, we define

$$\begin{aligned}\mathcal{E}(\Omega^{lin,n}) :&= \{x \in T \text{ for some } T \in \mathcal{T}_h \text{ such that } (T \times I_n) \cap Q^{lin,n} \neq \varnothing\},\\ \mathcal{E}(Q^{lin,n}) &= \mathcal{E}(\Omega^{lin,n}) \times I_n.\end{aligned}$$

For the interior domain, we introduce $\mathcal{I}(\Omega^{lin,n})$ for the interior region with the obvious according changes. The respective facet spaces will be denoted by $\mathcal{F}_R^{*,lin,n}, \mathcal{F}_R^{*,lin,n,ext}$, c.f. Assumption 3.1.

Next, we basically copy the definition of the respective function spaces from the construction of the method involving the assumption of an exact handling of geometries.

$$\begin{aligned}W_{h,space,lin,n}^{k_s} &= \{v|_{\mathcal{E}(\Omega^{lin,n})} \,|\, v \in W_{h,space}^{k_s}\},\\ W_{n,lin}^{k_s,k_t} &= W_{h,space,lin,n}^{k_s} \otimes W_{h,time,n}^{k_t},\\ W_{h,lin} &= \{v \in L^2(\mathcal{E}(Q^{lin})) \,|\, v|_{\mathcal{E}(Q^{lin,n})} \in W_{n,lin}^{k_s,k_t} \,\forall n = 1,\dots,N\}.\end{aligned}$$

This allows us define the final discrete space:

$$W_h = W_{h,lin} \circ (\Theta_h^{st})^{-1}.$$

Fix a function $u \in W_h$. Then, by definition of W_h, the function can be "factorised" over Θ_h^{st}: While u is of type $u\colon \Theta_h^{st}(Q^{lin}) = Q^h \to \mathbb{R}$, we can find a corresponding function $\hat{u}\colon Q^{lin} \to \mathbb{R}$ such that

$$u = \hat{u} \circ (\Theta_h^{st})^{-1}, \quad \text{or} \quad \hat{u} = u \circ \Theta_h^{st}.$$

This can also be written in coordinates as

$$\hat{u}(\hat{x},t) = u(\Theta_h(\hat{x},t),t) \quad \text{for } \hat{x} \in Q^{lin}.$$

That factorisation leads to an equivalence of norms:

Lemma 2.9. *For a discrete function $u \in W_h$, $u = \hat{u} \circ (\Theta_h^{st})^{-1}$, and each set $S \subseteq \tilde{\Omega}$ and time t there holds*

$$\|\hat{u}(t,\dots)\|_S \simeq \|u(t,\dots)\|_{\Theta_h^{st}(t,S)}.$$

We can exploit the representation of $\hat{u}$ as $\hat{u} = u \circ \Theta_h^{st}$ in order to calculate a formula for the partial derviative in time direction. In general, the multi-dimensional chain rule applied to this case yields

$$D(\hat{u}) = D(u \circ \Theta_h^{st})(x,t) = D(u)(\Theta_h^{st}(x,t)) \cdot D(\Theta_h^{st})(x,t).$$

Denoting the components of Θ_h^{st} as $\Theta_h^{st} = (\Theta_{h,x1}^{st}, \dots, \Theta_{h,xn}^{st}, \Theta_{h,t}^{st})^T$, we observe

$$D(\Theta_h^{st}) = \begin{pmatrix} \frac{\partial \Theta_{h,x1}^{st}}{\partial x_1} & \cdots & \frac{\partial \Theta_{h,x1}^{st}}{\partial x_n} & \frac{\partial \Theta_{h,x1}^{st}}{\partial t} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial \Theta_{h,xn}^{st}}{\partial x_1} & \cdots & \frac{\partial \Theta_{h,xn}^{st}}{\partial x_n} & \frac{\partial \Theta_{h,xn}^{st}}{\partial t} \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

The Jacobian $D(\hat{u})$ will be a row vector. For $\frac{\partial \hat{u}}{\partial t}$, we are interested in the last entry. In line with the general rule above, we obtain

$$\begin{aligned} \frac{\partial \hat{u}}{\partial t} &= (\frac{\partial u}{\partial x_1} \circ \Theta_h^{st}) \cdot \frac{\partial \Theta_{h,x_1}^{st}}{\partial t} + \cdots + (\frac{\partial u}{\partial x_n} \circ \Theta_h^{st}) \cdot \frac{\partial \Theta_{h,x_n}^{st}}{\partial t} + \frac{\partial u}{\partial t} \circ \Theta_h^{st} \\ &= (\nabla u \circ \Theta_h^{st}) \cdot \frac{\partial \Theta_h^{st}}{\partial t} + \frac{\partial u}{\partial t} \circ \Theta_h^{st} \end{aligned}$$

In the numerical method, the time derivative of $\frac{\partial \hat{u}}{\partial t}$ is assembled, while, physically speaking, we want to have $\frac{\partial u}{\partial t}$ in our bilinear form. Hence, we note for later purposes that

$$\frac{\partial u}{\partial t} \circ \Theta_h^{st} = \frac{\partial \hat{u}}{\partial t} - (\nabla u \circ \Theta_h^{st}) \cdot \frac{\partial \Theta_h^{st}}{\partial t}.$$

For the purposes of the analysis, we need a further directional derviative of a function u. To this end, we define:

Definition 2.10 (Derivative in mesh direction)**.** The derivative ∂_t^{Θ} is defined as follows

$$\partial_t^{\Theta_h^{st}} u = \partial_t u + \nabla u \cdot (\frac{\partial \Theta_h^{st}}{\partial t} \circ (\Theta_h^{st})^{-1}).$$

Often, we write ∂_t^{Θ} as a shorthand notation for $\partial_t^{\Theta_h^{st}}$.

From the definition of ∂_t^{Θ} and the calculations above, we obtain

Corollary 2.11. *In terms of the notation introduced above,*

$$\partial_t^{\Theta} u = (\partial_t \hat{u}) \circ (\Theta_h^{st})^{-1}.$$

This implies for some $S \subseteq \tilde{\Omega}$ and a time t

$$\|\partial_t^{\Theta} u\|_{\Theta_h^{st}(t,S)} \simeq \|\partial_t \hat{u}\|_S$$

Similar to the case of the time derivative, we can exploit the general version of the chain rule to derive a relation between the gradients of u and $\hat{u}$. For instance, for the first component of the gradient, we calculate

$$\frac{\partial \hat{u}}{\partial x_1} = (\frac{\partial u}{\partial x_1} \circ \Theta_h^{st}) \frac{\partial \Theta_{h,x1}^{st}}{\partial x_1} + \cdots + (\frac{\partial u}{\partial x_n} \circ \Theta_h^{st}) \frac{\partial \Theta_{h,xn}^{st}}{\partial x_1}.$$

With the notation

$$D_x(\Theta_h^{st}) = \begin{pmatrix} \frac{\partial \Theta_{h,x1}^{st}}{\partial x_1} & \cdots & \frac{\partial \Theta_{h,x1}^{st}}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Theta_{h,xn}^{st}}{\partial x_1} & \cdots & \frac{\partial \Theta_{h,xn}^{st}}{\partial x_n} \end{pmatrix},$$

this can be generalised as follows:

$$\nabla \hat{u} = D_x(\Theta_h^{st})^T (\nabla u \circ \Theta_h^{st})$$

With the help of the bound on Θ_h^{st} in Lemma 2.2, we can conclude

Corollary 2.12. *For a discrete function $u \in W_h$, $u = \hat{u} \circ (\Theta_h^{st})^{-1}$, and each set $S \subseteq \tilde{\Omega}$ there holds*

$$\|\nabla \hat{u}\|_S \simeq \|\nabla u\|_{\Theta_h^{st}(S)}.$$

With the help of the discrete regions and function spaces introduced above, we are now able to formulate the variational formulation of the DG method without assumptions of an

exact handling of geometries. We define

$$B(u,v) = \sum_{n=1}^{N} (\partial_t u + \mathbf{w}\cdot\nabla u, v)_{Q^{h,n}} + (\nabla u, \nabla v)_{Q^{h,n}} + \sum_{n=1}^{N-1} ([u]^n, v_+^n)_{\Omega^h(t_n)} + (u_+^0, v_+^0)_{\Omega^h(0)}.$$

Remark. At this point, we would like to remark that there is a difference between the bilinear form introduced here and analysed below and the actually implemented version concerning the $([u]^n, v_+^n)_{\Omega^h(t_n)}$ term. More specifically, the calculation of the jump $[u]^n$ poses a challenge as the mapping determining the discrete regions might be discontinuous in time. A full account of this would probably involve the definition of a suiting projection operator (c.f. [9, p. 54]). We leave this as a task for future research as far as this thesis is concerned.

The method is fully described by introducing a ghost penalty term in addition. To this end, we define in a first step the transformed facet patch region for some facet F from the uncurved mesh.

$$\omega_F^h(t) = \Theta_h^{st}(t)(\omega_F)$$

Note that the time dependence might actually cause the mapped domain to change within a time slab, such that the tensor-product structure of the space-time domain of integration known from Preuß does not hold any more.

In addition, we need a jump operation on these curved elements. Let us assume $u = \hat{u}\circ(\Theta_h^{st})^{-1}$, $\omega_F = T_1 \cup T_2$ and we interested on the jump for some point x in the mapped T_1:

$$[u]_{\omega_F^h(t)}|_{\Theta_h^{st}(t)(T_1)}(x) = u(x) - \mathcal{E}^p(\hat{u}|_{T_2})((\mathcal{E}^p\Theta_h^{st}(t)|_{T_2})^{-1}(x))$$

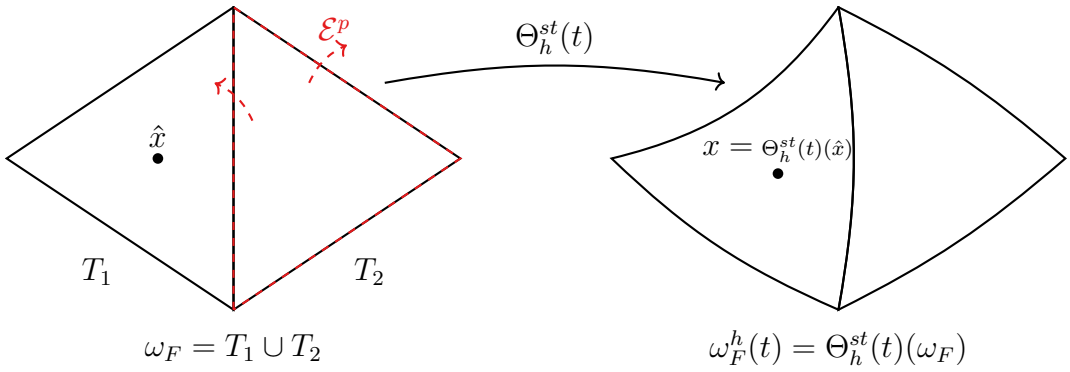
This situation is depicted in Fig. 9.

With that notation, we are in a position to introduce the Ghost penalty terms:

$$j_F^n(u,v) := \int_{t_{n-1}}^{t_n}\int_{\omega_F^h(t)} \frac{\tilde{\gamma}_J}{h^2}[u]_{\omega_F^h(t)}[v]_{\omega_F^h(t)}\mathrm{d}x\mathrm{d}t$$

$$j_h^n(u,v) = \sum_{F\in\mathcal{F}_R^{*,lin,n,ext}} j_F^n(u,v), \quad J(u,v) = \sum_{n=1}^{N} j_h^n(u,v).$$

Here, $\tilde{\gamma}_J = \left(1 + \frac{\Delta t}{h}\right)\gamma_J$.



$$[u]_{\omega_F^h}(x) = \underbrace{u(x)}_{=\hat{u}(\hat{x})} - \mathcal{E}^p(\hat{u}|_{T_2})((\mathcal{E}^p\Theta_h^{st}(t)|_{T_2})^{-1}(x))$$

Figure 9: Details of the definition of the jump operator on the curved elements.

In accordance to the previous definitions, the right-hand side f is given as

$$f(v) := \sum_{n=1}^{N} (f, v)_{Q^{h,n}} + (u_0, v_+^0)_{\Omega^h(0)}.$$

In total, the discrete problem reads: Find $u \in W_h$ such that

$$B(u, v) + J(u, v) = f(v) \quad \forall v \in W_h. \tag{3}$$

In addition, we define a variant of this bilinear form, which deviates by a term stemming from the numerical handling of the geometry in space.

$$\begin{aligned} B_{mc}(u, v) = &\sum_{n=1}^{N} (u, -\partial_t v - \mathbf{w} \cdot \nabla v)_{Q^{h,n}} \\ &+ (\nabla u, \nabla v)_{Q^{h,n}} - \sum_{n=1}^{N-1} (u_-^n, [v]^n)_{\Omega^h(t_n)} + (u_+^N, v_+^N)_{\Omega^h(T)}. \end{aligned}$$

The difference between those two formulations can be controlled by the following Lemma.

Lemma 2.13. *For all u, v the following holds*

$$B(u, v) - B_{mc}(u, v) = \int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x \, (\mathbf{w} \cdot \mathbf{n} - \mathcal{V}_n^h(t)) uv,$$

where $\mathcal{V}_n^h(t)$ denotes the velocity of the approximated boundary $\Omega^h(t)$ in normal direction.

Proof. For the proof, we just use partial integration as Preuß did in the proof of [9, Lemma 2.1]. □

With the help of this lemma, we can make an assumption on the maximum norm of the difference in the integrand:

Assumption 2.14. *We assume*

$$|\mathbf{w}\cdot\mathbf{n}-\mathcal{V}_n^h(t)|_{L^\infty(0,T,L^\infty(\partial\Omega^h(t)))}=\sup_{t\in[0,T]}\sup_{x\in\partial\Omega^h(t)}|\mathbf{w}\cdot\mathbf{n}-\mathcal{V}_n^h(t)|\leq C_{mc}(h,\Delta t),$$

where C_{mc} has the following property: For all $C>0$, there exists $\tilde{h}$ and $\tilde{\Delta}t$ such that for all $h<\tilde{h}$ and $\Delta t<\tilde{\Delta}t$,

$$C\cdot\frac{C_{mc}(h,\Delta t)}{\Delta t}\leq\frac{1}{4}.$$

The difference between B and B_{mc} reflects as following the Lemma about the positivity of the bilinear form summands:

Lemma 2.15 (Counterpart of P3.6)**.** *For all $u\in W_h$, and denoting $d(u,v)=\sum_{n=1}^N(\partial_t u+\mathbf{w}\cdot\nabla u,v)_{Q^{h,n}}$ and $b(u,v)=\sum_{n=1}^{N-1}([u]^n,v_+^n)_{\Omega^h(t_n)}+(u_+^0,v_+^0)_{\Omega^h(t_0)}$ we have*

(1) Testing d with u symmetrically yields

$$\begin{aligned}d(u,u)=&\frac{1}{2}\sum_{n=1}^{N-1}\left(\|u_-^n\|^2_{\Omega^h(t_n)}-\|u_+^n\|_{\Omega^h(t_n)}\right)+\frac{1}{2}\|u_-^N\|^2_{\Omega^h(T)}-\frac{1}{2}\|u_+^0\|^2_{\Omega^h(0)}\\&+\frac{1}{2}\int_0^T\mathrm{d}t\int_{\partial\Omega^h(t)}\mathrm{d}x\,(\mathcal{V}_n^h(t)-\mathbf{w}\cdot\mathbf{n})u^2.\end{aligned}$$

(2) In accordance with Preuß, we note

$$b(u,u)=\frac{1}{2}\sum_{n=1}^{N-1}\|[u]^n\|^2_{\Omega^h(t_n)}+\frac{1}{2}\sum_{n=1}^{N-1}\left(\|u_+^n\|^2_{\Omega^h(t_n)}-\|u_-^n\|^2_{\Omega^h(t_n)}\right)+\|u_+^0\|^2_{\Omega^h(0)}.$$

(3) Adding the two contributions up, we obtain

$$B(u,u)=\frac{1}{2}|||[u]|||^2+\sum_{n=1}^{N}(\nabla u,\nabla u)_{Q^{h,n}}+\frac{1}{2}\int_0^T\mathrm{d}t\int_{\partial\Omega^h(t)}\mathrm{d}x\,(\mathcal{V}_n^h(t)-\mathbf{w}\cdot\mathbf{n})u^2.$$

Proof. (1) For the first part, note that the generalisation of the formula relating $d(u,v)$ and

$d'(u,v)$ goes as follows:

$$d(u,v) = \sum_{n=1}^{N} \left\{ (u_-^n, v_+^n)_{\Omega^h(t_n)} - (u_+^{n-1}, v_+^{n-1})_{\Omega^h(t_{n-1})} \right\} + d'(u,v)$$
$$+ \int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x \; (\mathcal{V}_n^h(t) - \mathbf{w}\cdot\mathbf{n}) uv$$

Following Preuß, we consider $v = u$ and note that $d(u,u) = -d'(u,u)$ by definition in order to obtain the result.

□

Remark. This finishes the presentation of the DG method under a weakened assumption on the handling of geometries. We would like to note that it should be possible and an interesting direction for future research to extend the CG method in the same direction.

3 Analysis of Unfitted Space-Time DG method (including geometry errors)

One main ingredient for the numerical analysis of the suggested method are the relevant norms. First, we will prove stability in terms of the norm $|||\dots|||_j$, which is defined as the sum of a norm $|||\dots|||$ and the ghost penalty:

$$|||u|||^2 := \sum_{n=1}^{N} \Delta t (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} + |||[u]|||^2 + \sum_{n=1}^{N} (\nabla u, \nabla u)_{Q^{h,n}}$$
$$|||[u]|||^2 := \sum_{n=1}^{N-1} ([u]^n, [u]^n)_{\Omega^h(t_n)} + (u_+^0, u_+^0)_{\Omega^h(0)} + (u_-^N, u_-^N)_{\Omega^h(T)}$$
$$|||u|||_j^2 := |||u|||^2 + J(u,u)$$

For the continuity proof, another norm will be relevant which contains an integral over u instead of $\partial_t^\Theta u$ (with a different scaling in Δt) and a different expression for the jump:

$$|||u|||_*^2 := \sum_{n=1}^{N} (\frac{1}{\Delta t} u, u)_{Q^{h,n}} + |||[u]|||_*^2 + \sum_{n=1}^{N} (\nabla u, \nabla u)_{Q^{h,n}}$$
$$|||[u]|||_*^2 := \sum_{n=1}^{N} (u_-^n, u_-^n)_{\Omega^h(t_n)}$$
$$|||u|||_{j,*}^2 := |||u|||_*^2 + J(u,u)$$

The analysis of our method is based on a list of assumptions. Those are variants of the assumptions given by Preuß. The first assumption relates to a technical aspect of the Ghost penalty

Assumption 3.1 (Existence of an exterior-interior mapping for the Ghost penalty)**.** *There exists a mapping* $\mathcal{E}\colon \{T \in \mathcal{T}_h \,|\, T \subseteq \mathcal{E}(\Omega^{lin,n})\} \to \{T \in \mathcal{T}_h \,|\, T \subseteq \mathcal{I}(\Omega^{lin,n})\}$ *such that*

(1) For some element T *in the image of* $\mathcal{E}$*, the number of exterior elements which are mapped onto* T *is bounded*

$$\#\{\mathcal{B}^{-1}(T)\} \leq C \quad \text{for } T \subseteq \mathcal{I}(\Omega^{lin,n}).$$

(2) Fix some $T \in \mathcal{T}_h$ *such that* $T \subseteq \mathcal{E}(\Omega^{lin,n})\backslash\mathcal{I}(\Omega^{lin,n})$*. Then, there exists a path* $\{T_i\}_{i=0}^{M}$ *between* $T_0 = T$ *and* $T_M = \mathcal{E}(T)$ *such that all facets along it are in* $\mathcal{F}_R^{*,lin,n,ext}$*.*

(3) Moreover, the number of elements in $\mathcal{E}(\Omega^{lin,n})\backslash\mathcal{I}(\Omega^{lin,n})$ *is bounded as follows:*

$$\#\{T \in \mathcal{T}_h \,|\, T \subseteq \mathcal{E}(\Omega^{lin,n})\backslash\mathcal{I}(\Omega^{lin,n})\} \leq C_B \left(1 + \frac{\Delta t}{h}\right),$$

where C_B *is independent of* Δt *and* h*.*

Next, there are further more general assumptions

Assumption 3.2. *The convection velocity is bounded,* $\|w\|_\infty \leq C$*.*

Assumption 3.3. *Time step* Δt *and mesh size* h *satisfy*

$$\frac{h^2}{\Delta t} \leq C_G \qquad \Delta t \leq C_o.$$

We start the analysis (parallel to Preuß) by stating some results concerning the Ghost penalty.

3.1 Ghost-penalty related estimates

First, we recapitulate the Lemma about the general mechanism behind Ghost penalty in the case of an uncurved mesh:

Lemma (Counterpart of 3.1 Preuß)**.** *Let* T_1 *and* T_2 *be the elements corresponding to the patch* ω_F *of the facet* F*. The* v *be a piecewise polynomial of degree at most* $k_s \in \mathbb{N}$ *defined on the macro-element* ω_F*. Then there holds*

$$\|v\|_{T_1}^2 \leq C_{P.3.1} \left(\|[v]_{\omega_F}\|_{T_1}^2 + \|v\|_{T_2}^2\right)$$

Proof. First, note that the proof of this Lemma should be regarded as a preparation for the proof of the following Lemma. This Lemma is not needed for the analysis of our method in a strict sense, but its proof can be seen as a simple version of the proof of the more sophisticated following Lemma, which is why we want to present it here.

Fix a point $x \in T_1$. There it holds

$$v(x) = (v - \mathcal{E}^p(v|_{T_2}))(x) + \mathcal{E}^p(v|_{T_2})(x) = [v]_{\omega_F}(x) + \mathcal{E}^p(v|_{T_2})(x).$$

Then, we obtain by integration on all of T_1

$$\|v_1\|_{T_1}^2 \lesssim \|[v]_{\omega_F}\|_{T_1}^2 + \|\mathcal{E}^p(v|_{T_2})\|_{T_1}^2.$$

Because of shape regularity, the norms of the polynomial $\mathcal{E}^p(v|_{T_2})$ on T_1 and T_2 are equivalent and we can conclude

$$\|v_1\|_{T_1}^2 \lesssim \|[v]_{\omega_F}\|_{T_1}^2 + \|v\|_{T_2}^2. \quad \square$$

The Lemma can be extended to the mapped case as follows:

Lemma 3.4 (P.3.1M (mapped)). *Let T_1 and T_2 be the elements of the facet patch ω_F. Let $v = \hat{v} \circ (\Theta_h^{st}(t))^{-1}$, where $\hat{v}$ is a piecewise polynomial of degree at most k_s defined on ω_F. Then there holds*

$$\|v\|_{\Theta_h^{st}(t)(T_1)}^2 \leq C_{P.3.1M} \left(\|[v]_{\omega_F^h}\|_{\Theta_h^{st}(t)(T_1)}^2 + \|v\|_{\Theta_h^{st}(t)(T_2)}^2 \right)$$
$$\|\nabla v\|_{\Theta_h^{st}(t)(T_1)}^2 \leq C_{P.3.1M} \left(\frac{1}{h^2} \|[v]_{\omega_F^h}\|_{\Theta_h^{st}(t)(T_1)}^2 + \|\nabla v\|_{\Theta_h^{st}(t)(T_2)}^2 \right)$$

Proof. First note that we are considering a fixed time t and can define $T_i^h := \Theta_h^{st}(t)(T_i)$, $i = 1, 2$ relative to that. Then we have a norm equivalence between T_i and T_i^h and the corresponding discrete functions:

$$\|v\|_{T_i^h} \simeq \|\hat{v}\|_{T_i}, \quad \text{where } v = \hat{v} \circ (\Theta_h^{st}(t))^{-1}.$$

In particular, $\|v\|_{T_1^h} \lesssim \|\hat{v}\|_{T_1}$. That motivates to start our proof by deriving a representation for $\hat{v}|_{T_1}$ with the help of the facet patch jump: On T_1,

$$[v]_{\omega_F^h} \circ (\Theta_h^{st}(t)) = \hat{v} - \mathcal{E}^p(\hat{v}|_{T_2}) \circ (\mathcal{E}^p \Theta_h^{st}(t)|_{T_2})^{-1} \circ (\Theta_h^{st}(t)).$$

Rearranging and integrating over T_1 yields

$$\|\hat{v}\|_{T_1} \leq \|[v]_{\omega_F^h} \circ (\Theta_h^{st}(t))\|_{T_1} + \|\mathcal{E}^p(\hat{v}|_{T_2}) \circ (\mathcal{E}^p\Theta_h^{st}(t)|_{T_2})^{-1} \circ (\Theta_h^{st}(t))\|_{T_1}.$$

By Cauchy-Schwarz and Young's inequality,

$$\|\hat{v}\|_{T_1}^2 \lesssim \underbrace{\|[v]_{\omega_F^h} \circ (\Theta_h^{st}(t))\|_{T_1}^2}_{\simeq \|[v]_{\omega_F^h}\|_{T_1^h}^2} + \|\mathcal{E}^p(\hat{v}|_{T_2}) \circ (\mathcal{E}^p\Theta_h^{st}(t)|_{T_2})^{-1} \circ (\Theta_h^{st}(t))\|_{T_1}^2.$$

Hence, it remains to bound the right hand side term in terms of $v|_{T_2}$. To this end, we first exploit shape regularity to arrive at a maximum norm estimate

$$\underbrace{\|\mathcal{E}^p(\hat{v}|_{T_2}) \circ (\mathcal{E}^p\Theta_h^{st}(t)|_{T_2})^{-1} \circ (\Theta_h^{st}(t))\|_{T_1}}_{I} \lesssim h^{d/2}\|\mathcal{E}^p(\hat{v}|_{T_2}) \circ (\mathcal{E}^p\Theta_h^{st}(t)|_{T_2})^{-1} \circ (\Theta_h^{st}(t))\|_{\infty,T_1}.$$

Next, we define B to be a ball on the reference domain such that, $(\Theta_h^{st}(t)|_{T_2})^{-1} \circ (\Theta_h^{st}(t))(T_1) \cup T_2 \subseteq B$. Then, we can conclude

$$I \lesssim h^{d/2}\|\mathcal{E}^p(\hat{v}|_{T_2})\|_{\infty,B}.$$

Next, we note that our discrete function space is finite dimensional, so that all norms are equivalent. This suffices to establish

$$I \lesssim h^{d/2}\|\mathcal{E}^p(\hat{v}|_{T_2})\|_{\infty,T_2}.$$

Note that the constant in this estimate does not degenerate as we could find a ball $B_2 \subseteq T_2$, $B_2 \subseteq B$ and estimate the ∞-norm on B against that on B_2. As the radius of B and B_2 would only differ by a constant, the constant for the norm estimate is also bounded. Eventually, we move back to the T_2 norm, taking into account that we are on a polynomial space:

$$I \lesssim \|\mathcal{E}^p(\hat{v}|_{T_2})\|_{T_2} = \|\hat{v}\|_{T_2} \lesssim \|v\|_{T_2^h}.$$

That finished the overall proof for v. To see the result for ∇v we introduce $v^0 = \frac{1}{T_2}\int_{T_2} \hat{v} ds \in \mathbb{R}$ (with $[v^0]_{\omega_F^h} = 0$ and $\nabla v^0 = 0$) and use standard inverse inequalities and interpolation

results to obtain

$$\begin{aligned}
\|\nabla v\|^2_{\Theta_h^{st}(t)(T_1)} = \|\nabla(v-v^0)\|^2_{\Theta_h^{st}(t)(T_1)} &\lesssim \frac{1}{h^2}\|(v-v^0)\|^2_{\Theta_h^{st}(t)(T_1)} \\
&\lesssim \frac{1}{h^2}\|[v-v^0]_{\omega_F^h}\|^2_{\Theta_h^{st}(t)(T_1)} + \frac{1}{h^2}\|v-v^0\|^2_{\Theta_h^{st}(t)(T_2)} \\
&\lesssim \frac{1}{h^2}\|[v-v^0]_{\omega_F^h}\|^2_{\Theta_h^{st}(t)(T_1)} + \frac{1}{h^2}\|\hat{v}-v^0\|^2_{T_2} \\
&\lesssim \frac{1}{h^2}\|[v]_{\omega_F^h}\|^2_{\Theta_h^{st}(t)(T_1)} + \|\nabla\hat{v}\|^2_{T_2} \\
&\lesssim \frac{1}{h^2}\|[v]_{\omega_F^h}\|^2_{\Theta_h^{st}(t)(T_1)} + \|\nabla v\|^2_{\Theta_h^{st}(t)(T_2)}. \qquad \square
\end{aligned}$$

This Lemma allows us to "traverse" along the mesh and estimate norms of a function on the extended domain $\mathcal{E}(Q^{h,n})$ against the norm on $\mathcal{I}(Q^{h,n})$ at the expense of ghost penalty summands. This is summarised in the following Lemma:

Lemma 3.5 (Counterpart of P.3.2)**.** *There exists a constant $C_{3.2}$ such that for every $u \in W_h$ there holds*

$$\begin{aligned}
\|u\|^2_{\mathcal{E}(Q^{h,n})} &\leq C_{3.2}\left(\frac{h^2}{\gamma_J} j_h^n(u,u) + \|u\|^2_{\mathcal{I}(Q^{h,n})}\right) \\
\|\nabla u\|^2_{\mathcal{E}(Q^{h,n})} &\leq C_{3.2}\left(\frac{1}{\gamma_J} j_h^n(u,u) + \|\nabla u\|^2_{\mathcal{I}(Q^{h,n})}\right)
\end{aligned}$$

Proof. Basically, we can copy the proof of Preuß, taking into consideration that the changes between $\mathcal{I}(Q^{lin,n})$ and $\mathcal{I}(Q^n)$ and the respective extensions exactly accord to the modified facet set in the new ghost penalty.

In detail, the proof goes as follows: As $\mathcal{E}(Q^{h,n})$ was defined as the image of $\mathcal{E}(Q^{lin,n})$ under the mapping Θ_h^{st}, we can write

$$\begin{aligned}
\|u\|^2_{\mathcal{E}(Q^{h,n})} &= \sum_{T\in\mathcal{E}(\Omega^{lin,n})} \int_{I_n} \mathrm{d}t \|u\|^2_{\Theta_h^{st}(t,T)} \\
&= \sum_{T\in\mathcal{E}(\Omega^{lin,n})\setminus\mathcal{I}(\Omega^{lin,n})} \int_{I_n} \mathrm{d}t \|u\|^2_{\Theta_h^{st}(t,T)} + \sum_{T\in\mathcal{I}(\Omega^{lin,n})} \int_{I_n} \mathrm{d}t \|u\|^2_{\Theta_h^{st}(t,T)}
\end{aligned}$$

We can identify the summand $\sum_{T\in\mathcal{I}(\Omega^{lin,n})} \int_{I_n} \mathrm{d}t \|u\|^2_{\Theta_h^{st}(t,T)}$ with $\|u\|^2_{\mathcal{I}(Q^{h,n})}$ in the statement of the Lemma. Hence, it remains to bound the first term.

To this end, fix an element $T \subseteq \mathcal{E}(\Omega^{lin,n})\setminus\mathcal{I}(\Omega^{lin,n})$. By Assumption 3.1, we find an element $\mathcal{E}(T) \subseteq \mathcal{I}(\Omega^{lin,n})$ and a path between T and $\mathcal{E}(T)$ of finite length, $\{T_i\}_{i=0}^M$, $T_0 = T$, $T_M = \mathcal{E}(T)$. If we traverse along this path, we can apply Lemma 3.1M in each step in order

to obtain for each $T \subseteq \mathcal{E}(\Omega^{lin,n}) \backslash \mathcal{I}(\Omega^{lin,n})$ and time $t \in I_n$

$$\|u\|^2_{\Theta_h^{st}(t,T)} \lesssim \|u\|^2_{\Theta_h^{st}(t,\mathcal{E}(T))} + \sum_{F \in \mathcal{F}_R^{*,lin,n,ext}} \|[u]_{\omega_F^h}\|^2_{\omega_F^h(t)}.$$

If we integrate this equation in time over I_n and iterate over all elements from $\mathcal{E}(\Omega^{lin,n}) \backslash \mathcal{I}(\Omega^{lin,n})$, we end up with

$$\begin{aligned}
\|u\|^2_{\mathcal{E}(Q^{h,n})} \lesssim & \left(\sum_{T \in \mathcal{I}(\Omega^{lin,n})} (1 + \#\mathcal{B}^{-1}(T)) \int_{I_n} \mathrm{d}t \|u\|^2_{\Theta_h^{st}(t,T)} \right. \\
& \left. + \#\{T \in \mathcal{T}_h \,|\, T \subseteq \mathcal{E}(Q^{lin,n}) \backslash \mathcal{I}(Q^{lin,n})\} \int_{I_n} \mathrm{d}t \sum_{F \in \mathcal{F}_R^{*,lin,n,ext}} \|[u]_{\omega_F^h}\|^2_{\omega_F^h(t)} \right) \\
\lesssim & \|u\|^2_{\mathcal{I}(Q^{h,n})} + \left(1 + \frac{\Delta t}{h}\right) \sum_{F \in \mathcal{F}_R^{*,lin,n,ext}} \int_{I_n} \mathrm{d}t \|[u]_{\omega_F^h}\|^2_{\omega_F^h(t)} \\
\lesssim & \|u\|^2_{\mathcal{I}(Q^{h,n})} + \frac{h^2}{\gamma_J} j_h^n(u,u).
\end{aligned}$$

This finishes the proof for the first estimate of $\|u\|^2_{\mathcal{E}(Q^{h,n})}$. For the result about $\|\nabla u\|^2_{\mathcal{E}(Q^{h,n})}$ we proceed in the same way, taking into account the second result of Lemma 3.4. □

In the next Lemma, we will derive a bound for the norm of a discrete function on the temporal parts of the space-time boundary against the norm on the space-time "volume":

Lemma 3.6 (Counterpart of 3.3). *For any $u \in W_h$, it holds*

$$\begin{aligned}
(u_+^{n-1}, u_+^{n-1})_{\Omega^h(t_{n-1})} + (u_-^n, u_-^n)_{\Omega^h(t_n)} &\leq \frac{C}{\Delta t} \|u\|^2_{\mathcal{E}(Q^{h,n})} \\
&\leq \frac{C_{3.3}}{\Delta t} \left(\frac{h^2}{\gamma_J} j_h^n(u,u) + \|u\|^2_{\mathcal{I}(Q^{h,n})} \right)
\end{aligned}$$

Proof. For the proof of this Lemma, we transfer a result from the linear reference configuration to the curved geometry. By a similar argument to that in the proof of Lemma 3.3 of Preuß, it can be established that

$$\|\hat{u}_+^{n-1}\|^2_{\Omega^{lin}(t_{n-1})} \lesssim \frac{1}{\Delta t} \|\hat{u}\|^2_{\mathcal{E}(Q^{lin,n})}.$$

Then we can exploit the equivalence between Q^{lin} and Q^h to obtain

$$\|u_+^{n-1}\|^2_{\Omega^h(t_{n-1})} \lesssim \|\hat{u}_+^{n-1}\|^2_{\Omega^{lin}(t_{n-1})} \lesssim \frac{1}{\Delta t} \|\hat{u}\|^2_{\mathcal{E}(Q^{lin,n})} \lesssim \frac{1}{\Delta t} \|u\|^2_{\mathcal{E}(Q^{h,n})}$$

Combining this with the result of Lemma 3.2 gives the result for the first term. For the second term, one argues similary. □

This finishes the estimates for discrete functions u itself. Next, we will also consider a similar estimate for the time derivative of the discrete function. Both will be of importance for later application.

Lemma 3.7 (Counterpart of 3.4)**.** *For any $u \in W_h$, it holds*

$$\Delta t(\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} \leq \frac{C_{3.4}}{\Delta t}\left(\frac{h^2}{\gamma_J} j_h^n(u,u) + \|u\|_{\mathcal{I}(Q^{h,n})}\right),$$

and for the gradient

$$\Delta t(\partial_t^\Theta \nabla u, \partial_t^\Theta \nabla u)_{Q^{h,n}} \leq \frac{C_{3.4}}{\Delta t}\left(\frac{1}{\gamma_J} j_h^n(u,u) + \|\nabla u\|_{\mathcal{I}(Q^{h,n})}\right).$$

Proof. First, let us consider the first claim. Fix $u \in W_h$. Then, by Corollary 2.11

$$\Delta t\|\partial_t^\Theta u\|^2_{Q^{h,n}} \lesssim \Delta t\|\partial_t \hat{u}\|^2_{Q^{lin,n}}$$

By the construction according to that in the proof of Lemma 3.4 of Preuß, we note

$$\Delta t\|\partial_t \hat{u}\|^2_{Q^{lin,n}} \lesssim \frac{1}{\Delta t}\|\hat{u}\|^2_{\mathcal{E}(Q^{lin,n}} \lesssim \frac{1}{\Delta t}\|u\|^2_{\mathcal{E}(Q^{h,n})}$$

Applying Lemma 3.5 yields the result for $u \in W_h$.

Next, consider the second estimate about ∇u for $u \in W_h$. We observe

$$\Delta t\|\partial_t^\Theta \nabla u\|^2_{Q^{h,n}} \lesssim \Delta t\|\partial_t \nabla \hat{u}\|^2_{Q^{lin,n}} \leq \Delta t\|\partial_t \nabla \hat{u}\|^2_{\mathcal{E}(Q^{lin,n})}$$

The region $\mathcal{E}(Q^{lin,n})$ has tensor product structure and $\nabla \hat{u}$ is a polynomial elementwise. So we can apply an inverse inequality to obtain

$$\Delta t\|\partial_t^\Theta \nabla u\|^2_{Q^{h,n}} \lesssim \frac{1}{\Delta t}\|\nabla \hat{u}\|^2_{\mathcal{E}(Q^{lin,n})} \lesssim \frac{1}{\Delta t}\|\nabla u\|^2_{\mathcal{E}(Q^{h,n})}$$

Now we can apply the second result of Lemma 3.5 to arrive at the result. □

Finally, we also note that the relation between the ghost penalty of the time derivative and the function itself can be transfered:

Lemma 3.8 (Counterpart of 3.5). *For all $u \in W_h$, and for a constant $C_{3.5}$ only dependig on polynomial degree and shape regularity, it holds*

$$J(\partial_t^\Theta u, \partial_t^\Theta u) \leq \frac{C_{3.5}}{\Delta t^2} J(u,u) + \frac{C_{3.5}\tilde{\gamma}_J \cdot h^2}{\Delta t^2} \|\nabla u\|_{Q^h}^2.$$

Proof. We fix a time interval I_n and a facet F from $\mathcal{F}_R^{*,lin,n,ext}$ and observe

$$\begin{aligned}
j_F^n(\partial_t^\Theta u, \partial_t^\Theta u) &= \int_{t_{n-1}}^{t_n} \int_{\omega_F^h(t)} \frac{\tilde{\gamma}_J}{h^2} [\partial_t^\Theta u]_{\omega_F^h(t)}^2 \mathrm{d}x\mathrm{d}t \\
&= \int_{t_{n-1}}^{t_n} \sum_{\substack{i,j=1,2,\\ i\neq j}} \int_{\Theta_h^{st}(t)(T_i)} \frac{\tilde{\gamma}_J}{h^2} \left(\underbrace{\partial_t^\Theta u(x)}_{=(\partial_t \hat{u}|_{T_i}) \circ (\Theta_h^{st}(t)|_{T_i})^{-1}} - \mathcal{E}^p(\partial_t \hat{u}|_{T_j})((\mathcal{E}^p \Theta_h^{st}(t)|_{T_j})^{-1}(x)) \right)^2 \mathrm{d}x\mathrm{d}t \\
&\lesssim \int_{t_{n-1}}^{t_n} \sum_{\substack{i,j=1,2,\\ i\neq j}} \int_{T_i} \frac{\tilde{\gamma}_J}{h^2} \left(\partial_t \hat{u}|_{T_i} - \mathcal{E}^p(\partial_t \hat{u}|_{T_j}) \circ \underbrace{(\mathcal{E}^p \Theta_h^{st}(t)|_{T_j})^{-1} \circ (\Theta_h^{st}(t)|_{T_i})}_{:=\psi}(x) \right)^2 \\
&\lesssim \int_{t_{n-1}}^{t_n} \sum_{\substack{i,j=1,2,\\ i\neq j}} \int_{T_i} \frac{\tilde{\gamma}_J}{h^2} \left(\partial_t \hat{u}|_{T_i} - \mathcal{E}^p(\partial_t \hat{u}|_{T_j}) \right)^2 + \frac{\tilde{\gamma}_J}{h^2} \left(\mathcal{E}^p(\partial_t \hat{u}|_{T_j}) - \mathcal{E}^p(\partial_t \hat{u}|_{T_j}) \circ \psi \right)^2
\end{aligned}$$

The first expression $\partial_t \hat{u}|_{T_i} - \mathcal{E}^p(\partial_t \hat{u}|_{T_j})$ is a polynomial in time and we can apply an inverse inequality. For the second term, we can find a bound as follows:

$$\mathcal{E}^p(\partial_t \hat{u}|_{T_j}) - \mathcal{E}^p(\partial_t \hat{u}|_{T_j}) \circ \psi \lesssim \|\nabla \partial_t \mathcal{E}^p(\hat{u}|_{T_j})\|_{\infty,T_i} \cdot \underbrace{\|\mathrm{id} - \psi\|_{\infty,T_i}}_{\lesssim h^2}.$$

The term $\nabla \partial_t \hat{u}|_{T_j}$ is also a polynomial so that an inverse inequality holds and we overally have

$$j_F^n(\partial_t^\Theta u, \partial_t^\Theta u) \lesssim \int_{t_{n-1}}^{t_n} \sum_{\substack{i,j=1,2,\\ i\neq j}} \int_{T_i} \frac{\tilde{\gamma}_J}{h^2 \cdot \Delta t^2} \left(\hat{u}|_{T_i} - \mathcal{E}^p(\hat{u}|_{T_j}) \right)^2 + \frac{\tilde{\gamma}_J \cdot h^2}{\Delta t^2} \left(\|\nabla \mathcal{E}^p(\hat{u}|_{T_j})\|_{\infty,T_i} \right)^2$$

Concerning the expression $\left(\hat{u}|_{T_i} - \mathcal{E}^p(\hat{u}|_{T_j}) \right)^2$, we apply triangle inequality to yield

$$\left(\hat{u}|_{T_i} - \mathcal{E}^p(\hat{u}|_{T_j}) \right)^2 \lesssim \left(\hat{u}|_{T_i} - \mathcal{E}^p(\hat{u}|_{T_j}) \circ \psi \right)^2 + \underbrace{(\mathcal{E}^p(\hat{u}|_{T_j}) \circ \psi - \mathcal{E}^p(\hat{u}|_{T_j}))^2}_{\lesssim h^2 \|\nabla \mathcal{E}^p(\hat{u}|_{T_j})\|_{\infty,T_i}}.$$

Putting together the last two results yields

$$j_F^n(\partial_t^\Theta u, \partial_t^\Theta u) \lesssim \frac{1}{\Delta t^2} j_F^n(u,u) + \int_{t_{n-1}}^{t_n} \sum_{\substack{i,j=1,2,\\ i\neq j}} \int_{T_i} \frac{\tilde{\gamma}_J \cdot h^2}{\Delta t^2} (\underbrace{\|\nabla \mathcal{E}^p(\hat{u}|_{T_j})\|_{\infty,T_i}}_{\lesssim \|\nabla \mathcal{E}^p(\hat{u}|_{T_j})\|_{\infty,T_j}})^2$$
$$\lesssim \frac{1}{\Delta t^2} j_F^n(u,u) + \int_{t_{n-1}}^{t_n} \frac{\tilde{\gamma}_J \cdot h^2}{\Delta t^2} \|\nabla u\|^2_{\omega_F^h(t)}$$

Next, we consider the sum of all those facet contributions in the total ghost penalty term

$$J(\partial_t^\Theta u, \partial_t^\Theta u) = \sum_{n=1}^N \sum_{F \in \mathcal{F}_R^{*,lin,n,ext}} j_F^n(\partial_t^\Theta u, \partial_t^\Theta u) \lesssim \frac{1}{\Delta t^2} J(u,u) + \frac{\tilde{\gamma}_J \cdot h^2}{\Delta t^2} \|\nabla u\|^2_{\mathcal{E}(Q^h)}$$

We can exploit the second equation of Lemma 3.5 to obtain the bound

$$J(\partial_t^\Theta u, \partial_t^\Theta u) \lesssim \frac{1+h^2}{\Delta t^2} J(u,u) + \frac{\tilde{\gamma}_J \cdot h^2}{\Delta t^2} \|\nabla u\|^2_{Q^h}.$$

For h sufficiently small, the result follows. □

3.2 Special trace inequalities

In addition to the technical Ghost penalty lemmata, we also need specific special trace inequalities, which we present in this subsection.

Lemma 3.9 (Special trace inequality)**.** *For all $u \in W_h$ it holds*

$$\int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} u(t)^2 \leq \frac{C_{stra}}{\Delta t} |||u|||_j^2.$$

Proof. We start by fixing a time $t \in [t_{n-1}, t_n]$ for some n. Next, we define the image of the interior region on the reference configuration as follows:

$$\mathcal{I}(\Omega^{h,n}) := \Theta_h^{st}(\mathcal{I}(\Omega^{lin,n}), t).$$

In some contexts, the time dependence of this region will be also of importance. Then, we write $\mathcal{I}(\Omega^{h,n})(t)$ for this construction.

Next, note that if $\mathcal{I}(\Omega^{h,n}) \subseteq \Omega^h(t)$ has non-vanishing measure, trace and Poincare inequality yield

$$\|u(t)\|^2_{\partial\Omega^h(t)} \lesssim \|\nabla u\|^2_{\Omega^h(t)} + \left(\int_{\mathcal{I}(\Omega^{h,n})} u(t)\, \mathrm{d}x \right)^2$$

Next, we want to find an estimate for the right hand side term involving the directional time derivative. To this end, we observe that with the factorisation $u = \hat{u} \circ (\Theta_h^{st})^{-1}$, we obtain

$$\begin{aligned}\int_{\mathcal{I}(\Omega^{h,n})(t)} u(t)\,\mathrm{d}x &\lesssim \int_{\mathcal{I}(\Omega^{lin,n})} |\hat{u}(\hat{x},t)|\,\mathrm{d}\hat{x} \\ &\leq \int_{\mathcal{I}(\Omega^{lin,n})} |\hat{u}(\hat{x},t_{n-1})|\,\mathrm{d}\hat{x} + \int_{t_{n-1}}^{t}\int_{\mathcal{I}(\Omega^{lin,n})} |\frac{\partial \hat{u}}{\partial t}(\hat{x},\tau)|\,\mathrm{d}\tau\,\mathrm{d}\hat{x},\end{aligned}$$

using the fundamental theorem of calculus in the last inequality. Next, we apply the Corollary about the directional time derivative to obtain $(\partial_t^{\Theta} u) \circ \Theta_h^{st} = \partial_t \hat{u}$ and in turn

$$\int_{\mathcal{I}(\Omega^{h,n})(t)} u(t)\,\mathrm{d}x \lesssim \int_{\mathcal{I}(\Omega^{h,n})(t_{n-1})} |u(x,t_{n-1})|\,\mathrm{d}x + \int_{t_{n-1}}^{t}\int_{\mathcal{I}(\Omega^{h,n}(\tau))} |\partial_t^{\Theta} u(x,\tau)|\,\mathrm{d}\tau\,\mathrm{d}x. \quad (4)$$

The square of the left-hand-side term in this equation can be bound by the expression where the square is taken inside the integral, and this expression respectively can be bound by integrating this squared integrands of both right-hand-side summands. Finally, we extend the integration of the time interval over all of I_n. This yields together with the first equation

$$\begin{aligned}\int_{t_{n-1}}^{t_n} \|u(t)\|^2_{\partial\Omega^h(t)}\,\mathrm{d}t \lesssim& \|\nabla u\|^2_{Q^{h,n}} + \Delta t\|u^+(t_{n-1})\|^2_{\Omega^h(t_{n-1})} + \Delta t\|\partial_t^{\Theta} u\|^2_{Q^{h,n}} \\ &\overset{\Delta\text{-ineq}}{\lesssim} \|\nabla u\|^2_{Q^{h,n}} + \Delta t\|u^-(t_{n-1})\|^2_{\Omega^h(t_{n-1})} + \Delta t\|[u]^{n-1}\|^2_{\Omega^h(t_{n-1})} \\ &\quad + \Delta t\|\partial_t^{\Theta} u\|^2_{Q^{h,n}}\end{aligned}$$

Concerning the second summand, we apply the Lemma new 3.3 to obtain

$$\Delta t\|u^-(t_{n-1})\|^2_{\Omega^h(t_{n-1})} \leq C_{3.3}\left(\frac{h^2}{\gamma_J} j_h^{n-1}(u,u) + \|u\|^2_{\mathcal{I}(Q^{h,n-1})}\right)$$

Now, we can restate the result from equation (4) above as follows:

$$\|u\|^2_{\mathcal{I}(Q^{h,n})} \lesssim \Delta t\|u^+(t_{n-1})\|^2_{\Omega^h(t_{n-1})} + \Delta t\|\partial_t^{\Theta} u\|^2_{Q^{h,n}}.$$

Applying this to the term $\|u\|^2_{\mathcal{I}(Q^{h,n-1})}$ of our previous estimate, we obtain

$$\begin{aligned}\|u\|^2_{\mathcal{I}(Q^{h,n-1})} &\lesssim \Delta t\|u^+(t_{n-2})\|^2_{\Omega^h(t_{n-2})} + \Delta t\|\partial_t^{\Theta} u\|^2_{Q^{h,n-1}} \\ &\lesssim \Delta t\|u^-(t_{n-2})\|^2_{\Omega^h(t_{n-2})} + \Delta t\|[u]^{n-2}\|^2_{\Omega^h(t_{n-2})} + \Delta t\|\partial_t^{\Theta} u\|^2_{Q^{h,n-1}}\end{aligned}$$

Collecting the results from the previous inequalities, we arrive at

$$\begin{aligned}\int_{t_{n-1}}^{t_n} \|u(t)\|^2_{\partial\Omega^h(t)}\,\mathrm{d}t \lesssim \|\nabla u\|^2_{Q^{h,n}} &+ \Delta t\|u^-(t_{n-2})\|^2_{\Omega^h(t_{n-2})} + \Delta t\|[u]^{n-1}\|^2_{\Omega^h(t_{n-1})} \\ &+ \Delta t\|\partial_t^\Theta u\|^2_{Q^n} + \Delta t\|[u]^{n-2}\|^2_{\Omega^h(t_{n-2})} + \Delta t\|\partial_t^\Theta u\|^2_{Q^{h,n-1}} \\ &+ \frac{h^2}{\gamma_J} j_h^{n-1}(u,u)\end{aligned}$$

We apply this recipe until we arrive at $t = 0$ to obtain

$$\begin{aligned}\int_{t_{n-1}}^{t_n} \|u(t)\|^2_{\partial\Omega^h(t)}\,\mathrm{d}t \lesssim \|\nabla u\|^2_{Q^{h,n}} &+ \Delta t\|u^+(0)\|^2_{\Omega^h(0)} \\ &+ \sum_{j=1}^{n} \Delta t\|[u]^j\|^2_{\Omega^h(t_j)} + \Delta t\|\partial_t^\Theta u\|^2_{Q^{h,j}} + \frac{h^2}{\gamma_J} j_h^j(u,u)\end{aligned}$$

We can make the terms of the kind $\sum_{j=1}^n \dots$ independent of n by letting the sum run to $t = T$ in each case. Then, summing up the contributions will lead to $\frac{1}{\Delta t}$ summands, which finally gives

$$\begin{aligned}\int_0^T \|u(t)\|^2_{\partial\Omega^h(t)}\,\mathrm{d}t &\lesssim \|\nabla u\|^2_{Q^h} + \Delta t\|u^+(0)\|^2_{\Omega^h(0)} \\ &\quad + \sum_{j=1}^{n} \|[u]^j\|^2_{\Omega^h(t_j)} + \|\partial_t^\Theta u\|^2_{Q^h} + \frac{h^2}{\gamma_J} J(u,u) \\ &\lesssim \frac{1}{\Delta t}|||u|||_j^2,\end{aligned}$$

where the last inequality holds under the assumption that γ_J is sufficiently large. □

Lemma 3.10 (special trace inequality 2)**.** *For all $u \in W_h + H^1(Q^h), v \in W_h$ it holds*

$$\int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} u(t)v(t) \le \frac{C_{sstra}}{\sqrt{\Delta t}}|||u|||_*|||v|||_j.$$

Proof. We start by introducing an inner product notation for the left-hand-side expression:

$$(u,v)_\Gamma := \int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} u(t)v(t).$$

Then, we can apply Cauchy-Schwarz to split the contributions:

$$(u,v)_\Gamma^2 \le (u,u)_\Gamma \cdot (v,v)_\Gamma$$

As $v \in W_h$, the special trace inequality 1 can be applied to obtain $(v,v)_\Gamma \lesssim \frac{1}{\Delta t}|||v|||_j^2$. Concerning the u-term we observe that for each time

$$\|u\|^2_{\partial\Omega^h(t)} \lesssim \|\nabla u\|^2_{\Omega^h(t)} + \|u\|^2_{\Omega^h(t)}$$

due to the usual trace inequality. Integration over all $t \in [0,T]$ yields

$$(u,u)_\Gamma \lesssim \|\nabla u\|^2_{Q^h} + \|u\|^2_{Q^h} \lesssim |||u|||^2_*,$$

for Δt sufficiently small. Collecting all the results leads to the conclusion. □

3.3 Stability

In the stability proof, we will consider a linear combination between u and $\partial_t u$ as the test function v. The results of Lemma 3.6 help to gain control over the term stemming from u. Now, we also want to gain control over the pairing $(u, \partial_t u)$ stemming from the second summand. This is the aim of the following Lemma:

Lemma 3.11 (Counterpart of 3.7 Preuß)**.** *For any $u \in W_h$, there holds*

$$\begin{aligned} B(u,\Delta t\partial_t^\Theta u) + J(u,\Delta t\partial_t^\Theta u) \geq & \frac{1}{2}\sum_{n=1}^N \Delta t(\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} - C_{3.7}|||[u]|||^2 \\ & - C_{3.7}\sum_{n=1}^N (\nabla u, \nabla u)_{Q^{h,n}} - C_{3.7} J(u,u) \end{aligned}$$

Proof. We start by writing out $B(u,\Delta t\partial_t^\Theta u)$:

$$\begin{aligned} B(u,\Delta t\partial_t^\Theta u) = & \sum_{n=1}^N (\partial_t u, \Delta t\partial_t^\Theta u)_{Q^{h,n}} + \sum_{n=1}^N (\mathbf{w}\cdot\nabla u, \Delta t\partial_t^\Theta u)_{Q^{h,n}} + (\nabla u, \Delta t\nabla\partial_t^\Theta u)_{Q^{h,n}} \\ & + \sum_{n=1}^{N-1} \Delta t([u]^n, (\partial_t^\Theta u)_+^n)_{\Omega^h(t_n)} + \Delta t(u_+^0, (\partial_t^\Theta u)_+^0)_{\Omega^h(0)} \end{aligned}$$

From Definition 2.10, we obtain $\partial_t u = \partial_t^\Theta u - \nabla u \cdot \underbrace{(\frac{\partial \Theta_h^{st}}{\partial t} \circ (\Theta_h^{st})^{-1})}_{=:\tilde{w}}$, leading to

$$
\begin{aligned}
B(u, \Delta t \partial_t^\Theta u) - \sum_{n=1}^{N} \Delta t (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} = & \underbrace{\sum_{n=1}^{N} (-\nabla u \cdot \tilde{w}, \Delta t \partial_t^\Theta u)_{Q^{h,n}} + (\mathbf{w} \cdot \nabla u, \Delta t \partial_t^\Theta u)_{Q^{h,n}}}_{=:I} \\
& + \underbrace{\sum_{n=1}^{N-1} \Delta t ([u]^n, (\partial_t^\Theta u)_+^n)_{\Omega^h(t_n)} + \Delta t (u_+^0, (\partial_t^\Theta u)_+^0)_{\Omega^h(0)}}_{=:II} \\
& + \underbrace{\sum_{n=1}^{N} (\nabla u, \Delta t \nabla \partial_t^\Theta u)_{Q^{h,n}}}_{=:III}
\end{aligned}
$$

All those terms will now be estimated seperately. We start with I:

$$
\begin{aligned}
I &= \sum_{n=1}^{N} ((\mathbf{w} - \tilde{w}) \cdot \nabla u, \Delta t \partial_t^\Theta u)_{Q^{h,n}} \\
&\leq \frac{\epsilon_3}{2} \Delta t \sum_{n=1}^{N} ((\mathbf{w} - \tilde{w}) \cdot \nabla u, (\mathbf{w} - \tilde{w}) \cdot \nabla u)_{Q^{h,n}} + \frac{\Delta t}{2\epsilon_3} \sum_{n=1}^{N} (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} \\
&\leq \frac{\epsilon_3}{2} \Delta t \|(\mathbf{w} - \tilde{w})\|_\infty^2 \sum_{n=1}^{N} (\nabla u, \nabla u)_{Q^{h,n}} + \frac{\Delta t}{2\epsilon_3} \sum_{n=1}^{N} (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}}
\end{aligned}
$$

Next, we continue with II:

$$
\begin{aligned}
II &= \sum_{n=1}^{N-1} \Delta t ([u]^n, (\partial_t^\Theta u)_+^n)_{\Omega^h(t_n)} + \Delta t (u_+^0, (\partial_t^\Theta u)_+^0)_{\Omega^h(0)} \\
&\leq \frac{\epsilon_1}{2} |||[u]|||^2 + \frac{\Delta t^2}{2\epsilon_1} \sum_{n=1}^{N} ((\partial_t^\Theta u)_+^{n-1}, (\partial_t^\Theta u)_+^{n-1})_{\Omega^h(t_{n-1})} \\
&\leq \frac{\epsilon_1}{2} |||[u]|||^2 + \frac{C_{3.3} \cdot \Delta t}{2\epsilon_1} \left(\frac{h^2}{\gamma_J} J(\partial_t^\Theta u, \partial_t^\Theta u) + \sum_{n=1}^{N} (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} \right) \qquad \text{as } \partial_t^\Theta u \in W_h \\
&\leq \frac{\epsilon_1}{2} |||[u]|||^2 + \frac{C_{3.3} \cdot \Delta t}{2\epsilon_1} \left(\frac{h^2}{\gamma_J} J(\partial_t^\Theta u, \partial_t^\Theta u) + (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^h} \right) \\
&\leq \frac{\epsilon_1}{2} |||[u]|||^2 + \frac{C_{3.3}}{2\epsilon_1} \left(\frac{h^2}{\gamma_J} \frac{C_{3.5}}{\Delta t} J(u, u) + \frac{h^2}{\gamma_J} \frac{C_{3.5} \tilde{\gamma}_J \cdot h^2}{\Delta t} \|\nabla u\|_{Q^h}^2 + \Delta t (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^h} \right)
\end{aligned}
$$

Next, we estimate term III:

$$\begin{aligned} III &= \sum_{n=1}^{N} (\nabla u, \Delta t \nabla \partial_t^\Theta u)_{Q^{h,n}} \\ &\leq \frac{\epsilon_2}{2}(\nabla u, \nabla u)_{Q^h} + \frac{C_{3.4}}{2\epsilon_2} \sum_{n=1}^{N} \frac{1}{\gamma_J} j_h^n(u,u) + \|\nabla u\|_{\mathcal{I}(Q^{h,n})} \\ &\leq \frac{\epsilon_2}{2}(\nabla u, \nabla u)_{Q^h} + \frac{C_{3.4}}{2\epsilon_2 \gamma_J} J(u,u) + \frac{C_{3.4}}{2\epsilon_2} \|\nabla u\|_{Q^h} \\ &\leq \sqrt{C_{3.4}}(\nabla u, \nabla u)_{Q^h} + \frac{\sqrt{C_{3.4}}}{2\gamma_J} J(u,u), \end{aligned}$$

choosing $\epsilon_2 = \sqrt{C_{3.4}}$.

Finally, we consider the Ghost penalty term applying Lemma 3.8:

$$\begin{aligned} J(u, \Delta t \partial_t^\Theta u) &\leq \frac{\epsilon_4}{2} J(u,u) + \frac{\Delta t^2}{2\epsilon_4} J(\partial_t^\Theta u, \partial_t^\Theta u) \\ &\leq \frac{\epsilon_4}{2} J(u,u) + \frac{C_{3.5}}{2\epsilon_4} J(u,u) + \frac{C_{3.5}\tilde{\gamma}_J \cdot h^2}{2\epsilon_4} \|\nabla u\|_{Q^h}^2 \\ &= \sqrt{C_{3.5}} J(u,u) + \frac{\sqrt{C_{3.5}}\tilde{\gamma}_J \cdot h^2}{2} \|\nabla u\|_{Q^h}^2, \end{aligned}$$

choosing $\epsilon_4 = \sqrt{C_{3.5}}$.

Collecting all the subresults, we arrive at

$$\begin{aligned} &B(u, \Delta t \partial_t^\Theta u) + J(u, \Delta t \partial_t^\Theta u) \geq \Delta t (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^h} \left(1 - \frac{1}{2\epsilon_3} - \frac{C_{3.3}}{2\epsilon_1}\right) \\ &- J(u,u) \left(\frac{C_{3.3}}{2\epsilon_1} \frac{h^2}{\gamma_J} \frac{C_{3.5}}{\Delta t} + \frac{\sqrt{C_{3.4}}}{2\gamma_J} + \sqrt{C_{3.5}} \right) - \frac{\epsilon_1}{2} |||[u]|||^2 \\ &- (\nabla u, \nabla u)_{Q^h} \left(\frac{\epsilon_3}{2} \Delta t \|(\mathbf{w} - \tilde{w})\|_\infty^2 + \frac{C_{3.3}}{2\epsilon_1} \frac{h^2}{\gamma_J} \frac{C_{3.5}\tilde{\gamma}_J \cdot h^2}{\Delta t} + \sqrt{C_{3.4}} + \frac{\sqrt{C_{3.5}}\tilde{\gamma}_J \cdot h^2}{2} \right) \end{aligned}$$

Picking $\epsilon_1 = 2C_{3.3}$ and $\epsilon_3 = 2$ the claim follows for h sufficiently small. □

The next Lemma gives an estimate of a special function in the $|||\dots|||_j$ norm:

Lemma 3.12 (Counterpart of P.3.8)**.** *For any $u \in W_h$,*

$$|||\Delta t \partial_t^\Theta u|||_j \leq C_{3.8} |||u|||_j$$

Proof. The proof of this Lemma proceeds along the same lines we exploited in the last proof. Also compare to the proof of [9, Lemma 3.8]. □

With these Lemmata, we are able to prove the stability result:

Proposition 3.13 (Inf-sup stability, Counterpart of P.3.9)**.** *For all $u \in W_h$, there exists a $\tilde{v}(u) \in W_h$ such that*

$$B(u, \tilde{v}(u)) + J(u, \tilde{v}(u)) \geq C_{3.9} |||u|||_j \cdot |||\tilde{v}(u)|||_j.$$

Proof. Fix $u \in W_h$ and consider $\tilde{v}(u) := (2C_{3.7} + 1)u + \Delta t \partial_t^\Theta u \in W_h$. Then, we observe

$$\begin{aligned}
&B(u, \tilde{v}(u)) + J(u, \tilde{v}(u)) \\
&= (2C_{3.7} + 1)\left(B(u,u) + J(u,u)\right) + B(u, \Delta t \partial_t^\Theta u) + J(u, \Delta t \partial_t^\Theta u) \\
&\geq (2C_{3.7} + 1)\left(\frac{1}{2}|||[u]|||^2 + \sum_{n=1}^{N} (\nabla u, \nabla u)_{Q^{h,n}} + \frac{1}{2}\int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x\, (\mathcal{V}_n^h(t) - \mathbf{w}\cdot\mathbf{n})u^2\right. \\
&\qquad \left. + J(u,u)\right) + \frac{1}{2}\sum_{n=1}^{N} \Delta t (\partial_t^\Theta u, \partial_t^\Theta u)_{Q^{h,n}} - C_{3.7}|||[u]|||^2 \\
&\qquad - C_{3.7}\sum_{n=1}^{N} (\nabla u, \nabla u)_{Q^{h,n}} - C_{3.7} J(u,u) \\
&\geq \frac{1}{2}|||u|||_j^2 + \frac{2C_{3.7}+1}{2}\int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x\, (\mathcal{V}_n^h(t) - \mathbf{w}\cdot\mathbf{n})u^2 \\
&\geq \frac{1}{2}|||u|||_j^2 - \frac{2C_{3.7}+1}{2} \cdot C_{mc}(h, \Delta t) \cdot \int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x\, u^2 \\
&\geq \frac{1}{2}|||u|||_j^2 - \frac{2C_{3.7}+1}{2} \cdot C_{mc}(h, \Delta t) \cdot \frac{C_{stra}}{\Delta t}|||u|||_j^2 \\
&\geq \frac{1}{4}|||u|||_j^2, \quad \text{in line with Assumption 2.14.}
\end{aligned}$$

In addition, we note that because of 3.8. and the triangle inequality,

$$\begin{aligned}
|||\tilde{v}(u)|||_j &\leq (2C_{3.7}+1)|||u|||_j + |||\Delta t \partial_t^\Theta u||| \\
&\leq (2C_{3.7} + C_{3.8} + 1)|||u|||_j,
\end{aligned}$$

which remains unchanged in comparison to Preuß. Then, the result follows. □

3.4 Continuity

Keeping the general proof structure of Preuß, we continue by proving continuity after having obtained inf-sup stability in the last subsection. As the continuity proof exploits the alternative characterisation, we will have a new summand appearing in the proof, which fortunately can be bounded by an appropriate lemma.

Proposition 3.14 (Continuity, Counterpart of P.3.10)**.** *For all $u \in W_h + H^1(Q^h)$ and all $v \in W_h$ it holds*

$$B(u,v) \leq |||u|||_* \cdot |||v|||_j$$
$$J(u,v) \leq \|u\|_J \cdot \|v\|_J$$

Proof. Concerning the first equation, we note that

$$B(u,v) = d'(u,v) + b'(u,v) + a(u,v) + \int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x\ (\mathbf{w}\cdot\mathbf{n} - \mathcal{V}_n^h(t))uv$$

Following the proof of Lemma 3.10 of Preuß, we note that

$$d'(u,v) + b'(u,v) + a(u,v) \lesssim |||u|||_* \cdot |||v||| \leq |||u|||_* \cdot |||v|||_j.$$

It remains to bound the last summand accordingly. To this end, we apply the special trace inequality 2:

$$\begin{aligned}\int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x\ (\mathbf{w}\cdot\mathbf{n} - \mathcal{V}_n^h(t))uv &\leq C_{mc}(h,\Delta t) \cdot \int_0^T \mathrm{d}t \int_{\partial\Omega^h(t)} \mathrm{d}x\ (\mathbf{w}\cdot\mathbf{n} - \mathcal{V}_n^h(t))uv \\ &\lesssim \frac{C_{mc}(h,\Delta t)}{\sqrt{\Delta t}} |||u|||_* \cdot |||v|||_j,\end{aligned}$$

Taking into account Assumption 2.14, we arrive at the bound for $B(u,v)$. For the ghost penalty, the proof of Preuß can be applied for our slightly changed facet set in the same way. □

3.5 Strang-type analysis

The results of the previous subsections allow us now to proof a Strang-Lemma-like result. This is of high importance as it gives us a first upper bound on the numerical error.

As the solution of the continuous problem is defined on the exact geometry, our mapping Ψ^{st} translating between Q^{lin} and Q becomes relevant here. Remember that it was defined such that $\Psi^{st}(Q^{lin}) = Q$, paralleling $\Theta_h^{st}(Q^{lin}) = Q^h$. Last, also a mapping $\Phi^{st}\colon Q^h \to Q$ was defined.

We denote by $u \in H^2(Q)$ the solution to the exact problem and lift it to the discrete geometry:

$$u^l := u \circ \Phi^{st} = u \circ \Psi^{st} \circ (\Theta_h^{st})^{-1}.$$

In addition, we want to lift some functions q_h in the other direction, motivating the definition

$$q_h^{-l} := q_h \circ (\Phi^{st})^{-1} = q_h \circ \Theta_h^{st} \circ (\Psi^{st})^{-1}.$$

This allows us to measure the difference between u and the solution to the discrete problem u_h as follows:

Theorem 3.15 (Strang-type result)**.** *Denoting by u the solution to (2) and by u_h the solution to (3), we obtain*

$$\begin{aligned}|||u^l - u_h||| \lesssim \inf_{v_h \in W_h} |||u^l - v_h||| + |||u^l - v_h|||_* + |||v_h|||_J \\ + \sup_{q_h \in W_h} \frac{|f_h(q_h) - f(q_h^{-l})|}{|||q_h|||_J} + \sup_{q_h \in W_h} \frac{|B_h(u^l, q_h) - B(u, q_h^{-l})|}{|||q_h|||_J}\end{aligned}$$

Proof. Fix an arbitrary $v_h \in W_h$. Then, it holds by the triangle inequality

$$|||u^l - u_h||| \leq |||u^l - v_h||| + |||v_h - u_h|||$$

Concerning the second summand we observe that for some $q_h \in W_h$ chosen as $\tilde{v}(u_h - v_h)$ in line with Proposition 3.13

$$\begin{aligned}|||u_h - v_h||| \cdot |||q_h|||_J \leq |||u_h - v_h|||_J \cdot |||q_h|||_J \lesssim B_h(u_h - v_h, q_h) + J(u_h - v_h, q_h) \\ = \underbrace{B_h(u_h, q_h) + J(u_h, q_h)}_{= f_h(q_h)} - B_h(v_h, q_h) - J(v_h, q_h)\end{aligned}$$

As u was the solution to the continuous problem, we can add $0 = B(u, q_h^{-l}) - f(q_h^{-l})$ and $0 = B_h(u^l, q_h) - B_h(u^l, q_h)$ to obtain

$$\begin{aligned}|||u_h - v_h||| \cdot |||q_h|||_J \lesssim & f_h(q_h) + B(u, q_h^{-l}) - f(q_h^{-l}) + B_h(u^l, q_h) - B_h(u^l, q_h) \\ & - B_h(v_h, q_h) - J(v_h, q_h) \\ \lesssim & |f_h(q_h) - f(q_h^{-l})| + \underbrace{B_h(u^l - v_h, q_h)}_{\lesssim |||u^l - v_h|||_* \cdot |||q_h|||} \\ & + |B(u, q_h^{-l}) - B_h(u^l, q_h)| \underbrace{- J(v_h, q_h)}_{\lesssim |||v_h|||_J \cdot |||q_h|||_J}\end{aligned}$$

Dividing this inequality by $|||q_h|||_J$ and combining it with the first triangle inequality step of this proof, we arrive at the result. □

3.6 Geometrical consistency analysis

In the next two Lemmata, we will derive bounds for the summands in the previous result which are not related to interpolation.

Lemma 3.16 (Estimate on right-hand-side deviation in Strang)**.** *It holds*

$$\sup_{q_h \in W_h} \frac{|f_h(q_h) - f(q_h^{-l})|}{|||q_h|||_J} \lesssim \sup_{q_h \in W_h} \frac{\|q_h\|_{Q^h}}{|||q_h|||_J} \left(h^{k_s+1} \cdot \|f\|_{W^{1,\infty}(Q)} + h^{k_s} \cdot \|f\|_Q \right).$$

Proof. We start by writing out $f_h(q_h)$ and $f(q_h^{-l})$:

$$f(q_h^{-l}) = \int_Q f q_h^{-l} \,\mathrm{d}x = \int_{Q^h} |\mathrm{det} D\Phi^{st}| (f \circ \Phi^{st}) \underbrace{(q_h^{-l} \circ \Phi^{st})}_{=q_h}$$
$$f_h(q_h) = \int_{Q^h} f^e q_h,$$

where f^e denotes an extension of f to at least Q^h. For the difference we obtain

$$\begin{aligned} |f_h(q_h) - f(q_h^{-l})| &\lesssim \int_{Q_h} |(f^e - |\mathrm{det} D\Phi^{st}|(f \circ \Phi^{st}))q_h| \\ &\lesssim |\int_{Q_h} \underbrace{(1 - \mathrm{det} D\Phi^{st})}_{\lesssim h^{k_s}} f^e q_h| + |\int_{Q_h} \underbrace{|\mathrm{det} D\Phi^{st}|}_{\lesssim 1} q_h (f^e - f \circ \Phi^{st})|. \end{aligned}$$

For the expression $(f^e - f \circ \Phi^{st})$, we can derive an estimate as follows

$$\|f^e - f \circ \Phi^{st}\|_{\infty, Q^h} \lesssim \|f^e\|_{W^{1,\infty}(Q^h)} \cdot \underbrace{\|\mathrm{Id} - \Phi^{st}\|_{\infty, Q^h}}_{\lesssim h^{k_s+1}}$$

In total, we arrive at

$$|f_h(q_h) - f(q_h^{-l})| \lesssim h^{k_s} \|f^e\|_{Q^h} \cdot \|q_h\|_{Q^h} + h^{k_s+1} \|f^e\|_{W^{1,\infty}} \|q_h\|_{Q^h}$$

Factoring out the ratio between $\|q_h\|_{Q^h}$ and $\|q_h\|_J$, we arrive at the result. □

Lemma 3.17 (Estimate on left-hand-side deviation in Strang)**.** *It holds*

$$\sup_{q_h \in W_h} \frac{|B_h(u^l, q_h) - B(u, q_h^{-l})|}{|||q_h|||_J}$$
$$\lesssim S^* h^{k_s} \cdot \left(\|\frac{\partial u}{\partial t}\|_Q + (h + \frac{h^2}{\Delta t} + 1 + h\|w\|_{W^{1,\infty}} + \|w\|_{\infty,Q})\|\nabla u\|_Q + \|u^0\|_{\Omega^h(t_0)} \right)$$
$$\text{where } S^* = \max\left(1, \sup_{q_h \in W_h} \frac{\|q_h\|_{Q^h}}{|||q_h|||_J}\right).$$

Proof. We start by writing out the expressions for $B_h(u^l, q_h)$ and $B(u, q_h^{-l})$. Note that the solution u is continuous so that we can neglect the jump terms:

$$\begin{aligned}
B_h(u^l, q_h) &= \underbrace{\int_{Q^h} \partial_t(u^l) q_h}_{Ia} + \underbrace{\int_{Q^h} \mathbf{w}\nabla(u^l) q_h}_{IIa} + \underbrace{\int_{Q^h} \nabla(u^l)\nabla q_h}_{IIIa} + \underbrace{\int_{\Omega^h(t_0)} (u^l)^0_+ q_h^0}_{IVa} \\
B(u, q_h^{-l}) &= \underbrace{\int_{Q^h} |\det D\Phi^{st}| (\partial_t u \circ \Phi^{st}) q_h}_{Ib} + \underbrace{\int_{Q^h} |\det D\Phi^{st}| (\mathbf{w}\nabla u \circ \Phi^{st}) q_h}_{IIb} \\
&\quad + \underbrace{\int_{Q^h} |\det D\Phi^{st}| (\nabla u \circ \Phi^{st})(\nabla(q_h^{-l}) \circ \Phi^{st})}_{IIIb} \\
&\quad + \underbrace{\int_{\Omega^h(t_0)} |\det D_x\Phi^{st}(\ldots, t_0)| (u^0_+ \circ \Phi) q_h^0}_{IVb}
\end{aligned}$$

In order to estimate the respective differences, we start with a series of basically just applications of the chain rule:

$$\begin{aligned}
\frac{\partial u^l}{\partial t} &= \frac{\partial}{\partial t}(u \circ \Phi^{st}) = \frac{\partial u}{\partial t} \circ \Phi^{st} + (\nabla u \circ \Phi^{st}) \frac{\partial \Phi^{st}}{\partial t} \\
\nabla(u^l) &= D_x(\Phi^{st})^T (\nabla u \circ \Phi^{st}) \\
\nabla(q^{-l}) &= D_x(\Phi^{st})^{-T} (\nabla q \circ (\Phi^{st})^{-1})
\end{aligned}$$

Then we can estimate the differences of the respective terms

$$\begin{aligned}
Ia - Ib &= \int_{Q_h} q_h \left(\partial_t u \circ \Phi^{st} + (\nabla u \circ \Phi^{st}) \frac{\partial \Phi^{st}}{\partial t} - |\det D\Phi^{st}| (\partial_t u \circ \Phi^{st}) \right) \\
|Ia - Ib| &\leq \int_{Q_h} q_h (\nabla u \circ \Phi^{st}) \frac{\partial \Phi^{st}}{\partial t} q_h - \underbrace{|1 - \det D\Phi^{st}|}_{\|\ldots\|_\infty \lesssim h^k} (\partial_t u \circ \Phi^{st}) q_h
\end{aligned}$$

In addition, we observe $\|\frac{\partial \Phi^{st}}{\partial t}\|_\infty \lesssim h^{k_s}(h + h^2/\Delta t)$ leading to

$$|Ia - Ib| \lesssim h^{k_s}\|q_h\|_{Q^h}\left(\|\frac{\partial u}{\partial t}\|_Q + (h + \frac{h^2}{\Delta t})\|\nabla u\|_Q\right)$$

Next, we continue with the second term

$$IIa - IIb = \int_{Q^h}\left(\mathbf{w}D_x(\Phi^{st})^T - (\mathbf{w}\circ\Phi^{st})|\det D\Phi^{st}|\right)(\nabla u\circ\Phi^{st})q_h$$

We observe that for the inner term

$$\begin{aligned}
&|\mathbf{w}D_x(\Phi^{st})^T - (\mathbf{w}\circ\Phi^{st})|\det D\Phi^{st}||\\
&\leq \mathbf{w}D_x(\Phi^{st})^T - (\mathbf{w}\circ\Phi^{st}) + (\mathbf{w}\circ\Phi^{st})|1 - \det D\Phi^{st}|\\
&\leq \underbrace{|\mathbf{w} - (\mathbf{w}\circ\Phi^{st})|}_{\lesssim h^{k_s+1}\|w\|_{H^{1,\infty}}} + \underbrace{|(I - D_x(\Phi^{st})^T)|}_{\lesssim h^{k_s}}|\mathbf{w}| + (\mathbf{w}\circ\Phi^{st})\underbrace{|1 - \det D\Phi^{st}|}_{\lesssim h^{k_s}}.
\end{aligned}$$

In total we arrive at

$$|IIa - IIb| \lesssim \left(h^{k_s+1}\|w\|_{W^{1,\infty}} + h^{k_s}\|w\|_{\infty,Q}\right)\|\nabla u\|_Q\|q_h\|_{Q^h}.$$

Next, we estimate the third difference

$$\begin{aligned}
IIIa - IIIb = &\int_{Q^h}(D_x(\Phi^{st})^T(\nabla u\circ\Phi^{st}))\nabla q_h\\
&\quad - |\det D\Phi^{st}|(\nabla u\circ\Phi^{st})(D_x(\Phi^{st})^{-T}\circ\Phi^{st})\nabla q_h\\
&= \int_{Q^h}(\nabla u\circ\Phi^{st})\nabla q_h\left(D_x(\Phi^{st})^T - |\det D\Phi^{st}|(D_x(\Phi^{st})^{-T}\circ\Phi^{st})\right)
\end{aligned}$$

Concerning the inner term, we find

$$\begin{aligned}
&D_x(\Phi^{st})^T - |\det D\Phi^{st}|(D_x(\Phi^{st})^{-T}\circ\Phi^{st})\\
&= (D_x(\Phi^{st})^T(D_x(\Phi^{st})^T\circ\Phi^{st}) - I + I\underbrace{(1 - |\det D\Phi^{st}|)}_{\lesssim h^{k_s}})\underbrace{(D_x(\Phi^{st})^{-T}\circ\Phi^{st})}_{\|\ldots\|_\infty\lesssim 1}
\end{aligned}$$

In turn,

$$\begin{aligned} D_x(\Phi^{st})^T(D_x(\Phi^{st})^T \circ \Phi^{st}) - I =& (D_x(\Phi^{st})^T - I)((D_x(\Phi^{st})^T \circ \Phi^{st}) + I) \\ &+ ((D_x(\Phi^{st})^T \circ \Phi^{st}) - I) - (D_x(\Phi^{st})^T - I) \\ \lesssim& \|(D_x(\Phi^{st})^T - I)\|_\infty \lesssim h^{k_s}. \end{aligned}$$

In total, this allows us to conclude

$$|IIIa - IIIb| \lesssim h^{k_s} \|\nabla u\|_Q \cdot |||q_h|||.$$

For the last pair of summands, we note

$$\begin{aligned} |IVa - IVb| &\lesssim \int_{\Omega^h(t_0)} \underbrace{|1 - \det D_x \Phi^{st}(\dots, t_0)|}_{\lesssim h^{k_s}} (u_+^0 \circ \Phi) q_h^0 \\ &\lesssim h^{k_s} \|u^0\|_{\Omega^h(t_0)} |||q_h||| \end{aligned}$$

Collecting all the subresults and applying the triangle inequality, we arrive at the result. □

3.7 A priori error bounds

Our final step in the analysis consists of combining the upper bounds on two summands in Theorem 3.15 with bounds on the first terms. Fortunately, bounds of the latter kind, also called interpolation estimates, are already derived by [9]. We directly combine the results of his Propositions 3.23, 3.24, 3.26 in the following

Proposition 3.18 (Interpolation estimate)**.** *Choosing u and u_h as in Theorem 3.15, it holds with $k_{max} := \max(k_s, k_t)$*

$$\inf_{v_h \in W_h} |||u^l - v_h||| + |||u^l - v_h|||_* + |||v_h|||_J \lesssim \left(\Delta t^{k_t + 1/2} + \sqrt{1 + \frac{\Delta t}{h}} h^{k_s} \right) \|u\|_{H^{k_{max}+2}(Q)}$$

Proposition 3.19 (A priori error bound)**.** *Choosing u and u_h as in Theorem 3.15, it holds with*

$k_{max} := \max(k_s, k_t)$

$$
\begin{aligned}
|||u^l - u_h||| &\lesssim \left(\Delta t^{k_t+1/2} + \sqrt{1 + \frac{\Delta t}{h}} h^{k_s} \right) \|u\|_{H^{k_{max}+2}(Q)} \\
&+ S^* h^{k_s} \cdot \left(\|\frac{\partial u}{\partial t}\|_Q + (h + \frac{h^2}{\Delta t} + 1 + h\|w\|_{W^{1,\infty}} + \|w\|_{\infty,Q}) \|\nabla u\|_Q + \|u^0\|_{\Omega^h(t_0)} \right) \\
&+ \sup_{q_h \in W_h} \frac{\|q_h\|_{Q^h}}{|||q_h|||_J} \left(h^{k_s+1} \cdot \|f\|_{W^{1,\infty}(Q)} + h^{k_s} \cdot \|f\|_Q \right)
\end{aligned}
$$

4 Numerical investigations

In this section, we want to complement the theoretical findings of the previous section with some numerical examples. To this end, we will use the Finite Element library `ngsolve` and its extension `ngsxfem`. By and large, we will not discuss implementational aspects and refer the interested reader to either the sections of [9] or the report [3] for this.

4.1 Space-time DG

We want to start with an investigation of the Space-time DG method introduced in the previous section. We directly implement the higher-order variant exploiting the isoparametric mapping. In the following, we focus on combinations of equal polynomial order in space and time $k = k_s = k_t$, although it would be an interesting direction for further investigations to also consider other combinations. (See also [9, Chapter 6])

The first example is that of moving circle geometry in two spatial dimensions. The following levelset function is used for this purpose:

$$
\rho = \frac{1}{\pi} \sin 2\pi t, \quad \phi = \sqrt{x^2 + (y - \rho)^2} - r_0, \quad r_0 = \frac{1}{2}.
$$

The background domain is chosen to be $\tilde{\Omega} = [-0.6, 0.6] \times [-1, 1]$, such that the moving circle is fully contained in the mesh at all points in time. The spatial mesh is chosen to be an unstructured mesh with maximal mesh size $h = (\frac{1}{2})^{i_s+2}$ for appropriate choices of the parameter i_s. Accordingly, the time interval $[0, T] = [0, 0.5]$ is subdivided into 2^{i_t} steps for a parameter i_t. An example of the circle geometry in an example mesh can be seen in Fig. 10 on the left hand side.

To assess the convergence behaviour of the numerical method, we perform a series of refinements in space and time respectively. In each case, the other discretisation parameter

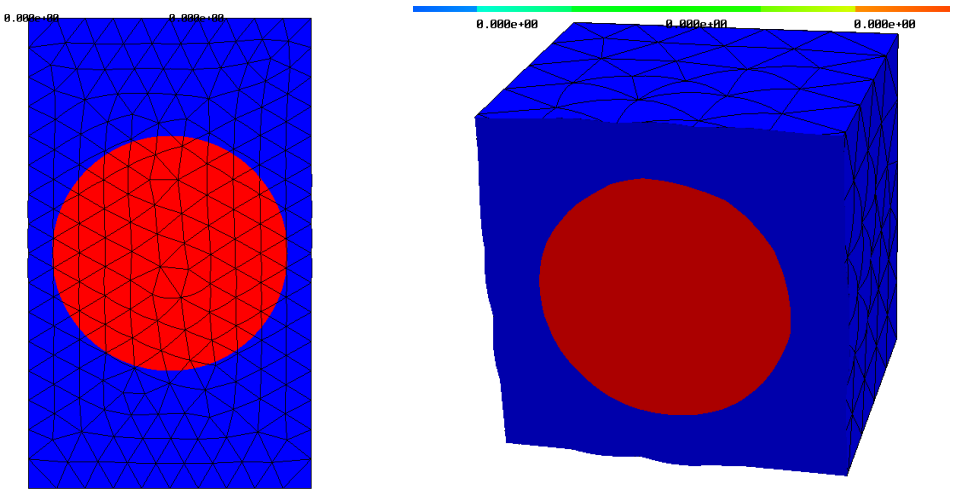


Figure 10: Snapshot of the circle and sphere geometry in the mesh (including deformation)

(i_s or i_t) is chosen sufficiently large such that the error is dominated by the discretisation under investigation. In order to calculate the numerical error from an exact solution, we pick a solution first and insert an according right-hand side f of the problem. In this case,

$$u = \cos(Qr) \cdot \sin(\pi t), \qquad Q = \frac{\pi}{r_0},$$
$$f = \left(\frac{Q}{r} \cdot \sin(Qr) + Q^2 \cdot \cos(Qr)\right) \cdot \sin(\pi t) + \pi \cdot \cos(Qr) \cdot \cos(\pi t).$$

The results for the two refinements studies for this case are shown in Fig. 11.

We observe numerically that the error decays with

$$\|u - u_h\|_Q \lesssim h^{k+1} + \Delta t^{k+1}.$$

To compare these findings with the results of the analysis of Section 3, we note that the error is measured in different norms respectively, as the $|||\dots|||$-norm does not contain a contribution $\|\dots\|_Q$. However, the approximation error measured in the $\|\dots\|_Q$-norm is known to scale with $h^{k+1} + \Delta t^{k+1}$ (compare with [9, Tab. 6.21]). Hence, we conclude that the numerical investigations confirm our theoretical results. However, it would be interesting to extend the analysis of section 3 towards an priori error bound measured in the $\|\dots\|_Q$ norm in order to compare the results directly.

In addition to the moving circle case, we also test our method with a more complex geometry, where the circle deforms over time into a kite shape. This behaviour is depicted in

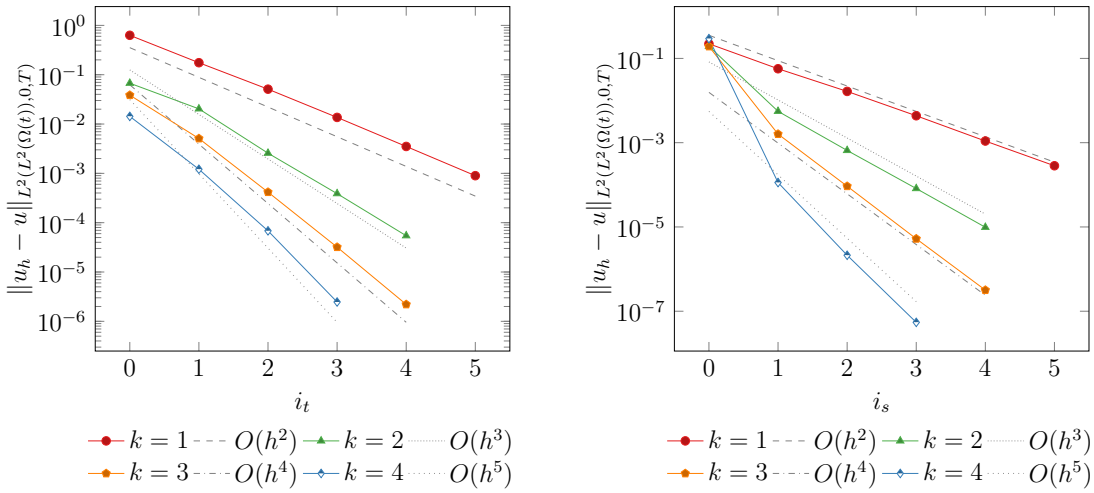


Figure 11: Convergence plots of the higher order DG method for the moving circle case. On the left-hand side, time refinements are performed for a sufficiently high i_s. On the right-hand side, space refinements are performed for a high i_t respectively. In each case, $k_t = k_s = k$.

Fig. 12 and can be described in terms of the following levelset function

$$\rho(t, y) = (1 - y^2) \cdot t, \quad \phi = \sqrt{(x - \rho)^2 + y^2} - r_0, \quad r_0 = 1.$$

Again, the convection field is calculated as

$$w = (\frac{\partial \rho}{\partial t}, 0),$$

and we calculate a manufactured right hand side to our solution $u = \cos(Qr) \cdot \sin(\pi t)$, where the radius is defined in terms of the new $\rho(t, y)$ as above in the levelset function. The numerical results are presented in Fig. 13. We observe the same experimental orders of convergence as with the moving circle case and hence conclude that our method can also be applied directly to more challenging geometries.

Because of implementational work described elsewhere [3], it is relatively straightforward to apply the methods of investigation in this text also to three-dimensional problem in space. To this end, we generalise the moving circle test case to a moving sphere test case and perform one exemplary refinement study in time. (See Fig.) Although it is computationally more expensive to generate results in three spatial dimensions, all generated data support our findings made in the context of two spatial dimensions.

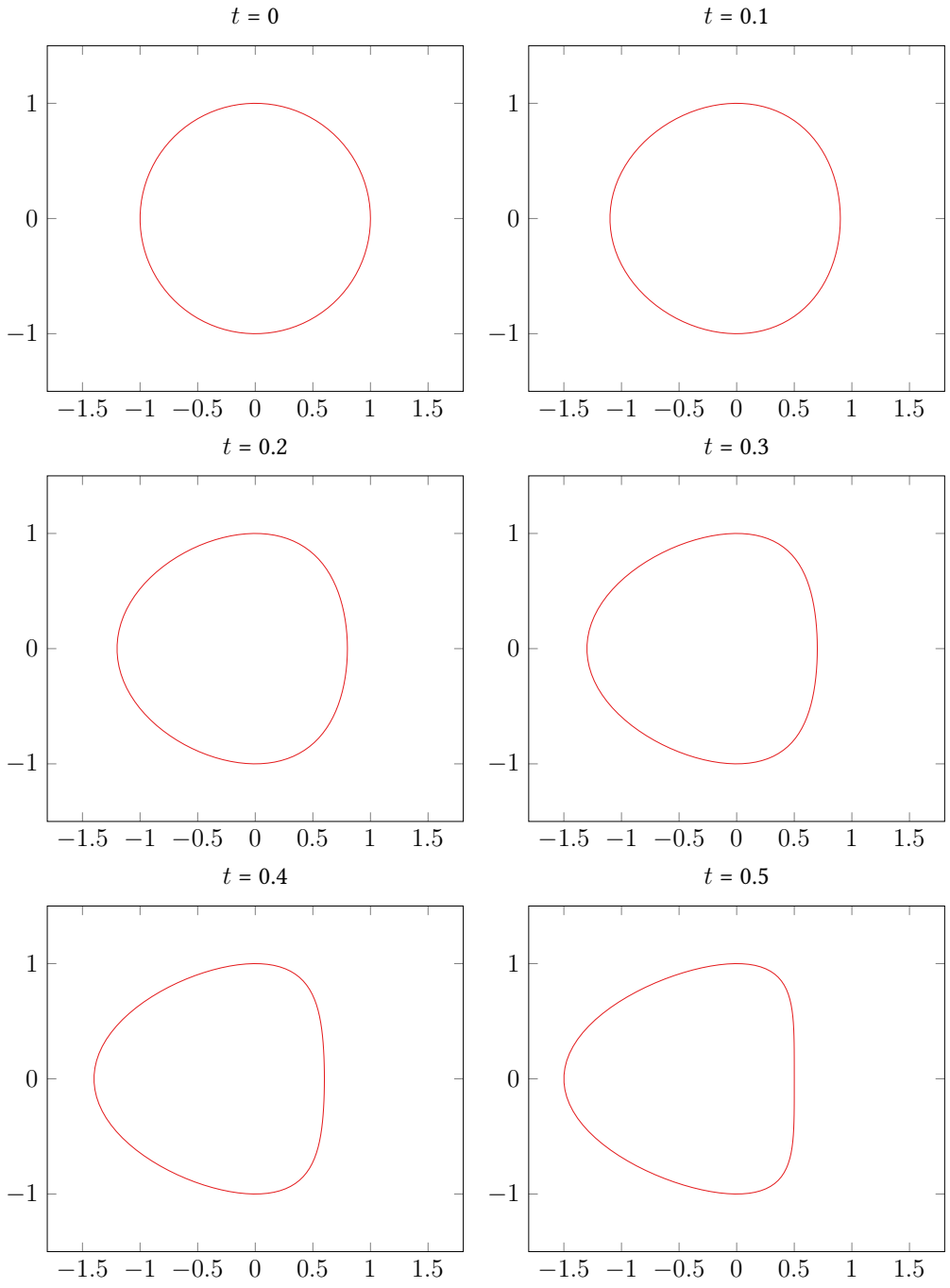


Figure 12: Kite geometry at different times.

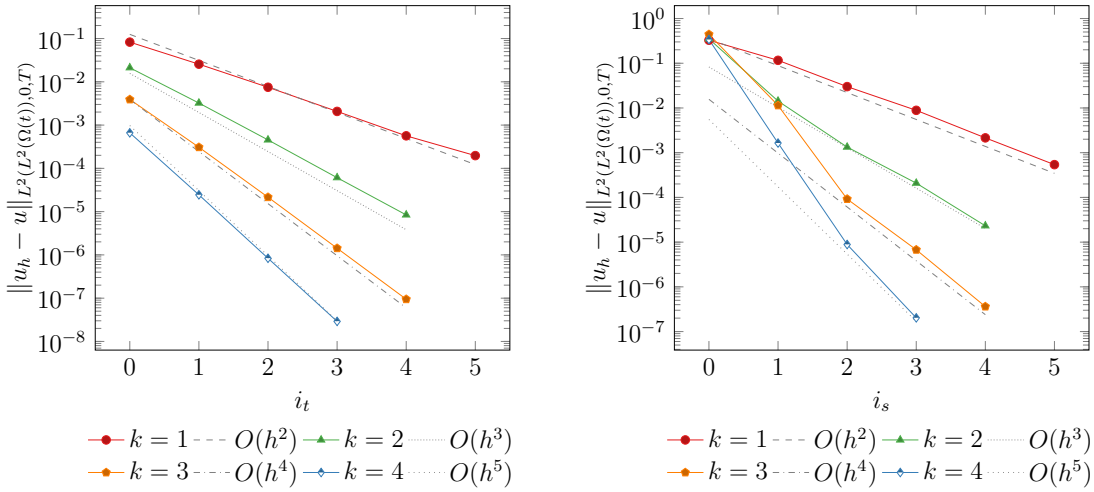


Figure 13: Convergence plots of the higher order DG method for the kite case. Presentation style as in Fig. 11.

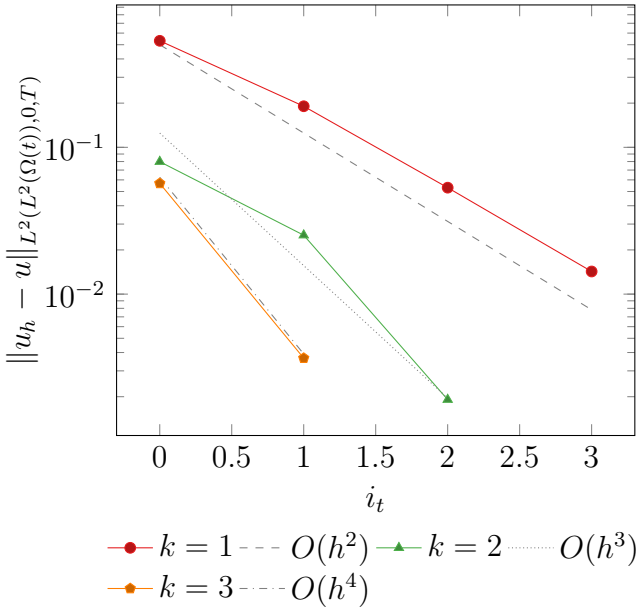


Figure 14: Convergence plots of the higher order DG method for the sphere case. Presentation style as in Fig. 11.

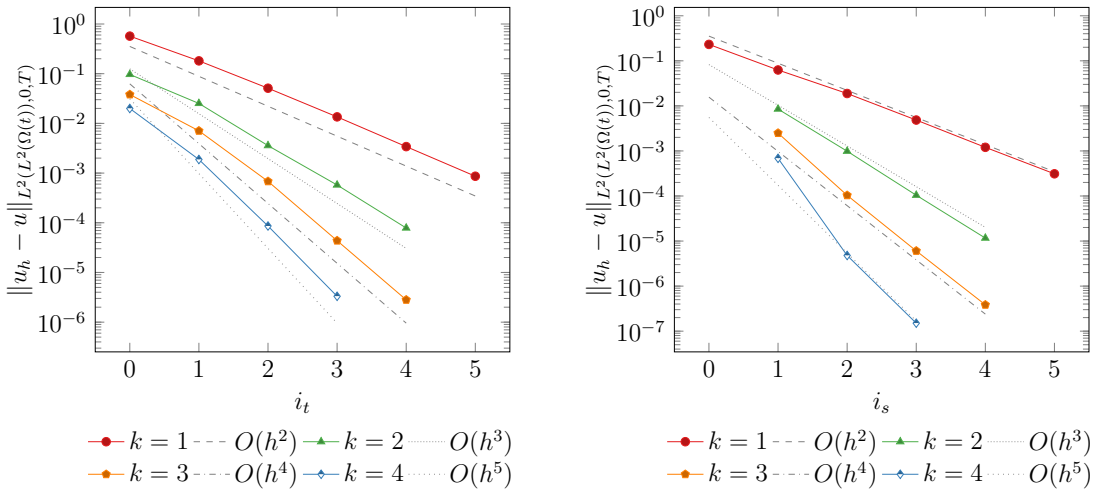


Figure 15: Convergence plots of the higher order CG method for the moving circle case. On the left-hand side, time refinements are performed for a sufficiently high i_s. On the right-hand side, space refinements are performed for a high i_t respectively. In each case, $k_t = k_s = k$.

4.2 Space-time CG

Next, we also want to investigate the Petrov-Galerkin variant of our method and compare it to the DG method numerically. To this end, we use the same moving circle setting as before. In particular, we generate the meshes by the same functions and also use the same notation for space and time refinement levels, i_s and i_t. In Fig. 15, results for spatial and time refinements for the method are displayed. By and large, the numerical errors are very similar and differ only slightly. Also, the general convergence behaviour for all investigated polynomial degrees remains the same. This is expected as there are only small differences concerning the Ghost penalty and the (dis)-continuity of the numerical solution along time slice boundaries. To stress this point, we also give a direct side-by-side comparison between the CG and the DG method for spatial refinements in Fig. 16.

Another interesting "norm" for comparing the CG and the DG method apart from numerical error is that of computational effort, or more specifically the number of non-zero entries of the matrix. That is why we also calculate these numbers for the numerical test cases of Fig. 16 and present it in Fig. 17.

The resulting numbers confirm the general expectation that the matrix size of the CG method would correspond to the matrix of the DG method of one polynomial order less in time. Accordingly, we see a considerable benefit from going to CG for low polynomial orders. However, for higher polynomial orders, the relative difference of "saving one order

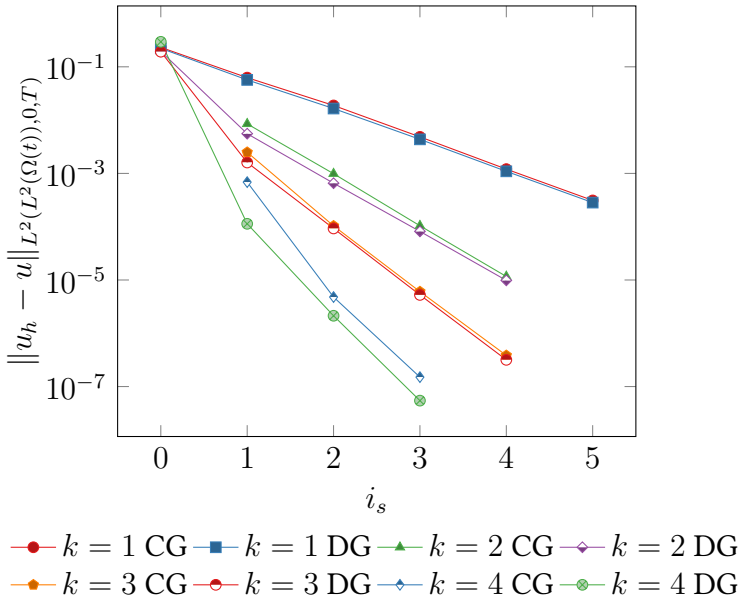


Figure 16: Direct comparison of the results for the CG and the DG method for the moving circle in terms of numerical error for space refinements.

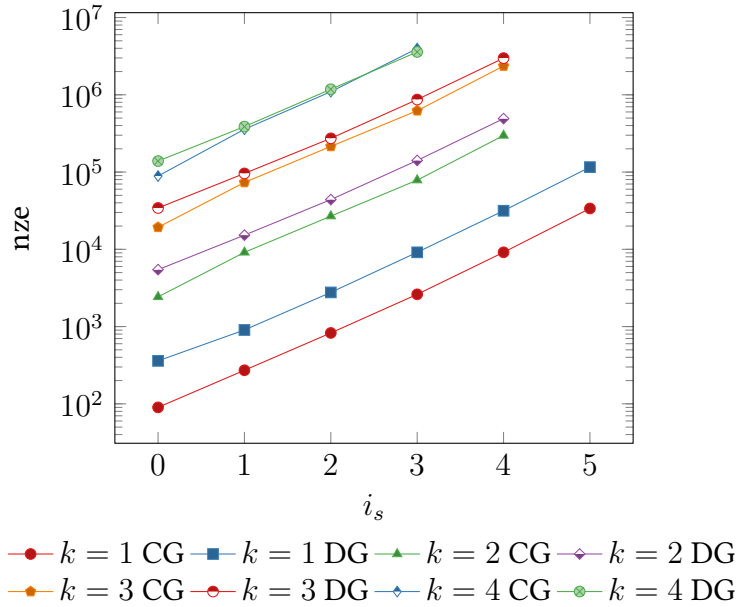


Figure 17: Direct comparison of the results for the CG and the DG method for the moving circle in terms of non-zero entries in the system matrix for space refinements.

in time" becomes less significant. We conclude that the CG method is interesting from a computational point of view in particular for low and moderate order application szenarios. One could also investigate non-matching pairs k_s, k_t in that regard in future research.

5 Conclusion and Outlook

We end this thesis by giving a summary and suggesting some directions for future research.

5.1 Summary

In this thesis, we extended results for Unfitted space-time Finite Element methods for the convection-diffusion problem in different directions. The first major analytical goal consisted in relaxing the assumption of an exact handling of the geometry in the analysis of [9]. In order to prepare for this analysis, we introduced the basic DG method of [9] and its variant under a weakened assumption on the handling of geometries carefully in Section 2. Afterwards, we presented a modified variant of the analysis in Section 3. In general, many proof techniques could be transfered to our case and most changes concern smaller Lemmata. We derived error bounds in a specific $|||\ldots|||$-norm which take the form of a sum of the terms in [9] and some new summands stemming from a Strang-type result. Numerical investigations confirm our theoretical results concerning the DG method. A second direction of generalisation was the introduction of a continuous Galerkin (CG) variant of the method. Its analysis remains work for future research, but we introduced the method and demonstrated its functionality numerically. In particular, the numerical errors differ only slightly from those of the DG method and the effect of a reduction of the system matrix size for small and moderate polynomial order was shown.

5.2 Directions for future research

The results of this thesis suggest several directions for future investigations:

Analysis

- The Continuous Galerkin method suggested in this thesis could be analysed along potentially similar lines as the DG methods presented in [9] and this text. Note that however, in the stability proof the issue appears that the time derivative is not in general contained in the according discrete space. As of now, we have not yet found a fix for this and leave it as a question for future research.

- The analysis of the DG method presented could be extended in the direction of an error bound in the "normal" Q-norm. That would be a measure of the error relating more naturally to applications of the numerical methods compared to the somewhat special $|||\dots|||$-norm.
- Concerning the isoparametric mapping, one could go one step further than our analysis and consider the discretisation in time in line with the presentation of 2.4.2. Roughly, we would expect terms involving the time discretisation in addition to the difference terms of our analysis here.
- Further work can be also done in an accurate definition and investigation of the technical aspects of the definition of the space-time version of the isoparametric mapping. To some extent, we circumvented such investigations by appropriate assumptions.
- In our estimates of the new Strang-type result, a factor $\sup_{q_h \in W_h} \frac{\|q_h\|_{Q_h}}{|||q_h|||_J}$ appeared. It would be desirable to bound this expression by an appropriate constant. We conjecture that doing so would be possible by e.g. imposing boundary conditions by a Nitsche-type approach or including a mass term in the problem.
- An analysis could be extended in the direction of the projection operator representing the `shifted_evaluate` function, c.f. the remark on Page 31.

Implementation

- Concerning implementation, one relatively straightforward next step would consist of investigating non-matching pair of polynomial orders for space and time, $k_s \neq k_t$.
- In addition, the error could be measured in more norms than just the $\|\dots\|_Q$-norm.
- Future investigations could also be made in the direction of three dimensional spatial problems. As of now, only very few results could be generated because of high memory consumption of the applied direct solvers. It might be interesting to try out iterative solvers to circumvent this issue.

Analysis and Implementation

- Finally, the methods of this thesis could be transfered to more complicated physical problems, such as a two phase flow problem. This concerns both analysis and implementation. For some first investigations in that direction see [9, Chapters 7,8].

References

[1] Erik Burman. Ghost penalty. *Comptes Rendus Mathematique*, 348(21-22):1217–1220, 2010.

[2] Erik Burman, Susanne Claus, Peter Hansbo, Mats G Larson, and André Massing. Cutfem: discretizing geometry and partial differential equations. *International Journal for Numerical Methods in Engineering*, 104(7):472–501, 2015.

[3] Fabian Heimann. Implementing cut space-time quadrature rules for three dimensions. report on the programming practial. April 12, 2019.

[4] Christoph Lehrenfeld. High order unfitted finite element methods on level set domains using isoparametric mappings. *Computer Methods in Applied Mechanics and Engineering*, 300:716 – 733, 2016.

[5] Christoph Lehrenfeld. Advanced finite element methods: Lecture notes (winter semester 2016/17). January 17, 2017.

[6] Christoph Lehrenfeld. Numerics of partial differential equations: Lecture notes (winter semester 2019/20 and summer semester 2020). July 10, 2020.

[7] Christoph Lehrenfeld and Maxim A. Olshanskii. An Eulerian finite element method for PDEs in time-dependent domains. *ESAIM: M2AN*, 53:585–614, 2019.

[8] Christoph Lehrenfeld and Arnold Reusken. Analysis of a high order unfitted finite element method for an elliptic interface problem. *IMA J. Numer. Anal.*, 38:1351–1387, 2018.

[9] Janosch Preuß. Higher order unfitted isoparametric space-time fem on moving domains. *Master's thesis, University of Gottingen*, 2018.

[10] Arnold Reusken. Analysis of an extended pressure finite element space for two-phase incompressible flows. *Computing and Visualization in Science*, 11(4):293–305, 2008.

Acknowledgements

First and foremost, I thank Christoph Lehrenfeld for his great support throughout the years, but also in particular within the supervision of this thesis. The very detailed and helpful discussions significantly shaped the structure of this thesis on a big scale, but also many technical aspects of the proofs.

Furthermore, I thank Prof. Gert Lube for agreeing to be the second assessor for this thesis.

Lastly, thanks to Alexander Osterkorn for providing some language improvement suggestions.