博士論文

論文題目: Mathematical foundation of Isogeometric Analysis for evolution problems
(発展問題に対する Isogeometric Analysis の数学的基礎)

氏名: 上田 祐暉

Contents

Introduction

Isogeometric Analysis (IGA) [14] is one of the numerical methods for partial differential equation (PDE). It is regarded as the Galerkin method using the NURBS (Non-uniform rational B-spline) for basis functions. Consequently, IGA is understood as the one of the finite element method (FEM) in an expanded sense. To compare the IGA with standard FEM, we consider the following PDE as an example,

$$\begin{cases} L(u) &= f \quad \text{in } \Omega, \\ \quad u &= g_D \quad \text{on } \Gamma, \end{cases} \tag{1}$$

where L is a partial differential operator, $\Omega \subset \mathbb{R}^d$ is a given bounded domain with Lipschitz boundary $\Gamma := \partial\Omega$. Letting V_h be an approximate finite dimensional space. We find the approximate solution $u_h \in V_h$ such that

$$u_h := \sum_{i=1}^{N} u_i \phi_i, \tag{2}$$

where we let $\{\phi_i\}_{i=1}^N$ be a set of basis function of V_h. Following the method of mean weighted residuals, we find $u_h \in V_h$ such that the weighted residue $(L(u_h) - f)\, w_i$ satisfies

$$\int_\Omega (L(u_h) - f)\, w_i = 0 \tag{3}$$

for all test function w_i. Here we can choose $w_i = \phi_i$, and then this method is called the Galerkin method, that is, we find $u_h \in V_h$ such that

$$(L(u_h), v_h)_{L^2(\Omega)} = (f, v_h)_{L^2(\Omega)} \text{ for all } v_h \in V_h. \tag{4}$$

Now we introduce the standard FEM and IGA. We construct partitions of the domain Ω and finite dimensional subspace $V_h \subset H^1(\Omega)$. The standard FEM employs triangulation and Lagrange finite element. For the basis functions $\{\phi_i\}_{i=1}^N$, there exist the nodes $\{\mathbf{x}_i\}_{i=1}^N$ such that $\phi_i(\mathbf{x}_i) = \delta_{ij}$, that is, the Lagrange interpolation can be defined. The standard FEM can treat complex geometry efficiently. However, it is necessary to approximate the exact domain to polygonal (or polyhedral) domains, or to fill the gaps between the exact domain and approximate one, because each element is often defined by the affine transform from the d-simplex or d-cube (see [13] or [10] for more detail).

On the other hand, the IGA describes the computational domain with NURBS [29]. This gives a partition of domain Ω immediately. Further we can get NURBS basis functions on Ω, and then we employ them for Galerkin method. The computer-aided design (CAD) system uses the NURBS for designing the geometric models of object. Therefore, IGA can treat the geometric models of industrial products directly. Further, NURBS can represent complex geometries or smooth approximate solutions of PDE using only a few degrees of freedom. This advantage can reduce the computer-storage cost in numerical simulations for large system. However, the set of NURBS basis functions does not satisfy the interpolation property, therefore the strong imposition of Dirichlet boundary condition is difficult in IGA.

It is easy to apply IGA for computational simulation instead of FEM; however, their mathematical treatment has some different points. Recent mathematical studies for IGA can be found in [9] and the references given there. They are mainly devoted to steady-state problems. The aim of our study is to consider the application of IGA to time-depending problems.

In this paper, two chapters provide some results of the application of NURBS in temporal representation, and spatial semi-discretization, respectively. In Chapter 1, the temporal representation with NURBS which gives smooth solution for PDE is discussed. The space-time computation technique with continuous representation in time (ST-C) [40] was introduced for reducing the computer-storage cost in the numerical computation with space-time (ST) method. The core technique for ST computational analysis has been applied successfully to many classes of fluid mechanics problems. The numerical solutions of ST methods are usually expressed by the discontinuous basis function in temporal variable. ST-C gives the option to the core technique, and this makes the computational result smooth in temporal variable.

Two versions of the ST-C method have been offered. In this chapter, we establish the mathematical justification of ST-C with successive projection technique (ST-C-SPT). This method extracts the continuous representation successively from a numerical solution that is previously computed and is discontinuous in time. It was originally proposed to represent the numerical solution with few degrees of freedoms and saving the computer-storage costs. We show that ST-C-SPT is stable and therefore the error estimate can be derived. Note that we consider the SPT for X-valued functions, where X is a real-valued Banach space.

Chapter 1 is composed of five sections. In Section 1.1, we give more detail about ST computational method and our motivation to study ST-C-SPT. Section 1.2 contains a brief summary of B-splines and the algorithm of SPT. Moreover, we show the mathematical formulation of SPT. Let

$$\Xi = \{\underbrace{t_0, \ldots, t_0}_{p+1 \text{ times}}, t_1, t_2, \ldots, t_{N-1}, \underbrace{t_N, \ldots, t_N}_{p+1 \text{ times}}\} \tag{5}$$

be p-open knot vector and $\widehat{B}_{i,p}, 1 \le i \le N+p$ be p-th degree B-spline basis functions. Without loss of generality, we let $t_0 = 0$ and $t_N = 1$. We define

$$S_n := \left\{ \sum_{i=1}^{n+p} x_i \ \widehat{B}_{i,p}(t)\Big|_{[t_0,t_n]} : x_i \in X \text{ for } i = 1 \cdots, n+p \right\}, \tag{6}$$

and then the mathematical formulation of SPT is described as follows: for given $\Pi_n(f) = \sum_{i=1}^{n+p} x_i^n \ \widehat{B}_{i,p}\Big|_{[t_0,t_n]} \in S_n$, find $\Pi_{n+1}(f) = \sum_{i=1}^{n+p+1} x_i^{n+1} \ \widehat{B}_{i,p}\Big|_{[t_0,t_{n+1}]} \in S_{n+1}$ such that

$$\int_{J_1} (\Pi_1(f) - f)\, \widehat{B}_{i,p} \ dt = 0 \text{ for all } i = 1, \cdots, p+1 \tag{7}$$

for $n = 0$, and

$$\begin{cases} x_i^{n+1} = x_i^n & \text{for } i = 1, \cdots, n \\ \displaystyle\int_0^{t_n} \widehat{B}_{i,p} \left(\Pi_{n+1}(f) - \Pi_n(f)\right) dt + \int_{t_n}^{t_{n+1}} \widehat{B}_{i,p} \left(\Pi_{n+1}(f) - f\right) dt = 0 & \text{for } i = n+1, \cdots, n+p \end{cases} \tag{8}$$

for $n = 1, \cdots, N-1$. Furthermore, this formulation has an alternative expression with matrix form. We set

$$M_{n+1} = \left(\int_{J_{n+1}} \widehat{B}_{i+n,p}(t) \widehat{B}_{j+n,p}(t) \ dt \right)_{1 \le i,j \le p+1} \in \mathbb{R}^{(p+1)\times(p+1)}, \tag{9}$$

$$F_{n+1} = \left(\int_{I_{n+1}} \widehat{B}_{i+n,p}(t) f(t) \ dt \right)_{1 \le i \le p+1} \in X^{p+1}, \tag{10}$$

$$\mathbf{x}_{n+1} = \left(x_{i+n}^{n+1} \right)_{1 \le i \le p+1} \in X^{p+1}, \tag{11}$$

for $n = 0, 1, \ldots, N-1$. Moreover, we introduce shift matrices U and L defined by

$$U = (\delta_{i+1,j})_{1 \le i,j \le p+1}, \qquad L = U^{\mathrm{T}}. \tag{12}$$

Then we can rewrite the algorithm: first we find $\mathbf{x}_1$ by

$$M_1 \mathbf{x}_1 = F_1, \tag{13}$$

and for $n = 1, \ldots, N$, we obtain

$$M_{n+1}\mathbf{x}_{n+1} = UM_nLU\mathbf{x}_n + F_{n+1}. \tag{14}$$

In Section 1.3, we prove the stability of Π_N under some assumptions; there exists a constant $C_1 > 0$ such that

$$\|\Pi_N(f)\|_{L^\infty(0,1;X)} \le C_1\|f\|_{L^\infty(0,1;X)} \tag{15}$$

for any $f \in L^\infty(0,1;X)$. Further we show that the stability yields the error estimate,

$$\|f - \Pi_N(f)\|_{L^\infty(0,1;X)} \le C_2 h^{\widetilde{k}} |f|_{W^{\widetilde{k},\infty}(0,1;X)} \tag{16}$$

for any $f \in W^{k,\infty}(J;X)$ and $k \ge 1$, where $\widetilde{k} = \min\{k, p+1\}$. The proof of theorems and numerical examples are given in Section 1.4.

In Chapter 2, we study the Nitsche method for parabolic problems. Let us consider the imposition of Dirichlet boundary condition, which is an important component of the well-posed problem. The traditional FEM employs the Lagrange interpolation, and the value at the boundary nodal points give the approximate boundary condition strongly. This strong imposition of Dirichlet boundary condition may causes numerical instability. For instance, numerical oscillations may appear if the boundary data is discontinuous. Furthermore, the Lagrange interpolation can not be constructed in IGA, because the NURBS basis functions do not satisfy the interpolation property.

It is natural to try to propose the non-standard imposition of Dirichlet boundary condition. One such study was given by Nitsche [27]. In his work, the penalty term is introduced, and the Dirichlet boundary condition is imposed weakly. We emphasize that the classical penalty method is not consistent; however, the Nitsche method satisfies consistency, that is, the Galerkin orthogonality follows. Therefore, combining this and the inf-sup condition yields the quasi-optimal error estimate in spatial semi-discretization.

This chapter is organized as follows. Section 2.1 is intended to describe the details of introduction. In Section 2.2, we review the Nitsche's classical paper briefly, and the Nitsche method for elliptic problems is explained. We mention the derivation of the coercivity for a finite dimensional approximation of the weak formulation, by taking the penalty parameter large enough. Section 2.3 provides the basic results about the weak formulation of the parabolic problems. We review the Banach-Nečas-Babuška theorem, which plays an important role in this chapter. Further, we show that the Banach-Nečas-Babuška theorem gives a unique weak solution of the advection-diffusion-reaction problem under some conditions. In Section 2.4, we introduce the classical FEM and IGA, and we mention our basic assumptions. Moreover, we check that these assumptions are satisfied for both of FEM and IGA. Section 2.5 contains our main results; first we let

$$V := \left\{ v \in H^1(\Omega) : v|_K \in H^2(K) \text{ for all } K \in \mathcal{T}_h \right\}, \tag{17}$$

where $\mathcal{T}_h$ is mesh for spatial semi-discretization. Moreover we let $V_h \subset V$ is finite dimensional subspace, and

$$X_V := \left\{ v \in W^{1,2,2}\left(0,T; H^1(\Omega), L^2(\Omega)\right) : v(t) \in V \text{ for a.e. } t \in (0,T) \right\}, \tag{18}$$

$$X_h := H^1\left(0,T;V_h\right),\ Y_h := L^2(0,T;V_h) \times V_h. \tag{19}$$

Then the Nitsche method for advection-diffusion-reaction problem is given as follows, find $u_{\varepsilon,h} \in X_h$ such that

$$b_{\varepsilon,h}(u_{\varepsilon,h}, \mathbf{v}_h) = F(\mathbf{v}_h) \text{ for all } \mathbf{v}_h := (v_h, \widetilde{v}_h) \in Y_h, \tag{20}$$

where

$$b_{\varepsilon,h}(w, \mathbf{v}_h) := \int_0^T \left((w', v_h)_{L^2(\Omega)} + a_\varepsilon(t; w, v_h) \right) dt + (w(0), \widetilde{v}_h)_{L^2(\Omega)} \tag{21}$$

for all $w \in X_V$ and $\mathbf{v}_h \in Y_h$, $a_\varepsilon(t;\cdot,\cdot) : V \times V_h \to \mathbb{R}$ is defined by

$$\begin{aligned} a_\varepsilon(t; w, v_h) &:= (A(t)w, v_h)_{L^2(\Omega)} - \sum_{E \in \mathcal{E}_h^e} (\mathbf{n} \cdot \mu \nabla v_h, w)_{L^2(E)} \\ &\qquad - (\mathbf{a} \cdot \mathbf{n} v_h, w)_{L^2(\Gamma_{\text{in}})} + \langle \varepsilon(t) v_h, w \rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} \\ &= \langle \widehat{A}(t) w, v_h \rangle_{(H^1(\Omega))^*, H^1(\Omega)} - \sum_{E \in \mathcal{E}_h^e} \left((\mathbf{n} \cdot \mu \nabla w, v_h)_{L^2(E)} + (\mathbf{n} \cdot \mu \nabla v_h, w)_{L^2(E)} \right) \\ &\qquad - (\mathbf{a} \cdot \mathbf{n} v_h, w)_{L^2(\Gamma_{\text{in}})} + \langle \varepsilon(t) v_h, w \rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} \end{aligned} \tag{22}$$

for all $w \in V$ and $v_h \in V_h$,

$$F(\mathbf{v}_h) = \int_0^T \mathcal{F}(t; v_h)dt + (u_0, \widetilde{v}_h)_{L^2(\Omega)} \tag{23}$$

for all $\mathbf{v}_h := (v_h, \widetilde{v}_h) \in Y_h$ and

$$\mathcal{F}(t; v_h) := (f, v_h)_{L^2(\Omega)} - \sum_{E \in \mathcal{E}_h^e} (\mathbf{n} \cdot \mu \nabla v_h, g_D)_{L^2(E)} - (\mathbf{a} \cdot \mathbf{n} v_h, g_D)_{L^2(\Gamma_{\text{in}})} + \langle \varepsilon(t) v_h, g_D \rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} \tag{24}$$

for all $v_h \in V_h$. Here $\varepsilon(t) : V_h \to H^{-1/2}(\Gamma)$ is given by

$$\langle \varepsilon(t) v_h, w \rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} := \varepsilon_0 \sum_{E \in \mathcal{E}_h^e} h_E^{-1} (w, v_h)_{L^2(E)} \tag{25}$$

for all $w \in V$ and $v_h \in V_h$, where ε_0 is a suitable constant. Then, we prove the inf-sup condition. That is, there exists a positive constant β such that

$$\inf_{0 \neq x_h \in X_h} \sup_{0 \neq \mathbf{y}_h \in Y_h} \frac{b_{\varepsilon,h}(x_h, \mathbf{y}_h)}{\|x_h\|_{X_h} \|\mathbf{y}_h\|_{Y_h}} \geq \beta. \tag{26}$$

This yields that we can apply the the Banach-Nečas-Babuška theorem, that is, the equation (20) has a unique solution $u_{\varepsilon,h}$. Furthermore, the inf-sup condition and the Galerkin orthogonality

$$b_{\varepsilon,h}(u - u_{\varepsilon,h}, \mathbf{v}_h) = 0 \text{ for all } \mathbf{v}_h \in Y_h. \tag{27}$$

give the quasi-optimal error estimate

$$\|u - u_{\varepsilon,h}\|_{X_h} \leq C \|u - w_h\|_{X_V} \tag{28}$$

for all $w_h \in X_h$. Combining this and interpolation (or quasi-interpolation) error estimates yields the following error estimate under the assumption $u \in X_{\ell,m} := W^{1,2,2}\left(0, T; H^\ell(\Omega), H^m(\Omega)\right)$; there exists a positive constant C such that

$$\|u - u_{\varepsilon,h}\|_{X_h}^2 \leq C \left(\int_0^T \left(h^{2(\ell-1)} \|u\|_{H^\ell(\Omega)}^2 + h^{2m} \|u'\|_{H^m(\Omega)}^2 \right) dt + h^{2j} \|u(0)\|_{H^k(\Omega)}^2 \right), \tag{29}$$

where $j := \min\{\ell, m\}$. Furthermore, we study the error estimate for full discrete problem in Section 2.6. We apply the implicit Euler scheme, and extend the approximate solution into the piecewise constant function in temporal variable. Then there exists a positive constant C such that for the extended approximate solution $u_{\varepsilon,h,\tau}$,

$$\begin{aligned} \|u - u_{\varepsilon,h,\tau}\|_{L^2(0,T;L^2(\Omega))}^2 \quad &\leq C \left(\int_0^T \left(h^{2(\ell-1)} \|u\|_{H^\ell(\Omega)}^2 + h^{2m} \|u'\|_{H^m(\Omega)}^2 \right) dt + h^{2j} \|u(0)\|_{H^k(\Omega)}^2 \right) \\ &\quad + C\tau^2 \|u'_{\varepsilon,h}\|_{H^1(0,T;L^2(\Omega))}^2 + \frac{T}{4\alpha^2} \tau^2 \|T_1 u''_{\varepsilon,h}\|_{L^\infty\left(0,T;V_h^*\right)}^2, \end{aligned} \tag{30}$$

where $j := \min\{\ell, m\}$. Especially, we let $k = 1$ and $\tau = h/10$, then we have

$$\|u - u_{\varepsilon,h,\tau}\|_{L^2(0,T;L^2(\Omega))} \approx Ch. \tag{31}$$

The numerical examples in Section 2.7 show that the rate of convergence is approximately equal to the unity that is actually expected by (31).

Acknowledgement

I would like to thank my supervisor, Professor Norikazu Saito, for his helpful advice and collaboration. I would also like to thank Professor Kenji Takizawa at Waseda University and Professor Tayfun E. Tezduyar at Rice University who brought the subject of Chapter 1 to my attention. I am grateful to the members of T*AFSM for their hospitality during my visit. I also thank to the secondary supervisor in the Program for Leading Graduate Schools, Professor Masahiro Yamamoto.

This work was supported by the Program for Leading Graduate Schools, MEXT, Japan.

Chapter 1

Analysis of space-time computation techniques with continuous representation in time: the successive projection technique

1.1 Introduction

The deforming-spatial-domain/stabilized space–time (DSD/SST) method [49, 50, 51] was developed for the computation of flows with moving boundaries and interfaces (MBIs), including fluid–structure interactions (FSIs). In the DSD/SST method, which is also called ST-SUPS because it is based on the streamline-upwind/Petrov–Galerkin (SUPG) [12] and pressure-stabilizing/Petrov–Galerkin (PSPG) [49] stabilizations, the discontinuous Galerkin method (DGM) is applied in time. The weak form of the governing equation is considered on one space–time (ST) "slab" at a time, where the "slab" is the slice of the ST domain between two time levels. The temporal variable is discretized by *piecewise* smooth basis functions, which are discontinuous from one ST slab to another, and consequently approximate solutions are discontinuous in time (see [7] for more details).

The ST-SUPS method and ST variational multiscale (ST-VMS) method [38, 39], which is the VMS (particularly residual-based VMS [24, 3]) version of the DSD/SST method, have an advantage in accuracy; therefore, they are desirable also when there is no MBI. In the computation of FSI and MBI problems, the arbitrary Lagrangian–Eulerian (ALE) method [26] and ALE-VMS method [4, 34, 8] are more commonly used moving-mesh methods (see [8] and references therein, and the references cited in [28]). The ST-SUPS and ST-VMS methods have also been applied successfully to many classes of fluid mechanics problems, including spacecraft parachute FSIs, wind-turbine aerodynamics, flapping-wing aerodynamics, cardiovascular fluid mechanics, spacecraft aerodynamics, thermo-fluid analysis of ground vehicles and their tires, thermo-fluid analysis of disk brakes, flow-driven string dynamics in turbomachinery, flow analysis of turbocharger turbines, flow around tires with road contact and deformation, ram-air parachutes, and compressible-flow parachute aerodynamics (see [7] and references therein, and the references cited in [28]).

The ST-SUPS and ST-VMS are core methods in the ST computational analysis (STCA), but computations of various classes of problems are performed with the core methods and their integration with other, special ST methods. For example, the ST topology change (ST-TC) method [42, 41] accomplishes computation of the problems with contact between moving interfaces in FSI, and the ST slip interface (ST-SI) method [45, 43] can be applied to FSI problems with slip interfaces such as spinning structures. Furthermore, we note that the ST isogeometric analysis (ST-IGA) [38, 36, 46] proposes using nonuniform rational basis spline (NURBS) for the representation of a moving domain and ST discretization. These methods have been successfully applied to the flow analysis of a turbocharger turbine [46, 28], a ram-air parachute [47], and a heart valve [48].

The IGA [25], developed originally by using B-spline or NURBS basis functions in space, has been widely applied in many fields of computational mechanics. It provides "smooth" approximate solutions of the target partial differential equations (PDEs) using only a few degrees of freedom (DOF) in comparison with the standard finite element method (FEM) and DGM. Moreover, it provides a more accurate representation of computational domains with complex shapes, that is, the geometric representation of a computational domain generated by a CAD system is handled directly. See [14] for more details. So far B-splines and NURBS have been used to discretize the spatial variables in IGA, and abundant mathematical studies have been reported (see [9] for a survey).

The ST-IGA with IGA basis functions in time has the advantage that higher-order NURBS in time provides a more accurate representation of the motion of a domain (see [38, 39, 36, 35, 37]). The ST/NURBS mesh update method (STNMUM) [36, 35, 37, 44] shows that the ST-IGA also provides more efficiency in representation of the motion and deformation of a mesh and remeshing. The application of NURBS in a temporal representation also enables continuous basis functions to be used for temporal discretization. In the ST computation techniques with continuous representation in time (ST-C) [40], we obtain a globally continuous or smooth representation in time. In general, if approximate solutions are expressed by globally continuous basis functions, the resulting system of algebraic equations becomes quite large. The ST-C method has overcome this difficulty by providing some successive projection algorithms. The ST-IGA computations with the STNMUM or ST-C have been demonstrated in many 3D problems, including flapping-wing aerodynamics, separation aerodynamics of spacecraft, wind-turbine aerodynamics, thermo-fluid analysis of ground vehicles and their tires, thermo-fluid analysis of disk brakes, flow-driven string dynamics, and flow analysis of turbocharger turbines (see the references cited in [28]).

Two versions of the ST-C method have been offered. The first is the ST-C with successive projection technique (ST-C-SPT). In the ST-C-SPT, the continuous representation is extracted successively from a numerical solution that is previously computed and discontinuous in time. One of motivations for considering the SPT is to save computer-storage cost, since the B-splines have few DOFs. In the SPT, each projection can be computed while the numerical solution for the next temporal interval is still being computed. The projected solution takes the place of the previously computed solution; therefore, the computed data can be compressed without storing a large amount of time-history data. The second version proposed in [40] is the direct computation technique (ST-C-DCT), which is an algorithm for temporal discretization of PDEs using NURBS basis functions directly. The motivation for considering the DCT is the same as that of the SPT.

The purpose of this study is to establish a mathematical justification of the SPT. That is, we prove its stability and error estimates. To this end, we consider the SPT for X-valued functions, where X is a real-valued Banach space. Note that the SPT was originally described for real scalar-valued functions in [40]. The DCT is also worthy of study; however, we postpone that to future work.

This chapter is organized as follows. In Section 1.2, we recall the notion of B-spline basis functions and state the algorithmic features of the SPT. Our results on the stability and error estimates in the $L^\infty(0,T;X)$ norm are presented in Section 1.3. Section 1.4 presents the proof of the results. Finally, we draw some conclusions in Section 1.5.

Notation. Throughout this chapter, we use the following notation:

- X denotes a (real-valued) Banach space equipped with the norm $\|\cdot\|_X$;
- $J=(0,T)$ for $T>0$;
- $p \geq 1$ is an integer.

For $1 \leq r \leq \infty$ and $0 \leq T_0 < T_1$, the space $L^r(T_0,T_1;X)$ denotes a Bochner space equipped with the norm

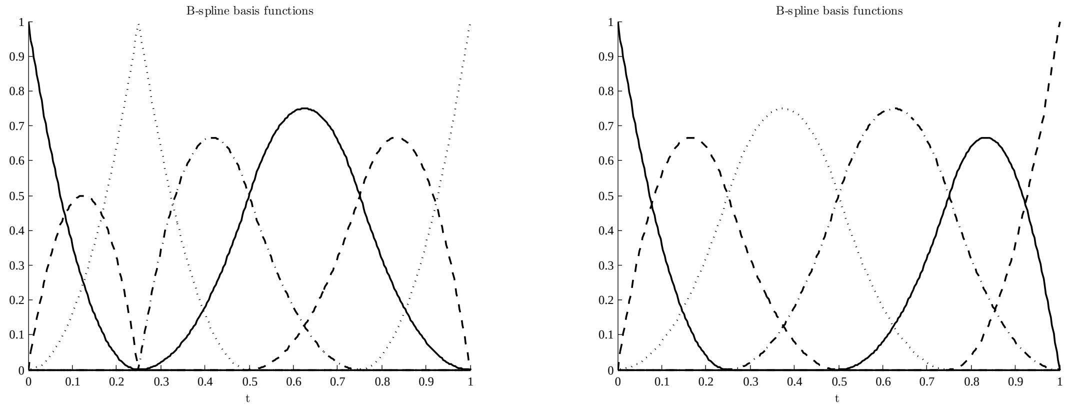
$$\|v\|_{L^r(T_0,T_1;X)} = \begin{cases} \left(\displaystyle\int_{T_0}^{T_1} \|v(t)\|_X^r \, dt\right)^{1/r} & (1 \leq r < \infty) \\ \operatorname{esssup}_{t\in(T_0,T_1)} \|v(t)\|_X & (r=\infty). \end{cases}$$

See [22, Appendix E.5] for example. We also use the so-called Bochner–Sobolev space $W^{k,\infty}(T_0,T_1;X)$ defined by

$$W^{k,\infty}(T_0,T_1;X) = \left\{ v \in L^\infty(T_0,T_1;X) \mid \frac{\mathrm{d}^i v}{\mathrm{d}t^i} \in L^\infty(T_0,T_1;X),\ i=0,1,\ldots,k \right\},$$

where $\frac{\mathrm{d}}{\mathrm{d}t}$ denotes the weak derivative for t. The space $W^{k,\infty}(T_0,T_1;X)$ is a Banach space with the seminorm

$$|v|_{W^{i,\infty}(T_0,T_1;X)} = \left\| \frac{\mathrm{d}^i v}{\mathrm{d}t^i} \right\|_{L^\infty(T_0,T_1;X)} = \operatorname{esssup}_{t\in(T_0,T_1)} \left\| \frac{\mathrm{d}^i v}{\mathrm{d}t^i}(t) \right\|_X$$



$\Xi = \{0, 0, 0, 1/4, 1/4, 1/2, 3/4, 1, 1, 1\}$ $\qquad$ $\Xi = \{0, 0, 0, 1/4, 1/2, 3/4, 1, 1, 1\}$

Figure 1.1: B-spline basis functions.

and the norm

$$\|v\|_{W^{k,\infty}(T_0,T_1;X)} = \max_{0 \le i \le k} |v|_{W^{i,\infty}(T_0,T_1;X)}.$$

For a matrix $A = (a_{ij})_{1 \le i,j \le N} \in \mathbb{R}^{N \times N}$,

$$\|A\|_\infty = \sup_{\boldsymbol{x} \in \mathbb{R}^N} \frac{\|A\boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} = \max_{1 \le i \le N} \sum_{j=1}^{N} |a_{ij}|$$

denotes its matrix norm induced by the ∞ norm $\|\boldsymbol{x}\|_\infty = \max_{1 \le i \le N} |x_i|$ for $\boldsymbol{x} = (x_i)_{1 \le i \le N} \in \mathbb{R}^N$. Moreover, for $F = (F_i)_{1 \le i \le N} \in X^N$, we write

$$\|F\|_\infty = \max_{1 \le i \le N} \|F_i\|_X.$$

1.2 SPT using B-splines

1.2.1 Review of B-spline basis functions

We introduce the *knot vector*

$$\Xi = \{\xi_1, \xi_2, \dots, \xi_m\},$$

where $0 = \xi_1 \le \xi_2 \le \cdots \le \xi_m = T$. Note that repetitions of knots are allowed. Suppose that $m \ge p + 2$. Then, the *univariate B-spline basis functions*

$$\widehat{B}_{i,p} : J \to \mathbb{R}, \quad i = 1, \dots, m - p - 1, \tag{1.1}$$

of degree p associated with the knot vector Ξ are successively defined by the Cox–de Boor algorithm (see [15, 17]).

Definition 1. Set

$$\widehat{B}_{i,0}(t) = \begin{cases} 1 & \text{if } t \in [\xi_i, \xi_{i+1}) \\ 0 & \text{otherwise} \end{cases} \qquad (i = 1, 2, \dots, m - 1). \tag{1.2}$$

Then, for $q = 1, 2, \dots, p$, set

$$\widehat{B}_{i,q}(t) = \frac{t - \xi_i}{\xi_{i+q} - \xi_i} \widehat{B}_{i,q-1}(t) + \frac{\xi_{i+q+1} - t}{\xi_{i+q+1} - \xi_{i+1}} \widehat{B}_{i+1,q-1}(t) \quad (i = 1, 2, \dots, m - q - 1). \tag{1.3}$$

Herein, $0/0$ should be replaced by 0.

B-spline basis functions are nonnegative

$$\widehat{B}_{i,p}(t) \geq 0 \quad (1 \leq i \leq m-p-1,\ t \in J), \tag{1.4}$$

and $\widehat{B}_{i,p} = 0$ for $t \notin [\xi_i, \xi_i + p + 1]$. Furthermore, if the knot vector Ξ is p-open, that is, $m \geq 2p+2$, the knots satisfy $\xi_1 = \cdots = \xi_{p+1}$ and $\xi_{m-p} = \cdots = \xi_m$, then the B-spline basis functions form a partition of unity:

$$\sum_{i=1}^{m-p-1} \widehat{B}_{i,p}(t) = 1 \qquad (t \in J). \tag{1.5}$$

B-spline basis functions are suitably smooth. To state them more concretely, we introduce an alternative representation of Ξ,

$$\Xi = \{\underbrace{t_0,\ \ldots,\ t_0}_{m_0 \text{ times}},\ \underbrace{t_1,\ \ldots,\ t_1}_{m_1 \text{ times}}, \ldots,\ \underbrace{t_N,\ \ldots,\ t_N}_{m_N \text{ times}}\}, \tag{1.6}$$

where $0 = t_0 < t_1 < \cdots < t_N = T$. Therein, by m_n, we denote the multiplicity of t_n. Assume that $m_n \leq p+1$ for all knots. Then, $\widehat{B}_{i,p}(t)$ is $p - m_n$ times continuously differentiable at a node t_n. Further details about these fundamental facts are explained in [29, Section 2].

1.2.2 Finite-dimensional subspace $S_{p,\Xi}(X)$ spanned by B-spline functions

In this study, we only consider the following special p-open knot vector,

$$\Xi = \{\underbrace{t_0, \ldots, t_0}_{p+1 \text{ times}}, t_1, t_2, \ldots, t_{N-1}, \underbrace{t_N, \ldots, t_N}_{p+1 \text{ times}}\}. \tag{1.7}$$

For the degree $p \geq 1$ and the knot vector Ξ defined by (1.7), we introduce a finite-dimensional subspace of X by setting

$$S_{p,\Xi}(X) = \left\{ \sum_{i=1}^{N+p} x_i \widehat{B}_{i,p}(t) \mid x_i \in X \ (1 \leq i \leq N+p) \right\}. \tag{1.8}$$

Then, $f \in S_{p,\Xi}(X)$ is a polynomial of degree p in $[t_{n-1}, t_n]$, $1 \leq i \leq N$, and is $p-1$ times continuously differentiable at nodes t_n, $1 \leq n \leq N-1$.

1.2.3 Algorithm of SPT

Set

$$I_n = [t_{n-1}, t_n], \quad J_n = [t_0, t_n] \quad (1 \leq i \leq N) \tag{1.9}$$

and define

$$S_n = S_{n,p,\Xi}(X) = \left\{ \sum_{i=1}^{n+p} x_i \widehat{B}_{i,p}(t)|_{J_n} \mid x_i \in X \ (1 \leq i \leq n+p) \right\} \tag{1.10}$$

for $n = 1, \ldots, N$. We note that $S_N = S_{p,\Xi}(X)$.

At this stage, we can state the SPT. Projections

$$\Pi_n : L^2(J_n; X) \to S_n \quad (1 \leq i \leq N) \tag{1.11}$$

are successively defined in the following way. Suppose that we are given

$$f \in L^1(J; X). \tag{1.12}$$

First, take $\Pi_1(f)$ as the standard L^2-projection of $f|_{J_1}$ onto S_1, that is,

$$\int_{J_1} (\Pi_1(f) - f)\, \widehat{B}_{i,p}\, dt = 0 \quad (1 \leq i \leq p+1). \tag{1.13}$$

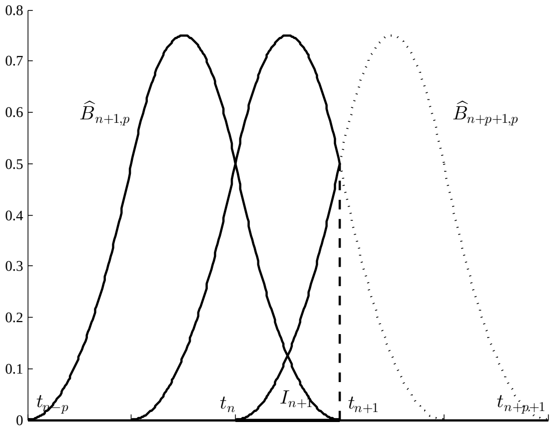


Figure 1.2: Denotations in the SPT algorithm.

Now, for $n \geq 1$, suppose that we are given

$$\Pi_n(f) = \sum_{i=1}^{n+p} x_i^n \widehat{B}_{i,p}|_{J_n} \in S_n. \tag{1.14}$$

Then,

$$\Pi_{n+1}(f) = \sum_{i=1}^{n+p+1} x_i^{n+1} \widehat{B}_{i,p}|_{J_{n+1}} \in S_{n+1} \tag{1.15}$$

is determined by

$$x_i^{n+1} = x_i^n \quad (1 \leq i \leq n) \tag{1.16}$$

and

$$\int_{J_n} \widehat{B}_{i,p} \left(\Pi_{n+1}(f) - \Pi_n(f)\right) dt + \int_{I_{n+1}} \widehat{B}_{i,p} \left(\Pi_{n+1}(f) - f\right) dt = 0$$
$$(n+1 \leq i \leq n+p+1). \tag{1.17}$$

Obviously, each $\Pi_n(f)$ is well defined and we finally obtain $\Pi_N(f) \in S_N = S_{p,\Xi}(X)$.

1.2.4 Alternative expression

It is convenient that we convert the above algorithm into the matrix form. We set

$$M_{n+1} = \left(\int_{J_{n+1}} \widehat{B}_{i+n,p}(t) \widehat{B}_{j+n,p}(t) \; dt \right)_{1 \leq i,j \leq p+1} \in \mathbb{R}^{(p+1)\times(p+1)},$$

$$F_{n+1} = \left(\int_{I_{n+1}} \widehat{B}_{i+n,p}(t) f(t) \; dt \right)_{1 \leq i \leq p+1} \in X^{p+1},$$

$$\mathbf{x}_{n+1} = \left(x_{i+n}^{n+1} \right)_{1 \leq i \leq p+1} \in X^{p+1},$$

for $n = 0, 1, \ldots, N-1$. Moreover, we introduce shift matrices U and L defined by

$$U = (\delta_{i+1,j})_{1 \leq i,j \leq p+1}, \qquad L = U^{\mathrm{T}}.$$

Then we can rewrite the algorithm: first we find $\mathbf{x}_1$ by

$$M_1 \mathbf{x}_1 = F_1,$$

and for $n = 1, \ldots, N$, we obtain

$$M_{n+1} \mathbf{x}_{n+1} = U M_n L U \mathbf{x}_n + F_{n+1}.$$

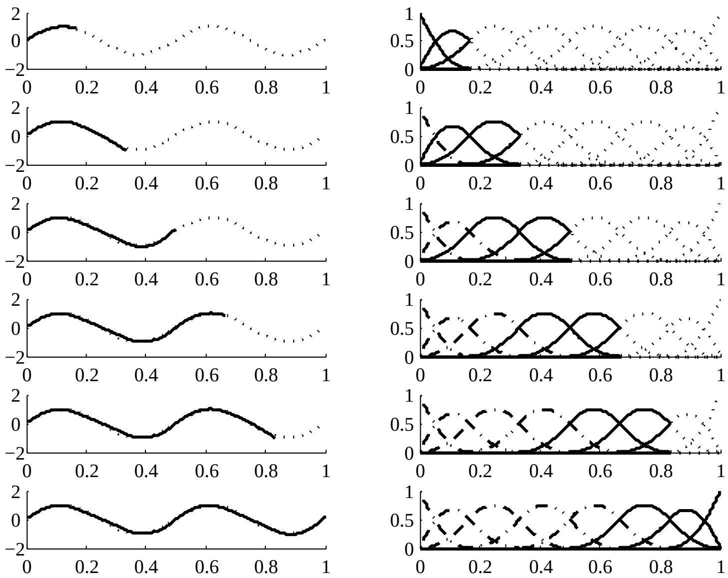


Figure 1.3: Each step of the SPT algorithm and basis functions. The left-hand side represents the projection at each time step. The right-hand side represents basis functions used to compute the projection.

It is apparent that the matrices M_{n+1} are symmetric and positive-definite; in particular, they are invertible. Therefore, (1.13)–(1.17) are equivalently written as

$$\mathbf{x}_1 = M_1^{-1} F_1, \tag{1.18}$$

$$\mathbf{x}_{n+1} = M_{n+1}^{-1} U M_n L U \mathbf{x}_n + M_{n+1}^{-1} F_{n+1}. \tag{1.19}$$

1.3 Stability and error estimates

In this section, we state the main results of this study. Let

$$h_n = |I_n| = t_n - t_{n-1}, \quad h = \max_{1 \le i \le N} h_n. \tag{1.20}$$

We make the following local quasi-uniformity assumption (see [9, Assumption 2.1]).

Assumption 1. There exists a constant $\theta > 0$ such that

$$\frac{1}{\theta} \le \frac{h_n}{h_{n+1}} \le \theta \quad (1 \le i \le N-1). \tag{1.21}$$

Further, we need to introduce a condition that implies stability results. Set

$$P_{n+1} = U M_{n+1}^{-1} U M_n L. \tag{1.22}$$

We prefer this form to $M_{n+1}^{-1} U M_n L U$ because the elements of P_{n+1} play an essential role in the computation. We offer the following as a useful sufficient condition for the stability.

Condition 1. There exist constants $C_0 > 0$ and $0 < r < 1$ such that

$$\left\| \prod_{\ell=1}^{k} P_{n+1-\ell} \right\|_\infty \le C_0 r^k, \tag{1.23}$$

for all $k = 1, \ldots, n-1$ and all $n = 2, \ldots, N-1$.

Theorem 2 (Stability)**.** Suppose that Assumption 1 and Condition 1 are satisfied. Then, there exists a constant $C_1 > 0$ depending only on p and θ such that

$$\|\Pi_N(f)\|_{L^\infty(J;X)} \le C_1 \|f\|_{L^\infty(J;X)} \tag{1.24}$$

for any $f \in L^\infty(J;X)$.

Theorem 3. (Error estimate) Suppose that Assumption 1 is satisfied. Further, assume that the stability inequality (1.24) of Π_N holds true. Then, there exists a constant $C_2 > 0$ depending only on p and θ such that

$$\|f - \Pi_N(f)\|_{L^\infty(J;X)} \le C_2 h^{\widetilde{k}} |f|_{W^{\widetilde{k},\infty}(J;X)} \tag{1.25}$$

for any $f \in W^{k,\infty}(J;X)$ and $k \ge 1$, where $\widetilde{k} = \min\{k, p+1\}$.

When $p = 1$, we can check whether Condition 1 holds.

Theorem 4 (Stability and error estimate for $p = 1$)**.** Suppose that Assumption 1 is satisfied. Then, for $p = 1$, we have the stability (1.24) for $f \in L^\infty(J;X)$ and error estimate (1.25) for $f \in W^{k,\infty}(J;X)$, $k \ge 1$.

If the knot vector Ξ is open uniform, we can verify that Condition 1 is satisfied for small p.

Theorem 5 (Stability and error estimate for open uniform Ξ)**.** Suppose that the knot vector Ξ is open uniform, that is, suppose that

$$\frac{1}{N} = h = h_n \quad (1 \le n \le N), \quad t_n = nh \quad (0 \le n \le N). \tag{1.26}$$

Then, for $p = 2, 3, 4$, we have the stability (1.24) for $f \in L^\infty(J;X)$ and error estimate (1.25) for $f \in W^{k,\infty}(J;X)$, $k \ge 1$.

Remark 6. Let Ξ be open uniform. We have proved that Condition 1 holds only for $p = 2, 3, 4$. However, we infer from numerical experiments that Condition 1 is true at least for $p = 5, 6, 7$. Note that the definition of B-spline basis functions implies that

$$\widehat{B}_{i,p}(t) = \widehat{B}_{i+1,p}(t - h) \quad (p + 1 \le i \le N - 1) \tag{1.27}$$

follows for a uniform knot vector. Therefore, there exists a matrix P such that

$$P = P_n \quad (p + 1 \le n \le N - p - 1). \tag{1.28}$$

Figure 1.4 shows the behaviors of $\alpha_k = \|P^k\|_\infty / \|P\|_\infty$. Hence, we conjecture that

$$\|P^k\|_\infty \le C_0 r^k \tag{1.29}$$

for large k with suitable C_0 and $0 < r < 1$ when $p = 5, 6, 7$. Then, we can obtain (1.23) in exactly the same way as the proof of Lemma 4 below. However, the proof of (1.29) is left for future study.

Example 1. We take $T = 1$, $X = \mathbb{R}$, $f(t) = \sin(4\pi t)$, and consider the uniform partition Ξ. Set $\mathcal{E}_h = \|f - \Pi_N(f)\|_{L^\infty(J;X)}$. We plot $(\log h, \log \mathcal{E}_h)$ for several h in Fig. 1.5. We observe that $O(h^{p+1})$-convergence actually takes place.

1.4 Proof of Theorems

This section is devoted to the proof of the theorems presented in the previous section. In the following lemmas, we assume that Assumption 1 always holds. We begin by proving the following.

Lemma 1. Let $0 \le n \le N - 1$. For $1 \le i, j \le p + 1$, there exists a constant $C_{i,j} > 0$ depending only on θ and p such that

$$\int_{I_{n+1}} \widehat{B}_{i+n,p}(t) \widehat{B}_{j+n,p}(t) \, dt = C_{i,j} h_{n+1}. \tag{1.30}$$

Proof. For simplicity, we state the proof only for $n \ge p$. Introducing $\tau = (t - t_n)/h_n$, we have

$$\int_{I_{n+1}} \widehat{B}_{i+n,p}(t) \widehat{B}_{j+n,p}(t) \, dt = h_{n+1} \int_0^1 \widehat{B}_{i+n,p}(t_n + \tau h_{n+1}) \widehat{B}_{j+n,p}(t_n + \tau h_{n+1}) \, d\tau.$$

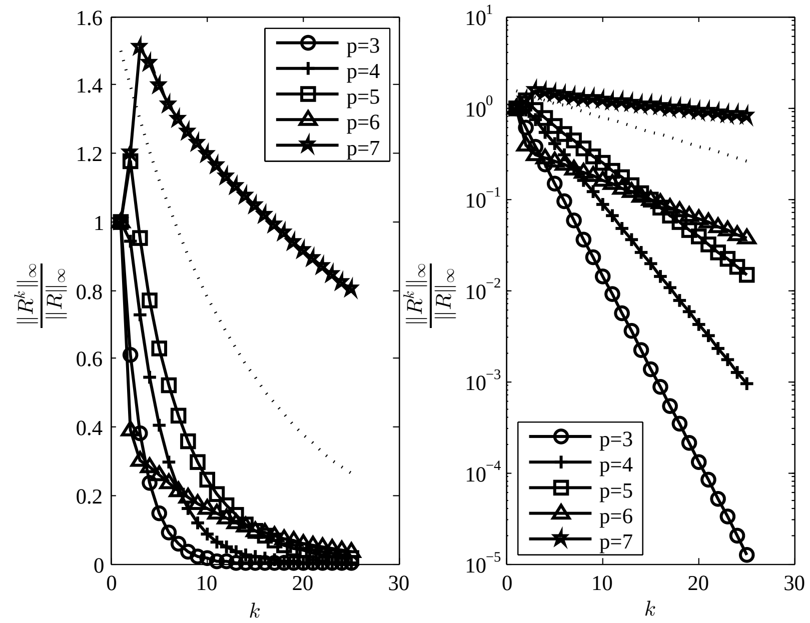


Figure 1.4: Plots of (k, α_k) (left) and $(k, \log \alpha_k)$ (right), where $\alpha_k = \|P^k\|_\infty / \|P\|_\infty$. The dashed line represents 0.9^k.

Using (1.3), we deduce

$$\begin{aligned}\widehat{B}_{i+n,p}(t_n + \tau h_{n+1}) &= \frac{t_n + \tau h_{n+1} - t_{n-p}}{t_n - t_{n-p}} \widehat{B}_{i+n,p-1}(t_n + \tau h_{n+1}) \\ &\quad + \frac{t_{n+1} - (t_n + \tau h_{n+1})}{t_{n+1} - t_{n-p+1}} \widehat{B}_{i+n+1,p-1}(t_n + \tau h_{n+1}) \\ &= \left(1 + \frac{\tau}{c_1}\right) \widehat{B}_{i+n,p-1}(t_n + \tau h_{n+1}) + \frac{1-\tau}{c_2} \widehat{B}_{i+n+1,p-1}(t_n + \tau h_{n+1}),\end{aligned}$$

where the constants c_1, c_2 depend only on θ. Hence, we find by induction that $\widehat{B}_{i+n,p}(t_n + \tau h_{n+1})$ is a continuous function of τ that depends only on θ and p, since $\widehat{B}_{i,0}(t_n + \tau h_n)$ is defined as (1.2). In particular, $\widehat{B}_{i+n,p}(t_n + \tau h_{n+1})$ is independent of h_n and n. Therefore, setting

$$C_{i,j} = \int_0^1 \widehat{B}_{i+n,p}(t_n + \tau h_{n+1}) \widehat{B}_{j+n,p}(t_n + \tau h_{n+1}) \, d\tau,$$

we obtain the expression (1.30). □

The following lemma is a direct consequence of the previous result.

Lemma 2. Let $0 \le n \le N-1$. The matrix M_{n+1} is expressed as $M_{n+1} = h_{n+1} A_{n+1}$, where $A_{n+1} \in \mathbb{R}^{(p+1)\times(p+1)}$ is a matrix whose entries depend only on θ and p. In particular, $P_{n+1} = U M_{n+1}^{-1} U M_n L$ is independent of h_{n+1}.

Lemma 3. There exists a positive constant C depending only on θ and p such that

$$\|M_{n+1}^{-1} F_{n+1}\|_\infty \le C \|f\|_{L^\infty(I_{n+1};X)}. \tag{1.31}$$

Proof. Equality (1.30) implies $\|\widehat{B}_{i+n,p}\|_{L^2(I_{n+1})} \le C h_{n+1}^{\frac{1}{2}}$, and therefore

$$\begin{aligned}\|F_{n+1}\|_\infty &\le \max_{1 \le i \le p+1} \|\widehat{B}_{i+n,p}\|_{L^2(I_{n+1})} \|f\|_{L^2(I_{n+1};X)} \\ &= C h_{n+1}^{\frac{1}{2}} \|f\|_{L^2(I_{n+1};X)} \\ &\le C h_{n+1} \|f\|_{L^\infty(I_{n+1};X)}.\end{aligned}$$

Lemma 2 gives $\|M_{n+1}^{-1}\|_\infty \le C h_{n+1}^{-1}$. Combining these results, we obtain (1.31). □

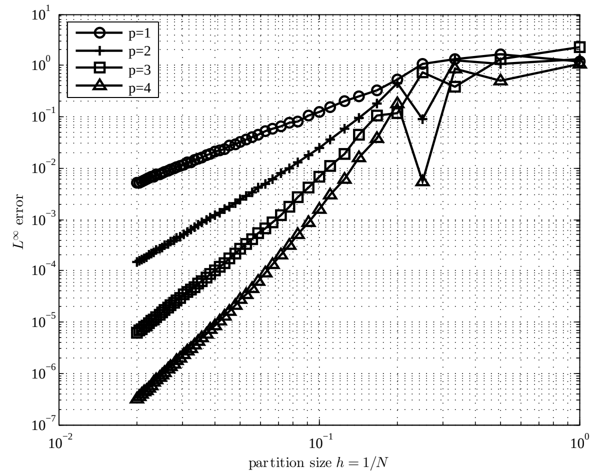


Figure 1.5: Plot of L^∞-error of SPT for $f(t) = \sin(4\pi t)$ $(0 \le t \le T = 1)$ using the uniform partition Ξ.

Now we can state the following proof of stability.

Proof of Theorem 2. We know $\|\widehat{B}_{i,p}\|_{L^\infty(J)} \le 1$ by (1.4) and (1.5). Hence,

$$\|\Pi_N(f)\|_{L^\infty(I;X)} \le (p+1) \max_{0\le n\le N-1} \|\mathbf{x}_{n+1}\|_\infty. \tag{1.32}$$

From equation (1.19), we obtain

$$\begin{aligned}\mathbf{x}_{n+1} &= M_{n+1}^{-1}F_{n+1} + \sum_{k=0}^{n-1}\left(\prod_{\ell=0}^{k} M_{n+1-\ell}^{-1}UM_{n-\ell}LU\right)M_{n-k}^{-1}F_{n-k}\\ &= M_{n+1}^{-1}F_{n+1} + \sum_{k=0}^{n-1} M_{n+1}^{-1}UM_nL\left(\prod_{\ell=1}^{k} P_{n+1-\ell}\right)UM_{n-k}^{-1}F_{n-k}.\end{aligned}$$

Using Condition 1, Lemma 2, and Lemma 3,

$$\begin{aligned}\|\mathbf{x}_{n+1}\|_\infty &\le C\|f\|_{L^\infty(I_{n+1};X)} + \sum_{k=0}^{n-1} Cr^k\|f\|_{L^\infty(I_{n-k};X)}\\ &\le C\|f\|_{L^\infty(J_{n+1};X)}\frac{1-r^n}{1-r},\end{aligned}$$

which, together with (1.32), implies the desired stability result (1.24) □

For the proof of the error estimate, we recall that there exists a *quasi-interpolant operator*

$$\pi_{p,\Xi} : L^\infty(J;X) \to S_{p,\Xi}(X)$$

satisfying the error estimate

$$\|f - \pi_{p,\Xi}(f)\|_{L^\infty(J;X)} \le Ch^s|f|_{W^{s,\infty}(J;X)} \quad (f \in W^{s,\infty}(J;X)) \tag{1.33}$$

for any positive integer $s \le p+1$, where $C > 0$ denotes a constant depending only on p. In fact, we know (see [9, Propositions 2.2 and 4.2])

$$\|f - \pi_{p,\Xi}(f)\|_{L^2(J;\mathbb{R})} \le Ch^s|f|_{W^{s,2}(J;X)} \quad (f \in W^{s,2}(J;\mathbb{R})). \tag{1.34}$$

The proof of (1.33) is essentially the same as that of (1.34).

$p=2$	$\lambda^2 + \frac{4823}{9166}\lambda + \frac{191}{9166}$
$p=3$	$-\lambda^3 - \frac{7658626644}{10698355585}\lambda^2 - \frac{2900186094}{53491777925}\lambda - \frac{3774802}{53491777925}$
$p=4$	$\lambda^4 + \frac{33907431272967929897}{34334740163403550366}\lambda^3 + \frac{6646444267949836917}{34334740163403550366}\lambda^2 + \frac{237693157497695699}{34334740163403550366}\lambda + \frac{474634492641025}{34334740163403550366}$

Table 1.1: Characteristic polynomial of P for $p = 2, 3, 4$.

Proof of Theorem 3. It is apparent that $\Pi_N(g) = g$ for $g \in S_{p,\Xi}(X)$. This property is sometimes called the *spline-preserving property*. According to (1.24), we deduce

$$\begin{aligned}\|f - \Pi_N(f)\|_{L^\infty(J;X)} &\le \|f - \pi_{p,\Xi}(f)\|_{L^\infty(J;X)} + \|\Pi_N(f - \pi_{p,\Xi}(f))\|_{L^\infty(J;X)} \\ &\le C\|f - \pi_{p,\Xi}(f)\|_{L^\infty(J;X)}.\end{aligned}$$

This, together with (1.33), provides the desired estimate. □

Now we consider the case $p = 1$.

Proof of Theorem 4. By a direct calculation, we find

$$\|P_{n+1}\|_\infty = \frac{2h_n}{4h_n + 3h_{n+1}} < 1$$

for all $n = 1, \dots, N-1$. Therefore, Condition 1 is satisfied. □

We finally consider the case where Ξ is open uniform. To prove Theorem 5, it suffices to show the following lemma.

Lemma 4. If Ξ is open uniform, Condition 1 is satisfied for $p = 2, 3, 4$.

Proof. Recall that there exists a matrix P such that

$$P = P_n \quad (p + 1 \le n \le N - p - 1) \tag{1.35}$$

when the partition is uniform. Moreover, we note that P_{n+1} depends only on p (see Lemma 2) and $\|P_{n+1}\|_\infty \le C$ for $n \le p$. Hence, Condition 1 is reduced to

$$\|P^k\|_\infty \le C_0 r^k \quad (k \ge 1). \tag{1.36}$$

At this stage, we admit

$$\operatorname{spr}(P) = \text{the spectrum radius of } P < 1. \tag{1.37}$$

Then, there exists a norm $|||\cdot|||$ of $\mathbb{R}^{p+1}$ such that the induced matrix norm satisfies $|||P||| \le \operatorname{spr}(P) + \varepsilon$, where $\varepsilon = (1 - \operatorname{spr}(P))/2$. Setting $r = (1 + \operatorname{spr}(P))/2$, we have $|||P||| \le r < 1$. Since the ∞ norm and $|||\cdot|||$ are equivalent in $\mathbb{R}^{p+1}$, we deduce $\|P^k\|_\infty \le C_0|||P^k||| \le C_0|||P|||^k \le C_0 r^k$ with a constant $C_0 > 0$ depending only on p. Therefore, we obtain (1.36).

It remains to show (1.37). Since

$$(P)_{i,p+1} = 0,\ (P)_{p+1,j} = 0 \quad (1 \le i, j \le p + 1),$$

all eigenvalues and eigenvectors of P are equal to those of

$$\widetilde{P} = \widetilde{U}M^{-1}UM\widetilde{L} \in \mathbb{R}^{p\times p},$$

where $\widetilde{U} = (\delta_{i+1,j})_{1\le i\le p,\ 1\le j\le p+1} \in \mathbb{R}^{p\times(p+1)}$ and $\widetilde{L} = \widetilde{U}^{\mathrm{T}}$. Herein, $\widetilde{P}$ also depends only on p, and we write $\widetilde{P}(p)$ to express the dependency on p. For $p = 2, 3, 4$, we are able to compute the characteristic

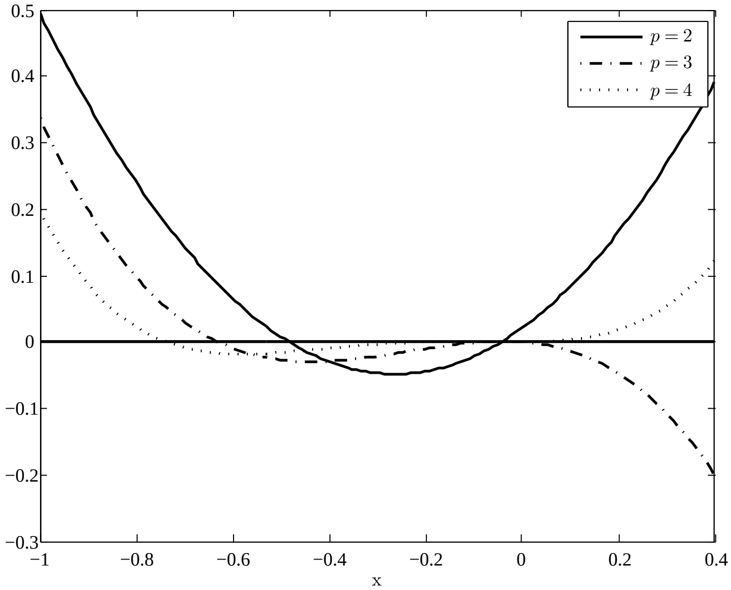


Figure 1.6: Graphs of characteristic polynomials for $p = 2, 3, 4$. This shows that all roots of characteristic polynomial are in the interval $(-1, 1)$.

polynomials of $\widetilde{P}(p)$. Table 1.1 shows the result using symbolic computation. Figure 1.6 shows that all roots are located in $(-1, 1)$. For example, let $p = 4$. Approximate roots of characteristic polynomial $\phi(\lambda)$ computed by MATLAB are given as

$$\begin{aligned}
\lambda_1^* &= -0.737897436525497,\\
\lambda_2^* &= -0.204310935198338,\\
\lambda_3^* &= -0.043224941317302,\\
\lambda_4^* &= -0.002121306903502.
\end{aligned}$$

These values are only approximations and contain rounding errors. However, we can calculate

$$\begin{aligned}
\phi\left(-\frac{737}{1000}\right) &= -\frac{4184230788864604364702796837}{171673700817017751830000000000000} < 0,\\
\phi\left(-\frac{739}{1000}\right) &= \frac{5188878821185878211588188803}{171673700817017751830000000000000} > 0.
\end{aligned}$$

Hence, there exists a root λ_1 such that $-0.739 < \lambda_1 < -0.737$. Similarly, we deduce that there exist three other roots (all real) $\lambda_2, \lambda_3, \lambda_4$ of $\phi(\lambda) = 0$ such that $-1 < \lambda_1 < \lambda_3 < \lambda_2 < \lambda_4 < 0$. This implies that $\mathrm{spr}(P(4)) < 1$. □

1.5 Conclusion

In this study, we described the mathematical formulation of the SPT with B-splines for X-valued functions and provided the stability and error estimates in the $L^\infty(0, T; X)$ norm. The quasi-uniformity of partition is always assumed. For $p = 1$, the stability holds true. We proved that, for $p = 2, 3, 4$, the uniformity of partition is a sufficient stability condition. For the case $5 \le p \le 7$, we inferred from numerical experiments that the SPT is stable in $L^\infty(0, T; X)$; the rigorous proof is left for future study. We also proved the error estimate using the spline-preserving property of the projector Π_N if the stability holds true.

Finally, we state an application of our results to the ST FEM. Let u and u_h denote the exact and approximate solutions for a time-dependent PDE. The function u_h is a piecewise smooth function in time variable, and we assume that u has sufficient regularity. Then, we can estimate

$$\|u - \Pi_N(u_h)\|_{L^\infty(0,T;X)} \le \|u - \Pi_N(u)\|_{L^\infty(0,T;X)} + C\|u - u_h\|_{L^\infty(0,T;X)}.$$

This estimate implies that if the SPT error is small enough, then the left-hand side can be controlled by $\|u - u_h\|_{L^\infty(0,T;X)}$.

Acknowledgement

We thank Professor K. Takizawa (Waseda University) and Professor T. E. Tezduyar (Rice University) who brought the subject to our attention.

Chapter 2

Analysis of the Nitsche method for evolution problem

2.1 Introduction

The boundary condition is an indispensable component of the well-posed problem of partial differential equation (PDE). It is not merely a side condition. In computational mechanics, the imposition of boundary conditions should be paid great attention, although it is sometimes understood as a simple and unambiguous task. The Neumann boundary condition or natural boundary condition is naturally taken into account in the variational equation so that it is handled directly in the Finite Element Method (FEM). On the other hand, there are several approaches to impose the Dirichlet boundary condition (DBC) numerically. The traditional finite element method (FEM) employs the Lagrange interpolation, that is, the nodal value at boundary nodal value give the discretized boundary condition. Then the problem can be reduced to homogeneous Dirichlet boundary value problem. It can treat complex geometry efficiently, however, there are difficulties in providing exact geometries in some cases. Moreover, the Lagrange interpolation requires that the basis functions have interpolation properties, which is not satisfied for the Isogeometric Analysis (IGA) [14].

IGA is the Galerkin method using the non-uniform rational B-spline (NURBS) for basis functions. Consequently, IGA is understood as one of the FEM. It can produce more exact geometric representations in the computational domain. Moreover, if the geometric models of object are designed by computer-aided design (CAD) systems, then we can use those models as the computational domains directly. Unfortunately, the NURBS basis functions don't satisfy the interpolation property, therefore the Lagrange interpolation is not applicable for IGA.

To surmount those shortcomings, Bazilevs et al. [5, 6] proposed a method of "weak imposition" of DBC by applying the methodology of the discontinuous Galerkin (DG) method and discussed its efficiency by numerical experiments in the non-stationary Navier–Stokes equations. Their method itself was originally proposed by Nitsche [27] and is commonly called the *Nitsche method*. The Nitsche method has been applied to mortaring on the interface (see [23] and the references given there), and stability and convergence of the Nitsche method for elliptic problems are well studied so far. Recently, it is also applied to IGA for Poisson equation, fourth-order problem [18] and steady Navier-Stokes problem [21] successfully.

In this chapter, we study the convergence analysis of the Nitsche method with FEM discretization including IGA for parabolic problems. Earlier studies of the Nitsche method are accomplished by formulating the method as a one-step method; see [52] for example. In contrast, we present a different approach: we study the Nitsche method using a variational approach. Consequently, the analysis becomes greatly simplified and optimal order error estimates in some appropriate norms are established. Such variational approach is recently successfully applied to the analysis of the DG time-stepping method for a wide class of parabolic equations in [32]. We will have that the Galerkin approximation for parabolic problems satisfy the inf-sup condition and Galerkin orthogonality under some conditions. Moreover, they yield the quasi-optimal error estimate in spatial semi-discretization.

This chapter is organized as follows. In Section 2.2, we review the Nitsche's classical paper [27] briefly, and the Nitsche method for elliptic problems are introduced. Section 2.3 presents some preliminaries and basic results for parabolic problems. We further define the advection-diffusion-reaction equation, and show that the equation has a weak solution under some conditions. In Section 2.4, we introduce the finite dimensional subspace for spatial semi-discretization with FEM and IGA. Our basic assumptions are mentioned first, and we check the traditional FEM and IGA both satisfy those assumptions. In Section 2.5, our main results are stated and proved. We show the formulation of the Nitsche method for parabolic problems, and we analyze that. The Banach-Nečas-Babuška theorem gives the necessary and sufficient condition for the unique existence of the approximate solution. Further, we have the error estimate by combining the inf-sup condition and Galerkin orthogonality. Section 2.6 establishes the error analysis of the full discrete problem. We apply the implicit Euler scheme to semi-discretized problem, and the solution is extended to piecewise constant function in temporal variable. The application of DG time-stepping is also interested, however it is our future work. Finnaly, in Section 2.7 we report numerical results to check the error.

In the following, $L^2(\Omega)$ and $H^s(\Omega)$ denote the usual Lebesgue and Sobolev spaces for domain $\Omega \subset \mathbb{R}^d$, respectively. Let V, H be two Banach spaces, then we write $V \hookrightarrow H$ if V is continuous embedded in H. V^* denotes the (continuous) dual of V.

2.2 Nitsche method for the elliptic problems

In this section, we briefly review the Nitsche method for the Ritz method applied to the Poisson equation with the Dirichlet boundary condition.

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary Γ. We consider the Poisson equation with non-homogeneous Dirichlet boundary condition. Letting $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$, then we find $u \in H^1(\Omega)$ such that

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega, \\ \quad u &= g \quad \text{on } \partial\Omega. \end{cases} \tag{2.1}$$

The variational principle gives that it is equivalent to the following minimization problem of Euler-Lagrange equation, find $u \in H^1_D(\Omega) =: \{v \in H^1(\Omega) : v = g \text{ on } \Gamma\}$ which minimizes the following functional,

$$J(u) := \frac{1}{2}|u|^2_{H^1(\Omega)} - (f, u)_{L^2(\Omega)}. \tag{2.2}$$

Here we impose the non-homogeneous Dirichlet boundary condition weakly, that is, we find $\widetilde{u} \in H^1(\Omega)$ which minimizes the following functional,

$$\widetilde{J}(\widetilde{u}) := \frac{1}{2}|\widetilde{u}|^2_{H^1(\Omega)} - (f, \widetilde{u})_{L^2(\Omega)} + \left(\widetilde{u} - g, L(\widetilde{u}) + c\right)_{L^2(\Gamma)}, \tag{2.3}$$

where L is a linear map to $L^2(\Gamma)$ and $c \in H^{\frac{1}{2}}(\Gamma)$, therefore $L(v) + c$ is an affine map. Further it is required that the following property is satisfied.

Assumption 2. We assume that L is continuous and

$$\begin{cases} (v - g, L(v) + c)_{L^2(\Gamma)} \geq 0 \text{ for all } v, \\ \dfrac{1}{2}|v|^2_{H^1(\Omega)} + (v, L(v))_{L^2(\Gamma)} \geq 0 \text{ for all } v, \end{cases} \tag{2.4}$$

where $u \in H^1(\Omega)$ minimizes the functional J.

Under this assumption, we have the following results.

Lemma 5. $w \in H^1(\Omega)$ minimizes the functional $\widetilde{J}$ if and only if w satisfies

$$\begin{aligned} \widetilde{a}(w, v) \quad &:= (\nabla w, \nabla v)_{L^2(\Omega)} + (w, L(v))_{L^2(\Gamma)} + (v, L(w))_{L^2(\Gamma)} \\ &= (f, v)_{L^2(\Omega)} + (g, L(v))_{L^2(\Gamma)} - (v, c)_{L^2(\Gamma)} \\ &=: \widetilde{F}(v) \end{aligned} \tag{2.5}$$

for all $v \in H^1(\Omega)$.

Proof. Let $v \in H^1(\Omega)$, then for all $\varepsilon > 0$, we have

$$\begin{aligned}
\widetilde{J}(w+\varepsilon v) &= \frac{1}{2}|w|^2_{H^1(\Omega)} + \varepsilon(\nabla w, \nabla v)_{L^2(\Omega)} + \frac{\varepsilon^2}{2}|v|^2_{H^1(\Omega)} - (f,w)_{L^2(\Omega)} - \varepsilon(f,v)_{L^2(\Omega)} \\
&\quad + (w-g, L(w)+c)_{L^2(\Gamma)} + \varepsilon\,(w-g, L(v))_{L^2(\Gamma)} \\
&\qquad + \varepsilon\,(v, L(w)+c)_{L^2(\Gamma)} + \varepsilon^2\,(v, L(v))_{L^2(\Gamma)} \\
&= \widetilde{J}(u) + \varepsilon^2\left(\frac{1}{2}|v|^2_{H^1(\Omega)} + (v, L(v))_{L^2(\Gamma)}\right) \\
&\quad + \varepsilon\Big((\nabla w, \nabla v)_{L^2(\Omega)} - (f,v)_{L^2(\Omega)} + (w-g, L(v))_{L^2(\Gamma)} + (v, L(w)+c)_{L^2(\Gamma)}\Big).
\end{aligned} \tag{2.6}$$

Applying the equation (2.5), we get $\widetilde{J}(w+\varepsilon v) = \widetilde{J}(w) + \varepsilon^2\left(\frac{1}{2}|v|^2_{H^1(\Omega)} + (v, L(v))_{L^2(\Gamma)}\right)$. Here the second equation in the Assumption 2 yields $\widetilde{J}(w+\varepsilon v) \geq \widetilde{J}(w)$ for all $v \in H^1(\Omega)$. On the other hand, if $w \in H^1(\Omega)$ minimizes the functional $\widetilde{J}$, then $\frac{\partial \widetilde{J}}{\partial \varepsilon}(w) = 0$, therefore the equation (2.5) holds. □

Lemma 6. There exists a unique element $\widetilde{u} \in H^1(\Omega)$ which minimizes the functional $\widetilde{J}$.

Proof. The Assumption 2 shows that $\widetilde{a}$ is continuous and coercive bilinear form, and the functional $\widetilde{F}$ is continuous, then the Lax-Milgram theorem results the conclusion. □

Our purpose is to define the linear map L and function c which satisfies the Assumption 2. One example of such a pair is give by

$$L(v) := \frac{\varepsilon}{2}v,\ c := -\frac{\varepsilon}{2}g, \tag{2.7}$$

where $\varepsilon > 0$ is a penalty parameter. We can check that for all $v \in H^1(\Omega)$,

$$(v-g, L(v)+c)_{L^2(\Gamma)} = \frac{\varepsilon}{2}\|v-g\|^2_{L^2(\Gamma)} \geq 0. \tag{2.8}$$

Moreover,

$$(v, L(v))_{L^2(\Gamma)} = \frac{\varepsilon}{2}\|v\|_{L^2(\Gamma)} \geq 0 \tag{2.9}$$

for all $v \in H^1(\Omega)$. Therefore this satisfies the Assumption 2. Actually, it agrees with the penalty method,

$$\widetilde{J}(\widetilde{u}) = \frac{1}{2}|\widetilde{u}|^2_{H^1(\Omega)} - (f, \widetilde{u})_{L^2(\Omega)} + \frac{\varepsilon}{2}\|\widetilde{u}-g\|^2_{L^2(\Gamma)}. \tag{2.10}$$

Here we have $u = \widetilde{u}$ as $\varepsilon \to \infty$, where u and $\widetilde{u}$ are minimizers of J and $\widehat{J}$ over $H^1_D(\Omega)$ and $H^1(\Omega)$, respectively.

On the other hand, the Nitsche method proposes to letting

$$L(v) := -\frac{\partial v}{\partial \mathbf{n}} + \frac{\varepsilon}{2}v,\ c := -\frac{\varepsilon}{2}g. \tag{2.11}$$

However, the map L is not well-defined as a map $H^1(\Omega) \to L^2(\Gamma)$. Therefore, we introduce the finite dimensional approximation. Let $\mathcal{T}_h$ be a partition of Ω and $V_h \subset H^1(\Omega)$ be a finite dimensional subspace such that

$$\frac{\partial v_h}{\partial \mathbf{n}} \in \left(L^2(\Gamma)\right)^* \simeq L^2(\Gamma) \text{ for all } v_h \in V_h. \tag{2.12}$$

Then we have $L : V_h \to L^2(\Gamma)$, and for all $v_h \in V_h$,

$$(v_h - g, L(v_h)+c)_{L^2(\Gamma)} = \frac{\varepsilon}{2}\|v_h - g\|^2_{L^2(\Gamma)} - \left(v_h - g, \frac{\partial v_h}{\partial \mathbf{n}}\right)_{L^2(\Gamma)}. \tag{2.13}$$

Moreover,

$$\begin{aligned}
\frac{1}{2}|v_h|^2_{H^1(\Omega)} + (v_h, L(v_h))_{L^2(\Gamma)} &= \frac{1}{2}|v_h|^2_{H^1(\Omega)} + \frac{\varepsilon}{2}\|v_h\|^2_{L^2(\Gamma)} - \left(v_h, \frac{\partial v_h}{\partial \mathbf{n}}\right)_{L^2(\Gamma)} \\
&\geq \frac{1}{2}|v_h|^2_{H^1(\Omega)} + \left(\frac{\varepsilon}{2} - C_0(x)\right)\|v_h\|^2_{L^2(\Gamma)} - \frac{1}{2C_0(x)}\left\|\frac{\partial v_h}{\partial \mathbf{n}}\right\|^2_{L^2(\Gamma)}
\end{aligned} \tag{2.14}$$

for all $v \in H^2(\Omega)$ and for some positive function $C_0 : \Gamma \to \mathbb{R}$.

Under some assumptions, we can prove that there exists a positive constant C_I such that

$$\left\| \frac{\partial v_h}{\partial \mathbf{n}} \right\|_{L^2(E)}^2 \leq C_I |v_h|_{H^1(\Omega)}^2 \sum_{E \in \mathcal{E}_h^e} h_E^{-1} \tag{2.15}$$

for all $v_h \in V_h$, where $\mathcal{E}_h^e := \{E \subset \Gamma :$ There exists $K \in \mathcal{T}_h$ such that E is a edge or face of $K\}$. Therefore, we let $C_0(x)$ be such that

$$C_0(x) \geq h_E^{-1} C_I \text{ for } x \in E \tag{2.16}$$

and $\varepsilon \geq 2C_0(x) \geq 2h_E^{-1}C_I$, then we have $1/2|v_h|_{H^1(\Omega)}^2 + (v_h, L(v_h))_{L^2(\Gamma)}$ for all $v_h \in V_h$. This gives the Nitsche method

$$\widetilde{J}(\widetilde{u}) = \frac{1}{2}|\widetilde{u}|_{H^1(\Omega)}^2 - (f, \widetilde{u})_{L^2(\Omega)} + \frac{\varepsilon}{2}\|\widetilde{u} - g\|_{L^2(\Gamma)}^2 - \left(\widetilde{u} - g, \frac{\partial \widetilde{u}}{\partial \mathbf{n}}\right)_{L^2(\Gamma)}. \tag{2.17}$$

Moreover, the equation (2.5) leads to Nitsche's formulation for Galerkin method, find $u_h \in V_h$ such that

$$\begin{aligned} \widetilde{a}(u_h, v_h) \quad &:= (\nabla u_h, \nabla v_h)_{L^2(\Omega)} - \left(\frac{\partial u_h}{\partial \mathbf{n}}, v_h\right)_{L^2(\Gamma)} - \left(u_h, \frac{\partial v_h}{\partial \mathbf{n}}\right)_{L^2(\Gamma)} + \varepsilon(u_h, v_h)_{L^2(\Gamma)} \\ &= (f, v_h)_{L^2(\Omega)} + \left(g, -\frac{\partial v_h}{\partial \mathbf{n}} + \varepsilon v_h\right)_{L^2(\Gamma)} = \widetilde{F}(v_h) \end{aligned} \tag{2.18}$$

for all $v_h \in V_h$. The solution $u_h \in V_h$ is an approximation of $\widetilde{u}$. Further, if $u \in H^2(\Omega)$, we have the Galerkin orthogonality

$$\widetilde{a}(u - u_h, v_h) = 0, \text{ for all } v_h \in V_h. \tag{2.19}$$

Note that the formulation of Nitsche method is similar to one of the discontinuous Galerkin method (DG method for short). The formulation of DG methods for the elliptic equation is referred in [1].

2.3 Weak formulation of abstract parabolic problems

In this section, we review the weak formulation of the parabolic problems. We apply the Banach-Nečas-Babuška theorem, which can be found in [19], and we will show that the parabolic problem has a unique solution.

2.3.1 Preliminary

We review some standard facts. First, the following lemma provides a necessary and sufficient condition for the unique existence of the weak solution of abstract parabolic problems.

Lemma 7 (Banach-Nečas-Babuška theorem, Theorem 2.6 of [19])**.** Let X be a Banach space and Y be a reflexive Banach space. For any continuous bilinear form $b : X \times Y \to \mathbb{R}$, the following (i) and (ii) are equivalent;

(i) For all $F \in Y^*$, there exists a unique solution $x \in X$ of

$$b(x, y) = \langle F, y \rangle_{Y^*,Y}, \text{ for all } y \in Y. \tag{2.20}$$

(ii) The bilinear form b satisfies the following.

(ii-a) There exists a positive constant β such that

$$\inf_{y \in Y} \sup_{x \in X} \frac{b(x, y)}{\|x\|_X \|y\|_Y} \geq \beta. \tag{2.21}$$

(ii-b) $b(x, y) = 0$, for all $x \in X \Rightarrow y = 0$.

Next, we define the following space, which may be called the Sobolev-Bochner space.

Definition 2. Let $1 \le p, q \le \infty$. For two Banach spaces V, H with norms $\|\cdot\|_V$, $\|\cdot\|_H$, respectively, we define

$$W^{1,p,q}(0,T;V,H) := \{v \in L^p(0,T;V) : \partial_t v \in L^q(0,T;H)\}. \tag{2.22}$$

This is a Banach space with norm

$$\|v\|^2_{W^{1,p,q}(0,T;V,H)} := \int_0^T \left(\|v\|^2_{L^p(0,T;V)} + \|\partial_t v\|^2_{L^q(0,T;H)} \right) dt \tag{2.23}$$

for all $v \in W^{1,2,2}(0,T;V,H)$. Further, the space $W^{1,2,2}(0,T;V,H)$ is a Hilbert space if V and H are Hilbert spaces.

Lemma 8 (Lemma 7.1. of [31])**.** Let $1 \le p, q \le \infty$ and V, H be two Banach spaces which satisfy $V \hookrightarrow H$, then

$$W^{1,p,q}(0,T;V,H) \hookrightarrow C^0([0,T];H). \tag{2.24}$$

Definition 3. Let V, H be two Hilbert spaces satisfying that $V \hookrightarrow H$ is dense. We identify H with its own dual $H \simeq H^*$. Then we have

$$V \hookrightarrow H \hookrightarrow V^*, \tag{2.25}$$

and this is called the Gelfand's evolution triple.

Lemma 9 (Theorem 1. of [16], Chapter XVIII)**.** Let $V \hookrightarrow H \hookrightarrow V^*$ be an evolution triple, then

$$W^{1,2,2}(0,T;V,V^*) \hookrightarrow C^0([0,T];H). \tag{2.26}$$

Corollary 1. We have

$$W^{1,2,2}\left(0,T;H^1_0(\Omega),H^{-1}(\Omega)\right) \hookrightarrow C^0\left([0,T];L^2(\Omega)\right), \tag{2.27}$$

where $H^{-1}(\Omega) := \left(H^1_0(\Omega)\right)^*$.

Definition 4. For a Banach space V with norm $\|\cdot\|_V$, we define

$$H^1(0,T;V) := W^{1,2,2}(0,T;V,V). \tag{2.28}$$

This is a Banach space with norm

$$\|v\|_{H^1(0,T;V)} := \|v\|_{W^{1,2,2}(0,T;V,V)} \tag{2.29}$$

for $v \in H^1(0,T;V)$. Further, this is a Hilbert space if V is a Hilbert space.

Corollary 2 (Theorem 2. of [22], Chapter 5.9)**.** Let V be a Banach space, then

$$H^1(0,T;V) \hookrightarrow C^0([0,T];V). \tag{2.30}$$

2.3.2 Remarks on the evolution triple

We consider the application of the Nitsche method to the following abstract parabolic problems. First, we recall the abstract parabolic problem with strongly-imposed Dirichlet boundary condition. Let V, L be two Hilbert spaces such that

$$V \hookrightarrow L \simeq L^* \hookrightarrow V^*. \tag{2.31}$$

Further, we let $A(t) : V \to V^*$ be a linear map for a.e. $t \in (0,T)$, and assume that there exists two positive constants M and α such that

$$\begin{cases} \langle A(t)u, v\rangle_{V^*,V} \le M\|u\|_V\|v\|_V & \text{for all } u,v \in V \text{ and for a.e. } t \in (0,T), \\ \langle A(t)v, v\rangle_{V^*,V} \ge \alpha\|v\|^2_V & \text{for all } v \in V \text{ and for a.e } t \in (0,T). \end{cases} \tag{2.32}$$

For $f \in L^2(0,T;V^*)$, and $u_0 \in L$, we find $u \in W^{1,2,2}(0,T;V,V^*)$ such that

$$\begin{cases} u' + A(t)u &= f \quad \text{in } L^2(0,T;V^*), \\ u(0) &= u_0 \quad \text{in } L. \end{cases} \tag{2.33}$$

We note that $W^{1,2,2}(0,T;V,V^*)$ is a Banach space with the norm

$$\|u\|^2_{W^{1,2,2}(0,T;V,V^*)} := \int_0^T \left(\|u\|_V^2 + \|u'\|_{V^*}^2\right)\, dt. \tag{2.34}$$

We also note that $W^{1,2,2}(0,T;V,V^*) \hookrightarrow C^0([0,T];L)$, therefore the initial condition is meaningful. The problem is called a parabolic problem.

We often let $V = H_0^1(\Omega)$ (therefore $V^* = \left(H_0^1(\Omega)\right)^* =: H^{-1}(\Omega)$) and $L = L^2(\Omega)$ for a bounded domain $\Omega \subset \mathbb{R}^d$. Further we note that Riesz representation theorem implies $\left(L^2(\Omega)\right)^* \simeq L^2(\Omega)$, and therefore if $u \in L^2(\Omega) \subset H^{-1}(\Omega)$ then we have

$$\langle u, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = u(v) = \langle u, v\rangle_{(L^2(\Omega))^*,L^2(\Omega)} = (u,v)_{L^2(\Omega)} \text{ for all } v \in H_0^1(\Omega). \tag{2.35}$$

The evolution triple $H_0^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ plays an important role even if we consider the non-homogeneous Dirichlet boundary condition, because it can be reduced to homogeneous Dirichlet boundary value problems. Here we note that $H_0^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ holds thanks to that $H_0^1 \hookrightarrow L^2(\Omega)$ is dense. In page 136 of [11], it is remarked that the restriction $T : \left(L^2(\Omega)\right)^* \to H^{-1}(\Omega)$, which is defined by

$$\langle T\phi, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} := \langle \phi, v\rangle_{(L^2(\Omega))^*,L^2(\Omega)} \tag{2.36}$$

for all $\phi \in \left(L^2(\Omega)\right)^*$ and $v \in H_0^1(\Omega)$, satisfies

$$\begin{cases} \|T\phi\|_{H^{-1}(\Omega)} \le C\|\phi\|_{(L^2(\Omega))^*}, \text{ for all } \phi \in \left(L^2(\Omega)\right)^*, \\ T \text{ is injective}, \\ R(T) \text{ (The range of } T\text{) is dense in } H^{-1}(\Omega), \text{ because } H_0^1(\Omega) \text{ is reflexive.} \end{cases} \tag{2.37}$$

Therefore we can consider T as a canonical embedding. This implies

$$H_0^1(\Omega) \hookrightarrow L^2(\Omega) \simeq \left(L^2(\Omega)\right)^* \simeq R(T) \hookrightarrow H^{-1}(\Omega). \tag{2.38}$$

Remark 7. On the other hand, let V, H be two Banach spaces where $V \hookrightarrow H$, but it is not dense. Then the restriction operator $T : H^* \to V^*$, which is defined by $T(\phi) = \phi|_V : V \to \mathbb{R}$, is not injective. This implies

$$H^* \not\simeq R(T) \subset V^* \tag{2.39}$$

in this case. However, we can define

$$\phi(v) = \phi|_V(v) = \langle \phi|_V, v\rangle_{V^*,V} =: \langle T(\phi), v\rangle_{V^*,V} \text{ for all } \phi \in H^*, v \in V, \tag{2.40}$$

even in this situation, that is, $\phi|_V : V \to \mathbb{R}$ is a linear and continuous functional for any linear and continuous functional $\phi : H \to \mathbb{R}$, because $V \hookrightarrow H$ implies

$$\|\phi|_V\|_{V^*} := \sup_{v\in V} \frac{\phi(v)}{\|v\|_V} \le \sup_{v\in H} \frac{\phi(v)}{\|v\|_H} \le \|\phi\|_{H^*}. \tag{2.41}$$

2.3.3 The weak formulation and the weak solution

We recall the unique existence of weak solution of parabolic problems. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary Γ. Remind that there exists a trace operator $\mathrm{Tr} : H^1(\Omega) \to H^{1/2}(\Gamma)$. This is surjective, and satisfies that there exists a constant C_{Tr} such that

$$\|\mathrm{Tr}\, v\|_{L^2(\Gamma)} \le C_{\mathrm{Tr}}\|v\|_{H^1(\Omega)} \text{ for all } v \in H^1(\Omega). \tag{2.42}$$

We let $A(t) : H^1(\Omega) \to H^{-1}(\Omega)$ be a linear map for a.e. $t \in (0,T)$.

We consider the following problem, find $u \in X := W^{1,2,2}\left(0,T;H^1(\Omega),H^{-1}(\Omega)\right)$ such that

$$\begin{cases} u' + A(t)u &= f(\mathbf{x},t) & \text{in } \Omega \times (0,T), \\ u &= g_D(\mathbf{x},t) & \text{in } \Gamma \times (0,T), \\ u(\mathbf{x},0) &= u_0(\mathbf{x}) & \text{for } \mathbf{x} \in \Omega, \end{cases} \tag{2.43}$$

where $f \in L^2\left(0,T;H^{-1}(\Omega)\right)$,

$$g_D \in G_0 := \left\{ g \in L^2\left(0,T;H^{1/2}(\Gamma)\right) : \begin{array}{c} \text{there exists } u_D \in X \text{ such that} \\ u_D(0) \in L^2(\Omega) \text{ and } \operatorname{Tr} u_D = g \text{ in } L^2\left(0,T;H^{1/2}(\Gamma)\right) \end{array} \right\}, \tag{2.44}$$

and $u_0 \in L^2(\Omega)$ are given.

We consider the weak formulation for the problem (2.43). The definition of G_0 implies there exists $u_D \in W^{1,2,2}\left(0,T;H^1(\Omega),H^{-1}(\Omega)\right)$ such that $\operatorname{Tr} u_D = g_D$ in $L^2\left(0,T;H^{1/2}(\Omega)\right)$ for the g_D. Here, we find $\widetilde{u} \in X_0 := W^{1,2,2}\left(0,T;H_0^1(\Omega),H^{-1}(\Omega)\right)$ such that

$$\begin{aligned} b_0(\widetilde{u},\mathbf{v}) &:= \int_0^T \langle \widetilde{u}' + A(t)\widetilde{u}, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\, dt + (\widetilde{u}(0),v_0)_{L^2(\Omega)} \\ &= \int_0^T \langle f - u_D' - A(t)u_D, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\, dt + (u_0 - u_D(0),v_0)_{L^2(\Omega)}, \\ &=: \langle F_0,\mathbf{v}\rangle_{Y_0^*,Y_0} \end{aligned} \tag{2.45}$$

for all $\mathbf{v} = (v,v_0) \in Y_0 := L^2\left(0,T;H_0^1(\Omega)\right) \times L^2(\Omega)$, and then we obtain $u := \widetilde{u} + u_D \in X$. Note that the Lemma 9 implies

$$X_0 := W^{1,2,2}\left(0,T;H_0^1(\Omega),H^{-1}(\Omega)\right) \subset C^0\left([0,T];L^2(\Omega)\right). \tag{2.46}$$

This and the definition of G_0 show that $\widetilde{u}(0), u_D(0) \in L^2(\Omega)$.

It follows easily that $b_0 : X_0 \times Y_0 \to \mathbb{R}$ is a bilinear form, F_0 is a linear functional, and X_0, Y_0 are Banach spaces with norms

$$\begin{aligned} \|x\|_{X_0}^2 &:= \int_0^T \left(\|x'\|_{H^{-1}(\Omega)}^2 + \|x\|_{H^1(\Omega)}^2 \right) dt + \|x(0)\|_{L^2(\Omega)}^2, \\ \|\mathbf{y}\|_{Y_0}^2 &= \|(y,y_0)\|_{Y_0}^2 \\ &:= \int_0^T \|y\|_{H^1(\Omega)}^2\, dt + \|y_0\|_{L^2(\Omega)}^2. \end{aligned} \tag{2.47}$$

Theorem 8 (Unique existence of the weak solution)**.** If there exist two positive constants M and α such that

$$\begin{cases} \langle A(t)w,v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \le M\|w\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} & \text{for all } w \in H^1(\Omega), v \in H_0^1(\Omega), \\ \langle A(t)v,v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \ge \alpha\|v\|_{H^1(\Omega)} & \text{for all } v \in H_0^1(\Omega) \end{cases} \tag{2.48}$$

for a.e. $t \in (0,T)$, then the problem (2.45) has a unique solution $\widetilde{u} \in X_0$ for all $u_D \in X$.

Furthermore, we have that the weak solution of problem (2.43) exists uniquely.

Proof. First, the assumption about $A(t) : H^1(\Omega) \to H^{-1}(\Omega)$ leads the bilinear form b_0 and the functional F_0 are both continuous. Therefore the Banach-Nečas-Babuška theorem implies there is a unique $\widetilde{u} \in X_0$ if and only if

(BNB1,inf-sup condition) There exists a positive constant β such that

$$\inf_{0\neq x\in X_0} \sup_{0\neq \mathbf{y}\in Y_0} \frac{b_0(x,\mathbf{y})}{\|x\|_{X_0}\|\mathbf{y}\|_{Y_0}} > \beta, \tag{2.49}$$

(BNB2) For all $x \in X_0$, $b_0(x,\mathbf{y}) = 0$ holds if and only if $\mathbf{y} = 0$.

Here we will show the (BNB1). By the continuity and coercivity of $A(t) \in \mathcal{L}\left(H_0^1(\Omega),H^{-1}(\Omega)\right)$, applying the Lax-Milgram theorem yield we have that there exists $A(t)^{-1} \in \mathcal{L}\left(H^{-1}(\Omega),H_0^1(\Omega)\right)$, for a.e. $t \in (0,T)$. The operator $A(t)^{-1}$ satisfies

$$\alpha\|A(t)^{-1}\phi\|_{H^1(\Omega)}^2 \le \left\langle A(t)\left(A(t)^{-1}\phi\right), A(t)^{-1}\phi\right\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \le \|\phi\|_{H^{-1}(\Omega)}\|A(t)^{-1}\phi\|_{H^1(\Omega)} \tag{2.50}$$

for all $\phi \in H^{-1}(\Omega)$. Therefore $A(t)^{-1}$ is continuous with

$$\|A(t)^{-1}\|_{\mathcal{L}\left(H^{-1}(\Omega),H_0^1(\Omega)\right)} = \sup_{0\neq\phi\in H^{-1}(\Omega)} \frac{\|A(t)^{-1}\phi\|_{H^1(\Omega)}}{\|\phi\|_{H^{-1}(\Omega)}} \le \alpha^{-1}. \tag{2.51}$$

Further,

$$\|\phi\|_{H^{-1}(\Omega)} = \|A(t)\left(A(t)^{-1}\phi\right)\|_{H^{-1}(\Omega)} \le M\|A(t)^{-1}\phi\|_{H^1(\Omega)} \tag{2.52}$$

for all $\phi \in H^{-1}(\Omega)$. This implies the coercivity of $A(t)^{-1}$,

$$\begin{aligned} \langle \phi, A(t)^{-1}\phi\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &= \langle A(t)\left(A(t)^{-1}\phi\right), A(t)^{-1}\phi\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \\ &\ge \alpha\|A(t)^{-1}\phi\|_{H^1(\Omega)}^2 \\ &\ge \alpha M^{-2}\|\phi\|_{H^{-1}(\Omega)}^2. \end{aligned} \tag{2.53}$$

Here we let $x \in X_0$ and $\mathbf{y} := \left(A(t)^{-1}x' + \mu x, \mu x(0)\right) \in Y_0$, where $\mu := \alpha^{-4}M^4$. Then we have

$$\begin{aligned} \|\mathbf{y}\|_{Y_0}^2 &= \int_0^T \|A(t)^{-1}x' + \mu x\|_{H^1(\Omega)}^2\, dt + \|\mu x(0)\|_{L^2(\Omega)}^2 \\ &\le C\int_0^T \left(\|x'\|_{H^{-1}(\Omega)}^2 + \|x\|_{H^1(\Omega)}^2\right)\, dt + \mu\|x(0)\|_{L^2(\Omega)}^2 \le C\|x\|_{X_0}^2, \end{aligned} \tag{2.54}$$

and

$$\begin{aligned} b_0(x,\mathbf{y}) &= \int_0^T \left(\langle x', A(t)^{-1}x' + \mu x\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} + \langle A(t)x, A(t)^{-1}x' + \mu x\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\right)\, dt \\ &\qquad + (x(0), \mu x(0))_{L^2(\Omega)} \\ &\ge \int_0^T \left(\alpha M^{-2}\|x'\|_{H^{-1}(\Omega)}^2 - \alpha^{-1}M\|x\|_{H^1(\Omega)}\|x'\|_{H^{-1}(\Omega)} + \mu\alpha\|x\|_{H^1(\Omega)}^2\right)\, dt \\ &\qquad + \mu/2\left(\|x(T)\|_{L^2(\Omega)}^2 - \|x(0)\|_{L^2(\Omega)}^2\right) + \mu\|x(0)\|_{L^2(\Omega)}^2 \\ &\ge \int_0^T \left(\alpha M^{-2}\|x'\|_{H^{-1}(\Omega)}^2 - \alpha^{-1}M\|x\|_{H^1(\Omega)}\|x'\|_{H^{-1}(\Omega)} + \mu\alpha\|x\|_{H^1(\Omega)}^2\right)\, dt \\ &\qquad + \mu/2\|x(0)\|_{L^2(\Omega)}^2 \\ &\ge \int_0^T \left(\alpha M^{-2}/2\|x'\|_{H^{-1}(\Omega)}^2 + \left(\mu\alpha - \alpha^{-3}M^4/2\right)\|x\|_{H^1(\Omega)}^2\right)\, dt + \mu/2\|x(0)\|_{L^2(\Omega)}^2 \\ &\ge C\left(\int_0^T \left(\|x'\|_{H^{-1}(\Omega)}^2 + \|x\|_{H^1(\Omega)}^2\right)\, dt + \|x(0)\|_{L^2(\Omega)}^2\right) = C\|x\|_{X_0}^2, \end{aligned} \tag{2.55}$$

because $\mu := \alpha^{-4}M^4$. Combining the above two inequalities, we conclude

$$b_0(x,\mathbf{y}) \ge C\|x\|_{X_0}^2 \ge \beta\|x\|_{X_0}\|\mathbf{y}\|_{Y_0} \tag{2.56}$$

for any $x \in X_0$ and $\mathbf{y} := \left(A(t)^{-1}x' + \alpha^{-4}M^4x, \alpha^{-4}M^4x(0)\right)$, which gives

$$\inf_{0\neq x\in X_0}\ \sup_{0\neq \mathbf{y}\in Y_0} \frac{b_0(x,\mathbf{y})}{\|x\|_{X_0}\|\mathbf{y}\|_{Y_0}} \ge \beta. \tag{2.57}$$

Therefore the inf-sup condition (BNB1) follows.

Next we show the condition (BNB2) is satisfied. Let $\mathbf{y}_1 = (y_1, y_{10}) \in Y_0$ be such that $b_0(x, \mathbf{y}_1) = 0$, for all $x \in X_0$. We let $x \in X_0$ be such that

$$x(0) = \widetilde{y}_{10}, \text{ and } x(t) = 0 \text{ for } t \ge \delta \tag{2.58}$$

for all $\delta > 0$. Then $b(x,\mathbf{y}_1) = 0$ yields

$$\int_0^\delta \langle x' + A(t)x, y_1\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\, dt + \|y_{10}\|_{L^2(\Omega)} = 0 \tag{2.59}$$

for all $\delta > 0$. Therefore we have $y_{10} = 0$.

Let $x \in C_0^\infty\left((0,T); H_0^1(\Omega)\right) \subset X_0$, then $b_0(x,\mathbf{y}_1) = 0$ gives

$$\begin{aligned} \left|\int_0^T \langle x', y_1\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\, dt\right| &= \left|\int_0^T \langle A(t)x, y_1\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\, dt\right| \\ &\le C\|x\|_{L^2(0,T;H^1(\Omega))}\|y_1\|_{L^2(0,T;H^1(\Omega))} < \infty, \end{aligned} \tag{2.60}$$

because of continuity of $A(t)$. Therefore the integration

$$-\int_0^T \langle y_1', x\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\,dt = \int_0^T \langle x', y_1\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\,dt \tag{2.61}$$

has value, and therefore we have $y_1' \in L^2\left(0,T;H^{-1}(\Omega)\right)$, that is, $y_1 \in X_0$. This implies

$$\int_0^T -\langle y_1', x\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} + \langle A(t)^* y_1, x\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\,dt = 0 \tag{2.62}$$

for all $x \in C_0^\infty\left((0,T);H_0^1(\Omega)\right)$, where $A(t)^*$ is adjoint operator of $A(t)$. By the density of $C_0^\infty\left((0,T);H_0^1(\Omega)\right) \subset L^2\left(0,T;H_0^1(\Omega)\right)$, the equation (2.62) holds for all $x \in L^2\left(0,T;H_0^1(\Omega)\right)$.

Further, for any $\phi \in H_0^1(\Omega)$, we let $x := t\phi \in X_0 \subset L^2\left(0,T;H_0^1(\Omega)\right)$. Then $b(x,\mathbf{y}_1) = 0$ and integration by part in equation (2.62) show that

$$\begin{aligned}
0 &= \int_0^T -\langle y_1', t\phi\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} + \langle A(t)^* y_1, t\phi\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\,dt \\
&= -(y_1(T), T\phi)_{L^2(\Omega)} + \int_0^T \langle (t\phi)' + A(t)(t\phi), y_1\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}\,dt \\
&= -(y_1(T), T\phi)_{L^2(\Omega)} + b_0(t\phi, \mathbf{y}_1) \\
&= -(y_1(T), T\phi)_{L^2(\Omega)}
\end{aligned} \tag{2.63}$$

for all $\phi \in H_0^1(\Omega)$, therefore we have $y_1(T) = 0$. If we let $x = (T-t)\phi$ for any $\phi \in H_0^1(\Omega)$, then we have $y_1(0) = 0$ by the same argument.

Finally, we let $x := y_1 \in L^2\left(0,T;H_0^1(\Omega)\right)$, then the partial integration and the coercivity of $A(t)$ give $y_1 = 0$. The above argument shows that (BNB2) holds.

Applying the Banach-Nečas-Babuška theorem, we can assert that there exists a unique $\widetilde{u} \in X_0$ for each $u_D \in X$ such that $Tu_D = g_D$. Here we let

$$u_{D1}, u_{D2} \in X \tag{2.64}$$

be such that $Tu_{D1} = Tu_{D2} = g_D$ and $u_{D1} \neq u_{D2}$. Then there exist $\widetilde{u}_1, \widetilde{u}_2 \in X_0$ satisfying the equation (2.45) for u_{D1}, u_{D2}, respectively. Now we have

$$u_1 := \widetilde{u}_1 + u_{D1}, \text{ and } u_2 := \widetilde{u}_2 + u_{D2}, \tag{2.65}$$

which are weak solution of problem (2.43). We will show that the two weak solution satisfy $u_1 = u_2$ by contradiction. If we assume $u_1 \neq u_2$, then $U := u_1 - u_2$ satisfies

$$\begin{cases} U' + A(t)U &= 0, \quad \text{in } \Omega \times (0,T), \\ U &= 0 \quad \text{in } \Gamma \times (0,T), \\ U(\mathbf{x},0) &= 0, \quad \text{for } \mathbf{x} \in \Omega. \end{cases} \tag{2.66}$$

This has a unique solution $U = 0$, which contradicts $U = u_1 - u_2 \neq 0$. □

In this paper, we will consider the following advection-diffusion-reaction equation. Let $f \in L^2\left(0,T;H^{-1}(\Omega)\right)$, $g_D \in G_0$ and $u_0 \in L^2(\Omega)$. Then we find $u \in W^{1,2,2}\left(0,T;H^1(\Omega),H^{-1}(\Omega)\right)$ such that

$$\begin{cases} u' + A(t)u &= f, \quad \text{in } \Omega \times (0,T), \\ u &= g_D, \quad \text{in } \Gamma \times (0,T), \\ u(\mathbf{x},0) &= u_0(\mathbf{x}), \quad \text{for } \mathbf{x} \in \Omega, \end{cases} \tag{2.67}$$

where the linear operator $A(t) : H^1(\Omega) \to H^{-1}(\Omega)$ is defined by

$$A(t)v := -\nabla(\mu(\mathbf{x},t)\nabla v) + \mathbf{a}(\mathbf{x},t)\cdot\nabla v + c(\mathbf{x},t)v \text{ for all } v \in H^1(\Omega), \tag{2.68}$$

$\mu = (\mu_{ij})_{1\le i,j\le d}$ is diffusivity tensor, $\mathbf{a} = (a_i)_{1\le i\le d}$ is advection vector, c is reaction. Henceforth we make the following assumption.

Assumption 3. For the diffusivity tensor μ, advection vector $\mathbf{a}$ and reaction c, we assume the following additional condition.

- $\mu \in \left(L^\infty\left(0,T;W^{1,\infty}(\Omega)\right)\right)^{d\times d}$ satisfies the ellipticity, that is, there exists a positive constant μ_0 such that
$$\boldsymbol{\xi}^T \mu(\mathbf{x},t)\boldsymbol{\xi} \geq \mu_0|\boldsymbol{\xi}|^2, \text{ for all } \boldsymbol{\xi}\in\mathbb{R}^d, (\mathbf{x},t)\in\Omega\times(0,T). \tag{2.69}$$
- $\mathbf{a} \in \left(L^\infty\left(0,T;W^{1,\infty}(\Omega)\right)\right)^{d}$ satisfies $\nabla\cdot\mathbf{a}\in L^\infty(\Omega)$ for a.e. $t\in(0,T)$.
- $c \in L^\infty\left(0,T;L^\infty(\Omega)\right)$.
- $p := \operatorname{ess\,inf}_{\mathbf{x}\in\Omega}\left\{c - \frac{1}{2}\nabla\cdot\mathbf{a}\right\} > 0$ for a.e. $t\in(0,T)$.

We can now show the following lemma under the above assumptions.

Lemma 10. The operator $A(t) : H^1(\Omega)\to H^{-1}(\Omega)$ satisfies the equation (2.48) for a.e. $t\in(0,T)$, that is, the problem (2.67) has a unique weak solution.

Proof. First we have
$$\langle -\nabla\left(\mu\nabla w\right), v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} := \int_\Omega (\mu\nabla w)\cdot\nabla v\ dx. \tag{2.70}$$
This implies
$$\langle A(t)w, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = \int_\Omega \left((\mu\nabla w)\cdot\nabla v + (\mathbf{a}\cdot\nabla w)v + cwv\right)\ dx \tag{2.71}$$
for $w\in H^1(\Omega)$ and $v\in H_0^1(\Omega)$. Therefore,
$$\begin{aligned}\langle A(t)w, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &\leq \max_{0\leq i,j\leq d}\|\mu_{ij}\|_{L^\infty(\Omega)}|w|_{H^1(\Omega)}|v|_{H^1(\Omega)}\\ &\quad+\max_{0\leq i\leq d}\|a_i\|_{L^\infty(\Omega)}|w|_{H^1(\Omega)}\|v\|_{L^2(\Omega)}\\ &\quad+\|c\|_{L^\infty(\Omega)}\|w\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}\\ &\leq C\|w\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}\end{aligned} \tag{2.72}$$
for all $w\in H^1(\Omega)$, $v\in H_0^1(\Omega)$ and for a.e. $t\in(0,T)$.

Next, we have
$$\begin{aligned}\int_\Omega(\mathbf{a}\cdot\nabla v)\,v dx &= \int_\Omega\left(\nabla\cdot(\mathbf{a}v)v - \nabla\cdot\mathbf{a}v^2\right)dx\\ &= \int_\Gamma \mathbf{a}\cdot\mathbf{n}v^2dx - \int_\Omega(\mathbf{a}\cdot\nabla v)\,vdx - \int_\Omega\nabla\cdot\mathbf{a}v^2dx\end{aligned} \tag{2.73}$$
for all $v\in H_0^1(\Omega)$. Here $v=0$ on Γ leads to
$$\int_\Omega(\mathbf{a}\cdot\nabla v)\,vdx = -\frac{1}{2}\int_\Omega\nabla\cdot\mathbf{a}v^2dx, \text{ for all } v\in H_0^1(\Omega). \tag{2.74}$$
From this equation, it follows that
$$\begin{aligned}\langle A(t)v, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} &\geq \mu_0|v|^2_{H^1(\Omega)} + \int_\Omega\left(c-\frac{1}{2}\nabla\cdot\mathbf{a}\right)v^2dx\\ &\geq \mu_0|v|^2_{H^1(\Omega)} + p\|v\|^2_{L^2(\Omega)}\end{aligned} \tag{2.75}$$
for all $v\in H_0^1(\Omega)$. As $p>0$ we have
$$\langle A(t)v, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq \min\{\mu_0,p\}\|v\|^2_{H^1(\Omega)} \text{ for all } v\in H_0^1(\Omega). \tag{2.76}$$
□

Remark 9. We review that the assumption $p>0$ is valid. First, we have the Gårding's inequality
$$\langle A(t)v, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \geq C\|v\|^2_{H^1(\Omega)} - \kappa\|v\|^2_{L^2(\Omega)} \text{ for all } v\in H_0^1(\Omega), \tag{2.77}$$

where

$$\kappa > \sup_{\Omega\times(0,T)} \frac{1}{2}\nabla\cdot\mathbf{a} - c. \tag{2.78}$$

In fact, we have

$$\begin{aligned}\langle A(t)v, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} + \kappa\|v\|^2_{L^2(\Omega)} &\geq \mu_0|v|_{H^1(\Omega)} + \int_\Omega \left(c - \frac{1}{2}\nabla\cdot\mathbf{a} + \kappa\right) v^2 dx \\ &\geq C\|v\|_{H^1(\Omega)} \text{ for all } v \in H_0^1(\Omega).\end{aligned} \tag{2.79}$$

Here we let

$$A_\kappa(t)w := A(t)w + \kappa w \text{ for all } w \in H_0^1(\Omega), \tag{2.80}$$

then we have $A_\kappa : H^1(\Omega) \to H^{-1}(\Omega)$ and the problem

$$\left\{\begin{array}{rll} \widehat{u}' + A_\kappa(t)\widehat{u} & = e^{-\kappa t} f, & \text{in } \Omega\times(0,T), \\ \widehat{u} & = e^{-\kappa t} g_D, & \text{in } \Gamma\times(0,T), \\ \widehat{u}(\mathbf{x},0) & = u_0(\mathbf{x}), & \text{for } \mathbf{x}\in\Omega \end{array}\right. \tag{2.81}$$

has a unique solution $\widehat{u} \in X$. Therefore, we have $u := e^{\kappa t}\widehat{u}$ is a unique weak solution of problem (2.43). This means that the problem results into the case $p > 0$.

Lemma 11 ([22], Theorem 5. of Chapter 7)**.** Further we assume that $f \in L^2\left(0,T;L^2(\Omega)\right)$,

$$g_D \in G := \left\{ g \in L^2\left(0,T;H^{1/2}(\Gamma)\right) : \begin{array}{l} \text{there exists } u_D \in W^{1,2,2}\left(0,T;H^2(\Omega),L^2(\Omega)\right) \\ \text{such that } \operatorname{Tr} u_D = g \text{ in } L^2\left(0,T;H^{1/2}(\Gamma)\right) \end{array} \right\}, \tag{2.82}$$

$\mu \in \left(L^\infty\left((0,T\times\Omega\right)\right)^{d\times d}$, $\mathbf{a} \in \left(L^\infty\left((0,T)\times\Omega\right)\right)^d$, $c \in L^\infty\left((0,T)\times\Omega\right)$, $f \in L^2\left(0,T;L^2(\Omega)\right)$ and $u_0 - u_D(0) \in H_0^1(\Omega)$, then $\widetilde{u} \in X_0$ satisfies

$$\widetilde{u} \in W^{1,2,2}\left(0,T;H^2(\Omega),L^2(\Omega)\right) \cap L^\infty\left(0,T;H_0^1(\Omega)\right). \tag{2.83}$$

Remark 10. We review Example 1.42 of [31]. Let $\Omega := [0,T]$ and

$$u(\mathbf{x},t) := \left\{\begin{array}{ll} 1 & \text{for } x \leq t, \\ 0 & \text{for } x < t. \end{array}\right. \tag{2.84}$$

Then this satisfies $u \in L^\infty\left((0,T)\times\Omega\right)$. However, u is not Bochner measurable, because $W^{k,\infty}(\Omega)$ is not separable. Therefore $u \notin L^\infty\left(0,T;L^\infty(\Omega)\right)$.

2.4 Finite dimensional subspace

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary Γ, then we have the trace operator $\operatorname{Tr} : H^1(\Omega) \to H^{1/2}(\Gamma)$. Hereinafter, $Tu \in H^{1/2}(\Gamma)$ is written as u for $u \in H^1(\Omega)$. Further, we let $\mathcal{T}_h$ be a partition of Ω, $\mathcal{E}_h$ be a set of all edges or faces of $\mathcal{T}_h$, and $\mathcal{E}_h^e := \{E \in \mathcal{E}_h : E \subset \Gamma\}$. We define $h_E := \operatorname{diam} E$ for $E \in \mathcal{E}_h$, and h_K be the mesh size of $K \in \mathcal{T}_h$. Further, we let $h := \max\{h_K : K \in \mathcal{T}_h\}$.

In the next chapter, we will state the Nitsche method for parabolic problems using the finite dimensional subspace $V_h \subset H^1(\Omega)$. The following assumptions will be needed throughout the chapter.

Assumption 4. There exists a positive constant C such that

$$h_{K_E} \leq C h_E, \text{ for all } E \in \mathcal{E}_h^e, \tag{2.85}$$

where K_E is an element such that $E \subset \partial K_E$. We note that $K_E \in \mathcal{T}_h$ is unique for each $E \in \mathcal{E}_h^e$.

Assumption 5. Let V_h be the finite dimensional subspace of $H^1(\Omega)$ which relates to $\mathcal{T}_h$. Then we assume that (A1)–(A3) hold, (A1) : Trace inequality.

$$\|v\|_{L^2(E)}^2 \le C\left(h_E^{-1}\|v\|_{L^2(K_E)}^2 + h_{K_E}|v|_{H^1(K_E)}^2\right), \text{ for all } v \in H^1(K_E), \tag{2.86}$$

where $K_E \in \mathcal{T}_h$ is an element such that $E \subset \partial K_E$, for $E \in \mathcal{E}_h^e$.

(A2) : Inverse inequality.

$$|v_h|_{H^1(K)} \le C h_K^{-1}\|v_h\|_{L^2(K)} \text{ for all } v_h \in V_h, K \in \mathcal{T}_h. \tag{2.87}$$

(A3) : Interpolation error estimate. Let k be the degree of elements of V_h, and ℓ be an integer satisfying $2 \le \ell \le k+1$. Then there exists a projection $\Pi_h : H^\ell(\Omega) \to V_h$ such that

$$\|w - \Pi_h w\|_{H^j(\Omega)} \le C h^{\ell-j}\|w\|_{H^\ell(\Omega)} \tag{2.88}$$

for $j = 0, 1, 2$ and for all $w \in H^\ell(\Omega)$.

Below we state that the finite element (FE) space and NURBS space satisfies the assumption 5 under some condition.

2.4.1 FE space

If we apply the finite element method for spatial semi-discretization, we first have to construct the mesh by the triangulation.

Definition 5 (see Section 3.1 of [30])**.** Let $\Omega \subset \mathbb{R}^d$ be a polyhedral domain. Then the space $\mathcal{T}_h$ is called a triangulation of Ω if

- $\overline{\Omega} = \bigcup_{\overline{K} \in \mathcal{T}_h} K$,
- each $K \in \mathcal{T}_h$ is a polyhedron with $\mathring{K} \neq \emptyset$,
- for all $K_1, K_2 \in \mathcal{T}_h$, then $K_1 \neq K_2 \Leftrightarrow K_1 \cap K_2 = \emptyset$,
- $\operatorname{diam} K =: h_K \le h$ for all $K \in \mathcal{T}_h$.

Moreover, if the triangulation $\mathcal{T}_h$ satisfies that

$$\overline{K}_1 \cap \overline{K}_2 \neq \emptyset \Rightarrow \overline{K}_1 \cap \overline{K}_2 \text{ is a common face, side or vertex of } K_1 \text{ and } K_2 \tag{2.89}$$

for all $K_1, K_2 \in \mathcal{T}_h$, then $\mathcal{T}_h$ is called an admissible triangulation.

Hereafter, we always assume that the triangulation $\mathcal{T}_h$ is admissible, and there exists a invertible affine map

$$F_K(\widehat{x}) := B_K \widehat{x} + \mathbf{b}_K \tag{2.90}$$

such that $K = F_K(\widehat{K})$ for all $K \in \mathcal{T}_h$, where $\widehat{K} \subset \mathbb{R}^d$ is a reference element, which is the unit d-simplex or d-cube.

Let $\mathbb{P}_k$ be the space of polynomials of degree less than or equal to k in variables, and $\mathbb{Q}_k$ be the space of polynomials of degree less than or equal to k with respect to each variable. When the reference element is the unit d-simplex, then we define the space of triangular finite elements V_h by

$$V_h := \{v_h \in C^0(\overline{\Omega}) : v_h|_K \in \mathbb{P}_k \text{ for all } K \in \mathcal{T}_h\}. \tag{2.91}$$

If the reference element is the unit d-cube, then we define the space of parallelepipedal finite elements X_h such that

$$V_h := \{v_h \in C^0(\overline{\Omega}) : v_h|_K \in \mathbb{Q}_k \text{ for all } K \in \mathcal{T}_h\}. \tag{2.92}$$

In both cases, we note that $V_h \subset H^1(\Omega)$ for all $k \ge 1$.

Let ρ_K be a radius of an inscribed circle,

$$\rho_K := \sup\{\operatorname{diam}(S) : S \text{ is a ball contained in } K\}. \tag{2.93}$$

Definition 6. The family of the triangulations $\{\mathcal{T}_h\}_h$ is said to be regular when there exists a positive constant σ such that

$$\frac{h_K}{\rho_K} \leq \sigma \text{ for all } K \in \bigcup_h \mathcal{T}_h. \tag{2.94}$$

Definition 7. The family of the triangulations $\{\mathcal{T}_h\}_h$ is said to be quasi-uniform when there exists a positive constant ν such that

$$\frac{h}{h_K} \leq \nu \text{ for all } K \in \bigcup_h \mathcal{T}_h. \tag{2.95}$$

Now we remind some inequalities which are bases of numerical analysis of partial differential equations. In this paper, we refer to [30] and only consider the Sobolev spaces $H^k(K)$. For the inequalities on $W^{k,p}(K)$, see [13].

Lemma 12 (Proposition 3.4.1 of [30])**.** For all $v \in H^m(K)$, the function $v \circ F_K$ belongs to $H^m(\widehat{K})$, and there exists two positive constant $C_1(n,m)$ and $C_2(n,m)$ such that

$$\begin{cases} |v \circ F_K|_{H^m(\widehat{K})} & \leq C_1 \|B_K\|^m |\det B_K|^{-1/2} |v|_{H^m(K)} \quad \text{for all } v \in H^m(K), \\ |\widehat{v} \circ F_K^{-1}|_{W^{m,p}(K)} & \leq C_2 \|B_K^{-1}\|^m |\det B_K|^{1/2} |\widehat{v}|_{H^m(\widehat{K})} \quad \text{for all } \widehat{v} \in H^m(\widehat{K}), \end{cases} \tag{2.96}$$

where

$$\|B_K\| := \sup_{\boldsymbol{\xi} \in \mathbb{R}^d} \frac{|B_K \boldsymbol{\xi}|}{|\boldsymbol{\xi}|}. \tag{2.97}$$

Lemma 13 (Proposition 3.4.2 of [30])**.**

$$\|B_K\| \leq \frac{h_K}{\rho_{\widehat{K}}} \text{ and } \|B_K^{-1}\| \leq \frac{h_{\widehat{K}}}{\rho_K}. \tag{2.98}$$

Lemma 14 (Trace inequality on element)**.** Let $\mathcal{T}_h$ be a triangulation, then

$$\|f\|_{L^2(\partial K)}^2 \leq C \rho_K^{-d} \left(h_K^{d-1} \|f\|_{L^2(K)}^2 + h_K^{d+1} |f|_{H^1(K)}^2 \right) \text{ for all } K \in \mathcal{T}_h, f \in H^1(K). \tag{2.99}$$

Moreover, if the family of the triangulations $\{\mathcal{T}_h\}_h$ is regular, then

$$\|f\|_{L^2(\partial K)}^2 \leq C \left(h_K^{-1} \|f\|_{L^2(K)}^2 + h_K |f|_{H^1(K)}^2 \right) \text{ for all } K \in \mathcal{T}_h, f \in H^1(K). \tag{2.100}$$

Proof. Fix $K \in \mathcal{T}_h$ arbitrary, and let $\widehat{K} := F_K^{-1}(K)$ be a reference element, then we have $\rho_{\widehat{K}}$ and $\mathrm{meas}(\widehat{K})$ are constant. Further,

$$|\det B_K|^{-1} = \frac{\mathrm{meas}(\widehat{K})}{\mathrm{meas}(K)} \leq C \rho_K^{-d}. \tag{2.101}$$

They yield

$$\begin{aligned} \|f\|_{L^2(\partial K)}^2 & \leq C h_K^{d-1} \|f \circ F_K\|_{L^2(\partial \widehat{K})}^2 \\ & \leq C h_K^{d-1} \|f \circ F_K\|_{H^1(\widehat{K})}^2 \\ & \leq C h_K^{d-1} \left(|\det B_K|^{-1} \|f\|_{L^2(K)}^2 + \|B_K\|^2 |\det B_K|^{-1} |f|_{H^1(K)}^2 \right) \\ & \leq C \rho_K^{-d} \left(h_K^{d-1} \|f\|_{L^2(K)}^2 + h_K^{d+1} |f|_{H^1(K)}^2 \right) \end{aligned} \tag{2.102}$$

for all $f \in H^1(K)$. If $\{\mathcal{T}_h\}_h$ is regular, then we have $\rho_K^{-d} \leq C h_K^{-d}$ for all $K \in \mathcal{T}_h \in \{\mathcal{T}_h\}_h$. □

Lemma 15 (Local inverse inequality, Lemma 1.138 of [19])**.** There exists a positive constant C such that

$$|v_h|_{H^1(K)} \leq C \rho_K^{-1} \|v_h\|_{L^2(K)} \text{ for all } K \in \mathcal{T}_h, v_h \in V_h. \tag{2.103}$$

Moreover, if the family of the triangulations $\{\mathcal{T}_h\}_h$ is regular, then

$$|v_h|_{H^1(K)} \leq C h_K^{-1} \|v_h\|_{L^2(K)} \text{ for all } K \in \mathcal{T}_h, v_h \in V_h. \tag{2.104}$$

Proof. Fix $K \in \mathcal{T}_h$ and $v_h \in V_h$ arbitrary, and let $\widehat{K} := F_K^{-1}(K)$ be a reference element. Then $v_h \circ F_K$ is polynomial on $\widehat{K}$, therefore we have

$$|v_h \circ F_K|_{H^1(\widehat{K})} \le C\|v_h \circ F_K\|_{L^2(\widehat{K})}. \tag{2.105}$$

We note that $h_{\widehat{K}}$ is a constant, and this gives

$$\begin{aligned} |v_h|_{H^1(K)} &\le C\|B_K^{-1}\||\det B_K|^{1/2}|v_h \circ F_K|_{H^1(\widehat{K})} \\ &\le C\rho_K^{-1}|\det B_K|^{1/2}\|v_h \circ F_K\|_{L^2(\widehat{K})} \\ &\le C\rho_K^{-1}\|v_h\|_{L^2(K)}. \end{aligned} \tag{2.106}$$

Further, if $\{\mathcal{T}_h\}_h$ is regular, then we have $\rho_K^{-1} \le Ch_K^{-1}$ for all $K \in \mathcal{T}_h \in \{\mathcal{T}_h\}_h$. □

Corollary 3 (Global inverse inequality)**.** Let the family of triangulation $\{\mathcal{T}_h\}_h$ be regular and quasi-uniform, then

$$|v_h|_{H^1(\Omega)} \le Ch^{-1}\|v_h\|_{L^2(\Omega)} \text{ for all } v_h \in V_h. \tag{2.107}$$

Let $\mathbf{x}_i, i = 1, \cdots, \dim V_h$ be nodes in $\overline{\Omega}$, then we can construct the functions $\phi_i \in V_h, i = 1, \cdots, \dim V_h$ such that

$$\phi_i(\mathbf{x}_j) = \delta_{ij} \text{ for } i, j = 1, \cdots, \dim V_h \text{ and } \operatorname{supp} \phi_i = \bigcup_{K \in \mathcal{T}_h, \mathbf{x}_i \in \overline{K}} \overline{K}, \tag{2.108}$$

where δ_{ij} is the Kronecker delta. These functions are called shape functions, and they form a basis of V_h. Further we have the Lagrange interpolation as follow, let $v \in C^0(\overline{\Omega})$, then we define

$$\Pi_h(v) := \sum_{i=1}^{\dim V_h} v(\mathbf{x}_i)\phi_i. \tag{2.109}$$

It is clear that $\Pi_h : C^0(\overline{\Omega}) \to V_h$ is a projection. Moreover, let $\mathbf{x}_i^K, i = 1, \cdots, M_K$ be the nodes in $\overline{K}$, and

$$\Pi_h^K(v) := \sum_{i=1}^{M_K} v(\mathbf{x}_i^K)\phi_i|_K \text{ for all } v \in C^0(\overline{\Omega}), \tag{2.110}$$

then we have

$$\Pi_h(v)|_K = \Pi_h^K(v) \text{ for all } K \in \mathcal{T}_h, v \in C^0(\overline{\Omega}). \tag{2.111}$$

Lemma 16 (Interpolation error estimate, Theorem 3.4.1 of [30])**.** Let s be a positive integer, $\ell := \min\{k+1, s\}$ and $0 \le m \le \ell$. Then there exists a positive constant C such that

$$|v - \Pi_h^K(v)|_{H^m(K)} \le C\frac{h_K^\ell}{\rho_K^m}|v|_{H^\ell(K)} \text{ for all } K \in \mathcal{T}_h, v \in H^s(K). \tag{2.112}$$

Moreover, if the family of the triangulations $\{\mathcal{T}_h\}_h$ is regular, then

$$|v - \Pi_h^K(v)|_{H^m(K)} \le Ch_K^{\ell-m}|v|_{H^\ell(K)} \text{ for all } K \in \mathcal{T}_h, v \in H^s(K). \tag{2.113}$$

2.4.2 NURBS space

The Isogeometric Analysis proposes to describe the computational domain by a NURBS geometry. Further, its mesh can define the finite dimensional subspace for the discretization in the Galerkin method. Here, we will review the definition and properties of NURBS.

Univariate B-spline basis functions on $[0, 1]$

We call a vector $\Xi := \{\xi_1, \xi_2, \ldots, \xi_m\}$ the *knot vector* if

$$\xi_1 \le \xi_2 \le \cdots \le \xi_m. \tag{2.114}$$

We note the repetition of the knots are allowed. Without loss of generality, we let $\xi_1 = 0$ and $\xi_m = 1$. Let k be a given positive integer. Then, the univariate B-spline functions of degree k associated with the knot vector Ξ are defined by the Cox-de Boor algorithm.

Definition 8. Let $\Xi = \{\xi_1, \dots, \xi_m\}$ be a knot vector. Then the k-th degree B-spline basis functions $\widehat{B}_{i,k}$ is defined by

$$\widehat{B}_{i,0}(\widehat{x}) := \begin{cases} 1 & \text{if } \xi_i \le \widehat{x} \le \xi_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = 0, \tag{2.115}$$

$$\widehat{B}_{i,k}(\widehat{x}) := \frac{\widehat{x} - \xi_i}{\xi_{i+k} - \xi_i} \widehat{B}_{i,k-1}(\widehat{x}) + \frac{\xi_{i+k+1} - \widehat{x}}{\xi_{i+k+1} - \xi_{i+1}} \widehat{B}_{i+1,k-1}(\widehat{x}), \text{ for } k \ge 0, \tag{2.116}$$

with $i = 1, \dots, m-k-1$. We note that $0/0 = 0$ should be replaced by 0 in this definition.

We state some properties of the B-spline basis functions of degree k. They are non-negative k-th degree piecewise polynomials such that $\widehat{B}_{i,k}(\widehat{x}) = 0$ for $\widehat{x} \notin [\xi_i, \xi_{i+k+1}]$. Now we introduce an alternative representation of Ξ to state the other properties. Let

$$\Xi = \{\underbrace{\zeta_0, \dots, \zeta_0}_{m_0 \text{ times}}, \underbrace{\zeta_1 \dots, \zeta_1}_{m_1 \text{ times}}, \dots, \underbrace{\zeta_N, \dots, \zeta_N}_{m_N \text{ times}}\}, \tag{2.117}$$

where $\zeta_0 \le \zeta_1 \le \dots \le \zeta_N$. Therein we denote the multiplicity of ζ_n by m_n. Assume that $m_n \le k+1$ for all knots, then $\widehat{B}_{i,k}$ has $k - m_n$ continuous derivatives at the internal node ζ_n. Furthermore, we say that the knot vector Ξ is k-open if $m_0 = m_N = k+1$. Let Ξ be a p-open knot vector, then $\widehat{B}_{i,k}$ form the partition of unity, and they also form the basis of *spline space*, that is, the space of piecewise polynomials of degree k with $k - m_n$ continuous derivatives at ζ_n, for $n = 1, \dots, n-1$.

Henceforth, we assume the knot vector Ξ is k-open. We define the *univariate spline* $S_k(\Xi)$ by

$$S_k(\Xi) := \operatorname{span}\{\widehat{B}_{i,k} : i = 1, \dots, m-k-1\}. \tag{2.118}$$

We also define a *quasi-interpolant* operator $\Pi_{k,\Xi} : L^\infty(I) \to S_k(\Xi)$ by

$$\Pi_{k,\Xi}(f) := \sum_{i=1}^{m-k-1} \lambda_{i,k}(f) \widehat{B}_{i,k}, \tag{2.119}$$

where the dual basis functions $\lambda_{i,k}$, which are given by

$$\lambda_{i,k}(f) := \int_{\operatorname{supp}\widehat{B}_{i,k}} f(s)\psi_i(s)\, ds, \tag{2.120}$$

the definition of ψ_i and the proof of following inequality are given in Chapter 4 of [33],

$$\begin{aligned} |\lambda_{i,k}(f)| &\le \|f\|_{L^q(\operatorname{supp}\widehat{B}_{i,k})} \|D^{k+1}\psi_i\|_{L^{q'}(\operatorname{supp}\widehat{B}_{i,k})} \\ &\le C |\operatorname{supp}\widehat{B}_{i,k}|^{-1/q} \|f\|_{L^q(\operatorname{supp}\widehat{B}_{i,k})} \text{ for all } q \in [1, \infty], \end{aligned} \tag{2.121}$$

where $\frac{1}{q} + \frac{1}{q'} = 1$, and the constant C depends only on k. We note $\lambda_{i,k}$ is a dual basis such that $\lambda_{i,k}\widehat{B}_{j,k} = \delta_{i,j}$, therefore the spline preserving property of $Pi_{k,\Xi}$ holds, that is,

$$\Pi_{k,\Xi}(f) = f \text{ for all } f \in S_{k,\Xi}. \tag{2.122}$$

Let

$$I_n := [\zeta_{n-1}, \zeta_n], \; h_n := |I_n|, \; \widetilde{I}_n := \bigcup \left\{ \operatorname{supp}\widehat{B}_{i,p} : \widehat{B}_{i,p}|_{I_n} \not\equiv 0 \right\}, \; \widetilde{h}_n := |\widetilde{I}_n|. \tag{2.123}$$

For the partition size h_n, the following assumption will be always needed in this chapter.

Assumption 6 (Local quasi-uniform)**.** The knot vector Ξ is locally quasi-uniform, that is, there exists a constant $\theta \ge 1$ such that

$$\frac{1}{\theta} \le \theta_n := \frac{h_n}{h_{n+1}} \le \theta \text{ for all } n = 1, \cdots, N-1. \tag{2.124}$$

Further, we may assume that the knot vector Ξ satisfies the following global quasi-uniformity.

Definition 9. A knot vector Ξ is called (global) quasi-uniform if there exists a positive constant θ such that

$$\frac{\max_n h_n}{\min_n h_n} \le \theta. \tag{2.125}$$

Now we state the stability estimate of quasi-interpolant operator.

Lemma 17 (Stability of quasi-interpolant operator, see Theorem 4.41 of [33] and Proposition 2.2 of [9])**.** For all $q \in [1, \infty]$, there exists a positive constant C such that

$$\|\Pi_{k,\Xi}(f)\|_{L^q(I_n)} \le C\|f\|_{L^q(\widetilde{I}_n)}, \tag{2.126}$$

for all $f \in L^\infty(I)$. Moreover, if Assumption 6 is satisfied,

$$|\Pi_{p,\Xi}(f)|_{W^{1,q}(I_n)} \le C|f|_{W^{1,q}(\widetilde{I}_n)}, \tag{2.127}$$

where C depends θ also.

Proof. The B-spline basis function form partition of unity, therefore

$$\begin{aligned} \|\Pi_{k,\Xi}(f)\|_{L^q(I_n)} &\le \left\| \sum_{\mathrm{supp}\widehat{B}_{i,k}\cap I_n \neq \emptyset} \lambda_{i,k}(f)\widehat{B}_{i,k} \right\|_{L^q(I_n)} \\ &\le \max_{\mathrm{supp}\widehat{B}_{i,k}\cap I_n \neq \emptyset} |\lambda_{i,k}(f)| \left\| \sum_{\mathrm{supp}\widehat{B}_{i,k}\cap I_n \neq \emptyset} \widehat{B}_{i,p} \right\|_{L^q(I_n)} \\ &\le h_n^{1/q} \max_{\mathrm{supp}\widehat{B}_{i,k}\cap I_n \neq \emptyset} |\lambda_{i,k}(f)|. \end{aligned} \tag{2.128}$$

Here we combine the equation (2.121) and $h_n \le |\mathrm{supp}\widehat{B}_{i,p}|$ for all i such that $\mathrm{supp}\widehat{B}_{i,p} \cap I_n \neq \emptyset$, then

$$\begin{aligned} \|\Pi_{k,\Xi}(f)\|_{L^q(I_n)} &\le C \max_{\mathrm{supp}\widehat{B}_{i,k}\cap I_n \neq \emptyset} \|f\|_{L^q(\mathrm{supp}\widehat{B}_{i,k})} \\ &\le C\|f\|_{L^q(\widetilde{I}_n)}. \end{aligned} \tag{2.129}$$

This estimate implies equation (2.127) as follows, we take a constant function c as

$$\|f - c\|_{L^q(\widetilde{I}_n)} \le C\widetilde{h}_n|f|_{W^{1,q}(\widetilde{I}_n)}, \tag{2.130}$$

then $\Pi_{k,\Xi}(f) - c = \Pi_{k,\Xi}(f - c)$ is a k-the degree polynomial on I_n, therefore

$$\begin{aligned} |\Pi_{k,\Xi}(f)|_{W^{1,q}(I_n)} &= |\Pi_{k,\Xi}(f-c)|_{W^{1,q}(I_n)} \\ &\le Ch_n^{-1}\|\Pi_{k,\Xi}(f-c)\|_{L^q(I_n)} \\ &\le Ch_n^{-1}\|f-c\|_{L^q(\widetilde{I}_n)} \le C|f|_{W^{1,q}(\widetilde{I}_n)}, \end{aligned} \tag{2.131}$$

because the assumption 6 leads $\widetilde{h}_n \le Ch_n$. □

Combining the stability estimate and spline preserving property of $\Pi_{k,\Xi}$, we obtain the following error estimate.

Lemma 18 (Error estimate)**.** Let s be a positive integer, $\ell := \min\{k+1, s\}$ and $0 \le m \le \ell$. Then there exists a positive constant C such that

$$\|f - \Pi_{k,\Xi}(f)\|_{L^q(I_n)} \le C\widetilde{h}_n^{\ell}|f|_{W^{\ell,q}(\widetilde{I}_n)} \text{ for all } f \in W^{s,q}(I) \text{ and } n = 1, \cdots, N. \tag{2.132}$$

Moreover, if Assumption 6 is satisfied, There exists a constant C such that

$$|f - \Pi_{k,\Xi}(f)|_{W^{m,q}(I_n)} \le C\widetilde{h}_n^{\ell-m}|f|_{W^{\ell,q}(\widetilde{I}_n)} \text{ for all } f \in W^{s,q}(I) \text{ and } n = 1, \cdots, N. \tag{2.133}$$

Proof. For $q=2$, we can see the Proposition 4.2 of [9], but here we show them for general q. We note $f|_{\widetilde{I}_n} \in W^{s,q}(\widetilde{I}_n)$ for all $f \in W^{s,q}(I)$. Let

$$T_n^\ell(f) = \sum_{i=0}^{\ell-1} \frac{f^{(\tau)}(t_{n-1})}{(\ell-1)!}(t-t_{n-1})^{\ell-1} d\tau \in \mathbb{P}_{\ell-1}(\widetilde{I}_n) \tag{2.134}$$

be the Taylor polynomial of degree $\ell-1$, then we can get easily

$$|f - T_n^\ell(f)|_{W^{m,q}(\widetilde{I}_n)} \leq C\widetilde{h}_n^{\ell-m}|f|_{W^{\ell,q}(\widetilde{I}_n)}. \tag{2.135}$$

Since $\ell \leq k+1$, we obtain $T_n^\ell(f)|_{I_n} \in \mathbb{P}_{\ell-1}(I_n) \subset S_k(\Xi)$. Here the spline preserving property yields

$$\begin{aligned}\|f - \Pi_{k,\Xi}(f)\|_{L^q(I_n)} &\leq \|f - T_n^\ell(f)\|_{L^q(I_n)} + \|\Pi_{k,\Xi}(f) - T_n^\ell(f)\|_{L^q(I_n)} \\ &\leq \|f - T_n^\ell(f)\|_{L^q(I_n)} + \|\Pi_{k,\Xi}(f - T_n^\ell(f))\|_{L^q(I_n)} \\ &\leq C\|f - T_n^\ell(f)\|_{L^q(\widetilde{I}_n)} \\ &\leq C\widetilde{h}_n^\ell |f|_{W^{\ell,q}(\widetilde{I}_n)}.\end{aligned} \tag{2.136}$$

Further, the inverse inequality in one dimension leads to

$$\begin{aligned}|f - \Pi_{k,\Xi}(f)|_{W^{m,q}(I_n)} &\leq |f - T_n^\ell(f)|_{W^{m,q}(I_n)} + |\Pi_{k,\Xi}(f - T_n^\ell(f))|_{W^{m,q}(I_n)} \\ &\leq |f - T_n^\ell(f)|_{W^{m,q}(\widetilde{I}_n)} + Ch_n^{-m}\|\Pi_{k,\Xi}(f - T_n^\ell(f))\|_{L^q(I_n)} \\ &\leq C(\widetilde{h}_n^{\ell-m} + \widetilde{h}_n^\ell h_n^{-m})|f|_{W^{\ell,q}(\widetilde{I}_n)},\end{aligned} \tag{2.137}$$

here the Assumption 6 implies $h_n^{-m} \leq C\widetilde{h}_n^{-m}$. □

Multivariate B-spline basis functions and NURBS basis functions

Let d be the dimension of space. For given degree k_j and k_j-open knot vector

$$\Xi_j := \{\xi_{j,1}, \ldots, \xi_{j,m_j}\} \tag{2.138}$$

$$= \{\underbrace{\zeta_{j,1}, \ldots, \zeta_{j,1}}_{k_j+1 \text{ times}}, \underbrace{\zeta_{j,2}, \ldots, \zeta_{j,2}}_{m_{j,2} \text{ times}}, \ldots, \underbrace{\zeta_{j,r_j}, \ldots, \zeta_{j,r_j}}_{k_j+1 \text{ times}}\}, \tag{2.139}$$

we get the k_j-th degree univariate B-spline basis functions

$$\widehat{B}_{i_j,k_j}(\widehat{x}_j), i_j = 1, 2, \ldots, m_j - k_j - 1, \tag{2.140}$$

and we define

$$\mathbf{k} := (k_1, \ldots, k_d), \tag{2.141}$$

$$\boldsymbol{\Xi} := \Xi_1 \times \cdots \times \Xi_d. \tag{2.142}$$

Furthermore, the knots without repetition provides the mesh on parametric domain $\widehat{\Omega} := [0,1]^d$, which is denoted by $\widehat{\mathcal{M}}_h$:

$$\widehat{\mathcal{M}}_h := \{Q_\mathbf{s} := I_{1,s_1} \times \cdots \times I_{d,s_d} : I_{j,s_j} := (\zeta_{i,s_j}, \zeta_{i,s_j+1}), s_j = 1, \ldots, r_j - 1\}. \tag{2.143}$$

Then we define the multivariate B-spline basis functions:

$$\widehat{\mathbf{B}}_{\mathbf{i},\mathbf{k}}(\widehat{\mathbf{x}}) := \widehat{B}_{i_1,k_1}(\widehat{x}_1) \cdots \widehat{B}_{i_d,k_d}(\widehat{x}_d) \tag{2.144}$$

for $\mathbf{i} = (i_1, \ldots, i_d)$, where $\widehat{\mathbf{x}} := (\widehat{x}_1, \ldots, \widehat{x}_d) \in \widehat{\Omega} := (0,1)^d$. We define the *multivariate spline* $S_\mathbf{k}(\boldsymbol{\Xi})$ by

$$S_\mathbf{k}(\boldsymbol{\Xi}) := S_{k_1}(\Xi_1) \otimes \cdots \otimes S_{k_d}(\Xi_d) = \operatorname{span}\{\widehat{\mathbf{B}}_{\mathbf{i},\mathbf{k}} : \mathbf{i} \in \mathbf{I}\}, \tag{2.145}$$

where $\mathbf{I} := \{\mathbf{i} = (i_1, \ldots, i_d) : i_j = 1, \ldots, m_j - k_j - 1\}$. The quasi-interpolation for multivariate B-spline is defined also by the tensor product;

$$\Pi_{\mathbf{k},\boldsymbol{\Xi}}(f) := \Pi_{k_1,\Xi_1} \otimes \cdots \otimes \Pi_{k_d,\Xi_d} : L^\infty(\widehat{\Omega}) \to S_\mathbf{k}(\boldsymbol{\Xi}). \tag{2.146}$$

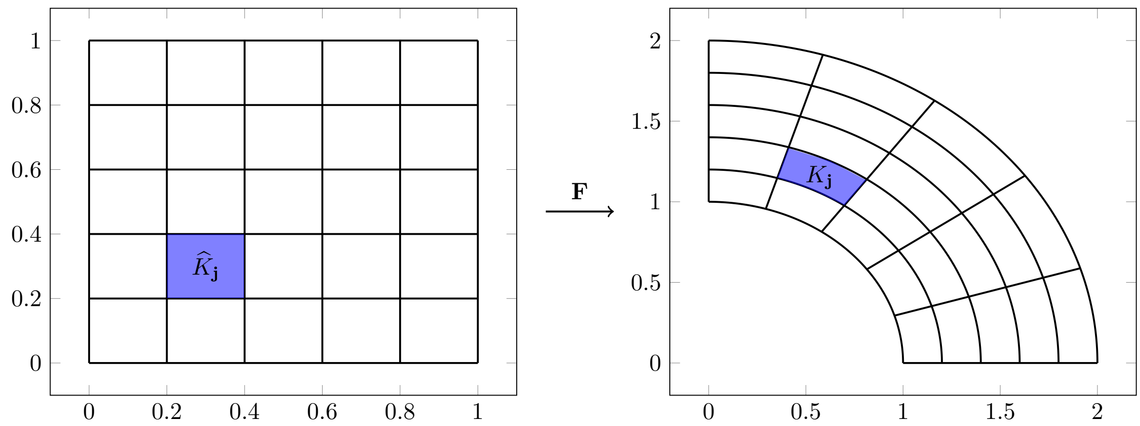


Figure 2.1: Parametric mesh and the mesh of physical domain Ω

Remark 11. In general, we can not find the stability of $\Pi_{\mathbf{k},\mathbf{\Xi}}$ because $L^\infty((0,1)^2) \neq L^\infty(0,1) \times L^\infty(0,1)$

Here, the definition of NURBS basis functions for given weight

$$W(\widehat{\mathbf{x}}) := \sum_{\mathbf{j} \in \mathbf{I}} w_{\mathbf{j}} \widehat{\mathbf{B}}_{\mathbf{j},\mathbf{k}}(\widehat{\mathbf{x}}) \tag{2.147}$$

are described as follows:

$$\widehat{\mathbf{N}}_{\mathbf{i},\mathbf{k}}(\widehat{\mathbf{x}}) := \frac{w_{\mathbf{i}} \widehat{\mathbf{B}}_{\mathbf{i},\mathbf{k}}(\widehat{\mathbf{x}})}{W(\widehat{\mathbf{x}})}, \tag{2.148}$$

where the positive constant $w_{\mathbf{j}} > 0$, $\mathbf{j} \in \mathbf{I}$ are called weights. Furthermore, a NURBS parametrization is given by a linear combination of NURBS basis functions. Let $P_{\mathbf{i}} \in \mathbb{R}^{\widehat{d}}$ be control points, then a NURBS parametrization $\mathbf{F}(\widehat{\mathbf{x}})$ is given by

$$\mathbf{F}(\widehat{\mathbf{x}}) := \sum_{\mathbf{i} \in \mathbf{I}} P_{\mathbf{i}} \widehat{\mathbf{N}}_{\mathbf{i},\mathbf{k}}(\widehat{\mathbf{x}}). \tag{2.149}$$

We denote the physical domain by $\Omega := \mathbf{F}(\widehat{\Omega})$, then the mesh on Ω is provided as the image of parametric mesh:

$$\mathcal{T}_h := \{K_{\mathbf{j}} := \mathbf{F}(\widehat{K}_{\mathbf{j}}), \widehat{K}_{\mathbf{j}} \in \widehat{\mathcal{M}}_h\} \tag{2.150}$$

The requirement on the map $\mathbf{F}$ is that it satisfies the following regularity.

Assumption 7. The map $\mathbf{F}$ is homeomorphism, and $\mathbf{F}^{-1}|_K$ and its inverse are smooth.

Under this assumption, we can define

$$V_h := \mathrm{span}\{\mathbf{N}_{\mathbf{i},\mathbf{k}}(\mathbf{x}) := \widehat{\mathbf{N}}_{\mathbf{i},\mathbf{k}} \circ \mathbf{F}^{-1}(\mathbf{x}), \mathbf{i} \in \mathbf{I}\} \tag{2.151}$$

where h is the mesh size $h := \max\{h_Q := \mathrm{diam}(Q) : Q \in \widehat{\mathcal{M}}_h\}$. Further, we define

$$h_K := \|\nabla \mathbf{F}\|_{L^\infty(Q)} h_Q \text{ for all } K \in \mathcal{T}_h, Q := \mathbf{F}^{-1}(K). \tag{2.152}$$

For the NURBS mesh, we define the regularity of the family of mesh $\{\mathcal{T}_h\}_h$ using $\{\widehat{\mathcal{M}}_h\}$.

Definition 10. The family of the mesh $\{\mathcal{T}_h\}_h$ is said to be regular when there exists a positive constant σ such that

$$\frac{h_Q}{h_{Q,\min}} \leq \sigma \text{ for all } Q \in \bigcup_h \widehat{\mathcal{M}}_h, \tag{2.153}$$

where $h_{Q,\min}$ denotes the length of the smallest edge of hypercube Q.

When we consider the Isogeometric Analysis, we assume the regularity of mesh. Further we always assume that the family of mesh $\{\mathcal{T}_h\}_h$ is locally quasi-uniform, that is, there exists a positive constant θ such that for all $\widehat{\mathcal{M}}_h \in \{\widehat{\mathcal{M}}_h\}_h$, the knots vectors $\Xi_1, \cdots, \Xi_d$ satisfies the local quasi-uniformity for θ. We may further assume that $\{\mathcal{T}_h\}_h$ satisfies global quasi-uniformity.

Now we review some results in previous researches.

Lemma 19 (Trace inequality, Theorem 3.2 of [20]). Let $K \in \mathcal{T}_h$ and $Q = \mathbf{F}^{-1}(K)$, then

$$\|f\|^2_{L^2(\partial K)} \leq C\lambda_Q\lambda_K \left(h_K^{-1}\|f\|^2_{L^2(K)} + h_K|f|^2_{H^1(K)}\right) \text{ for all } f \in H^1(K), \tag{2.154}$$

where λ_Q and λ_K are local shape regularity constant of Q, K, respectively and they are independent of h_K.

Lemma 20 (Inverse inequality, Theorem 4.2. of [2]). Let ℓ be an integer with $0 \leq k \leq \ell$, then we have

$$\|v_h\|_{H^\ell(K)} \leq C_{\text{shape}} h_K^{k-\ell} \sum_{i=0}^{k} \|\nabla \mathbf{F}\|^{i-k}_{L^\infty(\mathbf{F}^{-1}(K))} |v_h|_{H^\ell(K)} \text{ for all } K \in \mathcal{T}_h, v_h \in V_h. \tag{2.155}$$

Especially, we have

$$|v_h|_{H^1(K)} \leq \|v_h\|_{H^1(K)} \leq Ch_K^{-1}\|v_h\|_{L^2(K)} \text{ for all } K \in \mathcal{T}_h, v_h \in V_h. \tag{2.156}$$

Lemma 21 (Quasi-interpolation error estimate, Corollary 4.21 of [9]). Let the projection $\Pi_{V_h} : L^2(\Omega) \to V_h$ be

$$\Pi_{V_h} f(\mathbf{x}) := \frac{\Pi_{\mathbf{k},\boldsymbol{\Xi}}\left(W \circ \mathbf{F}^{-1}(\mathbf{x}) f(\mathbf{x})\right)}{W \circ \mathbf{F}^{-1}(\mathbf{x})} \text{ for } f \in L^2(\Omega). \tag{2.157}$$

Further we let s be an integer, $\ell := \min\{k_1+1, \cdots, k_d+1, s\}$ and $0 \leq m \leq \ell$. Then there exists a positive constant C such that

$$\|v - \Pi_{V_h} v\|_{H^m(K)} \leq Ch_{\widetilde{K}}^{\ell-m}\|v\|_{H^\ell(\widetilde{K})} \text{ for all } K \in \mathcal{T}_h, v \in H^s(\Omega), \tag{2.158}$$

where $\widetilde{K} := \mathbf{F}(\widetilde{Q})$ for $Q := \mathbf{F}^{-1}(K)$, and

$$\widetilde{Q} := \bigcup \left\{\operatorname{supp} \widehat{\mathbf{N}}_{\mathbf{i},\mathbf{k}} : \widehat{\mathbf{N}}_{\mathbf{i},\mathbf{k}}|_Q \not\equiv 0\right\}. \tag{2.159}$$

2.5 Application of the Nitsche method to parabolic problems

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary Γ, then we have the trace operator $\mathrm{Tr} : H^1(\Omega) \to H^{1/2}(\Gamma)$. Hereinafter, $Tu \in H^{1/2}(\Gamma)$ is written as u for $u \in H^1(\Omega)$.

In this section, we establish the inf-sup condition after introducing the Nitsche method for parabolic problems. We also introduce a subspace $V \subset H^1(\Omega)$ such that $A(t)w \in \left(L^2(\Omega)\right)^*$ for all $w \in V$, where $A(t) : H^1(\Omega) \to H^{-1}(\Omega)$ is elliptic operator. The important point we note here is that the Nitsche method for parabolic problems satisfies the Galerkin orthogonality if the weak solution u is in X_V, where

$$X_V := \left\{v \in W^{1,2,2}\left(0, T; H^1(\Omega), L^2(\Omega)\right) : v(t) \in V \text{ for a.e. } t \in (0,T)\right\}. \tag{2.160}$$

This provides the quasi-optimal error estimate.

2.5.1 The piecewise Sobolev space and normal derivative

We consider spacial semi-discretization with FEM or IGA. Let $\mathcal{T}_h$ be a partition of Ω, $\mathcal{E}_h$ be a set of all edges or faces of $\mathcal{T}_h$, and $\mathcal{E}_h^e := \{E \in \mathcal{E}_h : E \subset \Gamma\}$. We define $h_E := \operatorname{diam} E$ for $E \in \mathcal{E}_h$, and h_K be the mesh size of $K \in \mathcal{T}_h$. Further we let $h := \max\{h_K : K \in \mathcal{T}_h\}$.

Now we define the following piecewise Sobolev space.

Definition 11 (piecewise Sobolev space)**.**

$$V := \left\{ v \in H^1(\Omega) : v|_K \in H^2(K) \text{ for all } K \in \mathcal{T}_h \right\}, \tag{2.161}$$

$$\|v\|_V^2 := \|v\|_{H^1(\Omega)}^2 + \sum_{K\in\mathcal{T}_h} h_K^2 |v|_{H^2(K)}^2 + \sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|v\|_{L^2(E)}^2. \tag{2.162}$$

Now the elements of V satisfy the following lemma.

Lemma 22.

$$\mathbf{n}\cdot\mu\nabla w := \sum_{i=1}^{d} n_i \,\mathrm{Tr}\left(\sum_{j=1}^{d} \mu_{ij}\frac{\partial w}{\partial x_j}\right) \in L^2(E) \tag{2.163}$$

for all $\mu \in \left(W^{1,\infty}(\Omega)\right)^{d\times d}$, $w \in V$ and $E \in \mathcal{E}_h^e$. Note that $\mathrm{Tr} : H^1(\Omega) \to H^{1/2}(\Gamma)$ is a trace operator.

Proof. Note that $W^{1,\infty}(K) \subset H^1(K)$ for all $K \in \mathcal{T}_h$, because $K \in \mathbb{R}^d$ is bounded. Therefore we have

$$\mu_{ij}|_K \in H^1(K) \text{ and } \left.\frac{\partial w}{\partial x_j}\right|_K \in H^1(K) \tag{2.164}$$

for all $i, j = 1, \dots, d$ and $w \in V$. This implies

$$\left.\sum_{i=1}^{d} n_i \,\mathrm{Tr}\left(\sum_{j=1}^{d} \mu_{ij}\frac{\partial w}{\partial x_j}\right)\right|_E \in H^{1/2}(E) \subset L^2(E). \tag{2.165}$$

□

Lemma 23. There exists a positive constant C such that

$$\sum_{E\in\mathcal{E}_h^e} h_E \|\mathbf{n}\cdot\mu\nabla w\|_{L^2(E)}^2 \le C \sum_{E\in\mathcal{E}_h^e} \left(|w|_{H^1(K_E)}^2 + h_{K_E}^2 |w|_{H^2(K_E)}^2\right) \tag{2.166}$$

for all $\mu \in \left(W^{1,\infty}(\Omega)\right)^{d\times d}$ and $w \in V$, where $K_E \in \mathcal{T}_h$ satisfies $E \subset \partial K_E$.

Proof. We note that $\mu_{ij}\dfrac{\partial w}{\partial x_j} \in L^2(E)$, therefore we can apply the trace inequality (A2) and have

$$\begin{aligned}
\|\mathbf{n}\cdot\mu\nabla w\|_{L^2(E)}^2 &= \int_E \left(\sum_{i=1}^{d} n_i \sum_{j=1}^{d} \mathrm{Tr}\left(\mu_{ij}\frac{\partial w}{\partial x_j}\right)\right)^2 dx \\
&\le \int_E \left(\sum_{i=1}^{d} n_i^2\right)\left(\sum_{i=1}^{d}\left(\sum_{j=1}^{d} \mathrm{Tr}\left(\mu_{ij}\frac{\partial w}{\partial x_j}\right)\right)^2\right) dx \\
&\le d \sum_{i,j=1}^{d} \left\|\mathrm{Tr}\left(\mu_{ij}\frac{\partial w}{\partial x_j}\right)\right\|_{L^2(E)}^2 \\
&\le C \sum_{i,j=1}^{d} \left(h_{K_E}^{-1}\left\|\mu_{ij}\frac{\partial w}{\partial x_j}\right\|_{L^2(K_E)}^2 + h_{K_E}\left|\mu_{ij}\frac{\partial w}{\partial x_j}\right|_{H^1(K_E)}^2\right) \\
&\le C\left(h_{K_E}^{-1}|w|_{H^1(K_E)}^2 + h_{K_E}|w|_{H^2(K_E)}^2\right)
\end{aligned} \tag{2.167}$$

for all $E \in \mathcal{E}_h^e$. This and $h_E \le h_{K_E}$ imply the conclusion. □

Corollary 4. There exists a positive constant C such that

$$\sum_{E\in\mathcal{E}_h^e} h_E \|\mathbf{n}\cdot\mu\nabla w\|_{L^2(E)}^2 \le C\left(|w|_{H^1(\Omega)}^2 + \sum_{K\in\mathcal{T}_h} h_K^2 |w|_{H^2(K)}^2\right) \tag{2.168}$$

for all $\mu \in \left(W^{1,\infty}(\Omega)\right)^{d\times d}$ and $w \in V$.

2.5.2 Spatial semi-discretization

Here we let $V_h \subset V \subset H^1(\Omega)$ be a finite dimensional subspace. Further, we define $(\cdot,\cdot)_{V_h}$ and $\|\cdot\|_{V_h}$ by

$$\begin{cases} (w_h, v_h)_{V_h} & := (w_h, v_h)_{L^2(\Omega)} + (\nabla w_h, \nabla v_h)_{L^2(\Omega)} + \sum_{E\in\mathcal{E}_h^e} h_E^{-1}(w_h, v_h)_{L^2(E)}, \\ \|v_h\|_{V_h}^2 & := \|v_h\|_{H^1(\Omega)}^2 + \sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|v_h\|_{L^2(E)}^2 \end{cases} \tag{2.169}$$

for all $w_h, v_h \in V_h$. It is obvious that $(v_h, v_h)_{V_h} = \|v_h\|_{V_h}^2$, and therefore V_h is a Hilbert space.

The definition of $\|\cdot\|_{V_h}$ implies $V_h \hookrightarrow H^1(\Omega)$. Further we can define the dual space V_h^*, and check the following.

- V_h^* is a Banach space with norm

$$\|\phi\|_{V_h^*} := \sup_{0\neq v_h\in V_h} \frac{\langle \phi, v_h\rangle_{V_h^*,V_h}}{\|v_h\|_{V_h}} \text{ for } \phi \in V_h^*. \tag{2.170}$$

- We can define a map $T_1 : L^2(\Omega) \to V_h^*$ by

$$\langle T_1 f, v_h\rangle_{V_h^*,V_h} := (f, v_h)_{L^2(\Omega)} \text{ for all } f \in L^2(\Omega), v_h \in V_h. \tag{2.171}$$

  Especially, we have

$$(f, v_h)_{L^2(\Omega)} \leq \|T_1 f\|_{V_h^*}\|v_h\|_{V_h} \text{ for all } f \in L^2(\Omega), v_h \in V_h, \tag{2.172}$$

  and

$$\|T_1 f\|_{V_h^*} \leq \sup_{v\in H^1(\Omega)} \frac{(f, v)_{L^2(\Omega)}}{\|v\|_{L^2(\Omega)}} \leq \|f\|_{L^2(\Omega)} \text{ for all } f \in L^2(\Omega). \tag{2.173}$$

Here, we show the following lemmas.

Lemma 24.

$$\|v_h\|_{V_h} \leq \|v_h\|_V \leq C\|v_h\|_{V_h} \text{ for all } v_h \in V_h \subset V, \tag{2.174}$$

that is, The two norms $\|\cdot\|_{V_h}$ and $\|\cdot\|_V$ are equivalent in V_h.

Proof. First, the definition of the norms implies

$$\|v_h\|_V^2 = \|v_h\|_{V_h}^2 + \sum_{K\in\mathcal{T}_h} h_K^2 |v_h|_{H^2(K)}^2. \tag{2.175}$$

This leads $\|v_h\|_V \geq \|v_h\|_{V_h}$. Moreover, the assumption (A2) implies $\|v_h\|_V \leq C\|v_h\|_{V_h}$, for all $v_h \in V_h$. □

Lemma 25. There exists a positive constant C_I such that

$$\sum_{E\in\mathcal{E}_h^e} h_E \|\mathbf{n}\cdot\mu\nabla w_h\|_{L^2(E)}^2 \leq C_I |w_h|_{H^1(\Omega)}^2 \tag{2.176}$$

for all $w_h \in V_h$.

Proof. We apply the lemma 23 and the assumption (A2) for $w_h \in V_h \subset V$, then we have

$$\begin{aligned} \|\mathbf{n}\cdot\mu\nabla w_h\|_{L^2(E)}^2 &= \int_E \left(\sum_{i=1}^d n_i \sum_{j=1}^d T\left(\mu_{ij}\frac{\partial w_h}{\partial x_j}\right)\right)^2 dx \\ &\leq C\left(h_{K_E}^{-1}|w_h|_{H^1(K_E)}^2 + h_{K_E}|w_h|_{H^2(K_E)}^2\right) \\ &\leq C h_{K_E}^{-1}|w_h|_{H^1(K_E)}^2 \end{aligned} \tag{2.177}$$

for all $E \in \mathcal{E}_h^e$. This and $h_E \leq h_{K_E}$ implies

$$\sum_{E\in\mathcal{E}_h^e} h_E \|\mathbf{n}\cdot\mu\nabla w_h\|_{L^2(E)}^2 \leq C \sum_{E\in\mathcal{E}_h^e} |w_h|_{H^1(K_E)}^2 \leq C_I |w_h|_{H^1(\Omega)}^2 \tag{2.178}$$

for all $w_h \in V_h$. □

2.5.3 Elliptic operator in Nitsche method

The elliptic operator satisfies the following.

Lemma 26. Let $A(t) : H^1(\Omega) \to H^{-1}(\Omega)$ be defined by the equation (2.68), where μ, $\mathbf{a}$ and c satisfy the assumption 3. Then there exist an operator $\widehat{A}(t) : V \to \big(H^1(\Omega)\big)^*$, which is linear for a.e. $t \in (0,T)$, and two positive constant $\widehat{M}$ and $\widehat{\alpha}$ such that

$$\left\{\begin{array}{l} A(t)w \in \big(L^2(\Omega)\big)^* \subset H^{-1}(\Omega) \text{ for all } w \in V \text{ and} \\ \quad \langle A(t)w, v\rangle_{(L^2(\Omega))^*, L^2(\Omega)} = \langle \widehat{A}(t)w, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} - \displaystyle\sum_{E\in\mathcal{E}_h^e} (\mathbf{n}\cdot\mu\nabla w, v)_{L^2(E)} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for all } w \in V, v \in H^1(\Omega), \\ \langle \widehat{A}(t)w, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} \le \widehat{M}\|w\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \text{ for all } w \in V, v \in H^1(\Omega), \\ \langle \widehat{A}(t)v, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} \ge \widehat{\alpha}\|v\|^2_{H^1(\Omega)} + \dfrac{1}{2}(\mathbf{a}\cdot\mathbf{n}v, v)_{L^2(\Gamma_{\text{in}})} \text{ for all } v \in V \end{array}\right. \tag{2.179}$$

for a.e. $t \in (0,T)$, where $\Gamma_{\text{in}} \subset \Gamma$ is defined by

$$\Gamma_{\text{in}} = \Gamma_{\text{in}}(t) := \{\mathbf{x} \in \Gamma : \mathbf{a}(\mathbf{x},t)\cdot\mathbf{n}(\mathbf{x}) < 0\}. \tag{2.180}$$

Proof. We can define $A(t) : V \to \big(L^2(\Omega)\big)^*$ by

$$\langle A(t)w, v\rangle_{(L^2(\Omega))^*, L^2(\Omega)} = \sum_{K\in\mathcal{T}_h} (-\nabla(\mu\nabla w), v)_{L^2(K)} + (\mathbf{a}\cdot\nabla w, v)_{L^2(\Omega)} + (cw, v)_{L^2(\Omega)} \tag{2.181}$$

for all $w \in V$ and $v \in L^2(\Omega)$. Furthermore, we have

$$\begin{aligned} \langle A(t)w, v\rangle_{(L^2(\Omega))^*, L^2(\Omega)} &= (\mu\nabla w, \nabla v)_{L^2(\Omega)} - \sum_{E\in\mathcal{E}_h^e} (\mathbf{n}\cdot\mu\nabla w, v)_{L^2(E)} \\ &\quad + (\mathbf{a}\cdot\nabla w, v)_{L^2(\Omega)} + (cw, v)_{L^2(\Omega)} \end{aligned} \tag{2.182}$$

for all $w \in V$ and $v \in H^1(\Omega)$. Therefore we let

$$\langle \widehat{A}(t)w, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} := (\mu\nabla w, \nabla v)_{L^2(\Omega)} + (\mathbf{a}\cdot\nabla w, v)_{L^2(\Omega)} + (cw, v)_{L^2(\Omega)}. \tag{2.183}$$

This definition implies

$$\begin{aligned} \langle \widehat{A}(t)w, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} &\le \max_{0\le i,j\le d} \|\mu_{ij}\|_{L^\infty(\Omega)} |w|_{H^1(\Omega)} |v|_{H^1(\Omega)} \\ &\quad + \max_{0\le i\le d} \|a_i\|_{L^\infty(\Omega)} |w|_{H^1(\Omega)} \|v\|_{L^2(\Omega)} \\ &\quad + \|c\|_{L^\infty(\Omega)} \|w\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\le C\|w\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \end{aligned} \tag{2.184}$$

for all $w \in V$ and $v \in H^1(\Omega)$ and for a.e. $t \in (0,T)$. Further we have

$$\begin{aligned} \langle \widehat{A}(t)v, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} &\ge \mu_0 |v|^2_{H^1(\Omega)} + \frac{1}{2}\int_\Gamma \mathbf{a}\cdot\mathbf{n}v^2 dx + \int_\Omega \left(c - \frac{1}{2}\nabla\cdot\mathbf{a}\right) v^2 dx \\ &\ge \mu_0 |v_h|^2_{H^1(\Omega)} + p\|v_h\|^2_{L^2(\Omega)} + \frac{1}{2}\int_{\Gamma_{\text{in}}} \mathbf{a}\cdot\mathbf{n}v^2 dx \end{aligned} \tag{2.185}$$

for all $v \in V$. Here the assumption $p > 0$ leads to

$$\langle \widehat{A}(t)v, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} \ge \min\{\mu_0, p\}\|v\|^2_{H^1(\Omega)} + \frac{1}{2}(\mathbf{a}\cdot\mathbf{n}v, v)_{L^2(\Gamma_{\text{in}})} \text{ for all } v \in H^1_0(\Omega). \tag{2.186}$$

and the proof is complete. □

Remark 12. If we consider the advection and reaction term are both 0, then we have

$$\left\{\begin{array}{l} \langle \widehat{A}(t)w, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} \le \widehat{M}|w|_{H^1(\Omega)}|v|_{H^1(\Omega)} \text{ for all } w \in V, v \in H^1(\Omega), \\ \langle \widehat{A}(t)v, v\rangle_{(H^1(\Omega))^*, H^1(\Omega)} \ge \widehat{\alpha}|v|^2_{H^1(\Omega)} \text{ for all } v \in V \end{array}\right. \tag{2.187}$$

for a.e. $t \in (0,T)$. Therefore, we have to replace the definition of $\|\cdot\|_V$ with

$$\|w\|_V^2 = |w|_{H^1(\Omega)}^2 + \sum_{K\in\mathcal{T}_h} h_K^2 |w|_{H^2(K)}^2 + \sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|w\|_{L^2(E)}^2 \text{ for all } w \in V, \tag{2.188}$$

and $(\cdot,\cdot)_{V_h}$, $\|\cdot\|_{V_h}$ with

$$\begin{cases} (w,v)_{V_h} & := (\nabla w, \nabla v)_{L^2(\Omega)} + \sum_{E\in\mathcal{E}_h^e} h_E^{-1}(w,v)_{L^2(E)} \text{ for all } w,v \in H^1(\Omega), \\ \|v\|_{V_h}^2 & := |v|_{H^1(\Omega)}^2 + \sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|v\|_{L^2(E)}^2 \text{ for all } v \in H^1(\Omega). \end{cases} \tag{2.189}$$

In this chapter, we only consider the case when the advection vector $\mathbf{a}(\mathbf{x},t)$ is not zero.

Remark 13. Hereafter, we identify $L^2(\Omega)$ and $\left(L^2(\Omega)\right)^*$ by the Riesz representation theorem, and write

$$A(t)w \in L^2(\Omega) \text{ and } (A(t)w, v)_{L^2(\Omega)} \tag{2.190}$$

for $w \in V$ and $v \in L^2(\Omega)$.

2.5.4 The formulation of Nitsche method in parabolic problem

Let

$$X_V := \left\{ v \in W^{1,2,2}\left(0,T; H^1(\Omega), L^2(\Omega)\right) : v(t) \in V \text{ for a.e. } t \in (0,T) \right\}, \tag{2.191}$$

and

$$X_h := H^1\left(0,T;V_h\right),\ Y_h := L^2(0,T;V_h) \times V_h. \tag{2.192}$$

Further we define the norms

$$\|x\|_{X_V}^2 := \int_0^T \left(\|x\|_V^2 + \|T_1 x'\|_{V_h^*}^2 \right) dt + \|x(0)\|_{L^2(\Omega)}^2 \text{ for } x \in X_V, \tag{2.193}$$

$$\|x\|_{X_h}^2 := \int_0^T \left(\|x\|_{V_h}^2 + \|T_1 x'\|_{V_h^*}^2 \right) dt + \|x(0)\|_{L^2(\Omega)}^2 \text{ for } x \in W^{1,2,2}\left((0,T); H^1(\Omega), L^2(\Omega)\right), \tag{2.194}$$

$$\|\mathbf{y}_h\|_{Y_h}^2 = \|(y_h, \widetilde{y}_h)\|_{Y_h}^2 := \int_0^T \|y_h\|_{V_h}^2 dt + \|\widetilde{y}_h\|_{L^2(\Omega)}^2 \text{ for } \mathbf{y}_h \in Y_h. \tag{2.195}$$

It is easy to check that X_h and Y_h are Banach spaces with norms $\|\cdot\|_{X_h}$ and $\|\cdot\|_{Y_h}$, respectively. Moreover, Y_h is a Hilbert space with the inner product

$$(\mathbf{y}_{1h}, \mathbf{y}_{2h})_{Y_h} := \int_0^T (y_{1h}, y_{2h})_{V_h} dt + (\widetilde{y}_{1h}, \widetilde{y}_{2h})_{L^2(\Omega)} \tag{2.196}$$

for all $\mathbf{y}_{1h} = (y_{1h}, \widetilde{y}_{1h}), \mathbf{y}_{2h} = (y_{2h}, \widetilde{y}_{2h}) \in Y_h$. Note that the Lemma 8 implies

$$X_V \subset W^{1,2,2}\left((0,T); H^1(\Omega), L^2(\Omega)\right) \subset C^0\left([0,T]; L^2(\Omega)\right), \tag{2.197}$$

and the Corollary 2 gives

$$X_h \subset C^0\left([0,T]; V_h\right) \subset C^0\left([0,T]; L^2(\Omega)\right). \tag{2.198}$$

Let $u_0 \in L^2(\Omega)$, $f \in L^2\left(0,T; L^2(\Omega)\right) \subset L^2\left(0,T; H^{-1}(\Omega)\right)$, and $g_D \in G_0$. Here the theorem 8 implies there exists $u \in X := W^{1,2,2}\left(0,T; H^1(\Omega), H^{-1}(\Omega)\right)$ which is a unique weak solution of the problem (2.43).

Now we consider the following problem, find $u_{\varepsilon,h} \in X_h \subset X_V$ such that

$$b_{\varepsilon,h}(u_{\varepsilon,h}, \mathbf{v}_h) = F(\mathbf{v}_h) \text{ for all } \mathbf{v}_h := (v_h, \widetilde{v}_h) \in Y_h, \tag{2.199}$$

where

$$b_{\varepsilon,h}(w, \mathbf{v}_h) := \int_0^T \left((w', v_h)_{L^2(\Omega)} + a_\varepsilon(t; w, v_h) \right) dt + (w(0), \widetilde{v}_h)_{L^2(\Omega)} \tag{2.200}$$

for all $w \in X_V$ and $\mathbf{v}_h \in Y_h$, $a_\varepsilon(t;\cdot,\cdot) : V \times V_h \to \mathbb{R}$ is defined by (see [5] or page 56 of [52])

$$\begin{aligned} a_\varepsilon(t; w, v_h) \quad &:= (A(t)w, v_h)_{L^2(\Omega)} - \sum_{E\in\mathcal{E}_h^e} (\mathbf{n}\cdot\mu\nabla v_h, w)_{L^2(E)} \\ &\qquad -(\mathbf{a}\cdot\mathbf{n}v_h, w)_{L^2(\Gamma_{\text{in}})} + \langle\varepsilon(t)v_h, w\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \\ &= \langle\widehat{A}(t)w, v_h\rangle_{(H^1(\Omega))^*,H^1(\Omega)} - \sum_{E\in\mathcal{E}_h^e} \big((\mathbf{n}\cdot\mu\nabla w, v_h)_{L^2(E)} + (\mathbf{n}\cdot\mu\nabla v_h, w)_{L^2(E)}\big) \\ &\qquad -(\mathbf{a}\cdot\mathbf{n}v_h, w)_{L^2(\Gamma_{\text{in}})} + \langle\varepsilon(t)v_h, w\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \end{aligned} \tag{2.201}$$

for all $w \in V$ and $v_h \in V_h$,

$$F(\mathbf{v}_h) = \int_0^T \mathcal{F}(t; v_h)dt + (u_0, \widetilde{v}_h)_{L^2(\Omega)} \tag{2.202}$$

for all $\mathbf{v}_h := (v_h, \widetilde{v}_h) \in Y_h$,

$$\mathcal{F}(t; v_h) := (f, v_h)_{L^2(\Omega)} - \sum_{E\in\mathcal{E}_h^e} (\mathbf{n}\cdot\mu\nabla v_h, g_D)_{L^2(E)} - (\mathbf{a}\cdot\mathbf{n}v_h, g_D)_{L^2(\Gamma_{\text{in}})} + \langle\varepsilon(t)v_h, g_D\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \tag{2.203}$$

for all $v_h \in V_h$, and $\varepsilon(t) : V_h \to H^{-1/2}(\Gamma)$ is a given linear operator. Note that $X_V \subset C^0\big([0,T]; L^2(\Omega)\big)$ implies that the inner product $(w(0), \widetilde{v}_h)_{L^2(\Omega)}$ is meaningful.

Using the distribution theory, we have that the problem (2.199) is equivalent to finding $u_{\varepsilon,h} \in X_h$ such that

$$(u'_{\varepsilon,h}, v_h)_{L^2(\Omega)} + a_\varepsilon(t; u_{\varepsilon,h}, v_h) = \mathcal{F}(t; v_h) \tag{2.204}$$

for all $v_h \in V_h$ and for a.e. $t \in (0,T)$.

We have $a_\varepsilon(t;\cdot,\cdot)$ is a bilinear form for a.e. $t \in (0,T)$, and therefore $b_{\varepsilon,h} : X_V \times Y_h \to \mathbb{R}$ is also a bilinear form.

Hereafter we always assume the following.

Assumption 8. Let the operator $\varepsilon(t) : V_h \to H^{-1/2}(\Gamma)$ be such that

$$\langle\varepsilon(t)v_h, w\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} := \varepsilon_0 \sum_{E\in\mathcal{E}_h^e} h_E^{-1}(w, v_h)_{L^2(E)} \tag{2.205}$$

for all $w \in V$ and $v_h \in V_h$, where $\varepsilon_0 > 2\widehat{\alpha}^{-1}C_I$.

Then we have the following estimate.

Lemma 27 (Continuity of F)**.** $F : Y_h \to \mathbb{R}$ is linear functional and

$$F(\mathbf{v}_h) \le \big(\mathcal{M}_h + \|u_0\|_{L^2(\Omega)}\big)\, \|\mathbf{v}_h\|_{Y_h}, \tag{2.206}$$

where

$$\mathcal{M}_h^2 := 3\int_0^T \left(\|T_1 f\|_{V_h^*}^2 + C_{\text{Tr}}^2\|\mathbf{a}\cdot\mathbf{n}\|_{L^\infty(\Gamma_{\text{in}})}^2\|g_D\|_{L^2(\Gamma_{\text{in}})}^2 + (C_I+\varepsilon_0^2)\sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|g_D\|_{L^2(E)}^2 \right) dt. \tag{2.207}$$

Proof. It is clear that $F : Y_h \to \mathbb{R}$ is linear.

$$\begin{aligned}
F(\mathbf{v}_h) &= \int_0^T \Bigg((f, v_h)_{L^2(\Omega)} - \sum_{E\in\mathcal{E}_h^e} (\mathbf{n}\cdot\mu\nabla v_h, g_D)_{L^2(E)} \\
&\qquad - (\mathbf{a}\cdot\mathbf{n} v_h, g_D)_{L^2(\Gamma_{\text{in}})} + \langle \varepsilon(t) v_h, g_D\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \Bigg) dt + (u_0, \widetilde{v}_h)_{L^2(\Omega)} \\
&\le \int_0^T \Bigg(\|T_1 f\|_{V_h^*} \|v_h\|_{V_h} + C_I^{1/2} \left(\sum_{E\in\mathcal{E}_h^e} h_E^{-1} \|g_D\|_{L^2(E)}^2 \right)^{1/2} |v_h|_{H^1(\Omega)} \\
&\qquad + C_{\text{Tr}} \|\mathbf{a}\cdot\mathbf{n}\|_{L^\infty(\Gamma_{\text{in}})} \|g_D\|_{L^2(\Gamma_{\text{in}})} \|v_h\|_{H^1(\Omega)} \\
&\qquad + \varepsilon_0 \left(\sum_{E\in\mathcal{E}_h^e} h_E^{-1} \|v_h\|_{L^2(E)}^2 \right)^{1/2} \left(\sum_{E\in\mathcal{E}_h^e} h_E^{-1} \|g_D\|_{L^2(E)}^2 \right)^{1/2} \Bigg) dt \\
&\qquad + \|u_0\|_{L^2(\Omega)} \|\widetilde{v}_h\|_{L^2(\Omega)} \\
&\le \left(\int_0^T \left(\|v_h\|_{V_h}^2 + |v_h|_{H^1(\Omega)}^2 + \|v_h\|_{H^1(\Omega)}^2 + \sum_{E\in\mathcal{E}_h^e} h_E^{-1} \|v_h\|_{L^2(E)}^2 \right) dt \right)^{1/2} \mathcal{M}_h/\sqrt{3} \\
&\qquad + \|u_0\|_{L^2(\Omega)} \|\widetilde{v}_h\|_{L^2(\Omega)} \\
&\le (\mathcal{M}_h + \|u_0\|_{L^2(\Omega)}) \|\mathbf{v}_h\|_{Y_h}
\end{aligned} \tag{2.208}$$

for all $\mathbf{v}_h \in Y_h$. □

2.5.5 The continuity and coercivity of the bilinear form $a_\varepsilon(t;\cdot,\cdot)$

Here we show the bilinear form $a_\varepsilon(t;\cdot,\cdot) : V \times V_h \to \mathbb{R}$ is continuous and coercive for a.e. $t \in (0,T)$.

Lemma 28 (Continuity of $a_\varepsilon(t;\cdot,\cdot)$)**.** There exists a positive constant C such that

$$a_\varepsilon(t; w, v_h) \le C\|w\|_V \|v_h\|_{V_h} \tag{2.209}$$

for all $w \in V$, $v_h \in V_h$, for a.e. $t \in (0,T)$. Moreover, there exists a positive constant M such that

$$a_\varepsilon(t; w_h, v_h) \le M\|w_h\|_{V_h} \|v_h\|_{V_h} \tag{2.210}$$

for all $w_h, v_h \in V_h$, for a.e. $t \in (0,T)$.

Proof. The Cauchy-Schwarz inequality, Corollary 4, Lemma 25 and Lemma 26 show that

$$\begin{aligned}
a_\varepsilon(t; w, v_h) &\le \widehat{M} \|w\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)} + C\|w\|_{L^2(\Gamma)} \|v_h\|_{L^2(\Gamma)} \\
&\quad + \sum_{E\in\mathcal{E}_h^e} \Big(\|\mathbf{n}\cdot\mu\nabla w\|_{L^2(E)} \|v_h\|_{L^2(E)} + \|w\|_{L^2(E)} \|\mathbf{n}\cdot\mu\nabla v_h\|_{L^2(E)} \Big) \\
&\qquad + \langle \varepsilon(t) v_h, w\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \\
&\le \widehat{M} \|w\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)} + C\|w\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)} \\
&\quad + C \left(|w|_{H^1(\Omega)}^2 + \sum_{K\in\mathcal{T}_h} h_K^2 |w|_{H^2(K)}^2 \right)^{1/2} \left(\sum_{E\in\mathcal{E}_h^e} h_E^{-1} \|v_h\|_{L^2(E)}^2 \right)^{1/2} \\
&\quad + C \left(\sum_{E\in\mathcal{E}_h^e} h_E^{-1} \|w\|_{L^2(E)}^2 \right)^{1/2} |v_h|_{H^1(\Omega)} + \langle \varepsilon(t) v_h, w\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \\
&\le C\|w\|_V \|v_h\|_{V_h}
\end{aligned} \tag{2.211}$$

for all $w \in V$ and $v_h \in V_h$, for a.e. $t \in (0,T)$. Further, the Lemma 24 leads to

$$a_\varepsilon(t; w_h, v_h) \le C\|w_h\|_{V_h} \|v_h\|_{V_h} \le M\|w_h\|_{V_h} \|v_h\|_{V_h}, \tag{2.212}$$

for all $w_h, v_h \in V_h$, for a.e. $t \in (0,T)$. □

Lemma 29 (Coercivity of $a_\varepsilon(t;\cdot,\cdot)$)**.** There exists a positive constant α such that

$$a_\varepsilon(t; v_h, v_h) \geq \alpha \|v_h\|_{V_h}^2 \tag{2.213}$$

for all $v_h \in V_h$.

Proof. Since $-(\mathbf{a}\cdot\mathbf{n}v_h, v_h)_{L^2(\Gamma_{\text{in}})} \geq 0$ for all $v_h \in V_h$, the Lemma 25 and Lemma 26 give

$$\begin{aligned}
a_\varepsilon(t; v_h, v_h) &\geq \widehat{\alpha}\|v_h\|_{H^1(\Omega)}^2 - \frac{1}{2}(\mathbf{a}\cdot\mathbf{n}v_h, v_h)_{L^2(\Gamma_{\text{in}})} \\
&\quad -2\sum_{E\in\mathcal{E}_h^e}(\mathbf{n}\cdot\mu\nabla v_h, v_h)_{L^2(E)} + \langle\varepsilon(t)v_h, v_h\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \\
&\geq \widehat{\alpha}\|v_h\|_{H^1(\Omega)}^2 - 2C_I^{1/2}|v_h|_{H^1(\Omega)}\left(\sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|v_h\|_{L^2(E)}^2\right)^{1/2} \\
&\quad +\langle\varepsilon(t)v_h, v_h\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)} \\
&\geq \widehat{\alpha}\|v_h\|_{H^1(\Omega)}^2 - \left(\widehat{\alpha}/2|v_h|_{H^1(\Omega)}^2 + 2\widehat{\alpha}^{-1}C_I\sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|v_h\|_{L^2(E)}^2\right) \\
&\quad +\langle\varepsilon(t)v_h, v_h\rangle_{H^{-1/2}(\Gamma),H^{1/2}(\Gamma)}
\end{aligned} \tag{2.214}$$

The condition for $\varepsilon(t)$ leads to

$$a_\varepsilon(t; v_h, v_h) \geq \frac{\widehat{\alpha}}{2}\|v_h\|_{H^1(\Omega)}^2 + (\varepsilon_0 - 2\widehat{\alpha}^{-1}C_I)\sum_{E\in\mathcal{E}_h^e} h_E^{-1}\|v_h\|_{L^2(E)}^2 = \alpha\|v_h\|_{V_h}^2 \tag{2.215}$$

for all $v_h \in V_h$, for a.e. $t \in (0, T)$. □

Thanks to the above estimates, we have that the restriction of the bilinear form $a_\varepsilon(t;\cdot,\cdot) : V_h \times V_h \to \mathbb{R}$ is continuous and coercive, for a.e. $t \in (0, T)$. Then the Lax-Milgram theorem implies that for any $\phi(t) \in V_h^*$, there exists a unique solution $w_h \in V_h$ to the equation

$$a_\varepsilon(t; w_h, v_h) = \langle\phi, v_h\rangle_{V_h^*,V_h} \text{ for all } v_h \in V_h \text{ and for a.e. } t \in (0, T). \tag{2.216}$$

Here we apply the Riesz representation theorem for $V_h \subset L^2(\Omega)$, then there exists a $\psi(t) \in V_h$ such that

$$\langle\phi(t), v_h\rangle_{V_h^*,V_h} = (\psi(t), v_h)_{L^2(\Omega)} \text{ for all } \phi(t) \in V_h^*, v_h \in V_h \text{ and for a.e. } t \in (0, T). \tag{2.217}$$

Therefore, we define a linear operator $\mathcal{A}_\varepsilon(t) : V_h \to V_h$ by the following.

Definition 12.

$$(\mathcal{A}_\varepsilon(t)w_h, v_h)_{L^2(\Omega)} = a_\varepsilon(t; w_h, v_h), \text{ for all } w_h, v_h \in V_h, \text{ for a.e. } t \in (0, T). \tag{2.218}$$

The Lemma 28 and Lemma 29 imply

$$\begin{cases} (\mathcal{A}_\varepsilon(t)w_h, v_h)_{L^2(\Omega)} \leq M\|w_h\|_{V_h}\|v_h\|_{V_h} & \text{for all } w_h, v_h \in V_h, \\ (\mathcal{A}_\varepsilon(t)v_h, v_h)_{L^2(\Omega)} \geq \alpha\|v_h\|_{V_h}^2 & \text{for all } v_h \in V_h, \end{cases} \tag{2.219}$$

for a.e. $t \in (0, T)$. Again, thanks to the Lax-Milgram theorem, we have the operator $\mathcal{A}_\varepsilon(t) : V_h \to V_h$ is invertible. Further, we have the following result.

Lemma 30. The bijection $\mathcal{A}_\varepsilon(t) : V_h \to V_h$ satisfies

$$\begin{cases} \left\|\mathcal{A}_\varepsilon(t)^{-1}w_h\right\|_{V_h} \leq \alpha^{-1}\|T_1 w_h\|_{V_h^*}, \\ \left(w_h, \mathcal{A}_\varepsilon(t)^{-1}w_h\right)_{L^2(\Omega)} \geq \alpha M^{-2}\|T_1 w_h\|_{V_h^*}^2 \end{cases} \tag{2.220}$$

for all $w_h \in V_h$ and for a.e. $t \in (0, T)$.

Proof. First, we have

$$\begin{aligned} \alpha \left\|\mathcal{A}_\varepsilon(t)^{-1}\phi\right\|_{V_h}^2 &\le \left(\mathcal{A}_\varepsilon(t)\left(\mathcal{A}_\varepsilon(t)^{-1}w_h\right), \mathcal{A}_\varepsilon(t)^{-1}w_h\right)_{L^2(\Omega)} \\ &= \left(w_h, \mathcal{A}_\varepsilon(t)^{-1}w_h\right)_{L^2(\Omega)} \le \|T_1 w_h\|_{V_h^*} \left\|\mathcal{A}_\varepsilon(t)^{-1}w_h\right\|_{V_h} \end{aligned} \tag{2.221}$$

for all $w_h \in V_h$. Therefore we have

$$\left\|\mathcal{A}_\varepsilon(t)^{-1}w_h\right\|_{V_h} \le \alpha^{-1}\|T_1 w_h\|_{V_h^*}, \text{ for all } w_h \in V_h \tag{2.222}$$

for a.e. $t \in (0,T)$.

Next, we note that the constant M satisfies

$$\|T_1\mathcal{A}_\varepsilon(t)\|_{\mathcal{L}(V_h,V_h^*)} := \sup_{0\neq w_h\in V_h} \frac{\|T_1\mathcal{A}_\varepsilon(t)w_h\|_{V_h^*}}{\|w_h\|_{V_h}} = \sup_{0\neq w_h\in V_h}\sup_{0\neq v_h\in V_h} \frac{(\mathcal{A}_\varepsilon(t)w_h, v_h)_{L^2(\Omega)}}{\|w_h\|_{V_h}\|v_h\|_{V_h}} \le M. \tag{2.223}$$

Therefore,

$$\begin{aligned} \|T_1 w_h\|_{V_h^*} &= \left\|T_1\mathcal{A}_\varepsilon(t)\left(\mathcal{A}_\varepsilon(t)^{-1}w_h\right)\right\|_{V_h^*} \\ &\le \|T_1\mathcal{A}_\varepsilon(t)\|_{\mathcal{L}(V_h,V_h^*)}\left\|\mathcal{A}_\varepsilon(t)^{-1}w_h\right\|_{V_h} \le M\left\|\mathcal{A}_\varepsilon(t)^{-1}w_h\right\|_{V_h} \end{aligned} \tag{2.224}$$

for all $w_h \in V_h$. Moreover, this implies

$$\begin{aligned} \left(w_h, \mathcal{A}_\varepsilon(t)^{-1}w_h\right)_{L^2(\Omega)} &= \left(\mathcal{A}_\varepsilon(t)\left(\mathcal{A}_\varepsilon(t)^{-1}w_h\right), \mathcal{A}_\varepsilon(t)^{-1}w_h\right)_{L^2(\Omega)} \\ &\ge \alpha\left\|\mathcal{A}_\varepsilon(t)^{-1}w_h\right\|_{V_h}^2 \ge \alpha M^{-2}\|T_1 w_h\|_{V_h^*}^2 \end{aligned} \tag{2.225}$$

for all $w_h \in V_h$ and for a.e. $t \in (0,T)$. □

2.5.6 The continuity and the inf-sup condition of the bilinear form $b_{\varepsilon,h}$

The continuity and coercivity of $a_\varepsilon(t;\cdot,\cdot)$ imply that the bilinear form $b_{\varepsilon,h}$ is continuous and satisfies the inf-sup condition.

Lemma 31 (Continuity of $b_{\varepsilon,h}$)**.** There exists a positive constant C such that

$$b_{\varepsilon,h}(w, \mathbf{v}_h) \le C\|w\|_{X_V}\|\mathbf{v}_h\|_{Y_h} \text{ for all } w \in X_V, \mathbf{v}_h \in Y_h. \tag{2.226}$$

Especially, we have

$$b_{\varepsilon,h}(w_h, \mathbf{v}_h) \le C\|w_h\|_{X_h}\|\mathbf{v}_h\|_{Y_h} \text{ for all } w_h \in X_h, \mathbf{v}_h \in Y_h. \tag{2.227}$$

Proof. The Lemma 28 leads to

$$\begin{aligned} b_{\varepsilon,h}(w, \mathbf{v}_h) &:= \int_0^T \left((w', v_h)_{L^2(\Omega)} + a_\varepsilon(t; w, v_h)\right)\, dt + (w(0), \widetilde{v}_h)_{L^2(\Omega)} \\ &\le \int_0^T \left(\|T_1 w'\|_{V_h^*}\|v_h\|_{V_h} + C\|w\|_V\|v_h\|_{V_h}\right)\, dt + \|w(0)\|_{L^2(\Omega)}\|\widetilde{v}_h\|_{L^2(\Omega)} \\ &\le C\|w\|_{X_V}\|\mathbf{v}_h\|_{Y_h} \end{aligned} \tag{2.228}$$

for all $w \in X_V$ and $\mathbf{v}_h = (v_h, \widetilde{v}_h) \in Y_h$. Moreover, the Lemma 24 leads the equation (2.227). □

Theorem 14 (Inf-sup condition of $b_{\varepsilon,h}$)**.** There exists a positive constant β such that

$$\inf_{0\neq x_h\in X_h}\sup_{0\neq \mathbf{y}_h\in Y_h} \frac{b_{\varepsilon,h}(x_h, \mathbf{y}_h)}{\|x_h\|_{X_h}\|\mathbf{y}_h\|_{Y_h}} \ge \beta. \tag{2.229}$$

Proof. We let $x_h \in X_h$ and $\mathbf{y}_h := \left(\mathcal{A}_\varepsilon(t)^{-1}x_h' + \delta x_h, \delta x_h(0)\right) \in Y_h$, where $\delta := \alpha^{-4}M^4$. Then we have

$$\begin{aligned} \|\mathbf{y}_h\|_{Y_h}^2 &= \int_0^T \left\|\mathcal{A}_\varepsilon(t)^{-1}x_h' + \delta x_h\right\|_{V_h}^2\, dt + \|\delta x_h(0)\|_{L^2(\Omega)}^2 \\ &\le C\int_0^T \left(\|T_1 x_h'\|_{V_h^*}^2 + \|x_h\|_{V_h}^2\right)\, dt + \delta\|x_h(0)\|_{L^2(\Omega)}^2 \le C\|x_h\|_{X_h}^2, \end{aligned} \tag{2.230}$$

and

$$
\begin{aligned}
b_{\varepsilon,h}(x_h,\mathbf{y}_h) &= \int_0^T \left(\left(x_h', \mathcal{A}_\varepsilon(t)^{-1}x_h' + \delta x_h\right)_{L^2(\Omega)} + a_\varepsilon\left(t; x_h, \mathcal{A}_\varepsilon(t)^{-1}x_h' + \delta x_h\right)\right) dt \\
&\qquad\qquad\qquad\qquad\qquad\qquad + (x_h(0), \delta x_h(0))_{L^2(\Omega)} \\
&\ge \int_0^T \left(\alpha M^{-2} \|T_1 x_h'\|_{V_h^*}^2 - \alpha^{-1} M \|x_h\|_{V_h} \|T_1 x_h'\|_{V_h^*} + \delta\alpha \|x_h\|_{V_h}^2 \right) dt \\
&\qquad + \delta/2 \left(\|x_h(T)\|_{L^2(\Omega)}^2 - \|x_h(0)\|_{L^2(\Omega)}^2 \right) + \delta \|x_h(0)\|_{L^2(\Omega)}^2 \\
&\ge \int_0^T \left(\alpha M^{-2} \|T_1 x_h'\|_{V_h^*}^2 - \alpha^{-1} M \|x_h\|_{V_h} \|T_1 x_h'\|_{V_h^*} + \delta\alpha \|x_h\|_{V_h}^2 \right) dt \\
&\qquad\qquad\qquad\qquad\qquad\qquad + \delta/2 \|x_h(0)\|_{L^2(\Omega)}^2 \\
&\ge \int_0^T \left(\alpha M^{-2}/2 \|T_1 x_h'\|_{V_h^*}^2 + \left(\delta\alpha - \alpha^{-3}M^4/2\right) \|x_h\|_{V_h}^2 \right) dt + \delta/2 \|x_h(0)\|_{L^2(\Omega)}^2 \\
&\ge C \left(\int_0^T \left(\|T_1 x_h'\|_{V_h^*}^2 + \|x_h\|_{V_h}^2 \right) dt + \|x_h(0)\|_{L^2(\Omega)}^2 \right) = C \|x_h\|_{X_h}^2,
\end{aligned}
\tag{2.231}
$$

because we took $\delta := \alpha^{-4}M^4$. The above two inequalities implies

$$
b_{\varepsilon,h}(x_h,\mathbf{y}_h) \ge C\|x_h\|_{X_h}^2 \ge \beta \|x_h\|_{X_h} \|\mathbf{y}_h\|_{Y_h}, \tag{2.232}
$$

for any $x_h \in X_h$ and $\mathbf{y}_h := \left(\mathcal{A}_\varepsilon(t)^{-1}x_h' + \alpha^{-4}M^4 x_h, \alpha^{-4}M^4 x_h(0)\right)$. Therefore, we have

$$
\inf_{0\ne x_h \in X_h} \sup_{0 \ne \mathbf{y}_h \in Y_h} \frac{b_{\varepsilon,h}(x_h,\mathbf{y}_h)}{\|x_h\|_{X_h} \|\mathbf{y}_h\|_{Y_h}} \ge \beta, \tag{2.233}
$$

then the inf-sup condition follows. □

Furthermore, the inf-sup condition leads the unique existence of approximate solution.

Theorem 15. The problem (2.199) has a unique solution $u_{\varepsilon,h} \in X_h$.

Proof. According to the Lemma 7, we have the conclusion if

$$
b_{\varepsilon,h}(x_h,\mathbf{y}_h) = 0 \text{ for all } x_h \in X_h \Rightarrow \mathbf{y}_h = 0 \text{ in } Y_h. \tag{2.234}
$$

Let $\mathbf{y}_{1h} = (y_{1h}, \widetilde{y}_{1h}) \in Y_h$ be such that $b(x_h, \mathbf{y}_{1h}) = 0$, for all $x_h \in X_h$. First we set $x_h \in X_h$ be such that

$$
x_h(0) = \widetilde{y}_{1h}, \text{ and } x_h(t) = 0 \text{ for } t \ge \delta \tag{2.235}
$$

for all $\delta > 0$. Then $b(x_h, \mathbf{y}_{1h}) = 0$ implies

$$
\int_0^\delta \left((x_h', y_{1h})_{L^2(\Omega)} + a_\varepsilon(t; x_h, y_{1h}) \right) dt + \|\widetilde{y}_{1h}\|_{L^2(\Omega)} = 0 \tag{2.236}
$$

for all $\delta > 0$. Therefore we have $\widetilde{y}_{1h} = 0$.

Next we use the basis functions $\{\phi_i\}_{i=1}^N$ of finite dimensional subspace $V_h \subset H^1(\Omega)$, where $N := \dim V_h$, and we let

$$
y_{1h}(t) := \sum_{i=1}^N a_i(t)\phi_i. \tag{2.237}
$$

Here we set $x = \phi_i$ for $i = 1, \ldots, N$, then $b\left(x_h, (y_{1h}, 0)\right) = 0$ implies

$$
\mathbf{A}_\varepsilon(t)\mathbf{a}(t) = 0, \tag{2.238}
$$

where $\left(\mathbf{A}_\varepsilon(t)\right)_{i,j} := a_\varepsilon(t; \phi_i, \phi_j)$ and $\mathbf{a}(t) := (a_1(t), \ldots, a_N(t))^T$. Thanks to the Lemma 29, we have

$$
\begin{aligned}
\boldsymbol{\xi}^T \mathbf{A}_\varepsilon(t) \boldsymbol{\xi} &= \sum_{i,j=1}^d \xi_i \xi_j a_\varepsilon(t; \phi_i, \phi_j) \\
&= a_\varepsilon\left(t; \sum_{i=1}^d \xi_i\phi_i, \sum_{i=1}^d \xi_i\phi_i\right) \ge \alpha \left\| \sum_{i=1}^d \xi_i\phi_i \right\|_{V_h}^2 > 0
\end{aligned}
\tag{2.239}
$$

for all $\boldsymbol{\xi} = (\xi_i)_{1\le i\le d} \in \mathbb{R}^d$ for a.e. $t \in (0,T)$. Therefore $\mathbf{A}_\varepsilon(t)$ is a positive definite matrix for a.e. $t \in (0,T)$. This implies $\mathbf{a}(t) = \mathbf{0}$, for a.e. $t \in (0,T)$, that is, $y_{1h} = 0$. □

2.5.7 Galerkin orthogonality and error estimate

Henceforth, we always assume the following.

Assumption 9. We assume that unique weak solution of the problem (2.43) satisfies $u \in X_V$.

Then we have $A(t)u \in \left(L^2(\Omega)\right)^* \simeq L^2(\Omega)$, and therefore

$$u' + A(t)u = f \text{ in } L^2(\Omega) \text{ for a.e. } t \in (0,T). \tag{2.240}$$

Further, $u \in X_V \subset C^0\left([0,T]; H^1(\Omega)\right)$ implies $g_D(t) = \operatorname{Tr} u \in H^{1/2}(\Gamma)$ for all $t \in [0,T]$.
Now we have the Galerkin orthogonality as the following.

Lemma 32 (Galerkin orthogonality)**.** Let $u_{\varepsilon,h} \in X_h$ be a unique solution of the equation (2.199), then we have

$$b_{\varepsilon,h}(u - u_{\varepsilon,h}, \mathbf{v}_h) = 0 \text{ for all } \mathbf{v}_h \in Y_h. \tag{2.241}$$

Proof. First, we have

$$\begin{aligned} b_{\varepsilon,h}(u, \mathbf{v}_h) &= \int_0^T (u' + A(t)u, v_h)_{L^2(\Omega)} - \sum_{E \in \mathcal{E}_h^e} (\mathbf{n} \cdot \mu \nabla v_h, u)_{L^2(E)} \\ &\quad -(\mathbf{a} \cdot \mathbf{n} v_h, u)_{L^2(\Gamma_{\mathrm{in}})} + \langle \varepsilon(t) v_h, u \rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} dt + (u(0), \widetilde{v}_h)_{L^2(\Omega)} \end{aligned} \tag{2.242}$$

for all $\mathbf{v}_h \in Y_h$. Therefore we have

$$\begin{aligned} b_{\varepsilon,h}(u - u_{\varepsilon,h}, \mathbf{v}_h) &= \int_0^T \Bigg((u' + A(t)u - f, v_h)_{L^2(\Omega)} - \sum_{E \in \mathcal{E}_h^e} (\mathbf{n} \cdot \mu \nabla v_h, u - g_D)_{L^2(E)} \\ &\qquad -(\mathbf{a} \cdot \mathbf{n} v_h, u - g_D)_{L^2(\Gamma_{\mathrm{in}})} + \langle \varepsilon(t) v_h, u - g_D \rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} \Bigg) dt \\ &\qquad + (u(0) - u_0, \widetilde{v}_h)_{L^2(\Omega)} \\ &= 0 \end{aligned} \tag{2.243}$$

for all $\mathbf{v}_h \in Y_h$. □

The Galerkin orthogonality and the inf-sup condition conclude the following result.

Lemma 33 (quasi-optimal error estimate)**.** There exists a positive constant C such that

$$\|u - u_{\varepsilon,h}\|_{X_h} \le C \|u - w_h\|_{X_V} \tag{2.244}$$

for all $w_h \in X_h$.

Proof. We have

$$\|u - u_{\varepsilon,h}\|_{X_h} \le \|u - w_h\|_{X_h} + \|w_h - u_{\varepsilon,h}\|_{X_h} \le \|u - w_h\|_{X_V} + \|w_h - u_{\varepsilon,h}\|_{X_h} \tag{2.245}$$

for all $w_h \in V_h$. Moreover,

$$\begin{aligned} \|w_h - u_{\varepsilon,h}\|_{X_h} &\le \frac{1}{\beta} \sup_{0 \neq \mathbf{v}_h \in Y_h} \frac{b(w_h - u_{\varepsilon,h}, \mathbf{v}_h)}{\|\mathbf{v}_h\|_{Y_h}} \\ &\le \frac{1}{\beta} \sup_{0 \neq \mathbf{v}_h \in Y_h} \frac{b(w_h - u, \mathbf{v}_h)}{\|\mathbf{v}_h\|_{Y_h}} \\ &\le C \|w_h - u\|_{X_V} \end{aligned} \tag{2.246}$$

for all $w_h \in X_h \subset X_V$. Therefore we have the conclusion. □

Thanks to the assumption (A3), the following error estimate holds.

Theorem 16 (error estimate)**.** Assume that the family of mesh $\{\mathcal{T}_h\}_h$ is (globally) quasi-uniform. Further we assume that two integers ℓ, m satisfy $2 \le \ell, m \le k+1$, and the exact solution $u \in X_V$ satisfies $u \in X_{\ell,m} := W^{1,2,2}\left(0,T; H^\ell(\Omega), H^m(\Omega)\right)$. Then we have

$$\|u - u_{\varepsilon,h}\|_{X_h}^2 \le C\left(\int_0^T \left(h^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + h^{2m}\|u'\|_{H^m(\Omega)}^2\right)\, dt + h^{2j}\|u(0)\|_{H^k(\Omega)}^2\right), \tag{2.247}$$

where $j := \min\{\ell, m\}$ and $u_h \in X_h$ is a unique weak solution of problem (2.199).

Proof. First we have $X_{\ell,m} \subset X_V$, for all $\ell, m \ge 2$, therefore the Lemma 33 implies

$$\|u - u_h\|_{X_h} \le \inf_{w_h \in V_h} C\|u - w_h\|_{X_V} \le \|u - \Pi_h u\|_{X_V} \tag{2.248}$$

for all $w_h \in X_h$, especially we consider $w_h := \Pi_h u$. Here we note that $u(t) \in H^2(\Omega)$ and $u_h(t)|_K \in H^2(K)$ for all $K \in \mathcal{T}_h$, and therefore $u(t)|_E, u_h(t)|_E \in H^1(E)$, for all $E \in \mathcal{E}_h^e$ and for a.e. $t \in (0,T)$. Now we have

$$\begin{aligned}\|u - \Pi_h u\|_{V_h}^2 &= \|u - \Pi_h u\|_{H^1(\Omega)}^2 + \sum_{E \in \mathcal{E}_h^e} h_E^{-1}\|u - \Pi_h u\|_{L^2(E)}^2 \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + C\sum_{E \in \mathcal{E}_h^e} h_{K_E}^{-1}\left(h_{K_E}^{-1}\|u - \Pi_h u\|_{L^2(K_E)}^2 + h_{K_E}|u - \Pi_h u|_{H^1(K_E)}^2\right) \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + C\left(Ch^{-2}\|u - \Pi_h u\|_{L^2(\Omega)}^2 + |u - \Pi_h u|_{H^1(\Omega)}^2\right) \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2,\end{aligned} \tag{2.249}$$

and

$$\begin{aligned}\|u - \Pi_h u\|_V^2 &= \|u - \Pi_h u\|_{V_h}^2 + \sum_{E \in \mathcal{E}_h^e} h_E \left\|\mathbf{n} \cdot \mu\nabla(u - \Pi_h u)\right\|_{L^2(E)}^2 \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + \sum_{E \in \mathcal{E}_h^e} h_E|u - \Pi_h u|_{H^1(E)}^2 \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + \sum_{E \in \mathcal{E}_h^e} h_{K_E}\left(h_{K_E}^{-1}|u - \Pi_h u|_{H^1(K_E)}^2 + h_{K_E}|u - \Pi_h u|_{H^2(K_E)}^2\right) \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + |u - \Pi_h u|_{H^1(\Omega)}^2 + h^2|u - \Pi_h u|_{H^2(\Omega)}^2 \\ &\le Ch^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2.\end{aligned} \tag{2.250}$$

Next, we remark

$$(\Pi_h u)'(t) = \lim_{a \to 0} \frac{\Pi_h u(t+a) - \Pi_h u(t)}{a} = \Pi_h\left(\lim_{a \to 0} \frac{u(t+a) - u(t)}{a}\right) = \Pi_h u'(t). \tag{2.251}$$

This implies

$$\|T_1(u - \Pi_h u)'\|_{V_h^*} \le \|u' - \Pi_h u'\|_{L^2(\Omega)} \le Ch^m\|u'\|_{H^m(\Omega)}. \tag{2.252}$$

Finally, we have

$$X_{\ell,m} \subset H^1\left(0,T; H^k(\Omega)\right) \subset C^0\left([0,T]; H^j(\Omega)\right), \tag{2.253}$$

and therefore

$$\|u(0) - \Pi_h u(0)\|_{L^2(\Omega)} \le Ch^k\|u(0)\|_{H^j(\Omega)}. \tag{2.254}$$

The above estimates lead

$$\begin{aligned}\|u - \Pi_h u\|_{X_V}^2 &= \int_0^T \left(\|u - \Pi_h u\|_V^2 + \|T_1(u - \Pi_h u)'\|_{V_h^*}^2\right) dt + \|u(0) - \Pi_h u(0)\|_{L^2(\Omega)}^2 \\ &\le C\left(\int_0^T \left(h^{2(\ell-1)}\|u\|_{H^\ell(\Omega)}^2 + h^{2m}\|u'\|_{H^m(\Omega)}^2\right) dt + h^{2k}\|u(0)\|_{H^j(\Omega)}^2\right),\end{aligned} \tag{2.255}$$

which completes the proof. □

2.6 The full discrete problem

Let $N \in \mathbb{N}$ be the number of time steps, $\tau := T/N$ and $t_n := n\tau$. We now consider the temporal discretization with implicit Euler (backward Euler) method. We find $\{u^n_{\varepsilon,h,\tau}\}_{n=0}^N \in (V_h)^{N+1}$ such that

$$\begin{cases} \frac{1}{\tau}(u^n_{\varepsilon,h,\tau} - u^{n-1}_{\varepsilon,h,\tau}, v_h)_{L^2(\Omega)} + a_\varepsilon(t_n; u^n_{\varepsilon,h,\tau}, v_h) = \mathcal{F}^n(v_h) \text{ for all } v_h \in V_h, \\ (u^0_{\varepsilon,h,\tau} - u_0, \widetilde{v}_h)_{L^2(\Omega)} = 0 \text{ for all } \widetilde{v}_h \in V_h \end{cases} \tag{2.256}$$

for all $n = 1, \cdots, N$, where the functional $\mathcal{F}^n : V_h \to \mathbb{R}$ is defined by $\mathcal{F}^n(v_h) := \mathcal{F}(t_n; v_h)$ for all $v_h \in V_h$. It is clear that $u^0_{\varepsilon,h,\tau} = u_{\varepsilon,h}(0)$, where $u_{\varepsilon,h} \in X_h$ is a unique solution of the problem (2.199) or problem (2.204).

Lemma 34. The functinal $\mathcal{F}^n : V_h \to \mathbb{R}$ satisfies

$$\mathcal{F}^n(v_h) \le M^n_h \|v_h\|_{V_h} \text{ for all } v_h \in V_h, \tag{2.257}$$

where

$$(M^n_h)^2 := 3\left(\|T_1 f^n\|^2_{V_h^*} + C^2_{\mathrm{Tr}} \|\mathbf{a}^n \cdot \mathbf{n}\|^2_{L^\infty(\Gamma_{\mathrm{in}})} \|g^n_D\|^2_{L^2(\Gamma_{\mathrm{in}})} + (C_I + \varepsilon_0^2) \sum_{E \in \mathcal{E}^e_h} h_E^{-1} \|g^n_D\|^2_{L^2(E)} \right). \tag{2.258}$$

Proof.

$$\begin{aligned} \mathcal{F}^n(v_h) &\le \|T_1 f^n\|_{V_h^*} \|v_h\|_{V_h} + C_I^{1/2} |v_h|_{H^1(\Omega)} \left(\sum_{E \in \mathcal{E}^e_h} h_E^{-1} \|g^n_D\|^2_{L^2(E)} \right)^{1/2} \\ &\quad + \|\mathbf{a}^n \cdot \mathbf{n}\|_{L^\infty(\Gamma_{\mathrm{in}})} C_{\mathrm{Tr}} \|v_h\|_{H^1(\Omega)} \|g^n_D\|_{L^2(\Gamma_{\mathrm{in}})} \\ &\quad + \varepsilon_0 \left(\sum_{E \in \mathcal{E}^e_h} h_E^{-1} \|v_h\|^2_{L^2(E)} \right)^{1/2} \left(\sum_{E \in \mathcal{E}^e_h} h_E^{-1} \|g_D\|^2_{L^2(E)} \right)^{1/2} \\ &\le \left(\|v_h\|^2_{V_h} + |v_h|^2_{H^1(\Omega)} + \|v_h\|^2_{H^1(\Omega)} + \sum_{E \in \mathcal{E}^e_h} h_E^{-1} \|v_h\|^2_{L^2(E)} \right)^{1/2} M^n_h/\sqrt{3} \\ &\le M^n_h \|v_h\|_{V_h} \end{aligned} \tag{2.259}$$

□

Lemma 35. There exists a unique solution $\{u^n_{\varepsilon,h,\tau}\}_{n=0}^N \in (V_h)^{N+1}$ of the equation (2.256)

Proof. We let ϕ_i be a basis function of V_h, and $u^n_{\varepsilon,h,\tau} := \sum_{i=1}^{\dim V_h} u^n_{\varepsilon,h,\tau,i}\phi_i$. Here we set $v_h = \phi_i$ for $i = 1, \cdots, \dim V_h$. First, it is easy to show that there exists a unique $u^0_{\varepsilon,h,\tau} \in V_h$. Next, we have

$$(\mathbf{M} + \Delta t \mathbf{A}^n)\, \mathbf{u}^n_{\varepsilon,h,\tau} = \mathbf{M}\mathbf{u}^{n-1}_{\varepsilon,h,\tau} + \Delta t \mathbf{F}^n \tag{2.260}$$

for $n = 1, \cdots, N$, where $(\mathbf{M})_{ij} := (\phi_i, \phi_j)_{L^2(\Omega)}$, $(\mathbf{A}^n)_{ij} := a_\varepsilon(t_n; \phi_i, \phi_j)$, $(\mathbf{F}^n)_i := \mathcal{F}^n(\phi_i)$ and $\mathbf{u}^n_{\varepsilon,h,\tau} := (u^n_{\varepsilon,h,\tau,1}, \cdots, u_{\varepsilon,h,\tau,\dim V_h})^T$. We note that the bilinear form $a_\varepsilon(t_n; \cdot, \cdot)$ is coercive, and this gives $\mathbf{M}+\Delta t\mathbf{A}^n$ is a positive definite matrix. Therefore there exists a unique $\mathbf{u}^n_{\varepsilon,h,\tau}$ for $\mathbf{u}^{n-1}_{\varepsilon,h,\tau}$. □

Lemma 36 (stability)**.** Let $\{u^n_{\varepsilon,h,\tau}\}_{n=0}^N$ be a unique solution of the equation (2.256), then

$$\begin{cases} \|\{u^n_{\varepsilon,h,\tau}\}_{n=1}^N\|_{\ell^\infty([t_1,t_N];L^2(\Omega))} \le \|u_0\|_{L^2(\Omega)} + \sqrt{\frac{T}{2\alpha}} M_h, \\ \|\{u^n_{\varepsilon,h,\tau}\}_{n=1}^N\|_{\ell^2([t_1,t_N];V_h)} \le \frac{1}{\sqrt{\alpha}} \|u_0\|_{L^2(\Omega)} + \frac{\sqrt{T}}{\alpha} M_h, \end{cases} \tag{2.261}$$

where $M_h := \max_{1\le n\le N} M_h^n$, and

$$\begin{cases} \|\{\phi^n\}_{n=1}^N\|^2_{\ell^\infty([t_1,t_N];L^2(\Omega))} := \max_{n=1,\cdots,N} \|\phi^n\|_{L^2(\Omega)}, \\ \|\{\phi^n\}_{n=1}^N\|^2_{\ell^2([t_1,t_N];V_h)} := \tau \sum_{n=1}^N \|\phi^n\|^2_{V_h}. \end{cases} \tag{2.262}$$

Proof. We let $v_h = u^n_{\varepsilon,h,\tau}$ in the equation (2.256), then we have

$$\begin{aligned} M_h^n \|u^n_{\varepsilon,h,\tau}\|_{V_h} &\ge F^n(u^n_{\varepsilon,h,\tau}) \\ &= \frac{1}{\tau}(u^n_{\varepsilon,h,\tau} - u^{n-1}_{\varepsilon,h,\tau}, u^n_{\varepsilon,h,\tau})_{L^2(\Omega)} + a_\varepsilon(t_n; u^n_{\varepsilon,h,\tau}, u^n_{\varepsilon,h,\tau}) \\ &= \frac{1}{2\tau}\left(\|u^n_{\varepsilon,h,\tau} - u^{n-1}_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} + \|u^n_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} - \|u^{n-1}_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)}\right) \\ &\qquad + a_\varepsilon(t_n; u^n_{\varepsilon,h,\tau}, u^n_{\varepsilon,h,\tau}). \end{aligned} \tag{2.263}$$

The coercivity of a_ε implies

$$\begin{aligned} \|u^n_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} - \|u^{n-1}_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} + 2\tau\alpha\|u^n_{\varepsilon,h,\tau}\|^2_{V_h} &\le 2\tau M_h^n \|u^n_{\varepsilon,h,\tau}\|_{V_h} \\ &\le \tau\left(q(M_h^n)^2 + \frac{1}{q}\|u^n_{\varepsilon,h,\tau}\|^2_{V_h}\right) \end{aligned} \tag{2.264}$$

for any $q \in \mathbb{R}$. If $q = \dfrac{1}{2\alpha}$, then we have

$$\|u^n_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} - \|u^{n-1}_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} \le \frac{\tau}{2\alpha}(M_h^n)^2, \tag{2.265}$$

therefore we have

$$\|u^n_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} \le \|u^0_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} + \frac{n\tau}{2\alpha}(M_h)^2 \le \left(\|u_0\|_{L^2(\Omega)} + \sqrt{\frac{T}{2\alpha}} M_h\right)^2 \tag{2.266}$$

for all n $1, \cdots, N$. Further, if $q = \dfrac{1}{\alpha}$, then

$$\|u^n_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} - \|u^{n-1}_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} + \tau\alpha\|u^n_{\varepsilon,h,\tau}\|^2_{V_h} \le \frac{\tau}{\alpha}(M_h^n)^2. \tag{2.267}$$

Summing this from $n = 1$ to N gives

$$\tau\sum_{n=1}^N \|u^n_{\varepsilon,h,\tau}\|^2_{V_h} \le \frac{1}{\alpha}\|u^0_{\varepsilon,h,\tau}\|^2_{L^2(\Omega)} + \frac{\tau}{\alpha^2}\sum_{n=1}^N (M_h^n)^2 \le \frac{1}{\alpha}\|u_0\|^2_{L^2(\Omega)} + \frac{T}{\alpha^2}(M_h)^2. \tag{2.268}$$

□

Lemma 37. Assume that $u_{\varepsilon,h} \in C^2([0,T];V_h)$, then there exists $c_n \in (t_{n-1}, t_n)$ such that

$$\frac{1}{\tau}(e^n - e^{n-1}, v_h)_{L^2(\Omega)} + a_\varepsilon(t_n; e^n, v_h) = \frac{\tau}{2}(u''_{\varepsilon,h}(c_n), v_h)_{L^2(\Omega)} \text{ for all } v_h \in V_h, \tag{2.269}$$

where $e^n := u^n_{\varepsilon,h,\tau} - u_{\varepsilon,h}(t_n)$.

Proof. We note that the equation (2.204) and (2.256) yield

$$\begin{cases} \frac{1}{\tau}(u^n_{\varepsilon,h,\tau} - u^{n-1}_{\varepsilon,h,\tau}, v_h)_{L^2(\Omega)} + a_\varepsilon(t_n; u^n_{\varepsilon,h,\tau}, v_h) = F^n(v_h), \\ (u'_{\varepsilon,h}(t_n), v_h)_{L^2(\Omega)} + a_\varepsilon(t_n; u_{\varepsilon,h}, v_h) = \mathcal{F}(t_n; v_h) = F^n(v_h) \end{cases} \tag{2.270}$$

for all $v_h \in V_h$, and there exists $c \in (t-\tau, t)$ such that

$$u_{\varepsilon,h}(t-\tau) = u_{\varepsilon,h}(t) - \tau u'_{\varepsilon,h}(t) + \frac{\tau^2}{2}u''_{\varepsilon,h}(c). \tag{2.271}$$

Therefore we have that there exists $c_n \in (t_{n-1}, t_n)$ such that

$$\begin{aligned} \frac{1}{\tau}(e^n - e^{n-1}, v_h)_{L^2(\Omega)} + a_\varepsilon(t_n; e^n, v_h) &= \left(u'_{\varepsilon,h}(t_n) - \frac{1}{\tau}(u_{\varepsilon,h}(t_n) - u_{\varepsilon,h}(t_{n-1})), v_h \right)_{L^2(\Omega)} \\ &= \frac{\tau}{2}(u''_{\varepsilon,h}(c_n), v_h)_{L^2(\Omega)} \end{aligned} \tag{2.272}$$

for all $v_h \in V_h$, which is desired conclusion. □

Theorem 17 (Error estimate)**.** Assume that $u_{\varepsilon,h} \in C^2([0,T]; V_h)$, then

$$\begin{cases} \|\{u^n_{\varepsilon,h,\tau} - u_{\varepsilon,h}(t_n)\}_{n=1}^N\|_{\ell^\infty([t_1,t_N];L^2(\Omega))} \le \sqrt{\dfrac{T}{8\alpha}}\tau\|T_1 u''_{\varepsilon,h}\|_{L^\infty\left((t_1,t_N);V_h^*\right)}, \\ \|\{u^n_{\varepsilon,h,\tau} - u_{\varepsilon,h}(t_n)\}_{n=1}^N\|_{\ell^2([t_1,t_N];V_h)} \le \dfrac{\sqrt{T}}{2}\tau\|T_1 u''_{\varepsilon,h}\|_{L^\infty\left((t_1,t_N);V_h^*\right)}. \end{cases} \tag{2.273}$$

Proof. We let $e^n := u^n_{\varepsilon,h,\tau} - u_{\varepsilon,h}(t_n)$. We have $e^0 = 0$ and

$$\frac{1}{\tau}(e^n - e^{n-1}, e^n)_{L^2(\Omega)} + a_\varepsilon(t_n; e^n, e^n) \ge \frac{1}{2\tau}\left(\|e^n\|^2_{L^2(\Omega)} - \|e^{n-1}\|^2_{L^2(\Omega)}\right) + \alpha\|e^n\|^2_{V_h}. \tag{2.274}$$

On the other hand, the Lemma 37 implies

$$\frac{1}{\tau}(e^n - e^{n-1}, e^n)_{L^2(\Omega)} + a_\varepsilon(t_n; e^n, e^n) = \frac{\tau}{2}(u''_{\varepsilon,h}(c_n), e^n) \le \frac{\tau}{2}\|T_1 u''_{\varepsilon,h}(c_n)\|_{V_h^*}\|e^n\|_{V_h} \tag{2.275}$$

for some $c_n \in (t_{n-1}, t_n)$. Therefore we have

$$\begin{aligned} 0 &\ge \alpha\|e^n\|^2_{V_h} - \frac{\tau}{2}\|T_1 u''_{\varepsilon,h}(c_n)\|_{V_h^*}\|e^n\|_{V_h} + \frac{1}{2\tau}\left(\|e^n\|^2_{L^2(\Omega)} - \|e^{n-1}\|^2_{L^2(\Omega)}\right) \\ &\ge \alpha\|e^n\|^2_{V_h} - \left(q\|e^n\|^2_{V_h} + \frac{\tau^2}{16q}\|T_1 u''_{\varepsilon,h}(c_n)\|^2_{V_h^*}\right) + \frac{1}{2\tau}\left(\|e^n\|^2_{L^2(\Omega)} - \|e^{n-1}\|^2_{L^2(\Omega)}\right). \end{aligned} \tag{2.276}$$

If $q = \alpha$, then we have

$$\|e^n\|^2_{L^2(\Omega)} + \|e^{n-1}\|^2_{L^2(\Omega)} \le \frac{\tau^3}{8\alpha}\|T_1 u''_{\varepsilon,h}(c_n)\|^2_{V_h^*}. \tag{2.277}$$

Therefore

$$\|e^n\|^2_{L^2(\Omega)} \le \|e^0\|^2_{L^2(\Omega)} + \frac{n\tau^3}{8\alpha}\|T_1 u''_{\varepsilon,h}\|_{L^\infty\left((t_1,t_n);V_h^*\right)} \le \frac{T\tau^2}{8\alpha}\|T_1 u''_{\varepsilon,h}\|_{L^\infty\left((t_1,t_N);V_h^*\right)}. \tag{2.278}$$

for all $n = 1, \cdots, N$. Further if $q = \dfrac{\alpha}{2}$, this implies

$$\tau\|e^n\|^2_{V_h} + \frac{1}{\alpha}\left(\|e^n\|^2_{L^2(\Omega)} - \|e^{n-1}\|^2_{L^2(\Omega)}\right) \le \frac{\tau^3}{4}\|T_1 u''_{\varepsilon,h}(c_n)\|^2_{V_h^*}. \tag{2.279}$$

Summing this from $n = 1$ to N, then

$$\|\{e^n\}_{n=1}^N\|_{\ell^2([t_1,t_N];V_h)} \le \frac{\sqrt{T}}{2}\tau\|T_1 u''_{\varepsilon,h}\|_{L^\infty\left((t_1,t_N);V_h^*\right)}. \tag{2.280}$$

□

Let

$$\mathcal{S}_\tau := \left\{ x_h : [0,T] \to V_h : \begin{array}{c} \text{For all } n = 1, \cdots, N, \text{ there exists } v_h \in V_h \\ \text{such that } x_h|_{(t_{n-1},t_n]} = v_h, \text{ and } x_h(0) \in V_h \end{array} \right\} \tag{2.281}$$

and $v_{h,\tau}(t_n^+) := \lim_{t \to t_n + 0} v_{h,\tau}(t)$ for $v_{h,\tau} \in \mathcal{S}_\tau$. We note that $\mathcal{S}_\tau \subset L^2(0,T;V_h)$. Now we extend the solution of finite difference method to an element of $\mathcal{S}_\tau$. For the solution of implicit Euler scheme, we let

$$u_{\varepsilon,h,\tau}(t) := \begin{cases} u_0 & \text{if } t = 0, \\ u^{n+1}_{\varepsilon,h,\tau} & \text{if } t \in (t_n, t_{n+1}] \text{ for all } n = 0, \cdots, N-1. \end{cases} \tag{2.282}$$

Then $u_{\varepsilon,h,\tau} \in \mathcal{S}_\tau$, and this satisfies the following estimate.

Lemma 38. Let $\Pi_\tau u_{\varepsilon,h} \in \mathcal{S}_\tau$ be

$$\Pi_\tau u_{\varepsilon,h}(t) := \begin{cases} u_0 & \text{if } t = 0, \\ u_{\varepsilon,h}(t_{n+1}) & \text{if } t \in (t_n, t_{n+1}] \text{ for all } n = 0, \cdots, N-1, \end{cases} \tag{2.283}$$

where $u_{\varepsilon,h} \in X_h$ is a unique solution of the problem (2.199). Assume that $u_{\varepsilon,h} \in C^2([0,T]; V_h)$, then

$$\begin{cases} \max\limits_{n=0,\cdots,N-1} \|\rho(t_n^+)\|_{L^2(\Omega)} \le \sqrt{\dfrac{T}{8\alpha}}\tau \|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}, \\ \|\rho\|_{L^2(0,T;V_h)} \le \dfrac{\sqrt{T}}{2\alpha}\tau\|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}, \end{cases} \tag{2.284}$$

where $\rho := u_{\varepsilon,h,\tau} - \Pi_\tau u_{\varepsilon,h} \in \mathcal{S}_\tau$

Proof. First we note that $\rho(t_0) = 0$. Second, $u_{\varepsilon,h} \in C^2([0,T]; V_h)$ and the definitions yield

$$\begin{cases} u_{\varepsilon,h,\tau}^{n+1} = u_{\varepsilon,h,\tau}(t_n^+) = u_{\varepsilon,h,\tau}(t_{n+1}), \\ u_{\varepsilon,h}(t_{n+1}) = \Pi_\tau u_{\varepsilon,h}(t_n^+) = \Pi_\tau u_{\varepsilon,h}(t_{n+1}). \end{cases} \tag{2.285}$$

Here, as the implicit Euler scheme we have

$$\begin{aligned} 0 &= \left(u_{\varepsilon,h,\tau}^{n+1} - u_{\varepsilon,h,\tau}^{n}, v_{h,\tau}(t_n^+)\right)_{L^2(\Omega)} + \tau a_\varepsilon\left(t_{n+1}; u_{\varepsilon,h,\tau}^{n+1}, v_{h,\tau}(t_n^+)\right) - F^{n+1}\left(v_{h,\tau}(t_n^+)\right) \\ &= (u_{\varepsilon,h,\tau}(t_n^+) - u_{\varepsilon,h,\tau}(t_n), v_{h,\tau}(t_n^+))_{L^2(\Omega)} + \tau a_\varepsilon(t_{n+1}; u_{\varepsilon,h,\tau}(t_n^+), v_{h,\tau}(t_n^+)) \\ &\qquad -\tau\mathcal{F}(t_{n+1}; v_{h,\tau}(t_n^+)) \end{aligned} \tag{2.286}$$

for all $v_{h,\tau} \in \mathcal{S}_\tau$ and $n = 0, \cdots, N-1$. Now we consider the equation which $\Pi_\tau u_{\varepsilon,h} \in \mathcal{S}_\tau$ satisfies. Combining $u \in C^2([0,T]; V_h)$ and the equation (2.204) gives

$$(u'_{\varepsilon,h}, v_h)_{L^2(\Omega)} + a_\varepsilon(t; u_{\varepsilon,h}, v_h) = \mathcal{F}(t; v_h) \text{ for all } t \in [0,T], v_h \in V_h. \tag{2.287}$$

This gives

$$\begin{aligned} 0 &= \tau(u'_{\varepsilon,h}(t_{n+1}), v_{h,\tau}(t_n^+))_{L^2(\Omega)} + \tau a_\varepsilon(t_{n+1}; u_{\varepsilon,h}, v_{h,\tau}(t_n^+)) - \tau\mathcal{F}(t_{n+1}; v_{h,\tau}(t_n^+)) \\ &= \left(\Pi_\tau u_{\varepsilon,h}(t_n^+) - \Pi_\tau u_{\varepsilon,h}(t_n) + \frac{\tau^2}{2}u''_{\varepsilon,h}(c_n), v_{h,\tau}(t_n^+)\right)_{L^2(\Omega)} \\ &\qquad + \tau a_\varepsilon(t_{n+1}; \Pi_\tau u_{\varepsilon,h}(t_n^+), v_{h,\tau}(t_n^+)) - \tau\mathcal{F}(t_{n+1}; v_{h,\tau}(t_n^+)) \end{aligned} \tag{2.288}$$

for all $v_{h,\tau} \in \mathcal{S}_\tau$, $t \in (t_n, t_{n+1}]$ and for some $c_n \in (t_n, t_{n+1}]$. Therefore, let $v_{h,\tau} = \rho \in \mathcal{S}_\tau$ and then

$$\begin{aligned} 0 &= \tau a_\varepsilon(t_{n+1}; \rho, \rho) + (\rho(t_n^+) - \rho(t_n), \rho(t_n^+))_{L^2(\Omega)} - \frac{\tau^2}{2}(u''_{\varepsilon,h}(c_n), \rho(t_n^+))_{L^2(\Omega)} \\ &\ge \alpha\int_{t_n}^{t_{n+1}} \|\rho\|_{V_h}^2 \, dt + \frac{1}{2}\left(\|\rho(t_n^+) - \rho(t_n)\|_{L^2(\Omega)}^2 + \|\rho(t_n^+)\|_{L^2(\Omega)}^2 - \|\rho(t_n)\|_{L^2(\Omega)}^2\right) \\ &\qquad - \frac{\tau^{3/2}}{2}\|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}\|\rho\|_{L^2(t_n,t_{n+1};V_h)} \\ &\ge \alpha\|\rho\|_{L^2(t_n,t_{n+1};V_h)}^2 + \frac{1}{2}\left(\|\rho(t_n^+)\|_{L^2(\Omega)}^2 - \|\rho(t_n)\|_{L^2(\Omega)}^2\right) \\ &\qquad - \left(\frac{\tau^3}{16q}\|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}^2 + q\|\rho\|_{L^2(t_n,t_{n+1};V_h)}^2\right) \end{aligned} \tag{2.289}$$

for any $q \in \mathbb{R}$. If $q = \alpha$, then

$$\begin{aligned} \|\rho(t_n^+)\|_{L^2(\Omega)}^2 &\le \|\rho(t_n)\|_{L^2(\Omega)}^2 + \frac{\tau^3}{8\alpha}\|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}^2 \\ &\le \|\rho(t_0)\|_{L^2(\Omega)}^2 + \frac{\tau^3(n+1)}{8\alpha}\|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}^2 \\ &\le \frac{T}{8\alpha}\tau^2\|T_1 u''_{\varepsilon,h}\|_{L^\infty(0,T;V_h^*)}^2 \end{aligned} \tag{2.290}$$

for all $n = 0, \cdots, N = 1$. Further if $q = \dfrac{\alpha}{2}$, then

$$\alpha \|\rho\|^2_{L^2(t_n,t_{n+1};V_h)} + \|\rho(t_n^+)\|^2_{L^2(\Omega)} - \|\rho(t_n)\|^2_{L^2(\Omega)} \le \frac{\tau^3}{4\alpha} \|T_1 u''_{\varepsilon,h}\|^2_{L^\infty\left(0,T;V_h^*\right)}. \tag{2.291}$$

Summing this from $n = 0$ to $n = N - 1$, then

$$\begin{aligned} \|\rho\|^2_{L^2(0,T;V_h)} &\le \frac{1}{\alpha}\|\rho(t_0)\|^2_{L^2(\Omega)} + \frac{\tau^3 N}{4\alpha^2}\|T_1 u''_{\varepsilon,h}\|^2_{L^\infty\left(0,T;V_h^*\right)} \\ &\le \frac{T}{4\alpha^2}\tau^2 \|T_1 u''_{\varepsilon,h}\|^2_{L^\infty\left(0,T;V_h^*\right)}. \end{aligned} \tag{2.292}$$

□

Theorem 18 (Error estimate)**.** Let ℓ, m be two integers which satisfy $2 \le \ell, m \le k+1$. Assume $u \in X_{\ell,m}$ and $u_{\varepsilon,h} \in C^2\left([0,T]; V_h\right)$, then there exists a positive constant C such that

$$\begin{aligned} \|u - u_{\varepsilon,h,\tau}\|^2_{L^2(0,T;L^2(\Omega))} &\le C\left(\int_0^T \left(h^{2(\ell-1)}\|u\|^2_{H^\ell(\Omega)} + h^{2m}\|u'\|^2_{H^m(\Omega)}\right)\, dt + h^{2j}\|u(0)\|^2_{H^k(\Omega)}\right) \\ &\quad + C\tau^2\|u'_{\varepsilon,h}\|^2_{H^1(0,T;L^2(\Omega))} + \frac{T}{4\alpha^2}\tau^2\|T_1 u''_{\varepsilon,h}\|^2_{L^\infty\left(0,T;V_h^*\right)}, \end{aligned} \tag{2.293}$$

where $j := \min\{\ell, m\}$.

Proof. We have

$$\|u-u_{\varepsilon,h,\tau}\|^2_{L^2(0,T;L^2(\Omega))} \le \|u-u_{\varepsilon,h}\|^2_{X_h} + \|u_{\varepsilon,h}-\Pi_\tau u_{\varepsilon,h}\|^2_{L^2(0,T;L^2(\Omega))} + \|\Pi_\tau u_{\varepsilon,h} - u_{\varepsilon,h,\tau}\|^2_{L^2(0,T;V_h)}. \tag{2.294}$$

Here the Theorem 16, the Lemma 38 and the approximation error estimate for piecewise constant yield the result. □

2.7 Numerical examples

Here we show a numerical example of Nitsche method for parabolic problem. Our example is given as $\Omega := (0,1)^2$, and

$$\begin{cases} u' - \Delta u + (1,1)^{\mathrm{T}} \cdot \nabla u + u &= f(\mathbf{x}, t) \qquad \text{in } \Omega \times (0,T), \\ u &= 0 \qquad \text{on } \Gamma \times (0,T), \\ u(\mathbf{x}, 0) &= \sin(\pi x)\sin(\pi y) \quad \text{for } \mathbf{x} = (x,y) \in \Omega. \end{cases} \tag{2.295}$$

That is, we let $\mu(\mathbf{x},t) = 1$, $\mathbf{a}(\mathbf{x},t) = (1,1)^{\mathrm{T}}$ and $c(\mathbf{x},t) = 1$. We check at once that they satisfy the assumption 3, therefore this problem has a unique solution for any $f \in L^2\left(0,T; H^{-1}(\Omega)\right)$. We let $T = 4$ and

$$\begin{aligned} f(x,y,t) \;:= &\left((x+y+2t-2t^2+2\pi^2)\sin(\pi x)\sin(\pi y)\right. \\ &\left.+(\pi - 2\pi t)\cos(\pi x)\sin(\pi y) + (\pi - 2\pi t)\sin(\pi x)\cos(\pi y)\right) e^{(x+y-1)t}, \end{aligned} \tag{2.296}$$

then it follows easily that $f \in L^2\left(0,T; L^2(\Omega)\right)$ and

$$u(x,y,t) := \sin(\pi x)\sin(\pi y)e^{(x+y-1)t} \tag{2.297}$$

is the unique solution.

In Figure 2.2, we show the exact solution at different time steps.

We use $(1,1)$-th degree B-spline basis functions for spatial discretization and implicit Euler scheme in temporal discretization. Therefore we have the approximate problem (2.256) and error estimate (2.293). We let $\tau := h/10$, where h is the mesh size for uniform mesh, then we have

$$\|u - u_{\varepsilon,h,\tau}\|_{L^2(0,T;L^2(\Omega))} \approx Ch. \tag{2.298}$$

In Figure 2.3, we report the computational result of the Nitsche method, and we show the boundary value of numerical solution in Figure 2.4. The weak imposition of the Dirichlet boundary condition is actually observed because the boundary value does not vanish.

Further, we report the error for uniform mesh in Figure 2.5. This shows that the rate of convergence is approximately equal to the unity, that is expected by the theory.

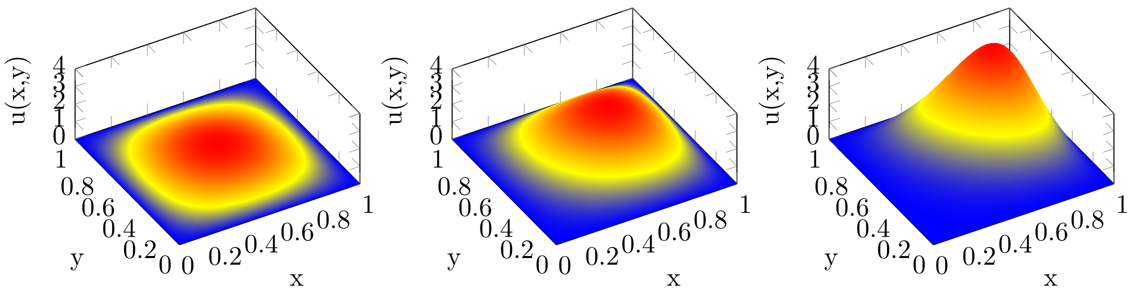


Figure 2.2: The exact solutions at different time steps

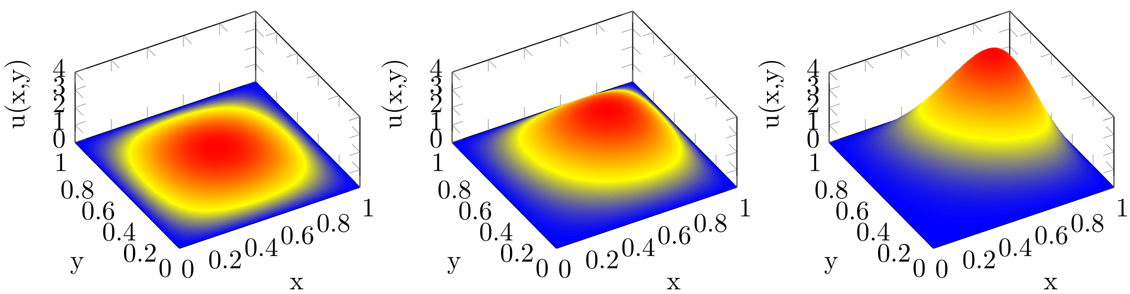


Figure 2.3: The numerical solutions of the Nitsche method at different time steps by considering the uniform mesh with mesh size $h = 1/30$.

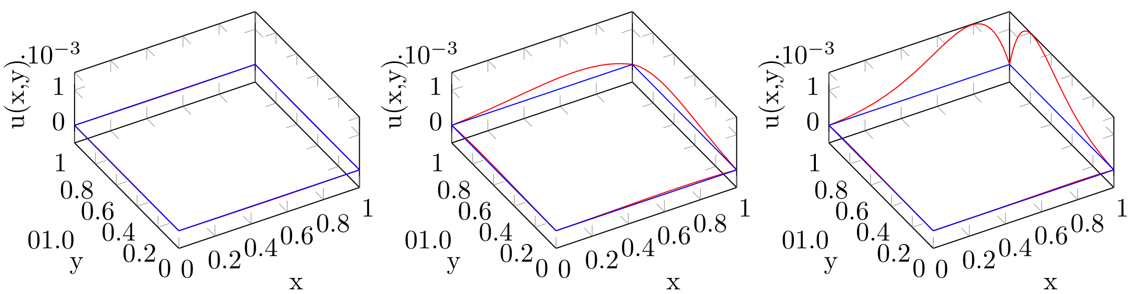


Figure 2.4: The numerical solutions of the Nitsche method on boundary Γ, at different time steps by considering the uniform mesh with mesh size $h = 1/30$.

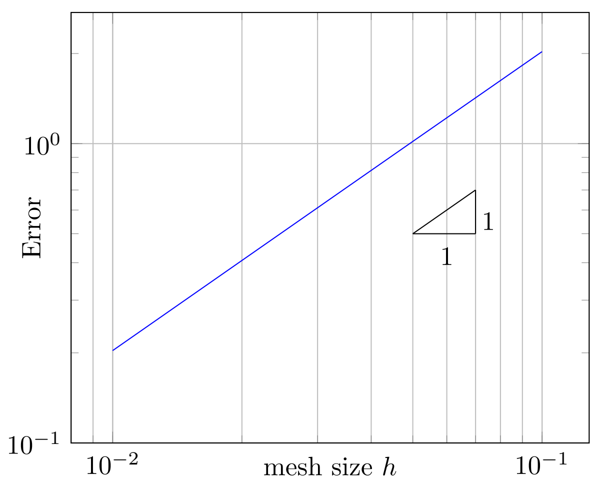


Figure 2.5: The $L^2\left(0,T;L^2(\Omega)\right)$ error on uniform mesh

Bibliography

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001/02.

[2] Y. Bazilevs, L. Beirão da Veiga, J. A. Cottrell, T. J. R. Hughes, and G. Sangalli. Isogeometric analysis: approximation, stability and error estimates for h-refined meshes. *Math. Models Methods Appl. Sci.*, 16(7):1031–1090, 2006.

[3] Y. Bazilevs, V. M. Calo, J. A. Cottrell, T. J. R. Hughes, A. Reali, and G. Scovazzi. Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 197:173–201, 2007.

[4] Y. Bazilevs, V. M. Calo, T. J. R. Hughes, and Y. Zhang. Isogeometric fluid–structure interaction: theory, algorithms, and computations. *Computational Mechanics*, 43:3–37, 2008.

[5] Y. Bazilevs and T. J. R. Hughes. Weak imposition of Dirichlet boundary conditions in fluid mechanics. *Comput. & Fluids*, 36(1):12–26, 2007.

[6] Y. Bazilevs, C. Michler, V. M. Calo, and T. J. R. Hughes. Weak Dirichlet boundary conditions for wall-bounded turbulent flows. *Comput. Methods Appl. Mech. Engrg.*, 196(49-52):4853–4862, 2007.

[7] Y. Bazilevs, K. Takizawa, and T. E. Tezduyar. *Computational Fluid-Structure Interaction: Methods and Applications*. Wiley, 2013.

[8] Y. Bazilevs, K. Takizawa, and T. E. Tezduyar. *Computational Fluid–Structure Interaction: Methods and Applications*. Wiley, February 2013.

[9] L. Beirão da Veiga, A. Buffa, G. Sangalli, and R. Vázquez. Mathematical analysis of variational isogeometric methods. *Acta Numer.*, 23:157–287, 2014.

[10] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[12] A. N. Brooks and T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32:199–259, 1982.

[13] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.

[14] J. A. Cottrell, T. J. R. Hughes, and Y. Bazilevs. *Isogeometric Analysis: Toward Integration of CAD and FEA*. Wiley, 2009.

[15] M. G. Cox. The numerical evaluation of B-splines. *J. Inst. Math. Appl.*, 10:134–149, 1972.

[16] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 5*. Springer-Verlag, Berlin, 1992. Evolution problems. I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig.

[17] C. de Boor. On calculating with *B*-splines. *J. Approximation Theory*, 6:50–62, 1972. Collection of articles dedicated to J. L. Walsh on his 75th birthday, V (Proc. Internat. Conf. Approximation Theory, Related Topics and their Applications, Univ. Maryland, College Park, Md., 1970).

[18] A. Embar, J. Dolbow, and I. Harari. Imposing Dirichlet boundary conditions with Nitsche's method and spline-based finite elements. *Internat. J. Numer. Methods Engrg.*, 83(7):877–898, 2010.

[19] A. Ern and J.-L Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.

[20] J. A. Evans and T. J. R. Hughes. Explicit trace inequalities for isogeometric analysis and parametric hexahedral finite elements. *Numer. Math.*, 123(2):259–290, 2013.

[21] J. A. Evans and T. J. R. Hughes. Isogeometric divergence-conforming B-splines for the steady Navier-Stokes equations. *Math. Models Methods Appl. Sci.*, 23(8):1421–1478, 2013.

[22] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[23] B. Heinrich and K. Pietsch. Nitsche type mortaring for some elliptic problem with corner singularities. *Computing*, 68(3):217–238, 2002.

[24] T. J. R. Hughes. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles, and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127:387–401, 1995.

[25] T. J. R. Hughes, J. A. Cottrell, and Y. Bazilevs. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput. Methods Appl. Mech. Engrg.*, 194(39-41):4135–4195, 2005.

[26] T. J. R. Hughes, W. K. Liu, and T. K. Zimmermann. Lagrangian–Eulerian finite element formulation for incompressible viscous flows. *Computer Methods in Applied Mechanics and Engineering*, 29:329–349, 1981.

[27] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971. Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

[28] Y. Otoguro, K. Takizawa, and T. E. Tezduyar. Space–time VMS computational flow analysis with isogeometric discretization and a general-purpose NURBS mesh generation method. *Computers & Fluids*, published online, DOI: 10.1016/j.compfluid.2017.04.017, 2017.

[29] L. Piegl and W. Tiller. *The NURBS book*. Springer Science & Business Media, 2012.

[30] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1994.

[31] T. Roubí ček. *Nonlinear partial differential equations with applications*, volume 153 of *International Series of Numerical Mathematics*. Birkhäuser/Springer Basel AG, Basel, second edition, 2013.

[32] N. Saito. Variational analysis of the discontinuous galerkin time-stepping method for parabolic equations, 2017.

[33] L. L. Schumaker. *Spline functions: basic theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2007.

[34] K. Takizawa, Y. Bazilevs, and T. E. Tezduyar. Space–time and ALE-VMS techniques for patient-specific cardiovascular fluid–structure interaction modeling. *Archives of Computational Methods in Engineering*, 19:171–225, 2012.

[35] K. Takizawa, B. Henicke, A. Puntel, N. Kostov, and T. E. Tezduyar. Space–time techniques for computational aerodynamics modeling of flapping wings of an actual locust. *Computational Mechanics*, 50:743–760, 2012.

[36] K. Takizawa, B. Henicke, A. Puntel, T. Spielman, and T. E. Tezduyar. Space–time computational techniques for the aerodynamics of flapping wings. *Journal of Applied Mechanics*, 79:010903, 2012.

[37] K. Takizawa, N. Kostov, A. Puntel, B. Henicke, and T. E. Tezduyar. Space–time computational analysis of bio-inspired flapping-wing aerodynamics of a micro aerial vehicle. *Computational Mechanics*, 50:761–778, 2012.

[38] K. Takizawa and T. E. Tezduyar. Multiscale space–time fluid–structure interaction techniques. *Computational Mechanics*, 48:247–267, 2011.

[39] K. Takizawa and T. E. Tezduyar. Space–time fluid–structure interaction methods. *Mathematical Models and Methods in Applied Sciences*, 22(supp02):1230001, 2012.

[40] K. Takizawa and T. E. Tezduyar. Space-time computation techniques with continuous representation in time (ST-C). *Comput. Mech.*, 53(1):91–99, 2014.

[41] K. Takizawa, T. E. Tezduyar, A Buscher, and S. Asada. Space–time fluid mechanics computation of heart valve models. *Computational Mechanics*, 54:973–986, 2014.

[42] K. Takizawa, T. E. Tezduyar, A Buscher, and S. Asada. Space–time interface-tracking with topology change (ST-TC). *Computational Mechanics*, 54:955–971, 2014.

[43] K. Takizawa, T. E. Tezduyar, T. Kuraishi, S. Tabata, and H. Takagi. Computational thermo-fluid analysis of a disk brake. *Computational Mechanics*, 57:965–977, 2016.

[44] K. Takizawa, T. E. Tezduyar, S. McIntyre, N. Kostov, R. Kolesar, and C. Habluetzel. Space–time VMS computation of wind-turbine rotor and tower aerodynamics. *Computational Mechanics*, 53:1–15, 2014.

[45] K. Takizawa, T. E. Tezduyar, H. Mochizuki, H. Hattori, S. Mei, L. Pan, and K. Montel. Space–time VMS method for flow computations with slip interfaces (ST-SI). *Mathematical Models and Methods in Applied Sciences*, 25:2377–2406, 2015.

[46] K. Takizawa, T. E. Tezduyar, Y. Otoguro, T. Terahara, T. Kuraishi, and H. Hattori. Turbocharger flow computations with the Space–Time Isogeometric Analysis (ST-IGA). *Computers & Fluids*, 142:15–20, 2017.

[47] K. Takizawa, T. E. Tezduyar, and T. Terahara. Ram-air parachute structural and fluid mechanics computations with the space–time isogeometric analysis (ST-IGA). *Computers & Fluids*, 141:191–200, 2016.

[48] K. Takizawa, T. E. Tezduyar, T. Terahara, and T. Sasaki. Heart valve flow computation with the integrated Space–Time VMS, Slip Interface, Topology Change and Isogeometric Discretization methods. *Computers & Fluids*, published online, DOI: 10.1016/j.compfluid.2016.11.012, November 2016.

[49] T. E. Tezduyar. Stabilized finite element formulations for incompressible flow computations. *Advances in Applied Mechanics*, 28:1–44, 1992.

[50] T. E. Tezduyar. Computation of moving boundaries and interfaces and stabilization parameters. *International Journal for Numerical Methods in Fluids*, 43:555–575, 2003.

[51] T. E. Tezduyar and S. Sathe. Modeling of fluid–structure interactions with the space–time finite elements: Solution techniques. *International Journal for Numerical Methods in Fluids*, 54:855–900, 2007.

[52] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1997.