Todos after you're done with writing

1. Run a spell checker!
2. Check proper punctuation for all displayed equations!

Master Thesis

Solving Cahn-Hilliard equation using a cut finite element method
WT: A cut finite element method for the biharmonic problem and its application to the Cahn-Hilliard equation

Fix title

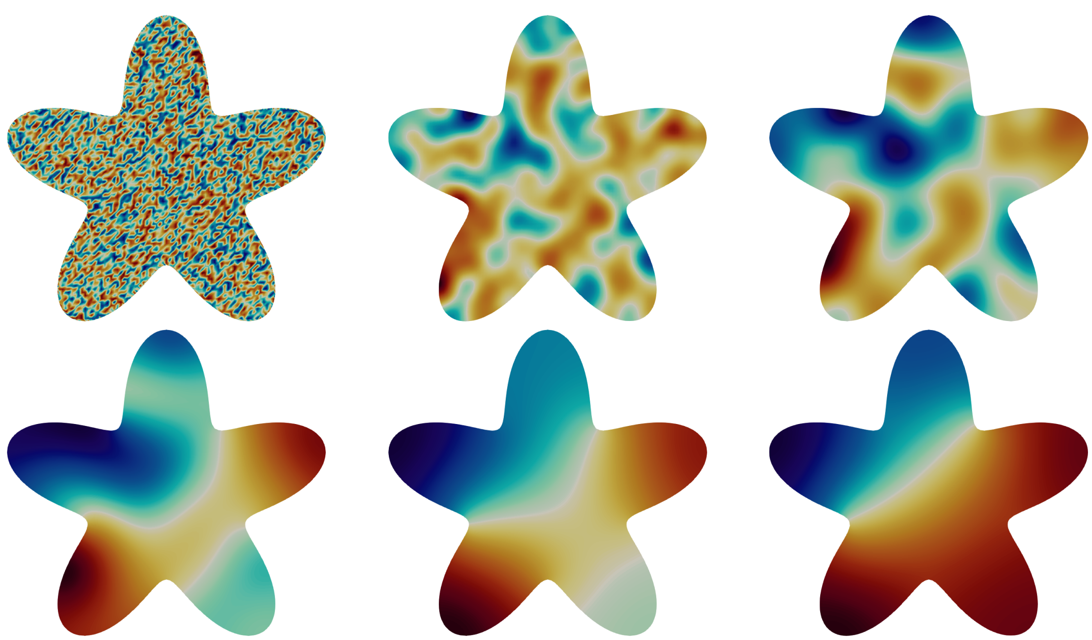
3. Check that bib is complete, no missing references
4. Scan for ?? to locate missing references

Isak Hammer

Supervisor: André Massing

July 6, 2023

5. Check spelling in Bibliography



Department of Mathematical Sciences
Norwegian University of Science and Technology

Contents

1 Introduction

The first application of the Cahn-Hilliard (CH) problem appeared when modelling phase separation of two-component incompressible fluids [1, 2, 3], but was quickly generalized to handle multi-component system as well [4, 5, 6, 7]. In engineering, CH is the critical component in the phase-field model, a mathematical framework to model transitions and interface dynamics in materials and fluid dynamics [8, 9]. From this, the equation found many interesting applications for a wide variety of problems. To mention a few, we have multiphase fluid dynamical problems [10, 11, 12, 13], solidification of binary or multi-component alloys [14, 15], and continuum modelling of fracture dynamics in materials [16, 17]. Perhaps an unexpected application is that CH can be used for in painting when recovering damaged parts of an image [18, 19, 4, 20] and modelling the origin of the irregular structure in Saturn's rings [21]. CH is also essential in many areas of biology and medicine. For example, from a macroscopic viewpoint, CH is a great tool to model tumour growth, wound healing and brain tumours [22, 23]. On the microscopic level on the biomembrane, there is an ongoing debate about the existence of the accumulation of lipids into so-called lipid rafts, which serve as a rigid platform for proteins with special properties such as signalling and intercellular trafficking [24, 25, 26, 27]. It turns out that the hypothesis can be tested by modelling the problem as a separation problem using CH [28, 29, 30].

write more objective, e.g. "CH has been used to model" That is then just stating a fact without coloring it by one's own opinion,

1.1 The physical Cahn Hilliard problem

The CH problem comes in many variants depending on its application, but we will in this report "focus" on the binary mixture version [7]. Let $\Omega \subset \mathbb{R}^d$ be a compact ~~set~~ domain for $d = 2, 3$ with a sufficiently smooth boundary Γ, see Figure 1. We define the time duration parameter $T \in [0, \infty)$ and the so-called unknown phase-field function as the mapping $u : [0, T] \times \Omega \to [-1, 1]$, which ~~is denoted as~~ denotes the local difference of a binary mixture of two concentrations $c_A, c_B \in [0, 1]$ ~~s.t.~~ i.e. $u = c_A - c_B$ and $c_A + c_B = 1$. ~~Remark~~ Note that if a local point exists so u ~~has~~ attains the extreme value $+1$, then it implies that the particular point has 100% concentration c_A and vice-versa for c_B. On the other hand, if u is zero, it implies that the mixture is 50% − 50%.

$+1 \Leftrightarrow c_A = 1, c_B = 0$

$-1 \Leftrightarrow c_B = 1, c_A = 0$

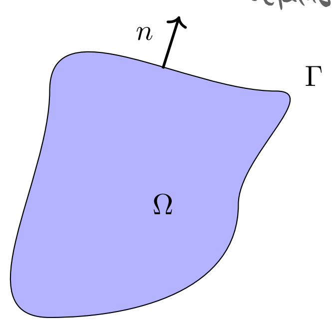


Figure 1: Illustration of the physical domain Ω for $d = 2$, the boundary Γ and the corresponding normal vector n.

Let $\varepsilon \ll 1$ be a small parameter. For an isotropic binary mixture non-uniform, the standard Ginzburg-Landay free energy functional is given by

$$E(u) = \int_\Omega \frac{\varepsilon}{2} |\nabla u|^2 + \frac{1}{\varepsilon} F(u) dx \,.$$

Punctuation!

The nonlinear function $F(u)$ ~~is denoted as~~ denotes the (Helmholtz) free energy density associated with the interaction dynamics between the components and thus comes in many forms depending on the thermodynamic properties, see [7]. However, we will in this article assume that $F(u) = (1/4)(u^2 - 1)^2$. ~~We choose to define~~ the chemical potential μ as the variational derivative, "is then defined"

no really "our choice"

$$\mu = \frac{\delta E(u)}{\delta u} = \frac{1}{\varepsilon} f(u) - \varepsilon \Delta u.$$

guarantee

where we used the notation $f(u) = F'(u) = u(u^2 - 1)$. First of all, to ~~require~~ local mass conservation, we may enforce the continuity equation; that is ,

$$\partial_t u + \nabla \cdot \mathcal{J} = 0,$$

where $\mathcal{J}$ denotes the flux governed by the physical dynamics. Hence, this naturally leads to the no-flux and the Neumann boundary conditions,

$$\mathcal{J} \cdot n = 0 \text{ on } \Gamma , \tag{1.1}$$
$$\partial_n u = 0 \text{ on } \Gamma . \tag{1.2}$$

A well-accepted law for the flux is that it is proportional to the chemical energy gradient, $\mathcal{J} = -M\nabla\mu$ for a parameter M. For the simplicity we assume $M = 1$. This implies we can rewrite the boundary condition (1.1) s.t. replace s.t. with "show that" everywhere. that

$$\mathcal{J} \cdot n = \partial_n \left(\varepsilon \Delta u - \frac{1}{\varepsilon} f(u) \right) = \varepsilon \partial_n \Delta u - \frac{1}{\varepsilon} f'(u) \partial_n u = \varepsilon \partial_n \Delta u,$$

for u evaluated on Γ. Here we used that $\partial_n f(u) = f'(u)\partial_n u = 0$ from the boundary condition (1.2). Hence, we now have an equivalent ~~a reasonable~~ set of boundary conditions and can finally write the strong form of the Cahn-Hilliard equation. Let $u(x, 0) = u_0$ then is the dynamics ~~on the form~~,

"of u is governed by"

$$\begin{aligned} \partial_t u + \Delta \left(\varepsilon \Delta u - \frac{1}{\varepsilon} f(u) \right) &= 0 \quad \text{in } \Omega \\ \partial_n u &= 0 \quad \text{on } \Gamma \\ \partial_n \Delta u &= 0 \quad \text{on } \Gamma \end{aligned} \tag{1.3}$$

In this report, we will refer to this formulation as the Physical CH formulation due to the physical characteristics it embodies. Based on these laws and the boundary conditions, it becomes evident that the energy functional serves as a Lyapunov function in the sense that its time derivative is monotonically decreasing and that the global mass concentration is conserved, i.e.

→ extreme case is constant stationary solution.

$$\frac{d}{dt} E(u) \leq 0 \text{ and } \frac{d}{dt} \int_\Omega u dx = 0. \tag{1.4}$$

Note

~~Remark~~ that the inequality computation utilizes the assumption of M to be constant, and both equations require the no-flux boundary condition, $\mathcal{J} \cdot n = 0$. For details, see [31, Equation 17] no space and [32, Equation 1.7]. This is useful since we expect $E(u(\cdot, t_2)) \leqslant E(u(\cdot, t_1))$ for $0 < t_1 < t_2$ and that the global mass is conserved,

$$\int_\Omega u(x, t) dx = \int_\Omega u_0(x) dx.$$

The properties serve as a theoretical foundation for establishing the existence, uniqueness, and long-term behaviour of the CH problem. Consequently, these properties are well-comprehended from a mathematical standpoint. For references, see [33, 34, 35].

1.2 Numerical methods

One of the key challenges with the CH problem is that it involves fourth-order spatial derivatives. It has, for simple domains, successfully been implemented using Finite Difference Methods [36, 37] and Spectral Methods [38, 39]. However, these methods are generally constrained to simple domains (with some notable exceptions [40, 41, 42]).

As a further evolution to address the CH problem, it is common to consider a corresponding biharmonic (BH) problem as a numerical testbed ~~in the spatial direction~~. This problem is defined as follows; Find $u : \Omega \to \mathbb{R}$ such that "for spatial discretization schemes".

$$
\begin{aligned}
\alpha u + \Delta^2 u &= f(x) \quad \text{in } \Omega, \\
\partial_n u &= g_1(x) \quad \text{on } \Gamma, \\
\partial_n \Delta u &= g_2(x) \quad \text{on } \Gamma,
\end{aligned}
\tag{1.5}
$$

for given

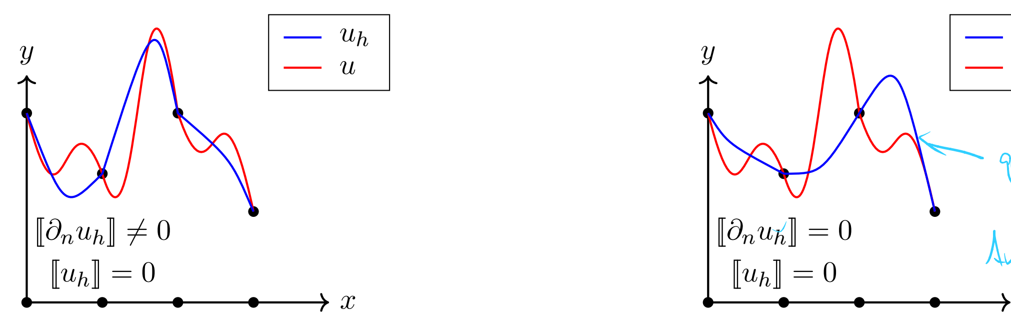
~~Here is the~~ functions ~~a mapping~~ $f, g_1, g_2 : \Omega \to \mathbb{R}$ and ~~the~~ constant α ~~is a real number s.t. α~~ > 0. The BH problem holds relevance since it is provide a proper spatial ~~integration~~ discretization test framework prior to moving on solving the non linearities and time integration.

The early Finite Element Methods (FEM) for CH were proposed in [43, 35] utilizing global C^1 and C^2 in one spatial dimension, but later it has been shown that making C^1 (or higher order) elements in multiple space dimensions ~~are~~ is far from being trivial. For reference, see [44, 45, 46].

Move to unfitted paragraph

There exist several promising alternative methods that guarantee C^1 continuity, and these have shown potential for solving the CH problem. A notable mention is isogeometric analysis (IGA), a technique that leverages Non-Uniform Rational B-Splines (NURBS) to efficiently handle complex geometries and smooth boundaries without the need ~~for mesh refinement.~~ for translating CAD-based geometry descriptions into classical FEM meshes. Thus, IGA ~~makes~~ presents a desirable alternative for problems dealing with intricate and smooth domains [47]. Specifically for the CH problem, has IGA successfully been implemented [48, 49]. Recent results have shown that investigations on unfitted versions of IGA [50] and its applicability to moving surfaces [51] also is possible. Another rising method is the virtual finite element method (VFEM), which has applied so-called virtual C^1 elements to handle the continuity requirement [52].

Skip   remove red "is" C^1



quite a bad approximation...? Are you sure that is the correct one?

Figure 2: Illustration of global C^0 continuous elements (left) and global C^1 elements (right) in 1 dimension. Here is u the exact solution and u_h the approximated solution. We define the jump between the elements as $[\![u_h]\!] = u_+ - u_-$. State what kind of elements (P^2) you have use for left and right example

An alternative approach is to avoid global C^1 continuity "and" weakend it to global C^0 continuity, see Figure 2 . As a result, this strategy has led to the development of two distinct families of methods for solving the CH problem. The first involves the Continuous Interior Penalty (CIP) methods, which uses the standard weak formulation but penalizes the discontinuity of the derivative between elements as a form of regularization. The method has been designed for several interesting stable variants for the BH problem, that is [53, 54, 55, 56, 57], and recently also for tri-harmonic problems [58]. This method is advantageous due to its symmetry and relationship with discontinuous Galerkin (DG) methods [59], renowned for their natural way of handling inhomogeneous boundary conditions, flexibility with unstructured meshes, efficient parallelization, and strong stability. This connection lends robust stability analysis tools, making the method highly suitable for intricate computational problems. The CIP formulation has also then been adapted to solve CH by applying the Newton-Raphson scheme to handle the non-linearities [60] or utilizing an implicit-explicit

(IMEX) time integration scheme, where the stiff part is treated implicitly (such as backward Euler) and the nonlinear part explicitly (such as the forward Euler or explicit Runge-Kutta) [61].

Another popular variant is to rewrite the BH problem as a system of second-order problems ~~in~~ leading to a mixed formulation. This strategy not only broadens the problem's flexibility but also provides a more natural means to incorporate boundary conditions, see [62, 63, 64, 65]. This approach also leverages the general saddle point theory for mixed FEM methods, which provide a mathematical framework to ensure numerical stability [66]. Moreover, this approach adapts well to the CH problem [60, 67, 68], and some methods even apply a so-called convex splitting scheme approach in a way that preserves the convexity of the energy functional, making the system easier to solve [69, 70]. A combination of these methods, the DG and mixed formulation for the CH problem, has recently been considered [71, 72].

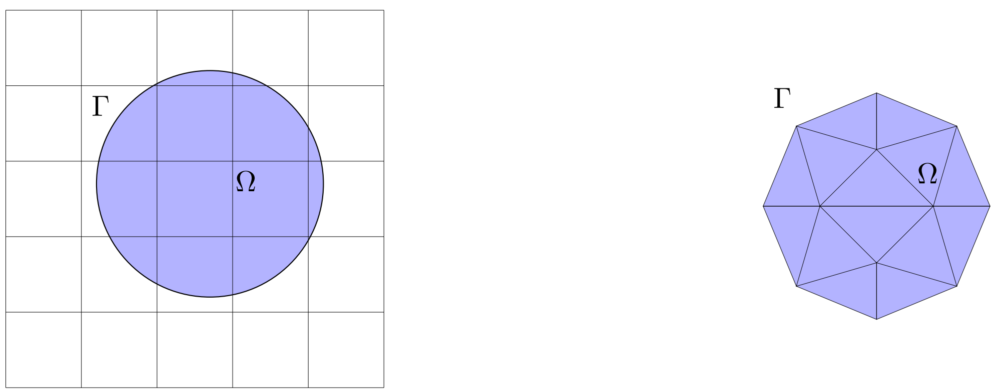


Figure 3: Mesh comparison: unfitted mesh (left) adheres to domain and boundary, while fitted mesh (right) employs a triangular mesh for polygonal approximation of the circular domain.

Break sentence in 2 here

two three dimension

Creating a high-quality mesh in 2- and 3-dimensional for realistic problems is a challenging task that can take a reasonable time in the simulation workflow and is hard to scale properly on distributed platforms and thus not that suitable for moving domains, very complex meshes or smooth boundaries. An interesting class to approach the problem is the so-called unfitted finite element method, which utilizes a background mesh and does not align with the physical boundary. For an illustration see Figure 3. This greatly reduces the need to generate an unstructured mesh and makes it very applicable for parallelization and moving domains since it avoids the need of remeshing entirely. However, without paying attention to the so-called cut cells, which are the elements intersecting with the boundary, the method quickly leads to instability and ill-conditioning. One of the methods to counter this is the cut finite element method (CutFEM), where the focus is to penalize the cut-cells weakly by adding an additional ghost penalty term to ensure ~~stability~~ ~~and~~ well-posedness [73]. Notably, there exist related methods, sometimes considered equivalent, named Extended FEM and Trace FEM [74, 75]. This has been successfully implemented for the BH problem for the mixed formulation [76] and the CIP formulation [77, 74]. However, both implementations are considering an interface problem between two domains. Specifically for the CH problem, the mixed formulation [78] has been shown to be successful. Aggregated unfitted finite element method (AgFEM) is a close relative to CutFEM and has also shown to be promising [79, 80]. The method is an alternative way to the ghost penalty, which instead applies a so-called cell aggregation with respect to a cut cell (assuming each cell has enough support with interior elements) and, thus, the badly cut cells are removed, ensuring robustness and well-posedness.

Extend a bit

well-posedness includes stability? Remove "optimal convergence properties"

1.3 Outline of the report

with Cahn-Hilliard type b.c.

In this article, we propose an novel stabilized unfitted cut continuous interior penalty method (CutCIP) specifically for the BH problem, which incorporates the CutFEM methodology in combination with a CIP formulation. ~~We will follow~~ ~~the~~ work theoretical procedure as presented in

"Our approach is inspired by the"

Bogner-Fox-Schmit elements.

① • Add a couple of references where CutFEM-based discretizations are considered, see Teams Chat.

• Make sure that you explain the key difference between previous work and your work:

⊕ you consider a CutFEM or ghost penalty stabilized unfitted CIP formulation of the BH with Cahn-Hilliard b.c.

⊕ you apply form to C-H giving the first C^0/CIP/CutFEM solver of C-H system.

"structure of" "into" "and extending the"

the DG Poisson formulation proposed by [81], but instead apply the CIP BH formulation while taking account ~~to~~ analytical results provided in [61, 54]. Mention Sect 2 as well.

The ~~main results in~~ this thesis is as follows. In Section 3 the basic construction of the CIP BH formulations and the related properties is ~~established~~. Subsequently, in Section 4 we propose the corresponding CutCIP BH method, ~~but also~~ and provide a complete theoretical ~~proof~~ analysis to ~~show~~ showing that the stability and convergence properties from the original CIP method are conserved. In Section ?? we ~~briefly~~ demonstrate "how to" extend the method to handle the CH problem. Finally, Section 4.6 is dedicated to numerical experiments.

"reviewed" Finally,

mention numerical results.

Mention that theoretical results are corroborated with detailed numerical evidence studies.

2 Mathematical Background

In this section, we revisit the established definitions and provide a brief overview of Sobolev spaces. Afterward, we briefly review the finite element method and discuss the necessary tools required for calculating a priori estimates.

Mention main sources you extracted the material from.

2.1 Sobolev spaces

We will in this report assume Ω to be a compact and open set in $\mathbb{R}^d$. Let $p \in \mathbb{R}$, $1 \leqslant p \leqslant \infty$, and define the space $L^p(\Omega)$ to be the set of all measurable functions $u : \Omega \mapsto \mathbb{R}$ such that $|f|^p$ is Lebesgue integrable, i.e,

$$L^p(\Omega) = \left\{ u : \Omega \mapsto \mathbb{R} \mid \int_\Omega |u|^p \, d\Omega < \infty \right\}.$$

Let $u \in L^p(\Omega)$. We define the integral norm of order p to be

$$\|u\|_{L^p(\Omega)} = \left(\int_\Omega |u|^p \, dx \right)^{\frac{1}{p}}.$$

The following definition for derivatives is employed. For d dimensions of order k we define the multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ with the absolute value $|\alpha| = \sum_{i=1}^d \alpha_i = k$ such that

$$\partial^\alpha u = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} u \quad \text{for } u \in C^{|\alpha|}(\Omega) \tag{2.1}$$

Let $k \geqslant 0$ be an integer and let $1 \leqslant p < \infty$ be a real number, then the Sobolev space $W^{k,p}(\Omega)$ is defined by

$$W^{k,p}(\Omega) = \{ u \in L^p(\Omega) \mid \partial^\alpha u \in L^p(\Omega) \, \forall \alpha : |\alpha| \leqslant k \}. \tag{2.2}$$

with the corresponding norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{j=0}^{k} |u|^p_{W^{j,p}(\Omega)} \right)^{\frac{1}{p}}. \tag{2.3}$$

Here the seminorm is defined such that $|u|_{W^{k,p}(\Omega)} = (\sum_{|\alpha|=k} \|\partial^\alpha u\|^p_{L^p(\Omega)})^{\frac{1}{p}}$.

For shorthand notation, we denote

$$\begin{aligned} \|\cdot\|_{k,p} &= \|\cdot\|_{k,p,\Omega} = \|\cdot\|_{W^{k,p}(\Omega)}, \\ |\cdot|_{k,p} &= |\cdot|_{k,p,\Omega} = |\cdot|_{W^{k,p}(\Omega)}, \end{aligned} \tag{2.4}$$

Since $p = 2$ is frequently used in this report, we also define for convenience a compact notation $\|u\|_\Omega = \|u\|_{L^2(\Omega)}$ and $H^m(\Omega) = W^{m,2}(\Omega)$ such that,

$$\begin{aligned} \|\cdot\|_{H^m(\Omega)} &= \|\cdot\|_{k,\Omega} = \|\cdot\|_{k,2,\Omega}, \\ |\cdot|_{H^m(\Omega)} &= |\cdot|_{k,\Omega} = |\cdot|_{k,2,\Omega}. \end{aligned} \tag{2.5}$$

Recall that $L^2(\Omega)$ and $H^m(\Omega)$ are Hilbert spaces equipped with the inner products

$$
\begin{aligned}
(u,v)_\Omega &= (u,v)_{L^2(\Omega)} = \int_\Omega uv dx, \quad \forall u,v \in L^2(\Omega) \\
(u,v)_{m,\Omega} &= \sum_{|\alpha|\leqslant m} (\partial^\alpha u, \partial^\alpha v)_\Omega, \quad \forall u,v \in H^m(\Omega).
\end{aligned}
\tag{2.6}
$$

let $u, v \in H^2(\Omega)$ be scalar functions. The following are the definition for their corresponding operator inner products.

$$
\begin{aligned}
(\nabla v, \nabla u)_\Omega &= \int_\Omega \nabla v \cdot \nabla u \ dx \\
(\Delta v, \Delta u)_\Omega &= \int_\Omega \Delta v \Delta u \ dx \\
(D^2 v, D^2 u)_\Omega &= \int_\Omega D^2 v : D^2 u \ dx
\end{aligned}
\tag{2.7}
$$

Also, $\|\nabla v\|_\Omega^2 = (\nabla v, \nabla v)_\Omega$, $\|\Delta v\|_\Omega^2 = (\Delta v, \Delta v)_\Omega$, and $\|D^2 v\|_\Omega^2 = (D^2 v, D^2 v)_\Omega$. Here, $\nabla v \cdot \nabla u$ and $D^2 v : D^2 u$ represent the inner product of the gradients and the Frobenius inner product [1] of Hessian matrices, respectively.

Let $v \in H^r(\Omega)$, we define $D^r v$ to be a tensor of order r such that

$$
[D^r v]_{i_1 \dots i_r} = \frac{\partial^r v}{\partial x_{i_1} \dots \partial x_{i_r}} \quad \forall i_1, \dots, i_r \in \{1, \dots, d\} \tag{2.8}
$$

where the norm $\|D^r v\|_\Omega^2 = \int_\Omega D^r v : D^r v \ dx$ is defined via the standard Frobenius inner product. Observe that this notations holds such that $D^0 v = v$, $D^1 v = \nabla v$ and $D^2 v = J(\nabla v) = \mathrm{Hess}(v)$, where J is the Jacobian operator .

Given the context of Sobolev spaces, we consider the functions $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$. We can denote Greens theorem, which links integrals over a volume and its boundary, as follows,

$$
(\Delta u, v)_\Omega = -(\nabla u, \nabla v)_\Omega + (u, \partial_n v)_\Gamma
$$

This identity serves as an essential tool for the calculations done.

2.2 Computational Domains

Assume that $\Omega \subset \mathbb{R}^d$ is an open and bounded domain with a boundary Γ. In standard FEM methods a key assumption is that the set Ω is a polyhedra. This is useful since a polyhedra can be fully covered by a collection of polyhedra and, hence, motivating us to define a fitted mesh. We define a fitted mesh $\mathcal{T}$ of the domain Ω to be a collection of closed polyhedra $\{T\}$ with disjoint interior forming a partition of Ω such that $\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$, for illustration see Figure 3. Here we say that each $T \in \mathcal{T}$ is a mesh element or an element. The mesh size is defined as the maximum diameter $h := h_{max}$ of any polyhedra in the mesh $\mathcal{T} = \{T\}$, that is, $h_{max} = \max_{T \in \mathcal{T}} h_T$ s.t. $h_T = \mathrm{diam}\,(T) = \max_{x_1, x_2 \in T} \mathrm{dist}(x_1, x_2)$ Hence, motivating us to use the notation $\mathcal{T}_h$ for a mesh $\mathcal{T}$ with size h.

For simplicity we restrict ourself to simplicial and quadrilateral elements. A mesh $\mathcal{T}_h$ in $\mathbb{R}^d$ is said to be matching if for all neighbouring elements $T_1, T_2 \in \mathcal{T}_h$ such that the intersection is non-empty, $T_1 \cap T_2 \neq \emptyset$, then $T_1 \cap T_2$ is for $d = 2$ either a common vertex, edge, and for $d = 3$ a common vertex, edge or a face.

[1] The Frobenius inner product for for two tensors $A, B \in \mathbb{R}^{\overbrace{n \times n \times \dots \times n}^{r \text{ times}}}$ is defined such that $A : B = \sum_{1 \leqslant i_1, \dots, i_r \leqslant d} A_{i_1 \dots i_r} B_{i_1 \dots i_r}$.

Let the chunkiness parameter $c_T := h_T/r_T$, where r_T is the largest ball that be inscribed inside a element $T \in \mathcal{T}_h$. A mesh is said to be shape regular if $c_T \leqslant c$ is independent of T and h. We also say that the mesh is quasi-uniform only if it is shape regular and $h_{\max} \leqslant c h_{\min}$. For a more complete description of meshes, see [82, Chapter 8].

In this thesis will we assume that a mesh $\mathcal{T}_h$ is matching, shape regular and quasi-uniform unless specified. The fact that the mesh is conform makes is a useful property since the interface between mesh elements has come into contact in the sense that it is either a vertex or a facet. This with the combination of shape regularity and quasi-uniformity is a major key to prove important inequalities in broken Sobolev spaces [83, Chapter 1.4.1]. Hence, the assumptions are very handy when proving convergence.

Let $\mathcal{T}_h = \{T\}$ be a mesh of $\Omega \subset \mathbb{R}^d$ consisting of polygons $T \in \mathbb{R}^d$. The set of all facets is the union of external and internal facets, $\mathcal{F}_h = \mathcal{F}_h^{ext} \cup \mathcal{F}_h^{int}$, where each are defined by

$$\mathcal{F}_h^{int} = \left\{F = T^+ \cap T^- \mid T^+, T^- \in \mathcal{T}_h\right\} \text{ and } \mathcal{F}_h^{ext} = \left\{F = \partial T \cap \partial\Omega \mid T \in \mathcal{T}_h\right\}.$$

Assume $T^+ \neq T^-$. Next, we define the following normal vectors.

1) We define $n = n_{\partial T}$ to be unit outward normal on ∂T for each $T \in \mathcal{T}_h$

2) Let $F \in \mathcal{F}_h^{int}$. We define n to be the facet normal $n = n_F = n|_{\partial T^+}$ from T^+ to T^-, illustrated in Figure 4.

3) Let $F \in \mathcal{F}_h^{ext}$. Then we define the facet normal $n|_F = n|_{\partial T}$ to be the unit outward normal.

Please note that we for convenience employ the notation n when it is clear what entity the normal is associated with.

Let $v \in L^2(\Omega)$ be a scalar function on Ω with a corresponding shape regular and quasi-uniform mesh $\mathcal{T}_h$. We will use the following definitions.

1) Let $F \in \mathcal{F}_h^{int}$ and $v^\pm|_F = \lim_{t\to 0^+} v(x \mp tn)$ for $x \in F$. We define the mean as $\{\!\{v\}\!\}\,|_F = \frac{1}{2}(v_F^+ + v_F^-)$ and the jump as $[\![v]\!]\,|_F = v_F^+ - v_F^-$.

2) Let $F \in \mathcal{F}_h^{ext}$ and let $v(x) = v(x)|_F$ for $x \in F$. We define the mean as $\{\!\{v\}\!\}\,|_F = v$ and the jump as $[\![v]\!]\,|_F = v$.

To simplify will we use the notation $\{\!\{v\}\!\} = \{\!\{v\}\!\}\,|_F$ and $[\![v]\!] = [\![v]\!]\,|_F$ for all $F \in \mathcal{F}_h$. Remark that if we have two functions u, v, for which $u^\pm(x)$ and $v^\pm(x)$ are defined, then the following identity holds $[\![uv]\!] = [\![u]\!]\,\{\!\{v\}\!\} + \{\!\{u\}\!\}\,[\![v]\!]$ along all facets $\mathcal{F}_h$ associated with the triangulation $\mathcal{T}_h$.

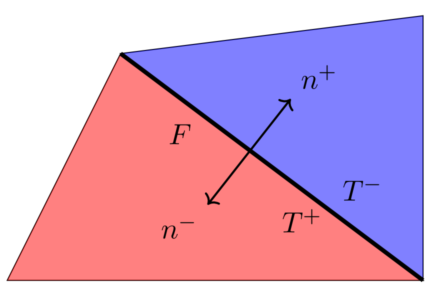


Figure 4: Facet $F \in \mathcal{F}_h^{int}$ shared by the triangles $T^+, T^- \in \mathcal{T}_h$ and the normal unit vector n^+ and n^-. If we pick $T = T^+$ and want to evaluate the normal vector n along a facet F, then we define $n = n\,|_F = n^+$.

2.3 Broken Sobolev spaces

In this work will we compute norms on discontinuous elements, thus, it will be necessary to define broken Sobolev spaces. Let $\mathcal{T}_h$ be a mesh and some integer $m \leqslant n$. Then we define the broken Sobolev space to be

$$\begin{aligned} H^m(\mathcal{T}_h) &:= \{v \in L^2(\Omega) \mid \; v|_T \in H^m(T) \quad \forall T \in \mathcal{T}_h\} \\ L^2(\mathcal{F}_h) &:= \{v \in L^2(\mathcal{T}_h) \mid \; v|_F \in L^2(F) \quad \forall F \in \mathcal{F}_h\}. \end{aligned}$$

This motivates us to define broken Sobolev norms and inner products using summation over mesh elements,

$$\|v\|^2_{H^m(\mathcal{T}_h)} = \sum_{T\in\mathcal{T}_h} \|v\|^2_{H^m(T)} \quad \text{and} \quad (v,w)_{m,\mathcal{T}_h} = \sum_{T\in\mathcal{T}_h} (v,w)_{m,T}.$$

As before, we use the shorthand notation, $\|v\|_{\mathcal{T}_h} = \|v\|_{L^2(\mathcal{T}_h)}$ and $(v,w)_{\mathcal{T}_h} = (v,w)_{L^2(\mathcal{T}_h)}$. That is,

$$\|v\|^2_{L^2(\mathcal{F}_h)} = \sum_{F\in\mathcal{F}_h} \|v\|^2_{L^2(F)} \quad \text{and} \quad (v,w)_{L^2(\mathcal{F}_h)} = \sum_{T\in\mathcal{F}_h} (v,w)_{L^2(F)}.$$

Again, we often use the more compact notation $\|v\|_{\mathcal{F}_h} = \|v\|_{L^2(\mathcal{F}_h)}$ and $(v,w)_{\mathcal{F}_h} = (v,w)_{L^2(\mathcal{F}_h)}$. Often it is needed to integrate over boundaries of elements $\partial\mathcal{T}_h = \{\partial T \mid \forall T \in \mathcal{T}_h\}$, hence, we also denote the notation $\|\cdot\|_{\partial\mathcal{T}_h} = \sum_{T\in\mathcal{T}_h} \|\cdot\|_{\partial T} = \sum_{T\in\mathcal{T}_h}\sum_{F\in\partial T} \|\cdot\|_F$.

A very useful lemma when working with estimates on broken Sobolev spaces is that a if a function is continuous, then the jump between the mesh elements is zero. Keep in mind a function $v \in H^1(\mathcal{T}_h)$ belongs to $H^1(\Omega)$ if and only if $[\![v]\!] = 0 \; \forall F \in \mathcal{F}_h^{int}$, see [83, Lemma 1.23].

Recall, the "$\lesssim$" symbol denotes an inequality up to a constant factor. That is, given $a, b > 0$, the statement $a \lesssim b$ is true if there exists a constant $C > 0$ such that $a \leq Cb$. Generally, this constant will contain information related to the properties of the mesh, such as shape regularity and quasi-uniformity, but it often also includes the maximum finite number or measure of a quantity. For instance, let $w, v_i \in L^2(\Omega)$, $a_i \in \mathbb{R}$ for $i = 1, \ldots, N$ and $\|w\|_\Omega = \|\sum_i^N a_i v_i\|_\Omega$, then is $\|w\|_\Omega \lesssim \sum_i^N \|v_i\|_\Omega$. To maintain clarity and avoid unnecessary complexity, we will not delve into this particular detail related to the constant.

We can express several general useful basic inequalities and estimates.

(i) A fundamental property of the inner-product the so-called Cauchy-Schwarz inequality

$$(u,v)_{m,\Omega} \leqslant \|u\|_{m,\Omega}\|v\|_{m,\Omega} \quad \forall u, v \in H^m(\Omega). \tag{2.9}$$

(ii) Let $v \in H^m(\Omega)$, then is

$$\|D^m v\|_\Omega \lesssim \|v\|_{m,\Omega}. \tag{2.10}$$

Similarly is $\|\Delta v\|_\Omega \lesssim \|v\|_{2,\Omega}$.

(iii) For all $u \in L^2(\mathcal{T}_h)$ we have,

$$\begin{aligned} \|\,[\![u]\!]\,\|_{\mathcal{F}_h} &\leqslant \|u^+\|_{\mathcal{F}_h} + \|u^-\|_{\mathcal{F}_h} \lesssim \|u\|^2_{\partial\mathcal{T}_h}, \\ \|\,\{\!\{u\}\!\}\,\|_{\mathcal{F}_h} &\leqslant \|u^+\|_{\mathcal{F}_h} + \|u^-\|_{\mathcal{F}_h} \lesssim \|u\|^2_{\partial\mathcal{T}_h}. \end{aligned} \tag{2.11}$$

(iv) For any $a, b > 0$ the well known Young's ε-inequality is on the form,

$$2ab \leqslant \varepsilon a^2 + \frac{1}{\varepsilon} b^2. \tag{2.12}$$

2.4 Lax-Milgram lemma

The intention is to introduce a abstract framework which can handle various types of partial differential equations (PDE). Let $\mathcal{A} : X \to Y$ be a abstract linear operator encoding the structure of any linear PDE, including boundary conditions and X, Y are spaces of functions. Then we denote the abstract strong formulation as the equation

for a given

$$\mathcal{A}u = f \tag{2.13}$$

~~Where the~~ function $f : \Omega \subset \mathbb{R}^d \mapsto \mathbb{R}$. We assume that the function $u : \Omega \to \mathbb{R}$ satisfies the relation (2.13) pointwise so that $\mathcal{A}u(x) = f(x) \forall x \in \Omega$. We will discover that Sobolev spaces are specifically engineered to study these kinds of problems.

Definition 2.1 (Linear bounded functional)**.** *Let V be a Hilbert space. Furthermore, we define the dual space V' to be the space of linear and bounded functionals $F : V \mapsto \mathbb{R}$, i.e.,*

$$V' = \{F : V \to \mathbb{R} \mid F \text{ is linear and bounded}\}$$

Problem 2.2 (Abstract linear problem)**.** *Assume X and Y to be two Hilbert spaces. Let the vector space $\mathcal{L}(X, Y)$ be all linear bounded operators spanned from V to Y. We define the abstract linear problem as follows; find $u \in V$ s.t.*

$$a(u, v) = l(v) := \langle f, v \rangle_{V',V} \, \forall v \in V$$

Where $a \in \mathcal{L}(V, V, \mathbb{R})$ is a bounded bilinear form and $f \in V' := \mathcal{L}(V, \mathbb{R})$ is a bounded linear form associated with the abstract strong formulation (2.13). Here we denote by $\langle \cdot, \cdot \rangle_{V',V}$ the duality pairing between V' and V.

Definition 2.3 (Coercivity and Boundedness)**.** *Let V be a Hilbert space and let $a(\cdot, \cdot) \in \mathcal{L}(V, V, \mathbb{R})$. Recall that the bilinear form $a(\cdot, \cdot)$ is coercive if*

$$a(v, v) \gtrsim \|v\|_V \quad \forall v \in V.$$

The bilinear form $a(\cdot, \cdot)$ is said to bounded if

v, w   v   w

$$a(u, v) \lesssim \|u\|_V \|v\|_V \quad \forall u, v \in V.$$

Lemma 2.4 (Lax-Milgram)**.** *The abstract linear problem 2.2 is well-posed if $a(\cdot, \cdot)$ is bounded and coercive. Moreover, the following a priori estimate holds true.*

$$\|v\|_V \lesssim \|f\|_{V'}$$

use u, u is the name of the solution

Proof. The problem can easily be proved using a special case of the Banach–Nečas–Babuška theorem. See [83, Lemma 1.4] □

2.5 Finite element method

point out that we consider non-conform

The finite element method (FEM) is a numerical method to solve partial differential equation by finding an approximation of the Problem 2.2. Let V_h be a finite-dimensional (polynomial) approximation space on the mesh $\mathcal{T}_h$. We say that a method is conform if $V_h \subset V$ and non-conform if $V_h \not\subset V$. We define the approximate problem as follows.

Problem 2.5 (The approximate problem)**.** *Let $V_h \not\subset V$ a non-conform finite-dimensional space. Find $u_h \in V_h$ such that,*

$$a_h(u_h, v) = l_h(v) := \langle f, v \rangle \, \forall v \in V_h.$$

We denote the functional $a_h : V_h \times V_h \to \mathbb{R}$ as an consistent approximation of $a : V \times V \to \mathbb{R}$, and similarly for the right-hand side $l_h : \times_h \to \mathbb{R}$ as an approximation of $l : V \to \mathbb{R}$.

non-conform ⇒ Bnd and Coercivity are not inherited automatically

⇒ Require - $\|\cdot\|_{V_h}$ ~ discrete weak problem and how you need require discrete bounded and coercivity.

Definition 2.6 (Local polynomial space)**.** *Let T be a element in a mesh $\mathcal{T}_h$, $x = [x_1, \dots, x_d]$ be a vector, and $\alpha = [\alpha_1, \dots, \alpha_d] \in \mathbb{N}^d$ be a multi index. The local polynomial space $\mathcal{P}^k(T)$ for a simplex is denoted as*

$$\mathcal{P}^k(T) = \operatorname{span}\{x^\alpha \text{ for } x \in T \text{ and } 0 \leqslant \alpha_i \leqslant k\}. \tag{2.14}$$

where x^α is a monomial such that $x^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d}$.

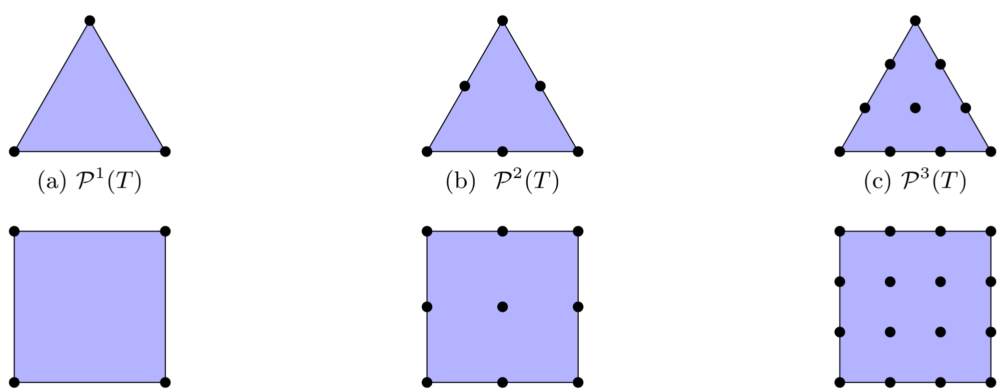
Let T be a cuboid, i.e., $T = \prod_{i=1}^d [z_i^-, z_i^+]$ where $z_i^- < z_i^+$ for $z_i^\pm \in \mathbb{R}$. Then the polynomial space $\mathcal{Q}^k(T)$ in $\mathbb{R}^d$ is defined as the tensor product of 1-dimensional finite elements, i.e.,

$$\mathcal{Q}^k(T) := \mathcal{P}^k([z_1^-, z_1^+]) \otimes \dots \otimes \mathcal{P}^k([z_d^-, z_d^+])$$

For more information about the local polynomial spaces, see [82, Chapter 6.4, 7.3]

Following Ciarlet [84, pp.93], the abstract definition of a finite element is defined as the triplet $(T, \mathcal{P}, \Sigma)$. In our case, T represents either a simplex or a quadrilateral geometry, and $\mathcal{P}$ denotes a finite-dimensional polynomial space consisting of N shape functions $\{\phi_i\}_{i\in\mathcal{I}}$, where $\mathcal{I} = \{1, \dots, N\}$, as depicted in Definition 2.6. On the other hand, Σ is the so-called dual of $\mathcal{P}$, that is, the set of linear forms $\{\sigma_i\}_{i\in\mathcal{I}}$ such that. $\sigma_j(\phi_i) = \delta_{ij}$ and $p(x) = \sum_{i\in\mathcal{I}} \sigma_i(p)p_i$. If there is a set of points $\{a_i\}_{i\in\mathcal{I}}$ in T such that $\sigma_i(p) = p(a_i)$ $\forall p \in \mathcal{P}$, then the triple $(T, \mathcal{P}, \Sigma)$ is called a Lagrangian finite element. The set of points $\{a_i\}_{i\in\mathcal{I}}$ is called nodes and is associated with the so-called nodal basis of $\mathcal{P}$ such that $\phi(a_i) = \delta_{ij}$ $\forall i, j \in \mathcal{I}$

As anticipated, the local node configuration of the polynomial space is influenced by the form of T. For the purpose of our discussion, let us represent the polynomial basis for a simplicial element and a quadrilateral element as $\mathcal{P}^k(T)$ and $\mathcal{Q}^k(T)$, both of polynomial order k. In Figure 5 is it illustrated for $k = 1, 2, 3$ in dimension $d = 2$ on how the node configuration evolve.



(a) $\mathcal{P}^1(T)$ (b) $\mathcal{P}^2(T)$ (c) $\mathcal{P}^3(T)$

(d) $\mathcal{Q}^1(T)$ (e) $\mathcal{Q}^2(T)$ (f) $\mathcal{Q}^3(T)$

Figure 5: Illustration of the nodes for the element of a simplex a quadrilateral for dimension $d = 2$ for polynomial orders $k = 1, 2, 3$.

We may introduce the reference element $\hat{T}$ in d dimensions. The reference for a quadrilateral is denoted as $\hat{T} = [0, 1]^d$. The reference for a simplex in is defined by the convex hull spanned by the points $(z_0, e_1, \dots, e_d)$ where $z_0 := 0$ is the origin and $\{e_i\}_{i=1}^d$ is the standard Cartesian unit basis in $\mathbb{R}^d$. A corresponding reference finite element is defined as $(\hat{T}, \hat{\mathcal{P}}, \hat{\Sigma})$.

Let the mapping $\mathcal{G} : \hat{T} \to T$ an affine mapping, i.e. $\mathcal{G}(x) = Ax + b$. The important property of affine transformations is the preservation of parallelism. Hence, for any two vectors $x, y \in \hat{T}$ that are parallel in the reference element, their images $\mathcal{G}(x)$ and $\mathcal{G}(y)$ will also be parallel. Generally speaking, an affine transformation of the reference simplex is a transformation to any another simplex of the same dimension. However, for any quadrilateral, an affine transformation preserves the parallelism of opposite sides, for an illustration see Figure 6 and for a counter example see Figure 7.

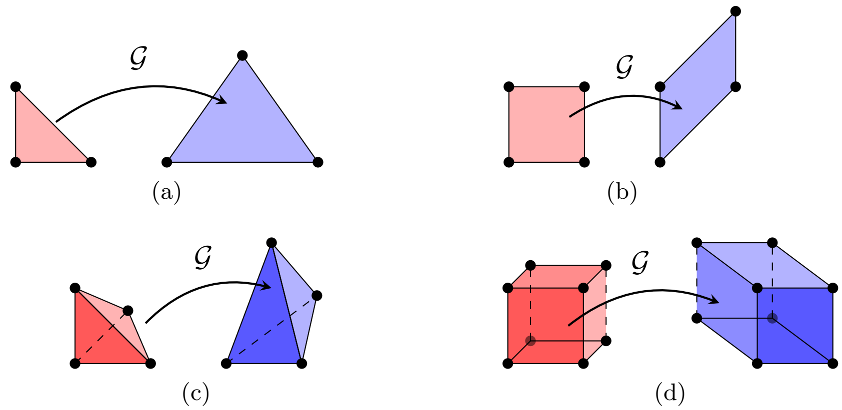


Figure 6: Illustration of affine mapping $\mathcal{G} : \hat{T} \to T$ in dimensions $d = 2, 3$ from a reference element $\hat{T}$ to element T for simplexes and and quadrilaterals.

Figure 7: Illustration of an affine mapping $\mathcal{G} : \hat{T} \mapsto T$ versus a non-affine transformation. The left figure preserves parallel lines before and after the transformation, indicating an affine transformation. However, the right figure does not maintain parallelism, making it a non-affine transformation.

Following [82, Example 9.4], the $(\hat{T}, \hat{\mathcal{P}}, \hat{\Sigma})$ is denoted as the reference finite element associated with the nodes $\{\hat{a}_i\}_{i\in N}$. Let the mapping ψ be function in $\mathcal{L}(\mathcal{P}^k(T), \mathcal{P}^k(\hat{T}))$ such that is an isomorphic from $\psi : \mathcal{P}^k(T) \to \mathcal{P}^k(\hat{T})$. Then $\sigma(p) = \hat{\sigma}(\psi(p))(a_i) = (p \circ \mathcal{G})(\hat{a}_i)$ for all $p \in \mathcal{P}^k(T)$. Then the Lagrange interpolation follows,

$$p(x) = \sum_{i\in N} \sigma(a_i)\phi_i(x) \quad \text{for } a_i = \mathcal{G}(\hat{a}_i) \quad \forall i \in N.$$

Hence, the finite element $(T, \mathcal{P}, \Sigma)$, associated with the notes $\{a_i\}_{i\in N}$, is reconstructed via the reference finite element $(\hat{T}, \hat{\mathcal{P}}, \hat{\Sigma})$. Thus, using the affine transformation can the definition of the local polynomial space be extended to dependent on the reference element. That is,

$$\mathcal{P}^k(T) = \left\{ \hat{v} \circ \mathcal{G}^{-1}(T) \mid \hat{v} \in \mathcal{P}^k(\hat{T}) \right\}$$
$$\mathcal{Q}^k(T) = \left\{ \hat{v} \circ \mathcal{G}^{-1}(T) \mid \hat{v} \in \mathcal{Q}^k(\hat{T}) \right\}$$

Working on shape-regular and affine geometries has shown to greatly simplify and generalise local interpolation estimates, see [82, Theorem 11.12], and thus is very useful for deriving a priori estimates. Please note that workarounds exist for proving nonaffine local interpolation estimates, but they require key assumptions on the relationship between the nodes a_i and the regularity of the mapping $\mathcal{G}$ [82, Chapter 13]. Hence, affine meshes is essential for the error analysis which utilize the interpolation estimates, but it limits us to work on structure mesh if we specifically choose on quadrilateral meshes.

Definition 2.7 (Broken polynomial spaces)**.** *Let $\mathcal{T}_h$ be a mesh of $\Omega \in \mathbb{R}^d$ and $\Omega_h = \bigcup_{T\in\mathcal{T}_h} T$. Let $\mathcal{P}^k(T)$ be the space of all polynomials of order k in the mesh element T in $\mathcal{T}_h$. We define the* *broken polynomial space and the global C^0 continuous polynomial space as*

$$\begin{aligned}\mathcal{P}^k(\mathcal{T}_h) &:= \left\{v \in L^2(\Omega_h) \mid v|_T \in \mathcal{P}^k(T) \quad \forall T \in \mathcal{T}_h\right\}. \\ \mathcal{P}_c^k(\mathcal{T}_h) &:= \left\{v \in C^0(\Omega_h) \mid v|_T \in \mathcal{P}^k(T) \quad \forall T \in \mathcal{T}_h\right\}.\end{aligned} \tag{2.15}$$

Similarly, for quadrilatural elements is the polynomial spaces defined as,

$$\begin{aligned}\mathcal{Q}^k(\mathcal{T}_h) &:= \left\{v \in L^2(\Omega_h) \mid v|_T \in \mathcal{Q}^k(T) \quad \forall T \in \mathcal{T}_h\right\}. \\ \mathcal{Q}_c^k(\mathcal{T}_h) &:= \left\{v \in C^0(\Omega_h) \mid v|_T \in \mathcal{Q}^k(T) \quad \forall T \in \mathcal{T}_h\right\}.\end{aligned} \tag{2.16}$$

In this thesis will we generally utilize the global C^0 continuity, hence, for the rest of the thesis do we define

$$V_h = \left\{P_c^k(\mathcal{T}_h) \text{ or } Q_c^k(\mathcal{T}_h)\right\} \tag{2.17}$$

Hence, all results holds for both polynomial spaces.

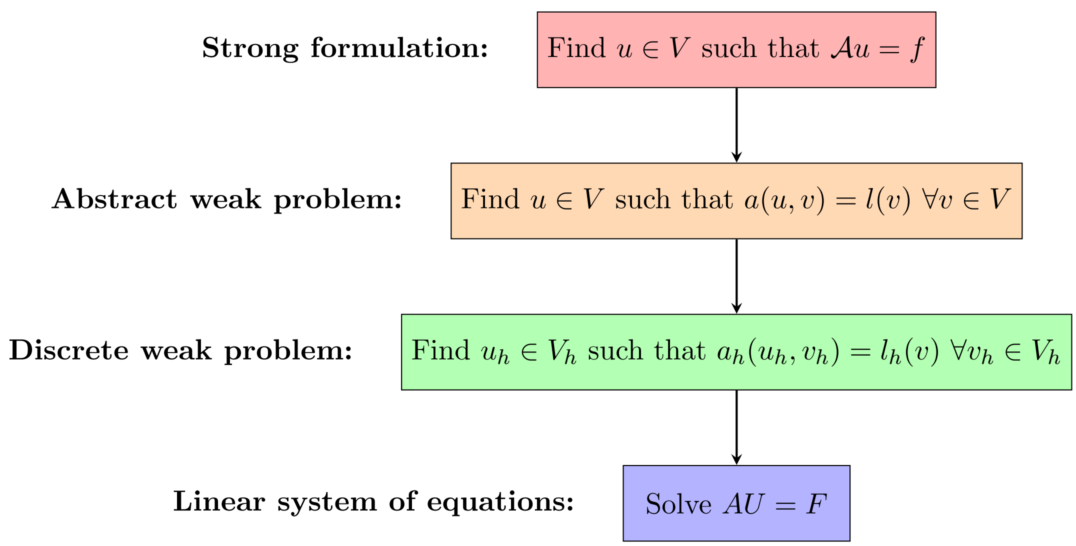


Figure 8: Workflow of solving linear PDEs using the FEM method.

We now have a well-defined discrete global space V_h consisting of the finite set of basis functions $\{\phi_i\}_{i=1}^N$ associated with the Lagriangian nodes $\{a_i\}_{i=1}^N$. The degree of freedoms, also known as ndofs, is denoted as $dim(V_h) = N$. Let $U_j = u_h(a_j)$, so that $u_h = \sum_{j=1}^N U_j \phi_j$. Then the Problem 2.5 is equivalent to

$$\sum_{j=1}^N u_j a_h(\phi_j, \phi_i) = l_h(\phi_i) \tag{2.18}$$

Hence, by letting $U = [U_j]$, $F = [(f, \phi_i)_\Omega]$ and $A = [a_h(\phi_j, \phi_i)]$ can we construct a linear system,

$$AU = F. \tag{2.19}$$

Ultimately, the matrix A is shown to be symmetric positive definite only if $a_h(\cdot, \cdot)$ is well-posed.

To summarize the workflow of solving linear PDEs using the FEM method, see Figure 8.

2.6 Condition number

Recall the discrete l^p norm for a vector,

$$\forall U \in \mathbb{R}^N, \quad \|U\|_p = \begin{cases} \left(\sum_{i=1}^N |U_i|^p\right)^{\frac{1}{p}}, & 1 \leqslant p < \infty \\ \max_i |U_i|, & p = \infty \end{cases} \tag{2.20}$$

Also recall the definition of the matrix norm,

$$\forall A \in \mathbb{R}^{N\times N}, 1 \leqslant p \leqslant \infty, \quad \|A\|_p = \max_{U\in\mathbb{R}^N\backslash 0} \frac{\|AU\|_p}{\|U\|_p}. \tag{2.21}$$

Remark that this notation is not to be confused with Sobolev norms. Assume that A is invertible, then we define the condition number for a matrix in l^p norm defined such that

$$\forall A \in \mathbb{R}^{N\times N}, 1 \leqslant p \leqslant \infty, \quad \kappa_p(A) = \|A\|_p \|A^{-1}\|_p. \tag{2.22}$$

From basic theory is it known that $\|A\|_2$ is equal to the maximum singular value of A , where singular values of A are the square roots of the eigenvalues of $A^T A$ [85, Theorem 2.9]. Because of the connection between $\|A\|_2$ norms and its singular values, is $k_2(A)$ is often in preferred numerical analysis. A challenge is that the computations generally involves performing Singular Value Decomposition (SVD) or power iteration, which can be quite expensive operations particularly for large sparse matrices. However, $\|A\|_\infty$ is computed as the maximum absolute row sum and ,hence, only necessary to compute the sum for the non-zero elements in each row. Thus, we seek to estimate $\kappa_2(A)$ using $\kappa_\infty(A)$.

It is well established that $\frac{1}{\sqrt{N}}\|A\|_\infty \leqslant \|A\|_2 \leqslant \sqrt{N}\|A\|_\infty$ for any matrix $A \in R^{N\times N}$ [2]. Applying this identity, we obtain the upper and lower bounds for $\kappa_2(A)$. Specifically, we have

$$\frac{1}{N}\kappa_\infty(A) \leqslant \kappa_2(A) \leqslant N\kappa_\infty(A) \tag{2.23}$$

Thus, since these norms are equivalent, will we in this thesis focus on $\kappa_\infty(A)$ because of the efficiency of computing $\|A\|_\infty$.

2.7 Cléments interpolation

Our goal is to to utilize interpolation estimates to compute convergence rates. An important tool in the process is the so-called Cléments interpolation operator, C_h. It is used for interpolation on non smooth functions by applying an regularization on so-called macroelements. However, we need to define affine operations on so-called macroelements before we can proceed with the error estimates.

A patch for a element $\omega(T)$ is denoted as the set of elements in $\mathcal{T}_h$ sharing at least one vertex with $T \in \mathcal{T}_h$. Similarly, a patch of a facet $\omega(F)$ is defined as the set of all elements in $\mathcal{T}_h$ sharing at least one vertex with $F \in \mathcal{F}_h$. For an illustrative example of patches in a two-dimensional triangular mesh, please refer to Figure 9.

Let the set $\{a_i\}_{i\in N}$ be all Lagrange nodes on the mesh $\mathcal{T}_h$. Associated with each node a_i we denote the macroelement A_i to consist of all elements containing a_i. Let n_{cf} be the number of configurations for the macroelement, then we define the index $j : \{1, \dots, N\} \to \{1, \dots, n_{cf}\}$ s.t. $j(i)$ is the index associated with the reference configuration $\widehat{A}_{j(i)}$ for corresponding macroelement A_i. Let us define a C^0-diffeomorphism $\mathcal{G}_{A_i} : \widehat{A}_{j(i)} \to A_i$ on the reference macroelements such that for all $\hat{T} \in \widehat{A}_{j(i)}$ is the restriction $\mathcal{G}_{A_i}|_{\hat{T}}$ affine. For an illustration of the reference macroelement $\widehat{A}_{j(i)}$ and how it related to a_i , see Figure 10.

[2]The identity naturally appears from the standard inequality $\|v\|_\infty \leqslant \|v\|_2 \leqslant \sqrt{N}\|v\|_\infty$ for $v \in R^N$, which simply comes from the fact that $\|v\|_\infty^2 = \max_i |v_i|^2 \leqslant \sum_i^n |v_i|^2 = \|v\|_2^2 \leqslant N \max_i |v_i|^2 = N\|v\|_\infty^2$. Now let $A \in \mathbb{R}^{N\times N}$ be any matrix. We can then deduce that $\|A\|_2 = \max_{v\in\mathbb{R}^N} \frac{\|Av\|_2}{\|v\|_2} \geqslant \max_{v\in\mathbb{R}^N} \frac{\|Av\|_\infty}{\sqrt{N}\|v\|_\infty} = \frac{1}{\sqrt{N}}\|A\|_\infty$ and $\|A\|_2 = \max_{v\in\mathbb{R}^N} \frac{\|Av\|_2}{\|v\|_2} \leqslant \max_{v\in\mathbb{R}^N} \frac{\sqrt{N}\|Av\|_\infty}{\|v\|_\infty} = \sqrt{N}\|A\|_\infty$. Hence, the identity $\frac{1}{\sqrt{N}}\|A\|_\infty \leqslant \|A\|_2 \leqslant \sqrt{N}\|A\|_\infty$ is proven.

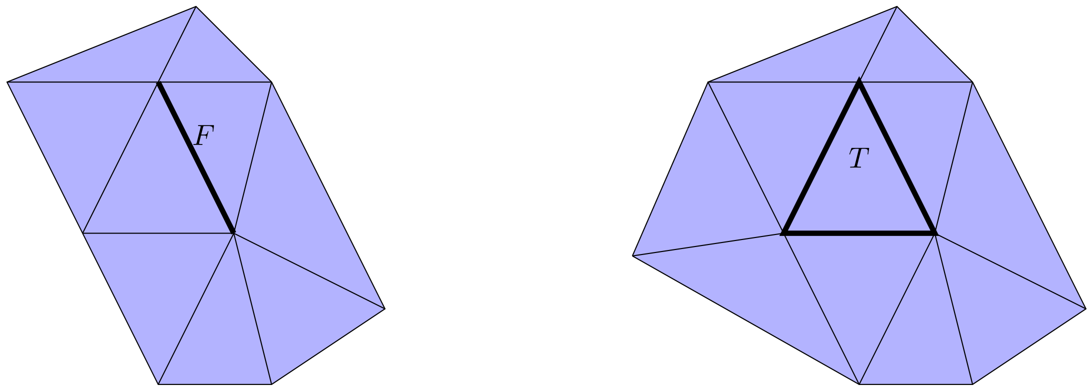


Figure 9: Illustration of the patch $\omega(F)$ on the left-hand side and $\omega(T)$ on the right-hand side.

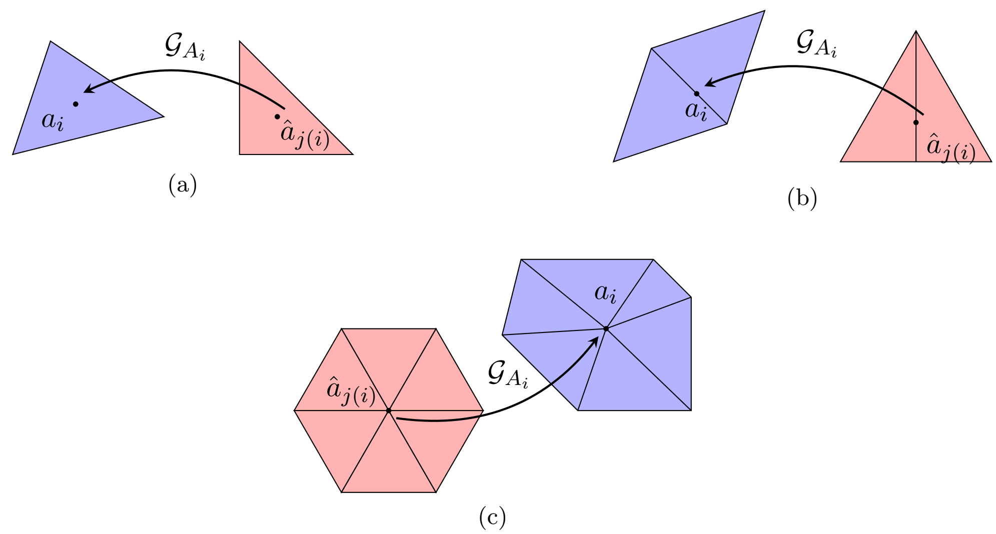


Figure 10: Illustration of the different cases when mapping from the reference macroelement $\widehat{A}_{j(i)}$ to the domain A_i, $\mathcal{G}_{A_i} : \widehat{A}_{j(i)} \to A_i$. Here we have defined $\hat{a}_{j(i)} \in \widehat{A}_{j(i)}$ s.t. $\mathcal{G}_{A_i}(\hat{a}_{j(i)}) = a_i$.

The Cléments interpolation operator C_h is the L^2-projection onto the macroelements. That is, given a reference macroelement $\widehat{A}_{j(i)}$ and a function $\hat{v} \in L^1(\widehat{A}_{j(i)})$, then $\widehat{C}_{j(i)}\hat{v}$ is the unique polynomial in $\mathcal{P}^k(\widehat{A}_{j(i)})$ s.t.

$$\int_{\widehat{A}_{j(i)}} (\widehat{C}_{j(i)}\hat{v} - \hat{v})p \; dx = 0 \quad \forall p \in \mathcal{P}^k(\widehat{A}_{j(i)})$$

Finally, the Cléments interpolator is defined as the mapping $C_h : L^1(\Omega) \to \mathcal{P}_c^k(\mathcal{T}_h)$ such that

$$C_h v = \sum_{i=1}^{N} \widehat{C}_{j(i)}(v(\mathcal{G}_{A_i})(\mathcal{G}_{A_i}^{-1}(a_i)))\phi_i,$$

where ϕ_i is the corresponding polynomial basis associated with the node a_i.

Finally, we have the following a priori estimate.

Lemma 2.8. *Let $v \in H^s(\Omega)$. We define the Clement interpolation as the mapping $C_h : L^2(\Omega) \to V_h$, where V_h has the order k. Then does the following stability estimate hold,*

$$\|C_h v\|_{s,\Omega} \lesssim \|v\|_{s,\Omega} \quad \forall v \in H^s(\Omega),$$

Let $r = \min(s, k+1)$. If the following conditions for an parameter l is satisfied, it exists error estimates such that

$$
\begin{aligned}
0 \leqslant l \leqslant r &\implies \|v - C_h v\|_{m,T} \lesssim h_T^{r-l} \|v\|_{l,\omega(T)} \quad \forall T \in \mathcal{T}_h, \forall v \in H^l(\omega(T)), \\
0 \leqslant l \leqslant r - \frac{1}{2} &\implies \|v - C_h v\|_{m,F} \lesssim h_T^{r-l-\frac{1}{2}} \|v\|_{l,\omega(F)} \quad \forall \partial T \in \mathcal{T}_h, \forall v \in H^l(\omega(F)).
\end{aligned}
\tag{2.24}
$$

Corollary 2.9. *Let $0 \leqslant l \leqslant k+1$ and let $0 \leqslant m \leqslant \min(1, l)$. Given Lemma 2.8 then does the following estimate hold*

$$
\inf_{v_h \in V_h} \|v - v_h\|_{m,\Omega} \lesssim h^{l-m} \|v\|_{l,\Omega} \quad \forall v \in H^l(\Omega).
$$

This result is very useful since it is now sufficient to show that a priori estimates holds to prove convergence rate. For further detailed information about the Cléments interpolation, please investigate [86, Chapter 1.6].

2.8 Useful local inverse estimates

Choose any element $T \in \mathcal{T}_h$ and let $v_h \in \mathcal{P}^m(T)$. Then does the local inverse estimate hold,

$$
|v_h|_{H^l(T)} \lesssim h^{m-l} |v_h|_{H^m(T)} \text{ for } l \leqslant m. \tag{2.25}
$$

For proof, see [82, Lemma 12.1]. An essential example is the following inequality.

$$
\|D^2 v_h\|_T \lesssim h^{-1} \|\nabla v_h\|_T \lesssim h^{-2} \|v_h\|_T \tag{2.26}
$$

Another very useful inequality is the so-called trace inequality which connect the relationship of evaluating the norm on element T and with any of the corresponding facets $F \in \partial T$. The general form is

$$
\|v_h\|_F \lesssim h^{-\frac{1}{2}} \|v_h\|_T, \tag{2.27}
$$

For proof, see [82, Lemma 12.8].

Let $\partial_n v = \nabla v \ n$ and $\partial_{nn} v = n^T D^2 v \ n$. Keeping in mind that the normal vector has a unit length and, thus, evidently applying the trace inverse inequality we have two present useful examples,

$$
\begin{aligned}
\|\partial_n v_h\|_F \leqslant \|\nabla v_h\|_F \leqslant h^{-\frac{1}{2}} \|\nabla v_h\|_T, \\
\|\partial_{nn} v_h\|_F \leqslant \|D^2 v_h\|_F \leqslant h^{-\frac{1}{2}} \|D^2 v_h\|_T.
\end{aligned}
\tag{2.28}
$$

Combining (2.25) and (2.27), we establish that

$$
|v_h|_{l,F} \lesssim h^{m-l-\frac{1}{2}} |v_h|_{m,T} \text{ for } l \leqslant m. \tag{2.29}
$$

2.9 Cea's lemma → adapt to non-conform

Assume that we have a discrete bilinear form $a_h : X_h \times X_h \to \mathbb{R}$ and let $X_h \subseteq V$ be the discrete conform polynomial space. We transition the problem to find a solution, $u_h \in \mathcal{V}_h$, so it holds that $a(u_h, v_h) = l(v_h) \quad \forall v_h \in X_h$. Since the method is conform, i.e., $X_h \subseteq V$, does it exists a exact solution, $u \in \mathcal{V}$, such that $a(u, v_h) = l(v_h) \quad \forall v_h \in X_h$. Furthermore, the problem is said to be strongly consistent since it fulfills the Galerkin orthogonality property, that is $a(u - u_h, v_h) = 0$. Thus, if $u_h, v_h \in X_h$, then

$$
\begin{aligned}
\alpha \|u - u_h\|_V^2 &\leqslant a(u - u_h, u - u_h) \\
&= a(u - u_h, u - v_h) - a(u - u_h, v_h - u_h). \\
&\leqslant M \|u - u_h\|_V \|u - v_h\|_V
\end{aligned}
\tag{2.30}
$$

$\cdot\ w \in V \cap W$ s.t. at $a_h(w, v_h)$ can be evaluated

Not $w \in H^2$ but in $w \in$

- Discrete Lax–Milgram

$\|\cdot\|_{V_h}$, Require that

$$a_h(v_h, w_h) \lesssim \|v_{h}\|_{V_h} \|w_h\|_{V_h}$$

$$a_h(v_h, v_h) \gtrsim \|v_h\|_{V_h}$$

$a_h(u, v_h) = l_h(v_h) \quad \forall v \in V_h$

- A priori error analysis requires that

$u \in V \cap W$ s.t. $a_h(u, v_h)$ makes sense and is consistent

$V \to H^2(\Omega)$, $W \to H^{\frac{3}{2}+\varepsilon}$ through regularity results.

- Cea's lemma

require that $\lesssim \|v_h - u_h\|_{V_h}$

triangle $\|u - u_h\|_{V,*} \lesssim \underbrace{\|u - v_h\|_{V,*}}_{\text{need interpolation estimates for that}} + \|v_h - u_h\|_{V,*}$

$$\|v_h - u_h\|_{V_h}^2 \lesssim a_h(v_h - u_h, v_h - u_h)$$

$$= a_h(v_h - u, v_h - u_h) + \underbrace{a_h(u - u_h, v_h - u_h)}_{\text{Galerkin Ortho} = 0}$$

$= 0$ since discrete form is consistent for $u \in V \cap W$.

$= 0$ Consistent

$$\lesssim \|v_h - u\|_{V,*} \|v_h - u_h\|_{V_h}$$

Require Boundness result w.r.t. to $\|\cdot\|_{V,*}$ $\|\cdot\|_{V_h}$

We require that this holds $a_h(v_h - u, w_h) \lesssim \|v_h - u\|_{V,*} \|w_h\|_{V_h}$

$\Rightarrow$ Cea's lemma

$$\|u - u_h\|_{V,*} \le C \inf_{v_h \in V_h} \|u - v_h\|_{V,*}.$$

Hence, we now have the so-called Ceas lemma,

$$\|u - u_h\|_V \lesssim \inf_{v_h \in X_h} \|v_h - u\|_V.$$

A useful property is that for a conformal numerical method to converge can we now simply require

$$\lim_{h \to 0} \inf_{v_h \in X_h} \|v - v_h\|_V = 0 \quad \forall v \in V.$$

In that case will $\|u - u_h\|_V \to 0$, $h \to 0$. Hence, if this requirement is fulfilled, the numerical methods will converge to a unique solution. For more information, see [87, p. 66].

In combination with Corollary 2.9, is Cea's lemma very handy to construct a priori estimates. However, Ceas lemma as it stands does assume conform methods $X_h \subset X$, hence, we must be cautious for broken Sobolev spaces.

↳ in the forthcoming chapters when non-conform broken Sobolev spaces need to be considered.

3 Continuous interior penalty methods for the biharmonic problem with Cahn-Hilliard type boundary conditions

One of the objective of this section is to discuss the strong formulation for the biharmonic problem. Following this, we will present both the continuous weak formulation and the derivation of the two proposed discrete weak formulations, specifically the continuous interior penalty methods. We then present a short discussion of the current status of the properties of the methods.

→ Discuss relevant References & Make sure that there is no misunderstand and that you only review the material from the literature.

3.1 The biharmonic equation

Let $\Omega \subseteq \mathbb{R}^d$ be a bounded polygonal domain and Γ be its corresponding boundary. Also let $\mathcal{T}_h = \{T\}$ be a shape-regular fitted mesh s.t. $\mathcal{T}_h = \Omega$. Let the biharmonic problem have the form,

$$\Delta^2 u + \alpha u = f(x) \quad \text{in } \Omega, \tag{3.1a}$$

$$\partial_n u = g_1(x) \quad \text{on } \Gamma, \tag{3.1b}$$

$$\partial_n \Delta u = g_2(x) \quad \text{on } \Gamma. \tag{3.1c}$$

Here is $\Delta^2 = \Delta(\Delta)$ the biharmonic operator, also known as the bilaplacian. We will assume for the strong form that $u \in H^4(\Omega)$, $\alpha > 0$ and $f \in L^2(\Omega)$. The functions $g_1, g_2 : \Omega \to \mathbb{R}$ are denoted as boundary conditions similar to the CH problem.

Remark. It is worth noting that the problem is closely related to the Kirchhoff's plate problem by changing the boundary conditions such that $u = \partial_n u = 0$ on Γ, which is in the literature known as so-called clamped boundary conditions. Many of the papers we refer to may consider clamped boundary condition and not the CH boundary conditions. The main difference relies on if the problem is treated with homogeneous or non-homogeneous boundary conditions and if the discrete space is imposing the Dirichlet and Neumann conditions strongly in the discrete solution space or weakly using the Nitsche's method [88].

We want to construct a weak form for the strong biharmonic problem (3.1). Let $v \in H^2(\Omega)$ Using Greens Theorem is it obvious that $\left(\Delta^2 u, v\right)_\Omega = (\partial_n \Delta u, v)_\Gamma - (\nabla(\Delta u), \nabla v)_\Omega$. Next, applying a new iteration of the Greens theorem we get $-(\nabla(\Delta u), \nabla v)_\Omega = (\Delta u, \Delta v)_\Omega - (\Delta u, \partial_n v)_\Gamma$. Hence, we obtain the identity

$$(\Delta^2 u, v)_\Omega = (\Delta u, \Delta v)_\Omega + (\partial_n \Delta u, v)_\Gamma - (\partial_n v, \Delta u)_\Gamma \tag{3.2}$$

Taking into account the boundary conditions, we end up with the following corresponding weak formulation of the biharmonic problem (3.1).

$$V := \{v \in H^s(\Omega),\ s \geqslant 5/2 + \varepsilon \mid \partial_n v = g_1\} \tag{3.3}$$

Consider the bilinear form $a : V \times V \to \mathbb{R}$ and the linear form $l : V \to \mathbb{R}$. We define the continuous weak formulation problem formulation as follows.

$$\text{Find } u \in V \text{ such that } a(u, v) = l(v) \forall v \in V \tag{3.4}$$

where

$$\begin{aligned} a(u, v) &= (\alpha u, v)_\Omega + (\Delta u, \Delta v)_\Omega - (\Delta u, \partial_n v)_\Gamma \\ l(v) &= (f, v)_\Omega - (g_2, v)_\Omega \end{aligned} \tag{3.5}$$

Note that the Neumann boundary condition g_1 is strongly imposed in V, and g_2 is naturally incorporated in the weak formulation.

3.2 Detailed construction of Hessian and Laplacian Formulations

The goal is to construct two CIP formulations for the problem (3.4), that is: Find $u_h \in V_h \notin V$ such that $a_h(u_h, v_h) = l_h(v_h)$ for all $v_h \in V_h$. We will follow the ideas presented in [53] and [61]. A property that is necessary is that when we have a bilinear form and replace the exact solution with $u_h \in V_h$, the system still remains consistent in V. Hence, we guarantee a consistency of the discrete weak formulation by assuming during the construction that $u \in H^4(\Omega)$ and $v_h \in V_h$. However, due to the nonconformal nature of V_h, it becomes necessary to introduce penalty terms to ensure the discrete system is well-posed when we replace with $u_h \in V_h$. Keep in mind that the C^1 continuity is imposed weakly and in the same way is the Neumann conditions also imposed weakly.

To achieve the objective of constructing the Hessian and Laplacian formulations, the following lemmas will be the primary components.

3.2.1 Construction of the Hessian formulation

Lemma 3.1. *Assume the homogeneous Neumann conditions $g_1 = 0$. Let $u \in H^4(\Omega)$ be the solution to (3.1), let $v_h \in V_h$ and a constant $\gamma > 0$. Then does the following identity hold.*

$$\begin{aligned} \left(\Delta^2 u, v_h\right)_\Omega = &\left(D^2 u, D^2 v_h\right)_\Omega + (g_2, v_h)_\Gamma \\ &- \left(\{\!\{\partial_{nn} u\}\!\}, [\![\partial_n v_h]\!]\right)_{\mathcal{F}_h^{int}} - \left([\![\partial_n u]\!], \{\!\{\partial_{nn} v_h\}\!\}\right)_{\mathcal{F}_h^{int}} + \frac{\gamma}{h}\left([\![\partial_n u]\!], \{\!\{\partial_{nn} v_h\}\!\}\right)_{\mathcal{F}_h^{int}} \\ &- (\partial_{nn} u, \partial_n v_h)_\Gamma - (\partial_n u, \partial_{nn} v_h)_\Gamma + \frac{\gamma}{h}(\partial_n u, \partial_n v_h)_\Gamma \end{aligned} \tag{3.6}$$

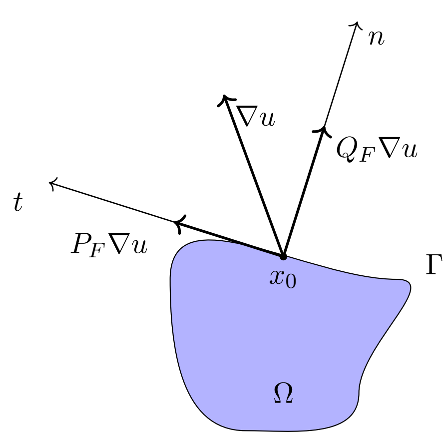


Figure 11: Let x_0 be a point at the boundary Γ for dimension $d = 2$. Here is a illustration of the gradient ∇u with the corresponding normal and tangential decomposition, $Q_F \nabla u$ and $P_F \nabla u$.

Proof. We will start constructing a local theory for a element T and then extend it to the full mesh $\mathcal{T}_h$. Using Greens Theorem is it obvious that

$$\left(\Delta^2 u, v_h\right)_T = (\partial_n \Delta u, v_h)_{\partial T} - \left(\nabla\left(\Delta u\right), \nabla v_h\right)_T \tag{3.7}$$

We can expand the second term in the following way.

$$\begin{aligned}
(\nabla(\Delta u), \nabla v_h)_T &= \sum_{i=1}^{d} (\Delta \partial_{x_i} u, \partial_{x_i} v_h)_T = \sum_{i=1}^{d} (\nabla \cdot (\nabla \partial_{x_i} u), \partial_{x_i} v_h)_T \\
&= \sum_{i=1}^{d} ((\partial_n \partial_{x_i} u, \partial_{x_i} v_h)_{\partial T} - (\nabla \partial_{x_i} u, \nabla \partial_{x_i} v_h)_T) \\
&= (\partial_n \nabla u, \nabla v_h)_{\partial T} - (D^2 u, D^2 v_h)_T
\end{aligned} \quad . \tag{3.8}$$

Hence, the normal flux of Δu appears naturally into the formulation. It can be denoted that D^2 is the Hessian matrix operator. Also remark that we apply the notation $(D^2 u, D^2 v_h)_\Omega = \int_\Omega D^2 u : D^2 v_h dx$ for the inner product $D^2 u : D^2 v_h$.

Next, we want to decompose the evaluation of ∇u on the boundary ∂T in the tangential and normal direction. Pick a facet $F \in \partial T$, then we define the following decomposition of linear transformation $\nabla u = P_F \nabla u + Q_F \nabla u$ s.t. the orthogonality, $P_F \nabla u \cdot Q_F \nabla u = 0$, holds. Here, the normal projection matrix is defined as $Q_F = n \otimes n$ and the tangential decomposition follows from $P_F = I - Q_F = I - n \otimes n = \sum_{i=1}^{d-1} t_i \otimes t_i$, where we defined a orthonormal basis t_i, $i = 1, \ldots, d-1$ for the space orthogonal to the outer normal vector n on a facet F. For demonstration in $d = 2$, see Figure 11. Let $a_1, a_2, a_3 \in \mathbb{R}^d$ be any vectors, then it is well known that the following identity holds $(a_1 \otimes a_2) a_3 = (a_2^T a_3) a_1$. Hence, we have

$$\begin{aligned}
Q_F \nabla u &= (n \otimes n) \nabla u = (n^T \nabla u) n \\
P_F \nabla u &= (I - n \otimes n) \nabla u = \nabla u - (n^T \nabla u) n = \sum_{i=1}^{d-1} (t_i^T \nabla u) t_i
\end{aligned} \tag{3.9}$$

Given that u is evaluated only on ∂T can we write $\nabla u = \left(n^T \nabla u\right) n + \sum_i^{d-1} \left(t_i^T \nabla u\right) t_i$ such that,

$$\begin{aligned}
(\partial_n \nabla u, \nabla v_h)_{\partial T} &= (\partial_n(\partial_n u \; n), \partial_n v_h \; n)_{\partial T} + \sum_{i,j=1}^{d-1} (\partial_n(\partial_{t_i} u \; t_i), \partial_{t_j} v_h \; t_j)_{\partial T} \\
&= (\partial_{nn} u, \partial_n v_h)_{\partial T} + \sum_{i=1}^{d-1} (\partial_{n t_i} u, \partial_{t_i} v_h)_{\partial T}
\end{aligned} \tag{3.10}$$

Here we used that $n^T n = 1$ and $t_i^T t_j = \delta_{ij}$. Remark that simple relation was applied,

$$\begin{aligned}
\partial_n(\partial_n u) &= n^T \nabla(\partial_n u) = n^T (D^2 u \; n) = n^T D^2 u \; n = \partial_{nn} u, \\
\partial_n(\partial_{t_i} u) &= t_i^T \nabla(\partial_n u) = t_i^T (D^2 u \; n) = n^T D^2 u \; t_i = \partial_{n t_i} u.
\end{aligned}$$

We may also deduce the relationship $\partial_{n t_i} u = \partial_{t_i n} u$ which arise from the fact that $n^T D^2 u \; t_i = (t_i^T D^2 u \; n)^T = t_i^T D^2 u \; n$, where we utilized the symmetry $D^2 u = (D^2 u)^T$ and that the product is a scalar. Combining (3.7), (3.8) and (3.10) we see that,

$$(\Delta^2 u, v_h)_T = (D^2 u, D^2 v_h)_T + (\partial_n \Delta u, v_h)_{\partial T} - (\partial_{nn} u, \partial_n v_h)_{\partial T} - \sum_{i=1}^{d-1} (\partial_{n t_i} u, \partial_{t_i} v_h)_{\partial T}$$

Since we aim construct a identity for the full mesh $\mathcal{T}_h$, we sum over the elements.

$$\left(\Delta^2 u, v_h\right)_\Omega = \sum_{T\in\mathcal{T}_h} (D^2u, D^2v_h)_T + (\partial_n \Delta u, v_h)_{\partial T} - (\partial_{nn}u, \partial_n v_h)_{\partial T} - \sum_{i=1}^{d-1}(\partial_{nt_i}u, \partial_{t_i}v_h)_{\partial T} \quad (3.11)$$

Our goal is to simplify the equation above so we can take account for discontinuities of the derivatives. By integrating over exterior facets $\mathcal{F}_h^{ext}$ and interior facets $\mathcal{F}_h^{int}$ we will get e more suitable formulation which makes it easier to control the jumps between the elements, hence makes it possible to penalize discontinuities.

$$\begin{aligned}
(\Delta^2 u, v_h)_\Omega &= \sum_{T\in\mathcal{T}_h} (D^2u, D^2v_h)_T + (\partial_n \Delta u, v_h)_{\partial T} - (\partial_{nn}u, \partial_n v_h)_{\partial T} - \sum_{i=1}^{d-1}(\partial_{t_i n}u, \partial_{t_i}v_h)_{\partial T} \\
&= \sum_{T\in\mathcal{T}_h} (D^2u, D^2v_h)_T + \sum_{F\in\mathcal{F}_h^{ext}} (\partial_n \Delta u, v_h)_F - (\partial_{nn}u, \partial_n v_h)_F - \sum_{i=1}^{d-1}(\partial_{t_i n}u, \partial_{t_i}v_h)_F \\
&\quad + \sum_{F\in\mathcal{F}_h^{int}} \underbrace{\left((\partial_{n^+}\Delta u^+, v_h^+)_F + \left(\partial_{n^-}\Delta u^+, v_h^-\right)_F\right)}_{(I)} \\
&\qquad - \underbrace{\left(\left(\partial_{n^+n^+}u^+, \partial_{n^+}v_h^+\right)_F + \left(\partial_{n^-n^-}u^-, \partial_{n^-}v_h^-\right)_F\right)}_{(II)} \\
&\qquad - \sum_{i=1}^{d-1} \underbrace{\left((\partial_{n^+t_i}u^+, \partial_{t_i}v_h^+)_F + \left(\partial_{n^-t_i}u^-, \partial_{t_i}v_h^-\right)_F\right)}_{(III)}
\end{aligned}.$$

Where integration of the interior facets is computed in the following fashion.

$$\begin{aligned}
(I) &= \left(\partial_{n^+}\Delta u^+, v_h^+\right)_F + \left(\partial_{n^-}\Delta u^-, v_h^-\right)_F \\
&= \int_F [\![\partial_n \Delta u \cdot v_h]\!] = \int_F \{\!\{\partial_n \Delta u\}\!\} \underbrace{[\![v_h]\!]}_{=0} + \underbrace{[\![\partial_n \Delta u]\!]}_{=0} \{\!\{v_h\}\!\} = 0 \\
(II) &= \left(\partial_{n^+n^+}u^+, \partial_{n^+}v_h^+\right)_F + \left(\partial_{n^-n^-}u^-, \partial_{n^-}v_h^-\right)_F \\
&= \int_F [\![\partial_{nn} u \cdot \partial_n v_h]\!] = \int_F \{\!\{\partial_{nn} u\}\!\} \underbrace{[\![\partial_n v_h]\!]}_{\neq 0} + \underbrace{[\![\partial_{nn} u]\!]}_{=0} \{\!\{\partial_n v_h\}\!\} \\
(III) &= \left(\partial_{n^+t_i}u^+, \partial_{t_i}v_h^+\right)_F + \left(\partial_{n^-t_i}u^-, \partial_{t_i}v_h^-\right)_F \\
&= \int_F [\![\partial_{nt_i} u \cdot \partial_{t_i} v_h]\!] = \int_F \{\!\{\partial_{nt_i} u\}\!\} \underbrace{[\![\partial_{t_i} v_h]\!]}_{=0} + \underbrace{[\![\partial_{nt_i} u]\!]}_{=0} \{\!\{\partial_{t_i} v_h\}\!\} = 0
\end{aligned} \quad . \quad (3.12)$$

Observe that the cancellations in the term (I) and term (III) appears of the continuity of $v_h \in V_h$ and $u \in H^4(\Omega)$ which makes the jumps and derivative jumps zero. On the other hand, the second term (II) does not vanish since the derivative of $v_h \in V_h$ has a nonzero jump. It can also be raised that $\{\!\{\partial_{nn}u\}\!\} = \partial_{nn}u$ holds of $H^4(\Omega)$.

Combining (3.12) and inserting the boundary condition $g_2 = \partial_n \Delta u$ is it clear that the formulation presented in (3.11) is equivalent to the following formulation.

$$\begin{aligned}
\left(\Delta^2 u, v_h\right)_\Omega = \left(D^2u, D^2v_h\right)_{\mathcal{T}_h} + (g_2, v_h)_\Gamma - \left(\{\!\{\partial_{nn}u\}\!\}, [\![\partial_n v_h]\!]\right)_{\mathcal{F}_h^{int}} \\
- (\partial_{nn}u, \partial_n v_h)_{\mathcal{F}_h^{ext}} - \sum_{i=1}^{d-1}(\partial_{t_i n}u, \partial_{t_i}v_h)_{\mathcal{F}_h^{ext}}
\end{aligned} \quad (3.13)$$

Under the assumption that $g_1 = 0$ on Γ, and given that the tangential decomposition is orthogonal to n, we can assert that $\partial_{t_i n} u = \partial_{t_i}(\partial_n u) = \partial_{t_i}(g_1) = 0$ holds for any $i = 1, \ldots, d-1$. This implies that the last term of the equation vanish.

We also note that we add consistent symmetry terms $(\{\!\!\{\partial_{nn} v_h\}\!\!\}, [\![\partial_n u]\!])_{\mathcal{F}_h^{int}}$ and $(\partial_{nn} v_h, \partial_n u)_\Gamma$ in addition to the penalty terms $\frac{\gamma}{h}(\partial_n u, \partial_n v)_\Gamma$ and $\frac{\gamma}{h}([\![\partial_n u]\!], [\![\partial_n v]\!])_{\mathcal{F}_h^{int}}$. Since $u \in H^4(\Omega)$ and the boundary condition, $\partial_n u = g_1 = 0$ on Γ, is each of these terms effectively zero, but does provide symmetry and will later be proven to be essential for well-posedness for the discrete problem. Finally, we have

$$
\begin{aligned}
\left(\Delta^2 u, v_h\right)_\Omega = &\left(D^2 u, D^2 v_h\right)_\Omega + (g_2, v_h)_\Gamma \\
&- (\{\!\!\{\partial_{nn} u\}\!\!\}, [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}} - ([\![\partial_n u]\!], \{\!\!\{\partial_{nn} v_h\}\!\!\})_{\mathcal{F}_h^{int}} + \frac{\gamma}{h}([\![\partial_n u]\!], [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}} \\
&- (\partial_{nn} u, \partial_n v_h)_\Gamma - (\partial_n u, \partial_{nn} v_h)_\Gamma + \frac{\gamma}{h}(\partial_n u, \partial_n v)_\Gamma
\end{aligned}
\tag{3.14}
$$

The proof is complete.

□

Note that since $V_h \not\subset V$ is it necessary to define the space $V \oplus V_h$, which essentially is the direct sum of these two spaces. This new space includes all elements from V and V_h and all possible linear combinations of these elements. i.e., let $u \in V$ and $u_h \in V_h$, then $u + u_h \in V \oplus V_h$.

We will now assemble the Hessian CIP formulation. Assume that the homogeneous boundary condition $g_1 = 0$. The discrete problem is as follows:

$$
\text{Find } u_h \in V_h \text{ such that } a^H(u_h, v_h) = l_h^H(v_h) \quad \forall v_h \in V_h. \tag{3.15}
$$

Here is the corresponding bilinear and linear form defined as,

$$
\begin{aligned}
a_h^H(u_h, v_h) = &(\alpha u_h, v_h)_\Omega + \left(D^2 u_h, D^2 v_h\right)_\Omega \\
&- (\{\!\!\{\partial_{nn} u_h\}\!\!\}, [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}} - ([\![\partial_n u_h]\!], \{\!\!\{\partial_{nn} v_h\}\!\!\})_{\mathcal{F}_h^{int}} + \frac{\gamma}{h}([\![\partial_n u_h]\!], [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}} \\
&- (\partial_{nn} u_h, \partial_n v_h)_\Gamma - (\partial_n u_h, \partial_{nn} v_h)_\Gamma + \frac{\gamma}{h}(\partial_n u_h, \partial_n v_h)_\Gamma \\
l_h^H(v_h) = &(f, v_h)_\Omega - (g_2, v_h)_\Gamma
\end{aligned}
\ . \tag{3.16}
$$

With the corresponding energy norms,

$$
\begin{aligned}
\|v_h\|_{a_h^H}^2 &= \alpha\|v_h\|_\Omega^2 + \|D^2 v_h\|_\Omega^2 + \|h^{-\frac{1}{2}} [\![\partial_n v_h]\!]\|_{\mathcal{F}_h^{int}}^2 + \|h^{-\frac{1}{2}} \partial_n v_h\|_\Gamma^2, \quad v_h \in V_h \\
\|v\|_{a_h^H,*}^2 &= \|v\|_{a_h}^2 + \|h^{\frac{1}{2}} \{\!\!\{\partial_{nn} v\}\!\!\}\|_{\mathcal{F}_h^{int}}^2 + \|h^{\frac{1}{2}} \partial_{nn} v\|_\Gamma^2, \quad v \in V \oplus V_h.
\end{aligned}
\tag{3.17}
$$

Remark. This formulation accommodates the nonconformity of V_h by factoring in the discontinuities among the facets, yet it preserves consistency when u exhibits sufficient regularity, specifically when $u \in H^s(\Omega), s \geqslant \frac{5}{2} + \varepsilon$. This implies that the solution u is continuous across the boundaries of interior elements, i.e., $[\![\partial_n u]\!] = 0$ and $\{\!\!\{\partial_{nn} u\}\!\!\} = \partial_{nn} u$ on any $F \in \mathcal{F}_h^{int}$.

It is noteworthy that we have the consistent terms $(\{\!\!\{\partial_{nn} u_h\}\!\!\}, [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}}$ and $(\partial_{nn} u_h, \partial_n v_h)_\Gamma$ naturally appear in the derivation. However, we also added two symmetry terms, $(\{\!\!\{\partial_{nn} u_h\}\!\!\}, [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}}$ and $(\partial_{nn} u_h, \partial_n v_h)_\Gamma$, and the so-called penalty terms, $\frac{\gamma}{h}([\![\partial_n u_h]\!], [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}}$ and $\frac{\gamma}{h}(\partial_n u_h, \partial_n v_h)_\Gamma$. These terms is essential for making the problem well-posed, hence, the name interior penalty method or symmetric interior penalty method. For more information of nonconformal CIP error analysis, see [83, Chapter 1.3].

3.2.2 Construction of the Laplacian formulation

Lemma 3.2. *Let $u \in H^4(\Omega)$ the solution of (3.1), $v_h \in V_h$ and a constant $\gamma > 0$. Then we have the following identity.*

$$
\begin{aligned}
(\Delta^2 u, v_h)_\Omega =& (\Delta u, \Delta v_h)_{\mathcal{T}_h} + (g_2, v_h)_\Gamma - (g_1, \Delta v_h)_\Gamma + \frac{\gamma}{h}(g_1, \partial_n v_h)_\Gamma \\
&- (\llbracket \partial_n u \rrbracket, \{\!\{\Delta v_h\}\!\})_{\mathcal{F}_h^{int}} - (\{\!\{\Delta u\}\!\}, \llbracket \partial_n v_h \rrbracket)_{\mathcal{F}_h^{int}} + \frac{\gamma}{h}(\llbracket \partial_n u \rrbracket, \llbracket \partial_n v_h \rrbracket)_{\mathcal{F}_h^{int}} \\
&- (\Delta u, \partial_n v_h)_\Gamma - (\partial_n u, \Delta v_h)_\Gamma + \frac{\gamma}{h}(\partial_n u, \partial_n v_h)_\Gamma.
\end{aligned}
$$

Proof. Similarly, we start by constructing integration by parts identities locally for a element T and then extend it to the full mesh $\mathcal{T}_h$. Utilizing (3.2) can we see that

$$
(\Delta^2 u, v_h)_T = (\Delta u, \Delta v_h) + (\partial_n \Delta u, v_h)_{\partial T} - (\partial_n v_h, \Delta u)_{\partial T}
$$

Now, summing over all elements we get

$$
\begin{aligned}
(\Delta^2 u, v_h)_\Omega &= \sum_{T \in \mathcal{T}_h} ((\Delta u, \Delta v_h)_T + (\partial_n \Delta u, v_h)_{\partial T} - (\partial_n v_h, \Delta u)_{\partial T}) \\
&= (\Delta u, \Delta v_h)_{\mathcal{T}_h} + \sum_{F \in \mathcal{F}_h^{ext}} (\overbrace{(\partial_n \Delta u, v_h)_F}^{=(g_2, v_h)_F} - (\partial_n v_h, \Delta u)_F) \\
&\quad + \sum_{F \in \mathcal{F}_h^{int}} \underbrace{((\partial_{n^+} \Delta u, v_h)_F + (\partial_{n^-} \Delta u, v_h)_F)}_{(I)} - \underbrace{((\partial_{n^+} v_h, \Delta u)_F + (\partial_{n^-} v_h, \Delta u)_F)}_{(II)}
\end{aligned}
$$

Decomposing the terms and utilizing the regularity of $u \in H^4(\Omega)$ and the C^0 continuity of $v_h \in V_h$ is it easy to see that,

$$
\begin{aligned}
(I) = (\partial_{n^+} \Delta u, v_h)_F + (\partial_{n^-} \Delta u, v_h)_F = \int_F [\partial_n \Delta u \ v_h] = (\{\!\{\partial_{n^+} \Delta u\}\!\}, \underbrace{\llbracket v_h \rrbracket}_{=0})_F + (\underbrace{\llbracket \partial_n \Delta u \rrbracket}_{=0}, \{\!\{v_h\}\!\})_F \\
(II) = (\partial_{n^+} v_h, \Delta u)_F + (\partial_{n^-} v_h, \Delta u)_F = \int_F \llbracket \partial_n v_h \ \Delta u \rrbracket = (\underbrace{\llbracket \partial_n v_h \rrbracket}_{\neq 0}, \{\!\{\Delta u\}\!\})_F + (\{\!\{\partial_n v_h\}\!\}, \underbrace{\llbracket \Delta u \rrbracket}_{=0})_F
\end{aligned}
.
$$

Hence, we end up with the identity,

$$
(\Delta^2 u, v_h)_\Omega = (\Delta u, \Delta v_h)_{\mathcal{T}_h} + (\llbracket \partial_n v_h \rrbracket, \{\!\{\Delta u\}\!\})_{\mathcal{F}_h} + (g_2, v_h)_\Gamma - (\partial_n v_h, \Delta u)_\Gamma.
$$

Similarly as for Lemma 3.1, we add consistent symmetry terms $(\llbracket \partial_n u \rrbracket, \{\!\{\Delta v_h\}\!\})_{\mathcal{F}_h^{int}}$ and $(\partial_n u, \Delta v_h)_\Gamma - (g_1, \Delta v_h)_\Gamma$ and the penalty terms $\frac{\gamma}{h}(\partial_n u, \partial_n v_h)_\Gamma - \frac{\gamma}{h}(g_1, \partial_n v_h)_\Gamma$ and $\frac{\gamma}{h}(\llbracket \partial_n u \rrbracket, \llbracket \partial_n v_h \rrbracket)_{\mathcal{F}_h^{int}}$. Effectively is the terms adding zero because of the regularity $u \in H^4(\Omega)$ and the boundary condition $\partial_n u = g_1$. Finally we have

$$
\begin{aligned}
(\Delta^2 u, v_h)_\Omega =& (\Delta u, \Delta v_h)_{\mathcal{T}_h} + (g_2, v_h)_\Gamma - (g_1, \Delta v_h)_\Gamma + \frac{\gamma}{h}(g_1, \partial_n v_h)_\Gamma \\
&- (\llbracket \partial_n u \rrbracket, \{\!\{\Delta v_h\}\!\})_{\mathcal{F}_h^{int}} - (\{\!\{\Delta u\}\!\}, \llbracket \partial_n v_h \rrbracket)_{\mathcal{F}_h^{int}} + \frac{\gamma}{h}(\llbracket \partial_n u \rrbracket, \llbracket \partial_n v_h \rrbracket)_{\mathcal{F}_h^{int}} \\
&- (\Delta u, \partial_n v_h)_\Gamma - (\partial_n u, \Delta v_h)_\Gamma + \frac{\gamma}{h}(\partial_n u, \partial_n v_h)_\Gamma.
\end{aligned}
$$

and the proof is complete. □

We will now assemble the Laplace CIP formulation. The discrete problem is as follows:

$$
\text{Find } u_h \in V_h \text{ such that } a^L(u_h, v_h) = l_h^L(v_h) \quad \forall v_h \in V_h. \tag{3.18}
$$

The corresponding bilinear and linear form is defined as,

$$
\begin{aligned}
a_h^L(u_h, v_h) = &(\alpha u_h, v_h)_\Omega + (\Delta u_h, \Delta v_h)_\Omega \\
&- (\{\!\{\Delta u_h\}\!\}, [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}} - (\{\!\{\Delta v_h\}\!\}, [\![\partial_n u_h]\!])_{\mathcal{F}_h^{int}} + \frac{\gamma}{h}([\![\partial_n u_h]\!], [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}} \\
&- (\Delta u_h, \partial_n v_h)_\Gamma - (\partial_n u_h, \Delta v_h)_\Gamma + \frac{\gamma}{h}(\partial_n u_h, \partial_n v_h)_\Gamma \\
l_h^L(v_h) = &(f, v_h)_\Omega - (g_2, v_h)_\Gamma - (g_1, \Delta v_h)_\Gamma + \frac{\gamma}{h}(g_1, \partial_n v_h)_\Gamma
\end{aligned}
\tag{3.19}
$$

With the corresponding energy norms

$$
\begin{aligned}
\|v\|_{a_h^L}^2 &= \alpha\|v\|_\Omega^2 + \|\Delta v\|_\Omega^2 + \|h^{-\frac{1}{2}}[\![\partial_n v]\!]\|_{\mathcal{F}_h^{int}}^2 + \|h^{-\frac{1}{2}}\partial_n v\|_\Gamma^2, \quad v \in V_h \\
\|v\|_{a_h^L,*}^2 &= \|v\|_{a_h}^2 + \|h^{\frac{1}{2}}\{\!\{\partial_{nn} v\}\!\}\|_{\mathcal{F}_h^{int}}^2 + \|h^{\frac{1}{2}}\partial_{nn} v\|_\Gamma^2, \quad v \in V \oplus V_h.
\end{aligned}
\tag{3.20}
$$

Remark. Again, note that we have the consistent terms $(\{\!\{\Delta u_h\}\!\}, [\![\partial_n v_h]\!])_{\mathcal{F}_h^{int}}$ and $(\Delta u_h, \partial_n v_h)_\Gamma$ naturally appearing in the derivation. Also recall the symmetry terms, $([\![\partial_n u_h]\!], \{\!\{\Delta v_h\}\!\})_{\mathcal{F}_h^{int}}$ and $(\partial_n u_h - g_1, \Delta v_h)_\Gamma$, with corresponding Nitsche penalty terms, $\frac{\gamma}{h}(\partial_n u_h - g_1, \partial_n v_h)_\Gamma$ and $\frac{\gamma}{h}([\![\partial_n u_h]\!], [\![\partial_n v_h]\!]_{\mathcal{F}_h^{int}}$, thus making the bilinear form $a^L(\cdot,\cdot)$ symmetric and well-posed.

3.2.3 Comments and earlier work

It should be noted that the Hessian formulation has a substantial limitation in that it is only valid for homogeneous Neumann conditions. This constraint arises from the challenges associated with imposing g_1 via the tangential derivative terms in Equation (3.13) during the proof of Lemma 3.1. From a physical perspective, this is not problematic as it aligns with the boundary conditions of the original CH problem (1.3). However, from the standpoint of numerical validation, the homogeneous Neumann condition enforces strict rules on the design of manufactured solutions on arbitrary domains. One way to fix this is to enforce tangential derivatives of g_1, i.e., inserting $\partial_n u = g_1$ for $(\partial_{t_i}(\partial_n u), \partial_n v)_\Gamma$ into (3.13). A downside with this method is that we must require g_1 in $H^{\frac{3}{2}}(\Gamma)$. Consequently, the examples illustrated in section 4.6 are only demonstrated on simple domains. This particular constraint does not apply to the Laplace formulation.

The Hessian formulation is well investigated by Susanne Brenner in several papers for [53, 54, 55] with a corresponding analysis and numerical validation. Similarly, variants of the Laplace formulation can be found here [61, 57]. In these article there also is good theoretical and experimental evidence that both formulation have the following expected a priori estimates. Let $u \in H^s(\Omega)$ for $s \geqslant \frac{5}{2} + \varepsilon$, and $u_h \in V_h$ of order $k \geqslant 2$. Then with $r = \min\{s, k+2\}$ the a priori estimates are

$$
\begin{aligned}
\|u - u_h\|_{a_h,*} &\lesssim h^{r-2}\|u\|_{r,\Omega} \\
\|u - u_h\|_\Omega &\lesssim h^{r-\max\{0,3-k\}}\|u\|_{r,\Omega}
\end{aligned}
$$

Be aware that the $\|\cdot\|_\Omega$ norm estimates is suboptimal for $k = 2$. It is worth noting that technically is the interior regularisation equivalent to do a Nitsche's method in all interior boundaries of the elements with boundary conditions of each element weakly imposed to zero. Thus, we expect the penalty parameter γ to be the same interior and exterior elements. Let where $k \geqslant 2$ is the polynomial order, then for the Hessian formulation is it theoretically proven that $\gamma = 2k(k-1)$ [54, 53]. However, we still prefer to experimentally verify the best parameter.

3.3 Note on the biharmonic mixed formulation

It is easy to see that the biharmonic problem can be rewritten into an equivalent mixed formulation , that is, to find $\sigma, \tau \in H^2(\Omega)$ s.t.

$$\begin{aligned} \Delta\sigma &= f \quad \text{in } \Omega \\ \sigma &= \Delta u \text{ in } \Omega \\ \partial_n \sigma &= g_1 \text{ on } \Gamma \\ \partial_n u &= g_2 \text{ on } \Gamma \end{aligned}$$

The goal is to obtain an useful weak formulation. Using Greens theorem on the first equation we get,

$$(\sigma, v)_\Omega = (\nabla u, \nabla v)_\Omega - (\nabla_n u, v)_\Gamma.$$

Similarly for the second equation we obtain

$$(\nabla\sigma, \nabla\varphi)_\Omega - (\partial_n \sigma, \varphi)_\Gamma = (f, \varphi)_\Omega$$

Putting it all together we have the following mixed weak formulation; Find $(u, \sigma) \in H^1(\Omega) \times H^1(\Omega)$ s.t.

$$\begin{aligned} (\nabla u, \nabla v)_\Omega - (\sigma, v)_\Omega &= (g_1, v)_\Gamma \quad \forall v \in H^1(\Omega) \\ (\nabla\sigma, \nabla\varphi)_\Omega &= (f, \varphi)_\Omega + (g_2, \varphi)_\Gamma \quad \forall\varphi \in H^1(\Omega) \end{aligned}$$

Now we want to relate this formulation to the abstract saddle point problem (SPP). Let $V = H^1(\Omega)$ and $W = H^1(\Omega)$ be Hilbert spaces and define the bilinear form $a : V \times V \to \mathbb{R}$ and $b : V \times W \to \mathbb{R}$ s.t. $a(\sigma, v) = -(\sigma, v)_\Omega$ and $b(u, v) = (\nabla u, \nabla v)_\Omega$. We also may define the linear forms, $G, F : V \to \mathbb{R}$ s.t. $G(v) = (g_1, v)_\Gamma$ and $F(\varphi) = (f, \varphi)_\Omega + (g_2, \varphi)_\Gamma$.

Hence, we obtain the following SPP. Find $(u, \sigma) \in V \times W$ s.t.

$$\begin{cases} a(\sigma, v) + b(u, v) & = G(v) \quad \forall v \in V \\ b(u, \varphi) & = F(\varphi) \quad \forall\phi \in W \end{cases}$$

This is useful since we can now apply standard saddle point theory to do an analysis for the problem. We will see that it is now easier to handle the boundary constraints naturally, but with the cost of a more challenging time discretization procedure. For more information about the biharmonic mixed formulation, see [89, 76]. However, in this thesis is the focus on solving the biharmonic equation avoiding the mixed formulation using the CIP formulation, which does in fact handle the downsides with the SPP problem.