MASTER'S THESIS 2023

FINITE ELEMENT MODELLING OF LINEAR ROLLING CONTACT PROBLEMS

LUIGI ROMANO

UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Computational Mathematics
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2023

Finite element modelling of linear rolling contact problems

LUIGI ROMANO

Supervisor: Axel Målqvist, Department of Mathematical Sciences
Examiner: Stig Larsson, Department of Mathematical Sciences

Master's Thesis 2023
Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Computational Mathematics
University of Gothenburg
SE-412 96 Gothenburg

Finite element modelling of linear rolling contact problems
LUIGI ROMANO
Department of Mathematical Sciences
University of Gothenburg

Abstract

The present work is devoted to the finite element modelling of linear hyperbolic rolling contact problems. The main equations encountered in rolling contact mechanics are reviewed in the first part of the thesis, with particular emphasis on applications from automotive and vehicle engineering. In contrast to the common hyperbolic systems found in the literature, such equations include integral and boundary terms, as well as time-varying transport velocities, that require special treatment. In this context, existence and uniqueness properties are discussed within the theoretical framework offered by the semigroup theory. The second part of the thesis is then dedicated to recovering approximated solutions to the considered problems, by combining discontinuous Galerkin finite element methods (DGMs) with explicit Runge-Kutta (RK) schemes of the first and second order for time discretisation. Under opportune assumptions on the smoothness of the sought solutions, and owing to certain generalised Courant-Friedrichs-Lewy (CFL) conditions, quasi-optimal error bounds are derived for the complete discrete schemes. The proposed algorithms are then tested on simple scalar equations in one space dimension. Numerical experiments seem to suggest the theoretical error estimates to be sharp.

Keywords

Rolling contact mechanics; hyperbolic rolling contact problems; numerical modelling; discontinuous Galerkin finite element methods (DGMs); Runge-Kutta schemes

Acknowledgements

Pursuing a PhD is difficult, and getting a Master's degree in Mathematics while pursuing a PhD is even more difficult. Fortunately, I was assisted by many wonderful people in this arduous journey. First, I would like to thank my Supervisor Axel Målqvist, who has always been available beyond what required, and Stig Larsson for agreeing to be the Examiner for this thesis project. Many people from Chalmers and GU exhorted me to apply for the Master's program in Mathematics: Håkan Andreasson, Simone Calogero, and of course my dear friend Milo Viviani. Thank you for your contagious passion and enthusiasm. To both my Supervisors at Chalmers M2, Fredrik Bruzelius and Bengt Jacobson, I am particularly indebted for allowing me to spend some time on my other interests than the COVER project. Finally, I would like to express my gratitude to my Mum, my Dad, my brother Francesco, and my beloved Yulia for their infinite support and love.

Luigi Romano
Gothenburg, 2023

Table of Contents

Chapter 1

Introduction

Linear hyperbolic partial differential equations (PDEs) are ubiquitous in physics and engineering [1, 2]. Particularly, in the field of contact mechanics, hyperbolic PDEs appear in numerous applications concerning rolling contact phenomena, where they typically describe the complex interactions occurring between wheel and rail [3–6], tyre and road asphalt [7, 8], but also the dynamics of roll bearing elements, belt-pulley mechanisms, and continuous automotive transmissions [9–11]. Especially in vehicle engineering, the most common formulations adopt brush-like representations of the contacting rolling bodies [12, 13], where dry friction is modelled according to the famous Coulomb-Amontons theory [14–16]. Advanced descriptions based on modified friction theories, such as the LuGre-brush model [17–25], are also able to accurately account for wet and lubricated friction and have been successfully employed for control design. In one space dimension, a typical example of hyperbolic PDEs encountered in rolling contact mechanics is as follows:

$$\frac{\partial u(x,t)}{\partial t} + a(x,t)\frac{\partial u(x,t)}{\partial x} = B(t)u(x,t) + C(t)u(1,t) + f(x,t), \quad \text{for } (x,t) \in (0,1)\times(0,T), \tag{1.1a}$$

$$u(0,t) = 0, \qquad \text{for } t \in (0,T), \tag{1.1b}$$

$$u(x,0) = u_0(x), \qquad \text{for } x \in (0,1), \tag{1.1c}$$

where $u(x,t) \in \mathbb{R}^n$ is the unknown solution, measuring a tangential deformation, $a(x,t) \in \mathbb{R}$ denotes the transport velocity, $f(x,t) \in \mathbb{R}^n$ is the external forcing term, and $C(t)$ and $B(t)$ are bounded operators.

Whilst numerous analytical steady-state models allowing for simple closed-form solutions may be encountered in the dedicated literature, understanding in detail unsteady effects usually requires solving the underlying PDEs numerically, which has limited the deal of effort devoted to such investigations. Restricting the attention to the field of railway dynamics, departing from the theory developed by Kalker [26, 27], the currently available numerical methods – including commercial software like FASTSIM® and CONTACT® – employ variational techniques to iteratively solve the transient problem until convergence is achieved [28–31]. Albeit being sufficiently accurate for the purpose they serve, such algorithms cannot be easily extended to cover problems arising from other subfields of mechanical engineering and contact mechanics. For example, the governing PDEs of the brush models classically adopted in tyre and road vehicle dynamics often incorporate nonlocal and boundary terms that are not accommodated by the existing methods [7, 8, 16, 32]. For this reason, unsteady phenomena are often disregarded in automotive engineering applications. Alternatively, transient dynamics are approximated in terms of first-order ordinary differential equations (ODEs): standard approaches include, for example, resorting to Padé approximants or other reduced-order descriptions [7, 8, 16, 32–34]. In the past years, such simplifications have been clearly motivated by the need of relying on computationally inexpensive models to be conveniently used in vehicle dynamics simulations and control applications. In fact, even simulating simple PDEs conflicted with the requirement for simplicity and real-time computational speed. On the one hand, the empirical validation of such simplified descriptions often requires expensive equipment and extensive experimental campaigns, which legitimates instead the adoption of higher-fidelity models allowing for comparison. In this way, approximated ODE-based representations may be validated directly against the PDE systems whose dynamical behaviour they aim to imitate. On the other hand, the limitations connected to the needs for simplicity and real-time performance may partially be overcome with the improved computational power at the disposal of modern vehicles. In this context, the present thesis is devoted to the numerical modelling of linear rolling contact phenomena, by combining PDE-based formulations with discontinuous Galerkin finite element methods (DGMs) for space approximation [35–38]. Since the ultimate ambition is to develop numerical algorithms (or extensions thereof) enabling fast calculations with performance close to real-time requirements, and easily implementable in virtual environments like MATLAB/Simulink®, time discretisation is then achieved using explicit schemes. More specifically, first and second-order Runge-Kutta (RK) algorithms, which represent a sufficiently good compromise between accuracy and computational speed [39, 40], are mainly explored in the present work. To the author's best knowledge, the numerical methods developed in this thesis are completely novel concerning the modelling of linear rolling contact phenomena described by brush-like formulations.

From the perspective of the pure mathematical analysis, the approach pursued in this thesis is heavily inspired by the research carried out in [39, 40], where the authors have rigorously investigated the rate of convergence of different finite element methods (FEMs) in conjunction with explicit RK schemes up to the third order. However, it is worth clarifying that the techniques developed in [39, 40] are not directly applicable to the initial-boundary-value problems (IBVPs) appearing in rolling contact mechanics and considered in the following chapters. Indeed, as already mentioned, PDEs describing rolling contact phenomena may contain integral and boundary terms whose presence has not been accounted for in previous studies. Additionally, handling time-varying data requires modifications of the analyses conducted in [39, 40], since the corresponding discrete equations for the error dynamics cannot be cast in the same form as that considered in [39, 40]. Therefore, apart from exploring new engineering applications of the DGMs combined with RK algorithms, the present work also delivers some fundamental results concerning the mathematical analysis of such numerical schemes. In this context, it should be also mentioned that the techniques presented in this thesis may be extended to the study of other interesting problems arising from different branches of physics and engineering. For example, typical equations that may be covered by the results advocated in the present work include those treated in [41].

In particular, the remainder of this thesis is organised as follows. In Chap. 2, the general structure for the considered hyperbolic PDEs is outlined concerning the one-dimensional and multi-dimensional cases. The existence and uniqueness of such equations are established according to the semigroup theory, which offers an adequate mathematical framework to analyse hyperbolic evolution problems. Chapter 3 recalls the generalities of the DGMs, including the notions of mesh and mesh elements, and introduces the space semi-discretisation approach, which enables recovering approximated solutions to the considered initial-boundary-value problems (IBVPs) within finite-dimensional functional spaces. More specifically, this is accomplished by opportunely choosing an appropriate polynomial behaviour on each mesh element and replacing the continuous operator appearing in the abstract formulation with a discrete counterpart, whose salient properties are investigated in detail. Then, Chap. 4 moves to the analysis of the complete discrete schemes, which, as already mentioned, cover explicit RK algorithms of the first and second order (RK1 and RK2, respectively). Owing to certain refined Courant-Friedrichs-Lewy (CFL) conditions, convergence results for the complete discrete schemes are derived under the assumption of sufficiently smooth exact solutions. Chapter 5 is dedicated to the numerical implementation of the proposed schemes. In particular, numerical experiments concerning the convergence of the error estimates derived theoretically are first conducted considering sufficiently smooth solutions. Regarding instead the applications, the adduced examples are borrowed from typical problems arising in rolling contact mechanics and are mainly focused on the one-dimensional IBVPs that are encountered in the automotive field. Finally, the main conclusions, together with some directions for future research, are summarised in Chap. 6.

Chapter 2

Hyperbolic equations in rolling contact mechanics: semigroup theory and well-posedness

The present Chapter is devoted to establishing existence and uniqueness results concerning the main hyperbolic equations, and the corresponding initial-boundary-value problems (IBVPs), arising in rolling contact mechanics. This is accomplished within the framework provided by the semigroup theory. In particular, some useful concepts and notions are recalled in Sect. 2.1 to the extent that is necessary to understand the results advocated in the thesis. Section 2.2 treats more specifically the main equations of interest, for which well-posedness is proved. Finally, Sect. 2.3 concludes the Chapter by providing a broader overview on the problem, and stating some essential results that fall outside the semigroup framework but might still be of interest.

2.1 Abstract hyperbolic evolution equations

This Section recalls some fundamental results that are needed to show existence and uniqueness for the main hyperbolic IBVPs considered in the thesis, and introduces the abstract formulation of hyperbolic evolution equations.

2.1.1 Preliminaries and notation

In this thesis, the set of real numbers is indicated with $\mathbb{R}$; $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ denote the set of positive real numbers and positive real numbers including zero, respectively. The sets of natural numbers excluding and including zero are indicated with $\mathbb{N}$ and $\mathbb{N}_0$.

Generic Banach spaces are conventionally denoted by X (respectively Y), and equipped with norm $\|\cdot\|_X$ (respectively $\|\cdot\|_Y$); the identity operator is denoted by I_X (respectively I_Y). Similarly, generic Hilbert spaces are indicated with V, and equipped with inner product $\langle\cdot,\cdot\rangle_V$ and norm $\|\cdot\|_V$. The corresponding identity operator is I_V. Specifically, given a domain $\Omega \subset \mathbb{R}^d$, the Hilbert space $L^2(\Omega;\mathbb{R}^n)$ is endowed with inner product and induced norm

$$\langle v, w\rangle_{L^2(\Omega;\mathbb{R}^n)} \triangleq \int_\Omega v(x)\cdot w(x)\,\mathrm{d}x = \int_\Omega v^{\mathrm{T}}(x)w(x)\,\mathrm{d}x, \tag{2.1a}$$

$$\|v(\cdot)\|^2_{L^2(\Omega;\mathbb{R}^n)} \triangleq \langle v, v\rangle_{L^2(\Omega;\mathbb{R}^n)} = \int_\Omega \|v(x)\|_2^2\,\mathrm{d}x, \tag{2.1b}$$

where $\|\cdot\|_2$ denotes the standard Euclidean norm in $\mathbb{R}^n$, respectively. A function $v(\cdot)$ is said to belong to the space $L^2(\Omega;\mathbb{R}^n)$, noted $v \in L^2(\Omega;\mathbb{R}^n)$, if its L^2-norm defined according to Eq. (2.1b) is finite. Similarly, the Hilbert space $H^1(\Omega;\mathbb{R}^n)$ is naturally equipped with seminorm and norm

$$\left|v(\cdot)\right|^2_{H^1(\Omega;\mathbb{R}^n)} \triangleq \sum_{i=1}^{n} \left\|\nabla v_i(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^d)}, \tag{2.2a}$$

$$\left\|v(\cdot)\right\|^2_{H^1(\Omega;\mathbb{R}^n)} \triangleq \left\|v(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left|v(\cdot)\right|^2_{H^1(\Omega;\mathbb{R}^n)}. \tag{2.2b}$$

A function $v(\cdot)$ is said to belong to the space $H^1(\Omega;\mathbb{R}^n)$, noted $v \in H^1(\Omega;\mathbb{R}^n)$, if its H^1-norm as in Eq. (2.2b) is finite.

The Banach space $C^0(\overline{\Omega};\mathbb{R}^n)$ is also endowed with norm

$$\left\|v(\cdot)\right\|_\infty \triangleq \max_{\Omega} \left\|v(x)\right\|_2. \tag{2.3}$$

Concerning a function $v(\cdot,\cdot)$ defined on the space-time cylinder $\Omega \times (0,T)$, it is often convenient to interpret $v(\cdot)$ as a function of the time variable with values in a Banach space X, spanned by functions of the space variables, i.e.,

$$v : (0,T) \ni t \mapsto v(t) \equiv v(\cdot,t) \in X. \tag{2.4}$$

For any integer $l \in \mathbb{N}_0$, the spaces $C^l([0,T];X)$ are also considered, spanned by functions that are l times continuously differentiable in the interval $[0,T]$. In particular, the space $C^0([0,T];X)$ is a Banach space when equipped with the norm

$$\left\|v(\cdot,\cdot)\right\|_{C^0([0,T];X)} \triangleq \max_{t\in[0,T]} \left\|v(\cdot,t)\right\|_X, \tag{2.5}$$

and space $C^l([0,T];X)$ is a Banach space when equipped with the norm

$$\left\|v(\cdot,\cdot)\right\|_{C^l([0,T];X)} \triangleq \max_{0\leq m\leq l} \left\|\frac{\partial^m v(\cdot,\cdot)}{\partial t^m}\right\|_{C^0([0,T];X)}. \tag{2.6}$$

In the following, given two Banach spaces X and Y, respectively, $\mathscr{L}(X;Y)$ denotes the space of (possibly unbounded) linear operators from X to Y, whereas $\mathscr{B}(X;Y)$ the space of bounded linear operators from X to Y, abbreviated $\mathscr{B}(X)$ whenever $Y = X$. For $l \in \mathbb{N}_0$, $C^l([0,T];\mathscr{L}(X;Y))$ and $C^l([0,T];\mathscr{B}(X;Y))$ denote the spaces of (possibly unbounded) and bounded linear operators from X to Y, respectively, whose coefficients are l times continuously differentiable on $[0,T]$.

Finally, $\mathbf{M}_{m\times n}(\mathbb{R})$, $\mathbf{Skew}_n(\mathbb{R})$, and $\mathbf{GL}_n(\mathbb{R})$ denote the groups of matrices, skew symmetric matrices, and invertible matrices, respectively, assuming values in $\mathbb{R}^{m\times n}$ and $\mathbb{R}^{n\times n}$; the identity matrix is denoted by $I_n \in \mathbf{GL}_n(\mathbb{R})$. $\mathbf{SO}_n(\mathbb{R})$ denotes the group of unitary rotations in $\mathbb{R}^n$.

2.1.2 Semigroup approach to abstract hyperbolic IBVPs

As already mentioned, the present Chapter is dedicated to the study of the well-posedness of hyperbolic IBVPs describing linear rolling contact phenomena, using a semigroup approach. In particular, given a Banach space X, these IBVPs may be recast in an abstract formulation as follows:

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = A(t)u(t) + f(t), \qquad \text{for} \, t \in (0,T), \tag{2.7a}$$

$$u(0) = u_0, \tag{2.7b}$$

where, for each $t \in [0,T]$, the (possibly) unbounded operator $(A, D(A))$ is the infinitesimal generator of a C_0-semigroup on X. In this context, the characterisation of a C_0-semigroup is briefly recalled below.

Definition 2.1.1 (C_0-semigroup). *Let X be a Banach space. A one-parameter family $\{T(t)\}_{t\in[0,\infty)}$ of bounded linear operators from X into X is a C_0-*semigroup *on X if*

(i) $T(0) = I_X$,

(ii) $T(t+\tilde{t}) = T(t)T(\tilde{t})$ *(semigroup property),*

(iii) $\lim_{t\downarrow 0} T(t)x = x$ *for every* $x \in X$.

The first two properties above are sufficient themselves to define a semigroup, whereas the last one is required for an operator to be a C_0-semigroup (on this matter, see, e.g., [42–46]).

Returning to the abstract IBVP (2.7), the main notion of solutions considered in the following is *strict*, according to [42, 43].

Definition 2.1.2 (Strict solution (Tanabe [42, 43])). *A* strict solution *is defined to be a function* $u \in C^1([0,T];X) \cap C^0([0,T];D(A))$ *solving the IBVP* (2.7) *for* $t \in [0,T]$.

For completeness, the more common notion of *classical solution*, according, e.g., to [44], is also given.

Definition 2.1.3 (Classical solution (Pazy [44])). *A* classical solution *is defined to be a function* $u \in C^0([0,T];X) \cap C^1((0,T];X)$*, with* $u(t) \in D(A)$*, and solving the IBVP* (2.7) *for* $t \in (0,T]$.

Comparing the Definitions 2.1.2 and 2.1.3, which are slightly more precise than that originally proposed in [45], it is clear that a strict solution is also classical. Both classical and strict solutions to the IBVP (2.7) may be deduced by constructing an evolution operator $U_A(t,\tilde{t})$ associated with the infinitesimal generator $(A, D(A))$. This is a rather delicate task, which involves a number of subtle technicalities [45, 46]. The first requirement consists in checking that $(A, D(A))$ generates, in fact, a C_0-semigroup for all $t \in [0,T]$. This thesis mainly considers evolution equations in Hilbert spaces V rather than more general Banach spaces X. In this context, the following alternative version of Lumer-Phillips' Theorem 2.1.1, due to Curtain and Zwart [47, 48], asserts necessary and sufficient conditions for the operator $(A, D(A))$ to be the infinitesimal generator of a C_0-semigroup $T(t)$ on a Hilbert space V.

Theorem 2.1.1 (Lumer-Phillips (Curtain and Zwart [47, 48])). *Necessary and sufficient conditions for a closed, densely defined operator $(A, D(A))$ on a Hilbert space V to be the infinitesimal generator of a C_0-semigroup satisfying* $\|T(t)\| \leq \mathrm{e}^{\omega t}$ *are*

$$\mathrm{Re}\langle Av, v\rangle_V \leq \omega \|v(\cdot)\|_V^2, \qquad \text{for } v \in D(A), \tag{2.8a}$$

$$\mathrm{Re}\langle A^* v, v\rangle_V \leq \omega \|v(\cdot)\|_V^2, \qquad \text{for } v \in D(A^*), \tag{2.8b}$$

where $(A^*, D(A^*))$ *denotes the adjoint of* $(A, D(A))$.

Proof. See Corollary 2.2.3 in [47] or Corollary 2.3.3 in [48]. □

If an operator $(A, D(A))$ verifies a relationship of the same type as in Eq. (2.8a), it is said to be *quasi-dissipative* with constant ω. Therefore, Theorem 2.1.1 requires both $(A, D(A))$ and its adjoint $(A^*, D(A^*))$ to be quasi-dissipative with the same constant. More generally, the group of operators on a Banach space X that are infinitesimal generators of a C_0-semigroup satisfying $\|T(t)\| \leq \mathrm{e}^{\omega t}$ is conventionally denoted by $\mathscr{G}(1,\omega)$ [45, 46]. Hence, Theorem 2.1.1 states necessary and sufficient conditions to verify that a closed, densely defined operator on a Hilbert V space fulfils $A \in \mathscr{G}(1,\omega)$. Whilst the density of the operator is often trivial to check, a useful criterion to deliberate whether an operator is closed is offered by the subsequent Theorem 2.1.2.

Theorem 2.1.2. *Assume that X and Y are Banach spaces and let $(A, D(A))$ be a linear operator with domain $D(A) \subset X$ and range Y. If, in addition, A is invertible with $A^{-1} \in \mathscr{L}(Y;X)$, then $(A, D(A))$ is a closed linear operator.*

Proof. See Theorem 4.2-C in [49]. □

When the unbounded operator $(A, D(A))$ appearing in Eq. (2.7a) is time-independent, i.e., $A(t) = A$ for all $t \in [0,T]$, the IBVP (2.7) may be studied by simply analysing the properties of the infinitesimal generator. For time-dependent evolution equations, where $A(t)$ actually depends on the time variable, the existence and uniqueness of strict solutions to the IBVP (2.7) need to be more generally addressed concerning the entire family $\{A(t)\}_{t\in[0,T]}$ of infinitesimal generators. In this context, a crucial role is played by the notion of *stable family*, according to the following Definition 2.1.4.

Definition 2.1.4 (Stable family of infinitesimal generators). *A family $\{A(t)\}_{t\in[0,T]}$ of infinitesimal generators of C_0-semigroups on a Banach space X is said to be stable with stability constants $M \geq 1$, ω if*

$$\rho\Big(A(t)\Big) \supset (\omega, \infty), \qquad \text{for} \, t \in [0,T] \tag{2.9}$$

for every finite sequence $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k \leq T$, $k \in \mathbb{N}$, and

$$\left\| \prod_{j=1}^{k} R\Big(\lambda, A(t_j)\Big) \right\| \leq M(\lambda-\omega)^{-k}, \qquad \text{for} \, \lambda > \omega, \tag{2.10}$$

where $\rho(A(t))$ denotes the resolvent set of $A(t)$, and $R(\lambda, A(t_j)) \triangleq (\lambda I_X - A(t_j))^{-1}$ the resolvent of $A(t)$ at the time $t = t_j$.

It is worth emphasising that, in Eq. (2.10), the products containing $\{t_j\}$ are always *time ordered*, that is, a factor with larger t_j stands to the left of ones with smaller t_j. It is also clear from Definition 2.1.4 that, if $A(t) \in \mathscr{G}(1,\omega)$ for all $t \in [0,T]$, then the corresponding family $\{A(t)\}_{t\in[0,T]}$ is stable with stability constants 1, ω. The following perturbation Theorem 2.1.3 provides an additional criterion to decide about the stability of a family of infinitesimal generators.

Theorem 2.1.3 (Perturbation theorem). *Let $\{A(t)\}_{t\in[0,T]}$ be a stable family of infinitesimal generators with stability constants M, ω, and $B(t) \in \mathscr{B}(X;Y)$, $t \in [0,T]$, a bounded linear operator on X. If $\|B(t)\| \leq B_{\max}$ for all $t \in [0,T]$, then $\{\tilde{A}(t)\}_{t\in[0,T]} \triangleq \{A(t) + B(t)\}_{t\in[0,T]}$ is a stable family of infinitesimal generators with stability constants M, $\omega + MB_{\max}$.*

Proof. See Theorem 7.4 in [43] or 5.2.3 in [44]. □

Finally, the next Theorem 2.1.4 establishes sufficient conditions to deduce the existence of strict solutions to the IBVP described by Eqs. (2.7), owing to the assumption that the domain $D(A(t)) = D(A(0))$ of the operator $(A, D(A))$ is independent of the time.

Theorem 2.1.4 (Existence and uniqueness of strict solutions)**.** *Suppose that* $\{A(t)\}_{t\in[0,T]}$ *is a stable family of infinitesimal generators of* C_0*-semigroups on a Banach space* X *such that* $D(A(t)) = D(A(0))$ *is independent of* t *and, for each* $u_0 \in D(A)$*,* $A(t)u_0$ *is continuously differentiable in* X*. If* $u_0 \in D(A)$ *and* $f \in C^1([0,T];X)$*, then the unique strict solution to the IBVP described by Eqs.* (2.7) *is given by*

$$u(t) = U_A(t,0)u_0 + \int_0^t U_A\left(t,t'\right) f\left(t'\right) \mathrm{d}t', \tag{2.11}$$

where $U_A(t,\tilde{t})$ *denotes the evolution operator associated with the infinitesimal generator* $(A, D(A))$*.*

Proof. See Theorem 4.5.4 in [42] or 7.6 in [43]. □

Remark 2.1.1. *The expression defined by Eq.* (2.11) *has meaning even if* $u_0 \in X$ *and* $f \in L^p((0,T);X)$*,* $p \geq 1$*. In such a case, the function in Eq.* (2.11) *is said to be a* mild solution*. This is formalised in the next Definition 2.1.5.*

Definition 2.1.5 (Mild solution)**.** *For every* $u_0 \in X$ *and* $f \in L^p((0,T);X)$*,* $p \geq 1$*, the continuous function defined according to Eq.* (2.11) *is a* mild solution *of the IBVP* (2.7)*.*

2.2 Linear hyperbolic IBVPs in rolling contact

The semigroup framework briefly illustrated in Sect. 2.1 may be conveniently applied to prove existence and uniqueness for hyperbolic equations arising from problems considered in rolling contact mechanics, after restating them in an abstract setting. In this context, the present Section considers two separate classes of IBVPs: equations in one space dimension, and equations in several space dimensions. Concerning systems evolving in one dimension, the presence of nonlocal and boundary terms is also taken into account, which is motivated by some recent works by the author.

In the remainder of this Chapter, the focus is primarily on regular (i.e., strict and classical) solutions, which enjoy peculiar smoothness properties that are required for the error analysis performed in Chap. 4; under more relaxed regularity assumptions, the existence of mild solutions to the considered IBVPs follows immediately from Definition 2.1.5. In particular, considering a generic domain $\Omega \subset \mathbb{R}^d$ with boundary $\Gamma \triangleq \partial\Omega$, regular solutions are sought in the Hilbert space $L^2(\Omega;\mathbb{R}^n)$, and it is assumed that the domain $D(A(t)) = D(A(0))$ of the operator $(A, D(A))$ is independent of the time variable. According to Definitions 2.1.3 and 2.1.2, classical and strict solutions therefore correspond to functions $u \in C^0([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^1((0,T];L^2(\Omega;\mathbb{R}^n))$ and $u \in C^1([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];D(A))$, respectively, with $D(A)$ opportunely defined and $A(t)$ satisfying $A \in C^1([0,T];\mathscr{L}(D(A);L^2(\Omega;\mathbb{R}^n)))$.

2.2.1 IBVPs in one space dimension

In this thesis, the considered IBVPs involving a single space dimension generalise those derived in, e.g., [32, 33]. More specifically, by setting explicitly $\Omega = (0,1)$, the same structure of that in

Eqs. (1.1) is assumed, recalled here for convenience:

$$\frac{\partial u(x,t)}{\partial t} + a(x,t)\frac{\partial u(x,t)}{\partial x} = B(t)u(x,t) + C(t)u(1,t) + f(x,t), \quad \text{for}\,(x,t) \in (0,1)\times(0,T), \tag{2.12a}$$

$$u(0,t) = 0, \quad \text{for}\,t \in (0,T), \tag{2.12b}$$

$$u(x,0) = u_0(x), \quad \text{for}\,x \in (0,1), \tag{2.12c}$$

where, again, $u(x,t) \in \mathbb{R}^n$ is the unknown solution, $a(x,t) \in \mathbb{R}$ denotes the transport velocity, $f(x,t) \in \mathbb{R}^n$ is the external forcing term, $C \in C^1([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$ is a matrix of coefficients, and $B \in C^1([0,T];\mathscr{B}(L^2((0,1);\mathbb{R}^n)))$ is a bounded operator, typically having the form

$$(Bv)(x,t) = \tilde{B}(t)v(x) + \int_0^1 K(x,t)v(x)\,\mathrm{d}x, \tag{2.13}$$

with $\tilde{B} \in C^1([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$ and $K \in C^1([0,1]\times[0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$.

It is worth mentioning that the presence of the boundary (or trace) term in Eq. (2.12a) makes the analysis slightly more involved compared to the case where no boundary term appears. The two situations are addressed separately in the following, mainly for pedagogical reasons.

IBVPs in one space dimension without boundary terms

Neglecting the boundary term in Eq. (2.12a), and considering for a moment the case $B(t) = 0$, the IBVP (2.12) simplifies to

$$\frac{\partial u(x,t)}{\partial t} + a(x,t)\frac{\partial u(x,t)}{\partial x} = f(x,t), \quad \text{for}\,(x,t) \in (0,1)\times(0,T), \tag{2.14a}$$

$$u(0,t) = 0, \quad \text{for}\,t \in (0,T), \tag{2.14b}$$

$$u(x,0) = u_0(x), \quad \text{for}\,x \in (0,1). \tag{2.14c}$$

The system described by Eqs. (2.14) evidently consists of n uncoupled, linear transport equations. Next, to deal with the presence of a variable transport velocity $a(x,t)$, the following Assumption 2.2.1 is introduced.

Assumption 2.2.1. *The transport velocity satisfies* $a \in C^1([0,1]\times[0,T];[a_{\min},a_{\max}])$, *with* $a_{\min} > 0$.

The above Assumption 2.2.1 ensures that $A \in C^1([0,T];\mathscr{L}(D(A);L^2((0,1);\mathbb{R}^n)))$ and additionally that $\inf_{t\in[0,T]} a(1,t) \geq a_{\min}$. Accordingly, the IBVP (2.14) may be reformulated as an abstract hyperbolic evolution equation in the same for of Eqs. (2.7), with the time-varying, unbounded operator $(A,D(A))$, $A(t) : D(A) \mapsto L^2((0,1);\mathbb{R}^n)$, defined as

$$(Av)(x,t) \triangleq -a(x,t)\frac{\partial v(x)}{\partial x}, \tag{2.15a}$$

$$D(A) \triangleq \left\{ v \in H^1((0,1);\mathbb{R}^n) \,\middle|\, v(0) = 0 \right\}. \tag{2.15b}$$

Here, the unbounded operator $(A,D(A))$ should be regarded as a continuously differentiable linear operator from $D(A)$ to $L^2((0,1);\mathbb{R}^n)$, i.e., $A \in C^1([0,T];\mathscr{L}(D(A);L^2(\Omega;\mathbb{R}^n)))$. However, for what follows, the time variable is often interpreted as a parameter $t \in [0,T]$ [41], which motivates to simply write $A(t) \in \mathscr{L}(D(A);L^2(\Omega;\mathbb{R}^n))$. Moreover, it is crucial to observe that, even though the operator $A(t)$ is time-dependent, its domain $D(A(t)) = D(A(0))$ is not, owing to Assumption 2.2.1. In the following, $(A(t), D(A(t)))$, regarded as an operator for $t \in [0,T]$ fixed but arbitrary, is often abbreviated as $(A, D(A))$ to alleviate the notation (and similarly for its adjoint $(A^*(t), D(A^*(t))$ and other time-dependent operators).

To prove the well-posedness of the IBVP described by Eqs. (2.14), the *iter* is standard: it must be shown that the operator $(A, D(A))$ defined as in Eq. (2.15) is dense, closed, and quasi-dissipative together with its adjoint $(A^*, D(A^*))$. The density of $(A, D(A))$ is actually straightforward to verify, whereas closedness and quasi-dissipativity are proved in Lemmata 2.2.1 and 2.2.2.

Lemma 2.2.1 (Closedness)**.** *The operator* $(A, D(A))$ *as defined in Eqs.* (2.15) *is closed.*

Proof. According to Theorem 2.1.2, to prove that $(A, D(A))$ is closed, it suffices to show that there exists $A^{-1}(t)$ for all $t \in [0,T]$. By setting

$$(Av)(x,t) = -a(x,t)\frac{\partial v(x)}{\partial x} = w(x), \tag{2.16}$$

it may be immediately deduced that

$$v(x) = (A^{-1}w)(x,t) = -\int_0^x \frac{w\big(x'\big)}{a\big(x',t\big)}\,\mathrm{d}x'. \tag{2.17}$$

It is easy to verify that $A^{-1}(t) \in \mathscr{L}(L^2((0,1);\mathbb{R}^n); H^1((0,1);\mathbb{R}^n))$ and that $A^{-1}(t)A(t) = I_{D(A)}$ and $A(t)A^{-1}(t) = I_{L^2((0,1);\mathbb{R}^n)}$. Hence, $(A, D(A))$ is closed. $\square$

Proposition 2.2.1 below allows deducing an expression for the adjoint operator.

Proposition 2.2.1 (Adjoint operator)**.** *The adjoint operator* $(A^*, D(A^*))$, $A^*(t) : D(A^*) \mapsto L^2((0,1);\mathbb{R}^n)$ *of the operator* $(A, D(A))$ *defined in Eqs.* (2.15) *is given by*

$$(A^*v)(x,t) \triangleq \frac{\partial}{\partial x}\Big(a(x,t)v(x)\Big), \tag{2.18a}$$

$$D(A^*) \triangleq \Big\{ v \in H^1((0,1);\mathbb{R}^n) \;\Big|\; v(1) = 0 \Big\}. \tag{2.18b}$$

Proof. Since $(A, D(A))$ admits a bounded inverse according to the above Lemma 2.2.1, it is sufficient to deduce an expression for $(A^{-1}(t))^* = (A^*(t))^{-1}$ (see, e.g., Lemma A.3.72 in [48]). From Eq. (2.17),

$$(A^{-1}v)(x,t) = -\int_0^x \frac{v\big(x'\big)}{a\big(x',t\big)}\,\mathrm{d}x'. \tag{2.19}$$

Therefore, an application of Fubini's Theorem provides

$$\begin{aligned}\Big\langle A^{-1}(t)v, w\Big\rangle_{L^2((0,1);\mathbb{R}^n)} &= -\int_0^1\int_0^x \frac{v^{\mathrm{T}}\big(x'\big)}{a\big(x',t\big)} w(x)\,\mathrm{d}x'\,\mathrm{d}x \\ &= -\int_0^1 \frac{v^{\mathrm{T}}\big(x'\big)}{a\big(x',t\big)}\int_{x'}^1 w(x)\,\mathrm{d}x\,\mathrm{d}x' = \Big\langle v, (A^{-1}(t))^* w\Big\rangle_{L^2((0,1);\mathbb{R}^n)},\end{aligned} \tag{2.20}$$

which implies

$$\left(\left(A^{-1}\right)^{*}v\right)(x,t) = \left(\left(A^{*}\right)^{-1}v\right)(x,t) = -\frac{1}{a(x,t)}\int_x^1 v\left(x'\right)\mathrm{d}x'. \tag{2.21}$$

Using Eqs. (2.18) and (2.21), it may be verified that $A^*(t)(A^*(t))^{-1} = I_{L^2((0,1);\mathbb{R}^n)}$ and also $(A^*(t))^{-1}A^*(t) = I_{D(A^*)}$, thus concluding the proof. □

Lemma 2.2.2 (Quasi-dissipativity). *The operator* $(A,D(A))$, *together with its adjoint* $(A^*,D(A^*))$, *is quasi-dissipative with constant* $\omega \triangleq \frac{1}{2}\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty$.

Proof. Considering the operator $(A,D(A))$, taking the inner product on $L^2((0,1);\mathbb{R}^n)$ and integrating by parts yields

$$\begin{aligned}\left\langle A(t)v,v\right\rangle_{L^2((0,1);\mathbb{R}^n)} &= -\int_0^1 a(x,t)\frac{\partial v^{\mathrm{T}}(x)}{\partial x}v(x)\,\mathrm{d}x = -\frac{1}{2}\int_0^1 a(x,t)\frac{\partial}{\partial x}\left\|v(x)\right\|_2^2\mathrm{d}x \\ &= -\frac{1}{2}a(1,t)\left\|v(1)\right\|_2^2 + \frac{1}{2}\int_0^1 \frac{\partial a(x,t)}{\partial x}\left\|v(x)\right\|_2^2\mathrm{d}x \\ &\le \frac{1}{2}\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty \left\|v(\cdot)\right\|^2_{L^2((0,1);\mathbb{R}^n)}, \quad \text{for } v\in D(A).\end{aligned} \tag{2.22}$$

Additionally, departing from Eqs. (2.18), similar manipulations as previously give

$$\begin{aligned}\left\langle A^*(t)v,v\right\rangle_{L^2((0,1);\mathbb{R}^n)} &= \int_0^1 \frac{\partial}{\partial x}\left(a(x,t)v^{\mathrm{T}}(x)\right)v(x)\,\mathrm{d}x \\ &= -a(0,t)\left\|v(0)\right\|_2^2 - \int_0^1 a(x,t)v^{\mathrm{T}}(x)\frac{\partial v(x)}{\partial x}\mathrm{d}x \\ &= -\frac{1}{2}a(0,t)\left\|v(0)\right\|_2^2 + \frac{1}{2}\int_0^1 \frac{\partial a(x,t)}{\partial x}\left\|v(x)\right\|_2^2\mathrm{d}x \\ &\le \frac{1}{2}\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty \left\|v(\cdot)\right\|^2_{L^2((0,1);\mathbb{R}^n)}, \quad \text{for } v\in D(A^*).\end{aligned} \tag{2.23}$$

By combining Eqs. (2.22) and (2.23), the result immediately follows. □

Owing to Assumption 2.2.1 and the two results advocated above, it is then easy to establish the well-posedness of the IBVP (2.14), by taking advantage of the abstract formulation (2.7) and invoking Theorem 2.1.4.

Theorem 2.2.1 (Existence and uniqueness). *If Assumption 2.2.1 holds, the IBVP (2.14) admits a unique strict solution* $u \in C^1([0,T];L^2((0,1);\mathbb{R}^n))\cap C^0([0,T];D(A))$ *as in Eq. (2.11) for all* $f\in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ *and* $u_0\in D(A)$.

Proof. Since $C_0^1([0,1];\mathbb{R}^n)\subset D(A)$, the operator $(A,D(A))$ as defined in Eqs. (2.15) is dense, i.e., $\overline{D(A)} = L^2((0,1);\mathbb{R}^n)$. Moreover, it is closed and quasi-dissipative together with its adjoint $(A^*,D(A^*))$ according to Lemmata 2.2.1 and 2.2.2. It follows from Lumer-Phillips' Theorem 2.1.1 that, for $t\in[0,T]$, $A(t)$ is the infinitesimal generator of a C_0-semigroup. In particular, $A(t)\in\mathscr{G}(1,\omega)$, with

$$\omega \triangleq \frac{1}{2}\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty. \tag{2.24}$$

Therefore, the family $\{A(t)\}_{t\in[0,T]}$ is stable. Since $D(A(t)) = D(A(0))$ is independent of t and, for every $u_0 \in D(A)$, $A(t)u_0$ is continuously differentiable in $L^2((0,1);\mathbb{R}^n)$ by Assumption 2.2.1, Theorem 2.1.4 ensures the existence of a unique strict solution to Eq. (2.7) reading as in (2.11) for all $f \in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ and $u_0 \in D(A)$. □

Theorem 2.2.1 above asserts the well-posedness for the simplified IBVP (2.14), in which the n linear transport equations are all uncoupled. Returning now to the more general case of $B(t) \neq 0$, the following IBVP may instead be derived from Eqs. (2.12):

$$\frac{\partial u(x,t)}{\partial t} + a(x,t)\frac{\partial u(x,t)}{\partial x} = B(t)u(x,t) + f(x,t), \qquad \text{for}\,(x,t) \in (0,1)\times(0,T), \tag{2.25a}$$

$$u(0,t) = 0, \qquad \text{for}\, t \in (0,T), \tag{2.25b}$$

$$u(x,0) = u_0(x), \qquad \text{for}\, x \in (0,1). \tag{2.25c}$$

Recalling that the operator $B \in C^1([0,T];\mathscr{B}(L^2((0,1);\mathbb{R}^n)))$ is in the form of Eq. (2.13), with $\tilde{B} \in C^1([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$ and $K \in C^1([0,1]\times[0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$, the well-posedness of the IBVP above follows directly from the perturbation theory of linear unbounded operators, according to Theorem 2.1.3. In this case, the well-posedness result is formalised in the following Corollary 2.2.1.

Corollary 2.2.1 (Existence and uniqueness ($B(t) \neq 0$))**.** *If Assumption 2.2.1 holds, the IBVP (2.25) admits a unique strict solution $u \in C^1([0,T];L^2((0,1);\mathbb{R}^n)) \cap C^0([0,T];D(A))$ for all $f \in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ and $u_0 \in D(A)$.*

Proof. Consider again the operator $(A,D(A))$ as in Eqs. (2.15) and define the operator $(\tilde{A},D(\tilde{A}))$, with $(\tilde{A}v)(x,t) = (Av)(x,t) + (Bv)(x,t)$ and $D(\tilde{A}(t)) = D(A(t)) = D(A(0))$ independent of t[1]. Since $A(t) \in \mathscr{G}(1,\omega)$ and $\|B(t)\| \leq B_{\max}$ for all $t \in [0,T]$, it follows that $\tilde{A}(t) \in \mathscr{G}(1,\omega + B_{\max})$, and consequently the family $\{\tilde{A}(t)\}_{t\in[0,T]}$ is stable according to Theorem 2.1.3. Moreover, for every $u_0 \in D(A)$, $\tilde{A}(t)u_0$ is continuously differentiable in $L^2((0,1);\mathbb{R}^n)$, and thus Theorem 2.1.4 ensures the existence of a unique strict solution for all $f \in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ and $u_0 \in D(A)$. □

The result advocated above concludes the study of IBVPs without trace terms. The next part of the analysis extends to the case were boundary contributions also appear.

IBVPs in one space dimension with boundary terms

In the context of rolling contact mechanics, transport equations with boundary terms have recently appear in [32, 33], concerning the study of problems involving partially flexible bodies. Therefore, the analysis conducted in this thesis considers IBVPs whose structure generalise that of the systems studied in [32, 33]. In particular, including the boundary term in Eq. (2.12a), and considering for a moment the case $B(t) = 0$, the IBVP (2.12) simplifies to

$$\frac{\partial u(x,t)}{\partial t} + a(x,t)\frac{\partial u(x,t)}{\partial x} = C(t)u(1,t) + f(x,t), \qquad \text{for}\,(x,t) \in (0,1)\times(0,T), \tag{2.26a}$$

$$u(0,t) = 0, \qquad \text{for}\, t \in (0,T), \tag{2.26b}$$

[1]Here, $D(A)$ is considered a Banach space equipped with the graph norm $\|v(\cdot)\|^2_{D(A)} = \|v(\cdot)\|^2_{L^2((0,1);\mathbb{R}^n)} + \|\tilde{A}(0)v(\cdot)\|^2_{L^2((0,1);\mathbb{R}^n)}$.

$$u(x,0) = u_0(x), \qquad \text{for}\, x \in (0,1). \tag{2.26c}$$

By retaining Assumption 2.2.1, the IBVP may be recast in abstract form as previously, but this time with the operator $(A, D(A))$, $A(t) : D(A) \mapsto L^2((0,1);\mathbb{R}^n)$, defined as

$$(Av)(x,t) \triangleq -a(x,t)\frac{\partial v(x)}{\partial x} + C(t)v(1), \tag{2.27a}$$

$$D(A) \triangleq \left\{ v \in H^1((0,1);\mathbb{R}^n) \;\middle|\; v(0) = 0 \right\}. \tag{2.27b}$$

The operator $(A, D(A))$ defined as in Eq. (2.27) includes the trace term, which makes the analysis slightly more complicated compared to that conducted in Sect. 2.2.1. The procedure to prove the well-posedness of the IBVP (2.26) is however analogous, and relies on showing closedness and quasi-dissipativity properties for the operator $(A, D(A))$. To this end, the additional Assumption 2.2.2 is introduced.

Assumption 2.2.2. *The matrix $\Sigma \in C^1([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$ defined as*

$$\Sigma(t) \triangleq I_n - \int_0^1 \frac{C(t)}{a(x,t)}\,\mathrm{d}x \tag{2.28}$$

is invertible for all $t \in [0,T]$, i.e., $\Sigma \in C^1([0,T];\mathbf{GL}_n(\mathbb{R}))$.

Assumption 2.2.2 above ensures the invertibility of the operator $(A, D(A))$, which facilitates the proof of Lemma 2.2.3 and Proposition 2.2.2 below. In this context, it is worth clarifying that the invertibility condition imposed on the matrix $\Sigma(t)$ appearing in Eq. (2.28) is not indispensable to derive existence and uniqueness results for the considered IBVPs; however, it usually holds for a vast majority of the rolling contact problems encountered in the literature, which legitimates its introduction in the present thesis.

Lemma 2.2.3 (Closedness)**.** *Under Assumption 2.2.2, the operator $(A, D(A))$ as defined in Eqs. (2.27) is closed.*

Proof. To prove that $(A, D(A))$ is closed, it suffices again to show that there exists $A^{-1}(t)$ for all $t \in [0,T]$. By setting

$$(Av)(x,t) = -a(x,t)\frac{\partial v(x)}{\partial x} + C(t)v(1) = w(x), \tag{2.29}$$

it may be deduced that

$$v(x) = \int_0^x \frac{C(t)}{a\left(x',t\right)}\,\mathrm{d}x'\,v(1) - \int_0^x \frac{w\left(x'\right)}{a\left(x',t\right)}\,\mathrm{d}x'. \tag{2.30}$$

Computing $v(1)$ in turn yields

$$v(x) = (A^{-1}w)(x,t) = -\tilde{\Sigma}(x,t)\int_0^1 \frac{w(x)}{a(x,t)}\,\mathrm{d}x - \int_0^x \frac{w\left(x'\right)}{a\left(x',t\right)}\,\mathrm{d}x', \tag{2.31}$$

where

$$\tilde{\Sigma}(x,t) \triangleq \int_0^x \frac{C(t)}{a\left(x',t\right)}\,\mathrm{d}x'\,\Sigma^{-1}(t), \tag{2.32}$$

with $\Sigma(t)$ defined as in Eq. (2.28). It is easy to verify that $A^{-1}(t) \in \mathscr{L}(L^2((0,1);\mathbb{R}^n);H^1((0,1);\mathbb{R}^n))$ and that $A^{-1}(t)A(t) = I_{D(A)}$ and $A(t)A^{-1}(t) = I_{L^2((0,1);\mathbb{R}^n)}$. Hence, $(A, D(A))$ is closed. □

Proposition 2.2.2 below allows deducing an expression for the adjoint operator.

Proposition 2.2.2 (Adjoint operator)**.** *The adjoint operator* $(A^*, D(A^*))$, $A^*(t) : D(A^*) \mapsto L^2((0,1);\mathbb{R}^n)$ *of the operator* $(A, D(A))$ *defined in Eqs.* (2.27) *is given by*

$$(A^*v)(x,t) = \frac{\partial}{\partial x}\Big(a(x,t)v(x)\Big), \tag{2.33a}$$

$$D(A^*) = \left\{ v \in H^1((0,1);\mathbb{R}^n) \;\middle|\; v(1) = \frac{C^{\mathrm{T}}(t)}{a(1,t)}\int_0^1 v(x)\,\mathrm{d}x \right\}. \tag{2.33b}$$

Proof. Since $(A, D(A))$ admits a bounded inverse according to the above Lemma 2.2.3, it is sufficient to deduce an expression for $(A^{-1}(t))^* = (A^*(t))^{-1}$ (see, e.g., Lemma A.3.72 in [48]). From Eq. (2.31),

$$(A^{-1}v)(x,t) = -\tilde{\Sigma}(x,t)\int_0^1 \frac{v(x)}{a(x,t)}\,\mathrm{d}x - \int_0^x \frac{v\big(x'\big)}{a\big(x',t\big)}\,\mathrm{d}x'. \tag{2.34}$$

Therefore, an application of Fubini's Theorem yields

$$\begin{aligned}\Big\langle A^{-1}(t)v, w\Big\rangle_{L^2((0,1);\mathbb{R}^n)} &= -\int_0^1 \frac{v^{\mathrm{T}}(x)}{a(x,t)}\,\mathrm{d}x\int_0^1 \tilde{\Sigma}^{\mathrm{T}}(x,t)w(x)\,\mathrm{d}x - \int_0^1\int_0^x \frac{v^{\mathrm{T}}\big(x'\big)}{a\big(x',t\big)}w(x)\,\mathrm{d}x'\,\mathrm{d}x \\ &= -\int_0^1 \frac{v^{\mathrm{T}}(x)}{a(x,t)}\,\mathrm{d}x\int_0^1 \tilde{\Sigma}^{\mathrm{T}}(x,t)w(x)\,\mathrm{d}x - \int_0^1 \frac{v^{\mathrm{T}}\big(x'\big)}{a\big(x',t\big)}\int_{x'}^1 w(x)\,\mathrm{d}x\,\mathrm{d}x' \\ &= \Big\langle v, (A^{-1}(t))^*w\Big\rangle_{L^2((0,1);\mathbb{R}^n)},\end{aligned} \tag{2.35}$$

which implies

$$\Big((A^{-1})^*v\Big)(x,t) = \Big((A^*)^{-1}v\Big)(x,t) = -\frac{1}{a(x,t)}\int_0^1 \tilde{\Sigma}^{\mathrm{T}}(x,t)v(x)\,\mathrm{d}x - \frac{1}{a(x,t)}\int_x^1 v\big(x'\big)\,\mathrm{d}x'. \tag{2.36}$$

Using Eqs. (2.33a) and (2.36), it may be verified that $A^*(t)(A^*(t))^{-1} = I_{L^2((0,1);\mathbb{R}^n)}$. Moreover, substituting for (2.33a) into (2.36) and recalling the definition of $\tilde{\Sigma}(x,t)$ as in Eq. (2.32) yields, after some manipulations,

$$\begin{aligned}\Big((A^*)^{-1}A^*v\Big)(x,t) &= -\frac{1}{a(x,t)}\Big(\Sigma^{-1}(t)\Big)^{\mathrm{T}}\left(\int_0^1 \frac{C^{\mathrm{T}}(t)}{a(x,t)}\,\mathrm{d}x - I_n\right)a(1,t)v(1) \\ &\quad - \frac{1}{a(x,t)}\Big(\Sigma^{-1}(t)\Big)^{\mathrm{T}}\left[a(1,t)v(1) - C^{\mathrm{T}}(t)\int_0^1 v(x)\,\mathrm{d}x\right] - \frac{a(1,t)}{a(x,t)}v(1) + v(x).\end{aligned} \tag{2.37}$$

With $D(A^*)$ and $\Sigma(t)$ reading according to Eqs. (2.33b) and (2.28), respectively, this finally provides $(A^*(t))^{-1}A^*(t) = I_{D(A^*)}$, thus concluding the proof. □

Lemma 2.2.4 (Quasi-dissipativity)**.** *The operator* $(A,D(A))$ *defined according to Eqs.* (2.27)*, together with its adjoint* $(A^*,D(A^*))$*, is quasi-dissipative with constant* ω *given by*

$$\omega \triangleq \frac{1}{2}\left(\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty + \frac{\sup_{t\in[0,T]}\left\|C(t)\right\|^2}{\inf_{t\in[0,T]} a(1,t)}\right). \tag{2.38}$$

Proof. Considering the operator $(A,D(A))$, taking the inner product on $L^2((0,1);\mathbb{R}^n)$ and integrating by parts yields

$$\begin{aligned}\left\langle A(t)v,v\right\rangle_{L^2((0,1);\mathbb{R}^n)} &= -\int_0^1 a(x,t)\frac{\partial v^{\mathrm{T}}(x)}{\partial x}v(x)\,\mathrm{d}x + v^{\mathrm{T}}(1)C^{\mathrm{T}}(t)\int_0^1 v(x)\,\mathrm{d}x \\ &= -\frac{1}{2}\int_0^1 a(x,t)\frac{\partial}{\partial x}\left\|v(x)\right\|_2^2\mathrm{d}x + v^{\mathrm{T}}(1)C^{\mathrm{T}}(t)\int_0^1 v(x)\,\mathrm{d}x \\ &= -\frac{1}{2}a(1,t)\left\|v(1)\right\|_2^2 + \frac{1}{2}\int_0^1 \frac{\partial a(x,t)}{\partial x}\left\|v(x)\right\|_2^2\mathrm{d}x \\ &\quad + v^{\mathrm{T}}(1)C^{\mathrm{T}}(t)\int_0^1 v(x)\,\mathrm{d}x, \quad \text{for } v\in D(A).\end{aligned} \tag{2.39}$$

Applying Cauchy-Schwarz' and then the generalised form of Young's inequality for products to the last term on the right-hand side of Eq. (2.39) gives

$$\begin{aligned}\left\langle A(t)v,v\right\rangle_{L^2((0,1);\mathbb{R}^n)} &\leq -\frac{1}{2}\left(\inf_{t\in[0,T]} a(1,t) - \frac{1}{\varepsilon}\sup_{t\in[0,T]}\left\|C(t)\right\|^2\right)\left\|v(1)\right\|_2^2 \\ &\quad + \frac{1}{2}\left(\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty + \varepsilon\right)\left\|v(\cdot)\right\|_{L^2((0,1);\mathbb{R}^n)}^2, \quad \text{for } v\in D(A).\end{aligned} \tag{2.40}$$

Therefore, selecting

$$\varepsilon \triangleq \frac{\sup_{t\in[0,T]}\left\|C(t)\right\|^2}{\inf_{t\in[0,T]} a(1,t)} \tag{2.41}$$

leads to $\langle A(t)v,v\rangle_{L^2((0,1);\mathbb{R}^n)} \leq \omega\left\|v(\cdot)\right\|_{L^2((0,1);\mathbb{R}^n)}^2$ with ω defined as in Eq. (2.38).

Additionally, departing from Eqs. (2.33), similar manipulations as previously give

$$\begin{aligned}\left\langle A^*(t)v,v\right\rangle_{L^2((0,1);\mathbb{R}^n)} &= \int_0^1 \frac{\partial}{\partial x}\left(a(x,t)v^{\mathrm{T}}(x)\right)v(x)\,\mathrm{d}x \\ &= a(1,t)\left\|v(1)\right\|_2^2 - a(0,t)\left\|v(0)\right\|_2^2 - \int_0^1 a(x,t)v^{\mathrm{T}}(x)\frac{\partial v(x)}{\partial x}\mathrm{d}x \\ &= \frac{1}{2}a(1,t)\left\|v(1)\right\|_2^2 - \frac{1}{2}a(0,t)\left\|v(0)\right\|_2^2 \\ &\quad + \frac{1}{2}\int_0^1 \frac{\partial a(x,t)}{\partial x}\left\|v(x)\right\|_2^2\mathrm{d}x, \quad \text{for } v\in D(A^*).\end{aligned} \tag{2.42}$$

Using the BC in Eq. (2.33b) and then applying the Cauchy-Schwarz' inequality yields

$$\left\langle A^*(t)v,v\right\rangle_{L^2((0,1);\mathbb{R}^n)} \leq \frac{1}{2}\left(\sup_{t\in[0,T]}\left\|\frac{\partial a(\cdot,t)}{\partial x}\right\|_\infty + \frac{\sup_{t\in[0,T]}\left\|C(t)\right\|^2}{\inf_{t\in[0,T]}a(1,t)}\right)\left\|v(\cdot)\right\|_{L^2((0,1);\mathbb{R}^n)}, \quad \text{for } v\in D(A^*). \tag{2.43}$$

By combining Eqs. (2.40) and (2.43), the result follows. □

The next Theorem 2.2.2 represents the analogous of 2.2.2 for the IBVP (2.26).

Theorem 2.2.2 (Existence and uniqueness)**.** *If Assumptions 2.2.1 and 2.2.2 hold, the IBVP (2.26) admits a unique strict solution* $u \in C^1([0,T];L^2((0,1);\mathbb{R}^n)) \cap C^0([0,T];D(A))$ *as in Eq. (2.11) for all* $f \in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ *and* $u_0 \in D(A)$.

Proof. Since $C_0^1([0,1];\mathbb{R}^n) \subset D(A)$, the operator $(A,D(A))$ as defined in Eqs. (2.27) is dense, i.e., $\overline{D(A)} = L^2((0,1);\mathbb{R}^n)$. Moreover, it is closed and quasi-dissipative together with its adjoint $(A^*,D(A^*))$ according to Lemmata 2.2.3 and 2.2.4. It follows from Lumer-Phillips' Theorem 2.1.1 that, for $t \in [0,T]$, $A(t)$ is the infinitesimal generator of a C_0-semigroup. In particular, $A(t) \in \mathscr{G}(1,\omega)$, with ω as in Eq. (2.38). Therefore, the family $\{A(t)\}_{t\in[0,T]}$ is stable. Since $D(A(t)) = D(A(0))$ is independent of t and, for every $u_0 \in D(A)$, $A(t)u_0$ is continuously differentiable in $L^2((0,1);\mathbb{R}^n)$ by Assumption 2.2.1, Theorem 2.1.4 ensures the existence of a unique strict solution reading as in Eq. (2.11) for all $f \in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ and $u_0 \in D(A)$. □

Concerning the complete formulation with $B(t) \neq 0$, Corollary 2.2.2 finally asserts the well-posedness of the IBVP (2.12).

Corollary 2.2.2 (Existence and uniqueness ($B(t) \neq 0$))**.** *If Assumptions 2.2.1 and 2.2.2 hold, the IBVP (2.12) admits a unique strict solution* $u \in C^1([0,T];L^2((0,1);\mathbb{R}^n)) \cap C^0([0,T];D(A))$ *for all* $f \in C^1([0,T];L^2((0,1);\mathbb{R}^n))$ *and* $u_0 \in D(A)$.

Proof. The proof is formally indentical to that of Corollary 2.2.1. □

Theorem 2.2.2 and Corollary 2.2.2 conclude the analysis of IBVPs in one space dimension.

2.2.2 Systems in several space dimensions

In the context of contact mechanics studies, transport equations in serveral space dimensions describe the dynamics of rolling and slipping bodies, including tyres, railway wheels, and elastic spheres [14, 15, 24].

More specifically, concerning systems in several space dimensions, the following structure for the IBVP is considered in this thesis:

$$\frac{\partial u(x,t)}{\partial t} + \Big(a(x,t)\cdot\nabla\Big)u(x,t) = B(t)u(x,t) + f(x,t), \qquad \text{for } (x,t) \in \mathit{\Omega}\times(0,T), \tag{2.44a}$$

$$u(x,t) = 0, \qquad \text{for } (x,t) \in \mathit{\Gamma}_- \times (0,T), \tag{2.44b}$$

$$u(x,0) = u_0(x), \qquad \text{for } x \in \mathit{\Omega}, \tag{2.44c}$$

where $u(x,t) \in \mathbb{R}^n$, $a(x,t) \in \mathbb{R}^d$, $d \geq 2$, represents the transport velocity, $f(x,t) \in \mathbb{R}^n$ is the external forcing term, $B \in C^1([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$, the open set $\Omega \subset \mathbb{R}^d$, represents the spatial domain and

$$\Gamma_+ = \Big\{x \in \Gamma \;\Big|\; a(x,t)\cdot \nu(x) > 0\Big\}, \tag{2.45a}$$

$$\Gamma_0 = \Big\{x \in \Gamma \;\Big|\; a(x,t)\cdot \nu(x) = 0\Big\}, \tag{2.45b}$$

$$\Gamma_- = \Big\{x \in \Gamma \;\Big|\; a(x,t)\cdot \nu(x) < 0\Big\}, \tag{2.45c}$$

being $\nu(x)$ the outward unit normal to Γ, defined almost everywhere on Γ. In the following, the set of points for which $\nu(x)$ is not defined is denoted by $\chi \subset \Gamma$.

The PDE (2.44a) describes a vector-valued transport equation whose scalar components are coupled via the matrix $B \in C^1([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$. It is obvious that the matrix $B(t)$ is the infinitesimal generator of a C_0-semigroup[2], with solution operator denoted by $U_B(t,\tilde{t})$. Therefore, substituting $u(x,t) \triangleq U_B(t,0)w(x,t)$ into Eq. (2.44a) yields

$$U_B(t,0)\left(\frac{\partial w(x,t)}{\partial t} + \Big(a(x,t)\cdot\nabla\Big)w(x,t)\right) = -\left(\frac{\partial U_B(t,0)}{\partial t} - B(t)U_B(t,0)\right)w(x,t) + f(x,t), \quad \text{for } (x,t) \in \Omega\times(0,T). \tag{2.46}$$

Since $U_B(0,0) = I_n$, the following matrix ODE is identically satisfied:

$$\frac{\partial U_B(t,0)}{\partial t} = B(t)U_B(t,0), \qquad \text{for } t \in [0,T]. \tag{2.47}$$

Additionally, by observing that the solution operator is invertible, i.e., $C^1([0,T];\mathbf{GL}_n(\mathbb{R})) \ni U_B^{-1}(t,0) \equiv U_B(0,t)$ exists for every $t \in [0,T]$, the original system described by Eqs. (2.44) may be recast in the following equivalent form:

$$\frac{\partial w(x,t)}{\partial t} + \Big(a(x,t)\cdot\nabla\Big)w(x,t) = g(x,t), \qquad \text{for } (x,t) \in \Omega\times(0,T), \tag{2.48a}$$

$$w(x,t) = 0, \qquad \text{for } (x,t) \in \Gamma_-\times(0,T), \tag{2.48b}$$

$$w(x,0) = w_0(x) = u_0(x), \qquad \text{for } x \in \Omega, \tag{2.48c}$$

being $g(x,t) \triangleq U_B(0,t)f(x,t)$. The technique outlined above may be conveniently applied to the rolling contact poblems considered, e.g., in [14, 15], which are or may be recast[3] in the same form as that described by Eqs. (2.44).

Concerning the study of the IBVPs (2.44) and (2.48), a common problem is that, when the velocity field $a(x,t)$ is time-dependent, the sets defined according to Eqs. (2.45) may also vary over time, that is, $\Gamma_+ = \Gamma_+(t)$, $\Gamma_0 = \Gamma_0(t)$ and $\Gamma_- = \Gamma_-(t)$. Moreover, the product $a(x,t)\cdot\nu(x)$ vanishes for all the points of the boundary belonging to Γ_- (characteristic condition). In Sect. 2.2.1, this possibility was conveniently excluded by introducing Assumption 2.2.1, which in turn ensured the domain of the considered unbounded operator $(A, D(A))$ to be also time-independent, i.e., $D(A(t)) = D(A(0))$. For general domains $\Omega \subset \mathbb{R}^d$, there is no simple

[2] This obviously remains true if $B \in C^0([0,T];\mathbf{M}_{n\times n}(\mathbb{R}))$.

[3] For example, IBVPs evolving on time-varying domains, such as those considered in [14, 15], may be restated according to Eqs. (2.44).

characterisation guaranteeing that the sets in Eqs. (2.45) are time-independent, whilst also ensuring that the noncharacteristic condition is never violated on the boundary. Albeit the well-posedness for transport equations of the type (2.44) and (2.48) may still be proved under very general assumptions, such results rely on different arguments than those presented so far (see the discussion in Sect. 2.3.1).

Therefore, mainly for the purpose of coherence, a simplified version of the IBVP (2.48), where the vector field $a(x,t) = a(x)$ is assumed to be time-independent, is considered in the following. Returning to the standard notation for a matter of convenience, the corresponding IBVP is stated as

$$\frac{\partial u(x,t)}{\partial t} + \Big(a(x)\cdot\nabla\Big)u(x,t) = f(x,t), \qquad \text{for}\,(x,t) \in \Omega\times(0,T), \tag{2.49a}$$

$$u(x,t) = 0, \qquad \text{for}\,(x,t) \in \Gamma_{-}\times(0,T), \tag{2.49b}$$

$$u(x,0) = u_0(x), \qquad \text{for}\,x \in \Omega. \tag{2.49c}$$

The motivation for considering the simplified IBVP (2.49) resides in that the semigroup theory may be more easily applied, since the sets defined in Eqs. (2.45) do not depend on time. In particular, the analysis proposed in this thesis relies on a classical result obtained by Bardos [50], according to which the IBVP (2.49) may be recast as follows in the abstract setting:

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = Au(t) + f(t), \qquad \text{for}\,t \in (0,T), \tag{2.50a}$$

$$u(0) = u_0, \tag{2.50b}$$

where the unbounded operator $(A,D(A))$, $A: D(A) \mapsto L^2(\Omega;\mathbb{R}^n)$, is defined as

$$(Av)(x) \triangleq -\Big(a(x)\cdot\nabla\Big)v(x), \tag{2.51a}$$

$$D(A) \triangleq \left\{ v \in L^2(\Omega;\mathbb{R}^n) \;\middle|\; (a\cdot\nabla)v \in L^2(\Omega;\mathbb{R}^n), \text{ and } v|_{\Gamma_{-}} = 0 \right\}. \tag{2.51b}$$

The characterisation of the unbounded operator $(A,D(A))$ defined according to Eqs. (2.51) relies on the two following Assumptions 2.2.3 and 2.2.4. Assumption 2.2.3, in particular, is the counterpart of Assumption 2.2.1 introduced in Sect. 2.2.1, proposed here in its original formulation due to Bardos [50].

Assumption 2.2.3. *The field $A = a(x)\cdot\nabla$ can be extended to a field $\bar{A} = \bar{a}(x)\cdot\nabla$ defined on an open set $\Omega^* \subset \mathbb{R}^d$ such that $\overline{\Omega} \subset \Omega^*$, with $\bar{a} \in C^1(\Omega^*;\mathbb{R}^d)$ bounded on Ω^* together with its derivatives.*

Assumption 2.2.4. *The boundary Γ of the domain Ω is piecewise C^1.*

Theorem 2.2.3 asserts the well-posedness for the equivalent IBVP (2.49), and consequently also for the original formulation (2.44). Its proof is only sketched and the reader is redirected to [50] for further details.

Theorem 2.2.3 (Existence and uniqueness (Bardos [50]))**.** *If Assumptions 2.2.3 and 2.2.4 hold, the IBVP (2.49) admits a unique strict solution $u \in C^1([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];D(A))$ for all $f \in C^1([0,T];L^2(\Omega;\mathbb{R}^n))$ and $u_0 \in D(A)$.*

Sketch of the proof. Since, under Assumptions 2.2.3 and 2.2.4, the trajectories of A intersecting $\Gamma_0 \cup \chi$ have zero measure, it can be shown that A is the infinitesimal generator of a C_0-semigroup (see [50]). In particular, $A \in \mathscr{G}(1,\omega)$, with $\omega \triangleq \frac{1}{2}\|\nabla \cdot a(\cdot)\|_\infty$. Therefore, for every $f \in C^1([0,T];L^2(\Omega;\mathbb{R}^n))$ and $u_0 \in D(A)$, the IBVP (2.49) has a unique strict solution in the form (2.11) ensured by Theorem 2.1.4. □

Mild solutions according to Definition 2.1.5 are easily obtained directly from Eq. (2.11).

Before concluding the analysis, it is perhaps worth mentioning that the result given in [50] (for scalar equations) is more general than that presented above, and extends to other $L^p(\Omega;\mathbb{R}^n)$, $1 \leq p < \infty$, spaces than $L^2(\Omega;\mathbb{R}^n)$. Moreover, the reader may easily realise that the extension of Theorem 2.2.3 to more general IBVPs, accounting for, e.g., the presence integral terms, may be worked out again within the mathematical framework provided by the perturbation theory of linear unbounded operators.

2.3 Additional results and remarks

The present Section contains some additional results and remarks on hyperbolic evolution equations that have not been covered in the main part of the Chapter.

2.3.1 A more general result for transport equations in several space dimensions

Concerning systems in several space dimensions, the well-posedness result obtained in Sect. 2.2.2 was restricted to the case of a time-independent vector field $a(x,t) = a(x)$. This limitation was mainly motivated by the fact that Theorem 2.1.4 requires the domain of the operator to also be independent of time. Such an assumption is actually standard in the mathematical literature dealing with hyperbolic PDEs. For example, within the semigroup framework, Massey has addressed the existence and uniqueness of regular solutions for systems of hyperbolic type by assuming that the product $a(x,t)\cdot\nu(x)$ never vanishes [41]: Considering instead other approaches, similar assumptions are introduced in the seminal works by Kreiss (see [51, 52] and references therein), and retained in the elegant monograph recently authored by Benzoni-Gavage and Serre [54].

However, in a very general setting, when, e.g., $B \in C^0([0,T];\mathbf{M}_{n\times n}(\mathbb{R}^n))$, existence and uniqueness for the IBVP (2.44) may be established by invoking the following Theorem 2.3.1, which targets the equivalent system (2.48).

Theorem 2.3.1 (Existence and uniqueness (Crippa, Donadello and Spinolo [53]))**.** *Let $\Omega \subseteq \mathbb{R}^d$ be an open set with uniformly Lipschitz boundary. Assume that the vector field $a \in L^\infty(\Omega \times (0,T);\mathbb{R}^d)$ satisfies $\nabla \cdot a \in L^\infty(\Omega \times (0,T))$ and, for every open and bounded set $\Omega_* \subseteq \Omega$, $a \in L^1_{\text{loc}}([0,T);\mathrm{BV}(\Omega_*;\mathbb{R}^d))$. If $g \in L^\infty(\Omega \times (0,T);\mathbb{R}^n)$, then the IBVP described by Eqs. (2.48) admits a unique weak or generalised solution $w \in L^\infty(\Omega \times (0,T);\mathbb{R}^n)$ for all ICs $w_0 \in L^\infty(\Omega)$.*

Proof. See [53]. □

Indeed, from Theorem 2.3.1 above, it is clear that the existence and uniqueness for the original IBVP (2.44) follow directly from the continuously invertible transformation $u(x,t) = U_B(t,0)w(x,t)$ introduced in Sect. 2.2.2. The main arguments used in the proof Theorem 2.3.1, which is omitted in this thesis for the sake of both brevity and consistency, build upon the famous renormalisation theory developed by Di Perna and Lions, and fall outside the semigroup framework outlined so far.

It is also worth observing that IBVPs on time-varying domains, such as those considered in [14, 15, 33], may be converted into IBVPs of the same type as that described by Eqs. (2.44) by performing a change of coordinates mapping the original domain into a fixed one, according to the methodologies outlined, e.g., in [33].

2.3.2 Existence of smoother solutions

Theorem 2.1.4 asserts the existence and uniqueness of strict solutions, which enjoy a certain type of regularity. To proceed with the error analysis in the context of DGMs, however, smoother solutions may need to be considered. In this context, the existence and uniqueness of more regular solutions may be proved by applying Theorem 2.1.4 iteratively. As an example, the one-dimensional IBVP described by Eqs. (2.12) may be considered. Defining $v(x,t) \triangleq \partial u(x,t)/\partial x$, the following IBVP may be deduced governing the time evolution of the new variable $v(x,t)$:

$$\frac{\partial v(x,t)}{\partial t} + a(x,t)\frac{\partial v(x,t)}{\partial x} = D(t)v(x,t) + g(x,t), \qquad \text{for}\,(x,t) \in (0,1)\times(0,T), \tag{2.52a}$$

$$v(0,t) = h(t), \qquad \text{for}\, t \in (0,T), \tag{2.52b}$$

$$v(x,0) = v_0(x) \triangleq \frac{\partial u_0(x)}{\partial x}, \qquad \text{for}\, x \in (0,1), \tag{2.52c}$$

where

$$(Dv)(x,t) \triangleq (Bv)(x,t) - \frac{\partial a(x,t)}{\partial x}v(x,t), \tag{2.53a}$$

$$g(x,t) \triangleq \frac{\partial f(x,t)}{\partial x}, \tag{2.53b}$$

and the BC $h(t)$ in Eq. (2.52b) may be obtained by differentiating Eq. (2.12b) with respect to the time variable and then inverting for $\partial u(0,t)/\partial t$ from Eq. (2.12a), yielding

$$h(t) = \frac{1}{a(0,t)}\Big(B(t)u(0,t) + C(t)u(1,t) + f(0,t)\Big), \tag{2.54}$$

with $a(0,t)$ being invertible owing to Assumption 2.2.1.

The problem described by Eqs. (2.52) may again be reformulated in an abstract setting according to Eqs. (2.7), clearly with the corresponding unbounded operator $(A, D(A))$ to be properly defined. It is clear, however, that, under opportune assumptions on the data, the domain of such an operator would still be time-independent, i.e., $D(A(t)) = D(A(0))$, and that moreover $D(A) \subset H^1((0,1);\mathbb{R}^n)$, thus delivering at least $H^1((0,1);\mathbb{R}^n)$-regularity for the space derivative $v(x,t)$, which in turn would imply $H^2((0,1);\mathbb{R}^n)$-regularity for the original variable.

Analogously, by defining $w \triangleq \partial u(x,t)/\partial t$, the following IBVP may be deduced:

$$\frac{\partial w(x,t)}{\partial t} + a(x,t)\frac{\partial w(x,t)}{\partial x} = B(t)w(x,t) + C(t)w(1,t) + g(x,t), \quad \text{for}\,(x,t) \in (0,1)\times(0,T), \tag{2.55a}$$

$$w(0,t) = 0, \qquad \text{for}\, t \in (0,T), \tag{2.55b}$$

$$w(x,0) = w_0(x), \qquad \text{for}\, x \in (0,1), \tag{2.55c}$$

with

$$g(x,t) \triangleq -\frac{\partial a(x,t)}{\partial t}\frac{\partial u(x,t)}{\partial x} + \frac{\mathrm{d}C(t)}{\mathrm{d}t}u(x,t) + \frac{\mathrm{d}B(t)}{\mathrm{d}t}u(1,t) + \frac{\partial f(x,t)}{\partial t}, \tag{2.56}$$

and where this time the IC $w_0(x)$ in Eq. (2.55c) may be derived directly from Eq. (2.12). In particular, it is important to observe that, concerning Eqs. (2.55), the unbounded operator $(A, D(A))$ appearing in the abstract IBVP (2.7) may be defined exactly as for the original evolution system governed by Eqs. (2.12). This remains clearly true for time derivatives of any order.

The problem becomes rapidly more complicated in several space dimensions, where typically the difficulty arises from the need of inverting the boundary matrix to derive a similar BC to that in Eq. (2.54). In fact, such a matrix might be unbounded, which often limits the analysis to problems where the product $a(x,t)\cdot\nu(x)$ cannot vanish. On this matter, the reader is redirected to, e.g., [41, 51, 52, 54].

Chapter 3

Space semi-discretisation: discontinuous Galerkin finite element methods (DGMs)

This Chapter is devoted to discussing the finite element modelling of hyperbolic IBVPs. More specifically, Sect. 3.1 provides an outline of the DGMs employed for space semi-discretisation, introducing the main concepts that are needed to build a mesh of the space domain, and the salient properties of the functional spaces used on each mesh element. Useful inequalities required to derive suitable error estimates are also recalled. Then, the first part of Sect. 3.2 is dedicated to the characterisation of the discrete operator that is chosen to approximate the time evolution of the considered hyperbolic IBVPs. The stability of the space semi-discrete problem, along with the convergence of the proposed DGMs, are instead analysed in the second part of Sect. 3.2. Most of the material presented in in the following is adapted, with opportune simplifications, from [39].

3.1 Discontinuous Galerkin finite element methods (DGMs)

The present Section illustrates the main concepts and assumptions that are required to construct finite-dimensional approximations of hyperbolic evolution equations within the (DGM) framework. The proposed method relies on the spatial discretisation of the considered domain Ω, using a mesh and choosing an appropriate local polynomial behaviour within each mesh element.

More specifically, in the remainder of the Chapter, the following hyperbolic IBVP is considered, which generalises those examined in Chap. 1:

$$\frac{\partial u(x,t)}{\partial t} + \Big(a(x,t)\cdot\nabla\Big)u(x,t) = B(t)u(x,t) + C(t)\int_{\Gamma} u(x,t)\,\mathrm{d}s + f(x,t), \quad \text{for } (x,t)\in\Omega\times(0,T), \tag{3.1a}$$

$$u(x,t) = 0, \qquad \text{for } (x,t)\in\Gamma_{-}\times(0,T), \tag{3.1b}$$

$$u(x,0) = u_0(x), \qquad \text{for } x\in\Omega, \tag{3.1c}$$

with the data $a(x,t)\in\mathbb{R}^d$, $f(x,t)\in\mathbb{R}^n$, $B(t)\in\mathscr{B}(L^2(\Omega;\mathbb{R}^n))$, and $C(t)\in\mathbf{M}_{n\times n}(\mathbb{R})$ assumed to be sufficiently regular for what follows. In particular, by setting the term $B(t)=0$ for simplicity[1]

[1] As in Chap. 1, when the operator $B(t)$ is explicitly considered, the problem becomes equivalent by replacing $(A,D(A))$ with $(\tilde{A},D(\tilde{A}))$, being $(\tilde{A}v)(x,t)\triangleq(Av)(x,t)+(Bv)(x,t)$.

and consistency of notation with Chap. 1, and assuming $\Gamma_-(t) = \Gamma_-$ to be constant over time, the above hyperbolic IBVP may be recast in abstract form by setting

$$(Av)(x,t) \triangleq -\Big(a(x,t)\cdot\nabla\Big)v(x) + C(t)\int_\Gamma v(x)\,\mathrm{d}s, \tag{3.2a}$$

$$D(A) \triangleq \Big\{ v \in L^2(\Omega;\mathbb{R}^n) \;\Big|\; A(t)v \in L^2(\Omega;\mathbb{R}^n), \text{ and } v|_{\Gamma_-} = 0 \Big\}, \tag{3.2b}$$

where it shall be supposed that $D(A(t)) = D(A(0))$. Such a condition is formalised according to the following Assumption 3.1.1, and is complemented with additional requirements imposed on the domain Ω and on matrix $C(t) \in \mathbf{M}_{n\times n}(\mathbb{R})$.

Assumption 3.1.1. *The following conditions are supposed to hold:*

1. *The domain $D(A(t)) = D(A(0))$ is independent of the time,*
2. *The domain Ω is bounded with compact boundary Γ,*
3. *The matrix $C(t) \in \mathbf{M}_{n\times n}(\mathbb{R}) = 0$ whenever the condition $a(x,t)\cdot\nu(x) = 0$ is satisfied for some $x \in \Gamma$.*

Concerning the characterisation of the the domain $D(A)$, it is worth emphasising that all the IBVPs analysed in Sect. 2.2 may be recast in the form (2.7) with $(A, D(A))$ according to Eqs. (3.2). In particular, with reference to the one-dimensional problems investigated in Sect. 5.2.1, it is clear that Assumptions 2.2.1 and 3.1.1 render the definition of the domain $D(A)$ in Eq. (2.27) equivalent to that in Eqs. (3.2). Finally, the exact solution to the IBVP (2.7), with $(A, D(A))$ defined as in Eqs. (3.2), is supposed to satisfy at least $u \in C^0([0,T];H^1(\Omega;\mathbb{R}^n))$.

Owing to the above premises, the objective consists then in approximating the solution to the continuous IBVP described by Eqs. (2.7), with $(A, D(A))$ defined as in Eq. (3.2), using functions $u_h(t) \in V_h$, and replacing the operator $(A, D(A)))$ with its discrete counterpart $(A_h, V_{h\star})$, where V_h is a finite-dimensional space to be opportunely selected, and $V_{h\star} \triangleq H^1(\Omega;\mathbb{R}^n) + V_h$ needs to be defined accordingly. In particular, the functional spaces V_h considered in this thesis are typically polynomial, which motivates the introduction of broken polynomial spaces as in Sect. 3.1.2. Furthermore, owing to the assumption $u \in C^0([0,T];H^1(\Omega;\mathbb{R}^n))$, it may be realised that the quantity $(u(t) - u_h(t)) \in V_{h\star}$. Consequently, the approximation error may be quantified by evaluating the norm difference $\|u(\cdot,t) - u_h(\cdot,t)\|_{L^2(\Omega;\mathbb{R}^n)}$.

3.1.1 Meshes

The first step to construct finite-dimensional approximations of the hyperbolic IBVPs discussed in Chap. 1 consists in discretising the spatial domain Ω using a mesh. To this end, the notions of *simplex* and *simplex faces* are preliminarly introduced, which are propaedeutic to the definition of *simplicial meshes*, which constitute the main type of mesh considered in this thesis.

Definition 3.1.1 (Simplex [39]). *Given a family $\{y_0,\ldots,y_d\}$ of $d+1$ points in $\mathbb{R}^d$ such that the vectors $\{y_1 - y_0,\ldots,y_d - y_0\}$ are linearly independent, the interior of the convex hull $\{y_0,\ldots,y_d\}$ is called a non-degenerate* simplex *of $\mathbb{R}^d$, and the points $\{y_0,\ldots,y_d\}$ are called its* vertices.

Definition 3.1.2 (Simplex faces [39]). *Let S be a non-degenerate simplex with vertices $\{y_0,\ldots,y_d\}$. For each $i \in \{0,\ldots,d\}$, the convex hull of $\{y_0,\ldots,y_d\}\setminus\{y_i\}$ is called a* face *of the simplex S.*

With the aid of Definitions 3.1.1 and 3.1.2, it is possible to introduce the notion of simplicial mesh, according to the following 3.1.3.

Definition 3.1.3 (Simplicial mesh [39]). *A* simplicial mesh $\mathscr{T}$ *of the domain* Ω *is a finite collection of disjoint non-degenerate simplices* $\mathscr{T} = \{T\}$ *forming a partition of* Ω*, i.e,*

$$\overline{\Omega} = \bigcup_{T \in \mathscr{T}} \overline{T}. \tag{3.3}$$

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

The concept of *meshsize* plays a crucial role in the analyses conducted in the Sect. 3.2.3 and Chap. 4.

Definition 3.1.4 (Element diameter, meshsize). *Let* $\mathscr{T}$ *be a mesh of the domain* Ω*. For all* $T \in \mathscr{T}$*,* h_T *denotes the* element diameter *of* T*, and the* meshsize *is defined as the real number*

$$h \triangleq \max_{T \in \mathscr{T}} h_T. \tag{3.4}$$

In the following, the notation $\mathscr{T}_h$ is used to indicate a mesh $\mathscr{T}$ with meshsize h. Finally, the notion of *element outward unit normal* is introduced, according to Definition 3.1.5.

Definition 3.1.5 (Element outward unit normal). *Let* $\mathscr{T}_h$ *be a mesh of the domain* Ω *and let* $T \in \mathscr{T}_h$*. The symbol* $\nu_{\partial T}(x)$ *denotes the* outward unit normal *to* ∂T*, defined almost everywhere.*

Mesh faces, averages, and jumps

The concepts of *mesh faces*, *averages*, and *jumps* play a crucial role in the design and analysis of DGMs.

Definition 3.1.6 (Mesh faces). *Let* $\mathscr{T}_h$ *be a mesh of the domain* Ω*. A (closed) subset* F *of* $\overline{\Omega}$ *is a* mesh face *if* F *has positive* $(d-1)$*-dimensional Hausdroff measure and if either one of the two following conditions are satisfied:*

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

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

Interfaces are collected in the set $\mathscr{F}_h^{\mathrm{i}}$ and boundary faces in the set $\mathscr{F}_h^{\mathrm{b}}$. Accordingly, faces are collected in the set $\mathscr{F}_h \triangleq \mathscr{F}_h^{\mathrm{i}} \cup \mathscr{F}_h^{\mathrm{b}}$. Moreover, for any mesh element $T \in \mathscr{T}_h$, the set

$$\mathscr{F}_T \triangleq \left\{ F \in \mathscr{F}_h \,\middle|\, F \subset \partial T \right\} \tag{3.5}$$

is defined collecting the mesh faces composing the boundary ∂T of T.

Finally, for any mesh face $F \in \mathscr{F}_h$, the set

$$\mathscr{T}_F \triangleq \left\{ T \in \mathscr{T}_h \,\middle|\, F \subset \partial T \right\}, \tag{3.6}$$

is introduced. It is worth observing that $\mathscr{T}_F$ consists of two mesh elements if $F \in \mathscr{F}_h^{\mathrm{i}}$, and of one mesh element if $F \in \mathscr{F}_h^{\mathrm{b}}$. Next, interface averages and jumps are defined.

Definition 3.1.7 (Interfaces, averages and jumps). *Let $v(x) \in \mathbb{R}$ be defined on Ω, and assume that $v(x)$ is smooth enough to admit, on all $F \in \mathscr{F}_h^{\mathrm{i}}$, a possibly two-valued trace. This means that, for all $T \in \mathscr{T}_h$, the restriction $v(x)|_T$ of $v(x)$ to the open set T can be defined up to the boundary ∂T. Then, for all $F \in \mathscr{F}_h^{\mathrm{i}}$, and almost every $x \in F$, the* average *of $v(x)$ is defined as*

$$\left\{\!\!\left\{ v(x) \right\}\!\!\right\} \triangleq \frac{1}{2}\Big(v(x)|_{T_1} + v(x)|_{T_2} \Big), \tag{3.7}$$

and the jump *as*

$$[\![v(x)]\!] \triangleq v(x)|_{T_1} - v(x)|_{T_2}. \tag{3.8}$$

For vector-valued functions $v(x) \in \mathbb{R}^n$, averages and jumps are defined component-wise.

The last concept to be introduced is that of *face unit normal*, according to the following Definition 3.1.8.

Definition 3.1.8 (Face unit normal). *For all $F \in \mathscr{F}_h$, and almost every $x \in F$, the* unit normal *$\nu_F(x)$ to F at x is defined as*

(i) $\nu_{\partial T_1}(x)$, the unit normal to F at x pointing from T_1 to T_2 if $F \in \mathscr{F}_h^{\mathrm{i}}$, with $F = \partial T_1 \cap \partial T_2$; the orientation of $\nu_F(x)$ is arbitrary depending on the choice of T_1 and T_2, but kept fixed in the following.

(ii) $\nu(x)$, the outward unit normal to Γ at x if $F \in \mathscr{F}_h^{\mathrm{b}}$.

3.1.2 Broken spaces

After constructing a mesh of the space domain Ω, it is necessary to choose a certain functional space within each element. In this thesis, such a choice for the space V_h falls on polynomial functions, which motivates the introduction of *broken polynomial spaces*, and also *broken Hilbert spaces*.

Broken polynomial spaces

Consider $k \in \mathbb{N}_0$ and the definition of A_d^k as

$$A_d^k \triangleq \left\{ \alpha \in \mathbb{N}^d \;\middle|\; |\alpha|_1 \leq k \right\}, \tag{3.9}$$

where $|\cdot|_1$ denotes the standard ℓ^1-norm on $\mathbb{R}^d$. Accordingly, the space of polynomials of d variables, of total degree at most k, is given by

$$\mathbb{P}_d^k \triangleq \left\{ p : \mathbb{R}^d \ni x \mapsto p(x) \in \mathbb{R} \;\middle|\; \exists (\gamma_\alpha)_{\alpha \in A_d^k} \in \mathbb{R}^{\mathrm{card} A_d^k},\, p(x) = \sum_{\alpha \in A_d^k} \gamma_\alpha x^\alpha \right\}, \tag{3.10}$$

with the convention that $x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$ for $x \in \mathbb{R}^d$. The dimension of the vector space is $\dim \mathbb{P}_d^k = \mathrm{card} A_d^k = (k+d)!/(k!d!)$. With $\mathbb{P}_d^k$ defined according to the above Eq. (3.10), it is possible to introduce the broken polynomial space

$$\mathbb{P}_d^k(\mathscr{T}_h) \triangleq \left\{ v \in L^2(\Omega) \;\middle|\; \text{for}\, T \in \mathscr{T}_h,\, v|_T \in \mathbb{P}_d^k(T) \right\}, \tag{3.11}$$

where $\mathbb{P}^k_d(T)$ is spanned by the restriction to T of polynomials in $\mathbb{P}^k_d$. Clearly, it holds that $\dim \mathbb{P}^k_d(\mathscr{T}_h) = \operatorname{card} \mathscr{T}_h \dim \mathbb{P}^k_d$. More generally, in the following, broken polynomial spaces $\mathbb{P}^k_d(T;\mathbb{R}^n)$ and $\mathbb{P}^k_d(\mathscr{T}_h;\mathbb{R}^n)$ are considered, whose characterisation is formally analogous to that provided by Eqs. (3.10) and (3.11).

Broken Hilbert spaces

Consider a mesh $\mathscr{T}_h$ of a domain Ω, and the Hilbert space $H^m(T;\mathbb{R}^n)$. The *broken Hilbert space* is then defined as

$$H^m(\mathscr{T}_h;\mathbb{R}^n) \triangleq \left\{ v \in L^2(\Omega;\mathbb{R}^n) \,\middle|\, \text{for}\, T \in \mathscr{T}_h, v|_T \in H^m(T;\mathbb{R}^n) \right\}, \tag{3.12}$$

where $m \in \mathbb{N}_0$ is an integer. Starting from Eq. (3.12), it comes natural to define a broken gradient operator acting on the broken Hilbert space $H^1(\mathscr{T}_h)$. Such an operator also acts on broken polynomial spaces.

Definition 3.1.9 (Broken gradient)**.** *The* broken gradient $\nabla_h : H^1(\mathscr{T}_h) \mapsto L^2(\Omega;\mathbb{R}^d)$ *is defined such that, for all* $T \in \mathscr{T}_h$*,*

$$\nabla_h v(x)|_T \triangleq \nabla\Big(v(x)|_T\Big), \qquad \text{for}\, v \in H^1(\mathscr{T}_h). \tag{3.13}$$

It is worth observing that the usual Hilbert spaces are subspaces of the corresponding broken counterparts and, in fact, the broken gradient coincides with the distributional one in $H^1(\Omega)$. This result is formalised in the following Lemma 3.1.1.

Lemma 3.1.1 (Broken gradient on $H^1(\Omega)$)**.** *Let* $m \in \mathbb{N}_0$ *and* $1 \leq p \leq \infty$*. There holds* $H^m(\Omega) \subset H^m(\mathscr{T}_h)$*. Moreover, for all* $v \in H^1(\Omega)$*,* $\nabla_h v = \nabla v$ *in* $L^2(\Omega;\mathbb{R}^d)$*.*

Proof. See Lemma 1.22 in [39]. □

An important characterisation of the Hilbert space $H^1(\Omega;\mathbb{R}^n)$ concerns jumps across interfaces: whereas functions in $H^1(\mathscr{T}_h;\mathbb{R}^n)$ are allowed to have discontinuities across interfaces, these jumps are always zero for functions in the usual Hilbert space $H^1(\Omega;\mathbb{R}^n)$.

Theorem 3.1.1 (Characterisation of $H^1(\Omega;\mathbb{R}^n)$)**.** *Let* $1 \leq p \leq \infty$*. A function* $v \in H^1(\mathscr{T}_h;\mathbb{R}^n)$ *belongs to* $H^1(\Omega;\mathbb{R}^n)$ *if and only if, for each* $F \in \mathscr{F}^{\mathrm{i}}_h$*, the following identity is satisfied:*

$$[\![v(x)]\!] = 0. \tag{3.14}$$

Proof. See Lemma 1.23 in [39]. □

3.1.3 Admissible mesh sequences

The objective of the present Section is to present some technical results to analyse the convergence of DGMs as the meshsize goes to zero. In particular, a simplicial mesh sequence $\mathscr{T}_{\mathscr{H}} \triangleq \{\mathscr{T}_h\}_{h\in\mathscr{H}}$, where $\mathscr{H}$ is a countable subset of $\mathbb{R}_{>0}$, having 0 as the only accumulation point, is considered in the following. The main tools recalled here are inverse and trace inequalities that allow asserting discrete stability and boundedness uniformly in h, as well as optimal polynomial approximation properties.

In this context, the first useful concept to be introduced is that of *matching simplicial mesh*. This notion is formalised according to Definition 3.1.10 below.

Definition 3.1.10 (Matching simplicial mesh)**.** *A mesh $\mathscr{T}_h$ is said to be a* matching simplicial mesh *if it is a simplicial mesh and, for any $T \in \mathscr{T}_h$ with vertices $\{y_0, \dots, y_d\}$, the set $\partial T \cap \partial T'$, for all $T' \in \mathscr{T}_h$ such that $T' \neq T$, is the convex hull of a (possibly empty) subset of $\{y_0, \dots, y_d\}$.*

Departing from Definition 3.1.10, the notion of *shape and contact regularity* for simplicial mesh sequences is given according to the following 3.1.11.

Definition 3.1.11 (Shape and contact regularity)**.** *A mesh sequence $\mathscr{T}_{\mathscr{H}}$ of matching simplicial meshes $\mathscr{T}_h$ is said to be* contact and shape regular *if there exists $\rho \in \mathbb{R}_{>0}$ independent of h such that, for all $T' \in \mathscr{T}_h$, $\rho h_{T'} \leq r_{T'}$, where $h_{T'}$ is the diameter of T' and $r_{T'}$ is the radius of the largest ball inscribed in T'.*

Inverse and trace inequalities

Inverse and trace inequalities are indispensable mathematical tools to analyse the convergence of DGMs. In the following, such inequalities are presented concerning the broken polynomial spaces $\mathbb{P}^k_d(T;\mathbb{R}^n)$ defined in Sect. 3.1.2. The first relationship presented below is the inverse trace inequality, which provides a local upper bound on the gradient of discrete functions, according to Lemma 3.1.2.

Lemma 3.1.2 (Inverse trace inequality)**.** *Let $\mathscr{T}_{\mathscr{H}}$ be a shape and contact-regular mesh sequence with parameters ρ. Then, for all $h \in \mathscr{H}$ and $T \in \mathscr{T}_{\mathscr{H}}$, the following estimate holds:*

$$\left\|\nabla v_h(\cdot)\right\|_{L^2(T;\mathbb{R}^{n\times d})} \leq \frac{C_{\mathrm{inv}}}{h_T}\left\|v_h(\cdot)\right\|_{L^2(T;\mathbb{R}^n)}, \qquad \text{for } v_h \in \mathbb{P}^k_d(\mathscr{T}_{\mathscr{H}};\mathbb{R}^n), \tag{3.15}$$

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

Proof. See Lemma 1.44 in [39]. □

As already mentioned, the inverse trace inequality in Eq. (3.15) holds locally on each mesh element. Consequently, it depends upon the shape of the mesh elements, but not on the way these come into contact.

The next relationship is the discrete trace inequality, which yields an upper bound on the face values of discrete functions. The result is formalised in Lemma 3.1.3.

Lemma 3.1.3 (Discrete trace inequality)**.** *Let $\mathscr{T}_{\mathscr{H}}$ be a shape and contact-regular mesh sequence with parameters ρ. Then, for all $h \in \mathscr{H}$ and $T \in \mathscr{T}_{\mathscr{H}}$, the following estimate holds:*

$$\left\|v_h(\cdot)\right\|_{L^2(F;\mathbb{R}^n)} \leq \frac{C_{\mathrm{tr}}}{\sqrt{h_T}}\left\|v_h(\cdot)\right\|_{L^2(T;\mathbb{R}^n)}, \qquad \text{for } v_h \in \mathbb{P}^k_d(\mathscr{T}_{\mathscr{H}};\mathbb{R}^n), \tag{3.16}$$

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

Proof. See Lemma 1.46 in [39]. □

Polynomial approximation

Definition 3.1.12 (Optimal polynomial approximation)**.** *A mesh sequence $\mathscr{T}_{\mathscr{H}}$ is said to enjoy* optimal polynomial approximation *properties if, for all $h \in \mathscr{H}$, all $T \in \mathscr{T}_h$, and all polynomial degree k, there is a linear interpolation operator $\mathscr{I}^k_T : L^2(T;\mathbb{R}^n) \mapsto \mathbb{P}^k_d(T;\mathbb{R}^n)$ such that, for all $s \in \{0, \dots, k+1\}$ and all $v \in H^s(T;\mathbb{R}^n)$, there holds*

$$\left|v(\cdot) - \mathscr{I}^k_T v(\cdot)\right|_{H^m(T;\mathbb{R}^n)} \leq C_{\mathrm{p}} h_T^{s-m}\left|v(\cdot)\right|_{H^s(T;\mathbb{R}^n)}, \qquad \text{for } m \in \{0, \dots, s\}, \tag{3.17}$$

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

As for the inverse trace inequality (3.15), the optimal polynomial approximation property holds locally on mesh elements and is also affected by the shape of such elements, but not on the way they come into contact. Owing to Definition 3.1.12, it is finally possible to enounce the concept of *admissible mesh sequence*, as stated in 3.1.13.

Definition 3.1.13 (Admissible mesh sequence)**.** *A mesh sequence $\mathscr{T}_{\mathscr{H}}$ is* admissible *if it is shape and contact regular and enjoys optimal polynomial approximation properties.*

In this thesis, the L^2-orthogonal projection $\pi_h : L^2(\Omega;\mathbb{R}^n) \mapsto \mathbb{P}_d^k(\mathscr{T}_h;\mathbb{R}^n)$ onto the broken polynomial space $\mathbb{P}_d^k(\mathscr{T}_h;\mathbb{R}^n)$ is often considered because of its simplicity, which delivers other two important technical results, according to the subsequent Lemmata 3.1.4 and 3.1.5.

Lemma 3.1.4 (Optimality of L^2-orthogonal projection)**.** *Let $\mathscr{T}_{\mathscr{H}}$ be an admissible mesh sequence. Let $\pi_h : L^2(\Omega;\mathbb{R}^n) \mapsto \mathbb{P}_d^k(\mathscr{T}_h;\mathbb{R}^n)$ be the L^2-orthogonal projection onto $\mathbb{P}_d^k(\mathscr{T}_h;\mathbb{R}^n)$. Then, for all $s \in \{0,\dots,k+1\}$ and all $m \in \{0,\dots,s\}$, it holds that*

$$\left|v(\cdot) - \pi_h v(\cdot)\right|_{H^m(T;\mathbb{R}^n)} \le C_{\mathrm{p}}' h_T^{s-m} \left|v(\cdot)\right|_{H^s(T;\mathbb{R}^n)}, \qquad \text{for } v \in H^s(T;\mathbb{R}^n), \tag{3.18}$$

where C_{p}' is independent of both T and h.

Proof. See Lemma 1.58 in [39]. □

Lemma 3.1.5 (Polynomial approximation on mesh faces)**.** *Under the same hypotheses of Lemma 3.1.4, assume additionally that $s \ge 1$. Then, for all $h \in \mathscr{H}$, all $T \in \mathscr{T}_h$, and all $F \in \mathscr{F}_T$, the following inequalities hold:*

$$\left\|v(\cdot) - \pi_h v(\cdot)\right\|_{L^2(F;\mathbb{R}^n)} \le C_{\mathrm{p}}'' h_T^{s-1/2} \left|v(\cdot)\right|_{H^s(T;\mathbb{R}^n)}, \qquad \text{for } s \ge 1, \tag{3.19a}$$

$$\sum_{i=1}^{n} \left\| \nabla \left(v_i(\cdot) - \pi_h v_i(\cdot) \right)\Big|_T \cdot \nu_{\partial T}(\cdot) \right\|_{L^2(F)} \le C_{\mathrm{p}}''' h_T^{s-3/2} \left|v(\cdot)\right|_{H^s(T;\mathbb{R}^n)}, \qquad \text{for } s \ge 2, \tag{3.19b}$$

where C_{p}'' and C_{p}''' are independent of both T and h.

Proof. The result is a direct consequence of the continuous trace inequality. □

3.2 Space semi-discretisation of linear hyperbolic IBVPs

Building upon the concepts developed above, the present Section deals with the space semi-discretisation of the IBVP described by Eqs. (2.7) and (3.2), by defining the discrete operator $(A_h, V_{h\star})$ and providing its characterisation. The mathematical properties of the operator $(A_h, V_{h\star})$ are then exploited to show stability and convergence of the space semi-discrete DGMs.

To this end, some preliminary assumptions are introduced in the following, together with the definition of appropriate norms and seminorms.

3.2.1 Preliminaries and assumptions

Here, the main assumptions concerning the regularity of the mesh and meshsize are enounced. Moreover, different norms and seminorms are introduced that are needed for the analyses conducted in Sect. 3.2.3 and later on in Chap. 4.

Assumptions and notation

As already mentioned, in this thesis, space semi-discretisation is achieved using an upwind DGM, which is particularly suited to treat hyperbolic IBVPs. In this context, the discrete space is assumed to coincide with the broken polynomial spaces discussed in Sect. 3.1.2, that is, $V_h = \mathbb{P}^k_d(\mathscr{T}_h;\mathbb{R}^n)$, defined with polynomial degree $k \in \mathbb{N}_0$ and with $\mathscr{T}_h$ belonging to an admissible mesh sequence. In particular, quasi-uniform mesh sequences are considered for simplicity, which essentially means that, for all $h \in \mathscr{H}$,

$$\max_{T\in\mathscr{T}_h} h_T \leq C \min_{T\in\mathscr{T}_h} h_T. \tag{3.20}$$

Moreover, the following reference quantities are introduced:

$$\frac{1}{t_\mathrm{c}} \triangleq \max\left\{ \sup_{t\in[0,T]} \left\|\nabla\cdot a(\cdot,t)\right\|_\infty, \varepsilon_h \right\}, \qquad \text{and} \qquad \eta_\mathrm{c} \triangleq \sup_{t\in[0,T]} \left\|a(\cdot,t)\right\|_\infty, \tag{3.21a}$$

with $\varepsilon_h \in \mathbb{R}_{\geq 0}$ satisfying

$$\varepsilon_h \triangleq \begin{cases} 2\psi_h \dfrac{\sup_{t\in[0,T]}\left\|C(t)\right\|^2}{\inf_{(x,t)\in\Gamma\times[0,T]}\left|a(x,t)\cdot\nu(x)\right|}, & \text{if } \inf_{(x,t)\in\Gamma\times[0,T]}\left|a(x,t)\cdot\nu(x)\right| \in \mathbb{R}_{>0}, \\ 0, & \text{otherwise}, \end{cases} \tag{3.22}$$

for some $\psi_h > 1$. From the definitions above, it may be realised that η_c scales as a velocity, whereas t_c scales as the reciprocal of a time only if $\varepsilon_h = 0$. The inconsistency is due to the fact that the term ε_h arises from the cross product between a boundary term and an integral over the physical domain. Moreover, whenever $1/t_\mathrm{c} = 0$, $t_\mathrm{c} = \infty$, which corresponds to the case of constant advection velocity, no reaction, and absence of boundary terms ($C(t) = 0$ by assumption). Another time scale, defined more specifically as

$$t_\star \triangleq \min\{T, t_\mathrm{c}\}, \tag{3.23}$$

is also introduced for what follows. It is essential to clarify that, in the subsequent analyses, expressions involving $1/t_\mathrm{c}$ are conventionally evaluated at zero whenever $t_\mathrm{c} = \infty$. Moreover, according to Eq. (3.23), the following Assumption 3.2.1 is supposed to hold.

Assumption 3.2.1 (Assumption on the meshsize)**.** *The meshsize h is chosen such as to verify*

$$h \leq \eta_\mathrm{c} t_\star. \tag{3.24}$$

Assumption 3.2.1 prevents the local Damköhler number from being too large, and allows the meshsize to resolve the spatial variations of the transport velocity [39]. Moreover, such choice of h implies that a particle advected at speed η_c crosses at least one mesh element over the finite time interval $(0,T)$.

Concerning instead the notation, considering a real number $y \in \mathbb{R}$, its positive and negative parts $y^\oplus$, $y^\ominus \in \mathbb{R}_{\geq 0}$ are also defined for convenience as

$$y^\oplus \triangleq \frac{1}{2}\Big(|y| + y\Big), \tag{3.25a}$$

$$y^\ominus \triangleq \frac{1}{2}\Big(|y| - y\Big). \tag{3.25b}$$

Moreover, to alleviate the notation, inequalities of the type $a \leq Cb$, where C is a constant independent of h and the problem data, are often abbreviated as $a \lesssim b$ in what follows.

Norms and seminorms

Inspired by [39], the following seminorms on $V_{h\star} \triangleq H^1(\Omega;\mathbb{R}^n) + V_h$ are introduced, which are needed for the analysis conducted in Sects. 3.2.2 and 3.2.3:

$$\left|v(\cdot)\right|_\eta^2 \triangleq \frac{1}{2}\int_\Gamma \left|a(x,t)\cdot\nu(x)\right|\left\|v(x)\right\|_2^2\,\mathrm{d}s + \frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F \left|a(x,t)\cdot\nu_F(x)\right|\left\|\left[\!\left[v(x)\right]\!\right]\right\|_2^2\,\mathrm{d}s, \tag{3.26a}$$

$$\left|v(\cdot)\right|_C^2 \triangleq \frac{1}{2}\left\|C(t)\right\|^2\left\|v(\cdot)\right\|_{L^2(\Gamma;\mathbb{R}^n)}^2, \tag{3.26b}$$

$$\left|v(\cdot)\right|_{\varepsilon_h}^2 \triangleq \begin{cases}\left|v(\cdot)\right|_\eta^2 - \dfrac{1}{\varepsilon_h}\left|v(\cdot)\right|_C^2, & \text{if } \varepsilon_h \in \mathbb{R}_{>0},\\ \left|v(\cdot)\right|_\eta^2 & \text{otherwise,}\end{cases} \tag{3.26c}$$

$$\left|v(\cdot)\right|_{\frac{\varepsilon_h}{2}}^2 \triangleq \begin{cases}\left|v(\cdot)\right|_{\varepsilon_h}^2 - \dfrac{1}{\varepsilon_h}\left|v(\cdot)\right|_C^2, & \text{if } \varepsilon_h \in \mathbb{R}_{>0},\\ \left|v(\cdot)\right|_\eta^2 & \text{otherwise.}\end{cases} \tag{3.26d}$$

It is worth clarifying, in particular, that the term $C(t)$ in Eq. (3.26b) represents the matrix of coefficient appearing in the IBVP (3.1), and hence the quantities defined according to Eqs. (3.26c) and (3.26d) are actually seminorms, owing to an appropriate choice of the parameter ε_h satisfying Eq. (3.22) with $\psi_h > 1$. Accordingly, the following norms, similar to those considered in [39], are also defined on $V_{h\star}$:

$$\left\|v(\cdot)\right\|_h^2 \triangleq \frac{1}{t_{\mathrm{c}}}\left\|v(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \left|v(\cdot)\right|_\eta^2, \tag{3.27a}$$

$$\left\|v(\cdot)\right\|_{h\star}^2 \triangleq \left\|v(\cdot)\right\|_h^2 + \sum_{T\in\mathscr{T}_h}\eta_{\mathrm{c}}\left\|v(\cdot)\right\|_{L^2(\partial T;\mathbb{R}^n)}^2. \tag{3.27b}$$

3.2.2 The discrete operator $(A_h, V_{h\star})$

The next step consists in replacing the operator $(A, D(A))$ appearing in Eqs. (2.7) and (3.2) with its discrete counterpart. In this thesis, based predominantly on [39], a discrete operator with upwind regularisation is proposed. More specifically, the discrete operator $(A_h, V_{h\star})$, $A_h(t) : V_{h\star} \mapsto V_h$, is defined such that

$$\begin{aligned}\left\langle A_h(t)v, w_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)} \triangleq & -\int_\Omega\left[\left(a(x,t)\cdot\nabla_h\right)v(x)\right]^{\mathrm{T}} w_h(x)\,\mathrm{d}x + \int_\Gamma v^{\mathrm{T}}(x)C^{\mathrm{T}}(t)\,\mathrm{d}s\int_\Omega w_h(x)\,\mathrm{d}x\\ & -\int_\Gamma\left(a(x,t)\cdot\nu(x)\right)^{\ominus} v^{\mathrm{T}}(x)w_h(x)\,\mathrm{d}s\\ & +\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F a(x,t)\cdot\nu_F(x)\left[\!\left[v(x)\right]\!\right]^{\mathrm{T}}\left\{\!\left\{w_h(x)\right\}\!\right\}\mathrm{d}s\\ & -\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F\left|a(x,t)\cdot\nu_F(x)\right|\left[\!\left[v(x)\right]\!\right]^{\mathrm{T}}\left[\!\left[w_h(x)\right]\!\right]\mathrm{d}s, \quad \text{for } (v,w_h)\in V_{h\star}\times V_h.\end{aligned} \tag{3.28}$$

In this way, $(A_h, V_{h\star})$ may be used to formulate the equivalent space semi-discrete problem of the IBVP (2.7) as

$$\frac{\mathrm{d}u_h(t)}{\mathrm{d}t} = A_h(t)u_h(t) + f_h(t), \qquad \text{for } t\in(0,T), \tag{3.29a}$$

$$u_h(0) = \pi_h u_0, \tag{3.29b}$$
$$f_h(t) = \pi_h f(t), \qquad \text{for} \, t \in (0,T), \tag{3.29c}$$

where $\pi_h : L^2(\Omega;\mathbb{R}^n) \mapsto V_h$ denotes once again the L^2-projection onto $V_h = \mathbb{P}_d^k(\mathscr{T}_h;\mathbb{R}^n)$. By choosing a suitable basis for the space V_h, the semi-discrete IBVP (3.29) may be transformed into a system of linear ODEs for the time-varying components $u_h(t)$ on the selected basis [39].

Remark 3.2.1. *In defining $(A_h, V_{h\star})$ according to Eq. (3.28), it has been assumed that $B(t) = 0$ for simplicity. The extension to the case where $B(t)$ is not identically zero is trivial. It is also worth clarifying that the following analyses are not affected by such a simplification, and the definition of the seminorms and norms as in Eqs. (3.26) and (3.27) can easily accommodate the more general case, by simply modifying the quantities appearing in Eq. (3.21). On this matter, see also [39].*

Characterisation of the discrete operator $(A_h, V_{h\star})$

Some technical results are preliminary needed to establish the stability and convergence of the proposed DGMs.

In particular, the first result advocated here, formalised in Lemma 3.2.1, delivers a stronger version of the consistency property considered in [39], which accounts for time-shifting. Consistency and discrete quasi-dissipativity properties for the operator $(A_h, V_{h\star})$ are also asserted in the classical version enounced in [39] by the subsequent Corollary 3.2.1 and Lemma 3.2.2, respectively. The latter, in particular, may be interpreted as the discrete counterpart of the dissipativity property discussed in Chap. 2 concerning the continuous hyperbolic IBVPs under investigation.

Lemma 3.2.1 (Time-shifted consistency)**.** *Under Assumption 3.1.1, the discrete operator $(A_h, V_{h\star})$ is* time-shifted consistent. *That is, for any exact solution $u \in C^1([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];H^1(\Omega;\mathbb{R}^n))$ to the IBVP described by Eqs. (2.7) and (3.2), it satisfies:*

$$A_h(t)u\left(t'\right) = A(t)u\left(t'\right), \qquad \text{for} \left(t,t'\right) \in [0,T]^2, \tag{3.30a}$$
$$A_h(t)u\left(t'\right) = \pi_h A(t)u\left(t'\right), \qquad \text{for} \left(t,t'\right) \in [0,T]^2. \tag{3.30b}$$

Proof. Since it solves the IBVP described by Eqs. (2.7) and (3.2), $u \in C^0([0,T];H^1(\Omega;\mathbb{R}^n)) \cap C^0([0,T];D(A))$ by assumption. Moreover, since $D(A(t)) = D(A(0))$ is constant over time, the third term on the right-hand side of Eq. (3.28) vanishes for all $(t,t') \in [0,T]^2$, whereas the identity $[\![u(x,t')]\!] = 0$ is valid for all $F \in \mathscr{F}_h^{\mathrm{i}}$ according to Theorem 3.1.1. Hence, Eq. (3.28) simplifies to

$$\begin{aligned} \left\langle A_h(t)u\left(t'\right), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} &= -\int_\Omega \left[\left(a(x,t)\cdot\nabla\right)u\left(x,t'\right)\right]^{\mathrm{T}} w_h(x)\,\mathrm{d}x \\ &\quad + \int_\Gamma u^{\mathrm{T}}\left(x,t'\right)C^{\mathrm{T}}(t)\,\mathrm{d}s \int_\Omega w_h(x)\,\mathrm{d}x \\ &= \left\langle A(t)u\left(t'\right), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)}, \end{aligned} \tag{3.31}$$

therefore proving Eq. (3.30a). Moreover, recalling the definition of $\pi_h : L^2(\Omega;\mathbb{R}^n) \mapsto V_h$ provides

$$\left\langle \pi_h A(t)u\left(t'\right), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} = \left\langle A(t)u\left(t'\right), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)}, \tag{3.32}$$

which, combined with the above Eq. (3.31), gives (3.30b). □

The classical notion of consistency, as stated in [39], follows immediately from Lemma 3.2.1, and is reported below for completeness.

Corollary 3.2.1 (Consistency (Di Pietro and Ern [39])). *The discrete operator* $(A_h, V_{h\star})$ *is* consistent. *That is, for any exact solution* $u \in C^1([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];H^1(\Omega;\mathbb{R}^n))$*, it satisfies*

$$\pi_h \frac{\mathrm{d}u(t)}{\mathrm{d}t} = A_h(t)u(t) + f_h(t), \qquad \text{for } t \in [0,T]. \tag{3.33}$$

Proof. The proof is almost identical to that of Lemma 3.4 in [39]. Taking the inner product of Eq. (2.7) on $L^2(\Omega;\mathbb{R}^n)$ with $w_h \in V_h$ yields

$$\left\langle \frac{\mathrm{d}u(t)}{\mathrm{d}t}, w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} = \left\langle A(t)u(t), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} + \left\langle f(t), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)}, \quad \text{for } t \in [0,T]. \tag{3.34}$$

From Lemma 3.2.1, it follows immediately that $\langle A_h(t)u(t), w_h\rangle_{L^2(\Omega;\mathbb{R}^n)} = \langle A(t)u(t), w_h\rangle_{L^2(\Omega;\mathbb{R}^n)}$. Moreover, owing to the definition of $\pi_h : L^2(\Omega;\mathbb{R}^n) \mapsto V_h$, $\langle f(t), w_h\rangle_{L^2(\Omega;\mathbb{R}^n)} = \langle \pi_h f(t), w_h\rangle_{L^2(\Omega;\mathbb{R}^n)}$, and similarly for $\mathrm{d}u(t)/\mathrm{d}t$. Hence,

$$\left\langle \pi_h \frac{\mathrm{d}u(t)}{\mathrm{d}t}, w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} = \left\langle A_h(t)u(t), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} + \left\langle f_h(t), w_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)}, \tag{3.35}$$

which proves the desired result. □

Lemma 3.2.2 (Discrete quasi-dissipativity). *For all* $v_h \in V_h$*, the operator* $(A_h, V_{h\star})$ *is quasi-dissipative with constant*

$$\omega_h \triangleq \frac{1}{2}\left(\sup_{t\in[0,T]} \left\|\nabla \cdot a(\cdot,t)\right\|_\infty + \varepsilon_h \right), \tag{3.36}$$

where ε_h *may be chosen abitrarily to satisfy Eq.* (3.22)*. In particular,*

$$\left\langle A_h(t)v_h, v_h \right\rangle_{L^2(\Omega;\mathbb{R}^n)} \leq -\left|v_h(\cdot)\right|^2_{\varepsilon_h} + \omega_h \left\|v_h(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}, \qquad \text{for } v_h \in V_h, \tag{3.37}$$

with the seminorm $|\cdot|_{\varepsilon_h}$ *defined on* $V_{h\star}$ *according to Eq.* (3.26c).

Proof. The proof is almost identical to that of Lemma 3.4 in [39]. Integrating by parts the advective term in Eq. (3.28) on each mesh element yields

$$\begin{aligned}
\int_\Omega \left[\Big(a(x,t)\cdot\nabla_h\Big) v_h(x) \right]^{\mathrm{T}} v_h(x)\,\mathrm{d}x &= \int_{T\in\mathscr{T}_h} \left[\Big(a(x,t)\cdot\nabla\Big) v_h(x) \right]^{\mathrm{T}} v_h(x)\,\mathrm{d}x \\
&= -\frac{1}{2}\int_\Omega \Big(\nabla\cdot a(x,t)\Big) \left\|v_h(x)\right\|_2^2 \mathrm{d}x \\
&\quad + \frac{1}{2}\sum_{T\in\mathscr{T}_h} \int_{\partial T} a(x,t)\cdot \nu_{\partial T}(x) \left\|v_h(x)\right\|_2^2 \mathrm{d}s, \quad \text{for } v_h \in V_h.
\end{aligned} \tag{3.38}$$

By exploiting the continuity of the normal component of the transport velocity $a(x,t)$ across the interfaces, the second term on the right-hand side may be reformulated as

$$\begin{aligned}\frac{1}{2}\sum_{T\in\mathscr{T}_h}\int_{\partial T} a(x,t)\cdot\nu_T(x)v_h^{\mathrm{T}}(x)v_h(x)\,\mathrm{d}s &= \frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F a(x,t)\cdot\nu_F(x)\left[\!\left[\|v_h(x)\|_2^2\right]\!\right]\mathrm{d}s\\ &\quad+\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{b}}}\int_F a(x,t)\cdot\nu_F(x)\|v_h(x)\|_2^2\,\mathrm{d}s\\ &= \sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F a(x,t)\cdot\nu_F(x)\left[\!\left[v_h(x)\right]\!\right]^{\mathrm{T}}\left\{\!\!\left\{v_h(x)\right\}\!\!\right\}\mathrm{d}s\\ &\quad+\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{b}}}\int_F a(x,t)\cdot\nu_F(x)\|v_h(x)\|_2^2\,\mathrm{d}s.\end{aligned}\tag{3.39}$$

Therefore

$$\begin{aligned}\left\langle A_h(t)v_h,v_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)} &= \frac{1}{2}\int_\Omega\Big(\nabla\cdot a(x,t)\Big)\|v_h(x)\|_2^2\,\mathrm{d}x+\int_\Gamma v_h^{\mathrm{T}}(x)C^{\mathrm{T}}(t)\,\mathrm{d}s\int_\Omega v_h(x)\,\mathrm{d}x\\ &\quad-\int_\Gamma\Big(a(x,t)\cdot\nu(x)\Big)^{\ominus}\|v_h(x)\|_2^2\,\mathrm{d}s-\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{b}}}\int_F a(x,t)\cdot\nu_F(x)\|v_h(x)\|_2^2\,\mathrm{d}s\\ &\quad-\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F|a(x,t)\cdot\nu_F(x)|\left\|\left[\!\left[v_h(x)\right]\!\right]\right\|_2^2\mathrm{d}s,\quad \text{for } v_h\in V_h,\end{aligned}\tag{3.40}$$

and combining the third and fourth terms on the right-hand side leads to

$$\begin{aligned}\left\langle A_h(t)v_h,v_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)} &= \frac{1}{2}\int_\Omega\Big(\nabla\cdot a(x,t)\Big)\|v_h(x)\|_2^2\,\mathrm{d}x+\int_\Gamma v_h^{\mathrm{T}}(x)C^{\mathrm{T}}(t)\,\mathrm{d}s\int_\Omega v_h(x)\,\mathrm{d}x\\ &\quad-\frac{1}{2}\int_\Gamma|a(x,t)\cdot\nu(x)|\|v_h(x)\|_2^2\,\mathrm{d}s\\ &\quad-\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F|a(x,t)\cdot\nu_F(x)|\left\|\left[\!\left[v_h(x)\right]\!\right]\right\|_2^2\mathrm{d}s,\quad \text{for } v_h\in V_h.\end{aligned}\tag{3.41}$$

The estimate (3.37) follows from an application of the Cauchy-Schwarz' and generalised Young's inequalities to the second term on the right-hand side of Eq. (3.41). □

The last technical result concerns a bound on orthogonal subscales on the discrete operator $(A_h, V_{h\star})$, according to Proposition 3.2.1 below.

Proposition 3.2.1 (Boundedness on orthogonal subscales)**.** *The discrete operator* $(A_h, V_{h\star})$ *satisfies*

$$\left|\left\langle A_h(t)(v-\pi_h v),w_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)}\right| \lesssim \|v(\cdot)-\pi_h v(\cdot)\|_{h\star}\|w_h(\cdot)\|_h,\quad \text{for } (v,w_h)\in H^1(\Omega;\mathbb{R}^n)\times V_h,\tag{3.42}$$

where the norms $\|\cdot\|_h$ *and* $\|\cdot\|_{h\star}$ *are defined according to Eqs.* (3.27a) *and* (3.27b)*, respectively.*

Proof. By observing that

$$\begin{aligned}
\int_\Gamma \Big(v(x)-\pi_h v(x)\Big)^{\mathrm{T}} C^{\mathrm{T}}(t)\,\mathrm{d}s \int_\Omega w_h(x)\,\mathrm{d}x &\le \sqrt{\sup_{t\in[0,T]}\big\|C(t)\big\|^2}\big\|v(\cdot)-\pi_h v(\cdot)\big\|_{L^2(\Gamma;\mathbb{R}^n)}\big\|w_h(\cdot)\big\|_{L^2(\Omega;\mathbb{R}^n)} \\
&= \sqrt{\inf_{(x,t)\in\Gamma\times[0,T]}\big|a(x,t)\cdot\nu(x)\big|}\big\|v(\cdot)-\pi_h v(\cdot)\big\|_{L^2(\Gamma;\mathbb{R}^n)} \\
&\quad\times\sqrt{\frac{\sup_{t\in[0,T]}\big\|C(t)\big\|^2}{\inf_{(x,t)\in\Gamma\times[0,T]}\big|a(x,t)\cdot\nu(x)\big|}}\big\|w_h(\cdot)\big\|_{L^2(\Omega;\mathbb{R}^n)} \\
&\lesssim \big|v(\cdot)-\pi_h v(\cdot)\big|_\eta\sqrt{\varepsilon_h}\big\|w_h(\cdot)\big\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \big\|v(\cdot)-\pi_h v(\cdot)\big\|_{h\star}\big\|w_h(\cdot)\big\|_h,
\end{aligned} \tag{3.43}$$

the result is a direct consequence of Lemma 2.30 in [39]. □

3.2.3 Stability and convergence of the space semi-discrete DGMs

The present Section delivers some preliminary results concerning the stability of the space semi-discrete problem obtained by introducing the discrete operator $(A_h, V_{h\star})$ according to Eq. (3.28), as well as the convergence of the space semi-discrete DGMs. The fully discrete DGMs are instead investigated in Chap. 4 adopting RK1 and RK2 approximations in time.

Stability of the space semi-discrete problem

The following Theorem 3.2.1 asserts the main stability result regarding the space semi-discrete problem constructed according to Eqs. (3.29) and (3.28).

Theorem 3.2.1 (Stability of the space semi-discrete problem)**.** *Consider the space semi-discrete problem* (3.29) *and the discrete operator* $(A_h, V_{h\star})$ *defined according to Eq.* (3.28)*. It holds*

$$\big\|u_h(\cdot,t)\big\|_{L^2(\Omega;\mathbb{R}^n)} \le \mathrm{e}^{t/t_\star}\left(\big\|u_h(\cdot,0)\big\|_{L^2(\Omega;\mathbb{R}^n)} + \int_0^t \Big\|f\Big(\cdot,t'\Big)\Big\|_{L^2(\Omega;\mathbb{R}^n)}\mathrm{d}t'\right), \quad \text{for } t\in[0,T]. \tag{3.44}$$

Proof. Taking the inner product of Eq. (3.29a) on $L^2(\Omega;\mathbb{R}^n)$ and invoking the discrete quasi-dissipativity property (3.37) proved in Lemma yields (3.2.2)

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big\|u_h(\cdot,t)\big\|^2_{L^2(\Omega;\mathbb{R}^n)} + \big|u_h(\cdot,t)\big|^2_{\varepsilon_h} \le \omega_h\big\|u_h(\cdot,t)\big\|^2_{L^2(\Omega;\mathbb{R}^n)} + \big\|f(\cdot,t)\big\|_{L^2(\Omega;\mathbb{R}^n)}\big\|u_h(\cdot,t)\big\|_{L^2(\Omega;\mathbb{R}^n)}, \quad \text{for } t\in(0,T). \tag{3.45}$$

By observing that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big\|u_h(\cdot,t)\big\|^2_{L^2(\Omega;\mathbb{R}^n)} = \big\|u_h(\cdot,t)\big\|_{L^2(\Omega;\mathbb{R}^n)}\frac{\mathrm{d}}{\mathrm{d}t}\big\|u_h(\cdot,t)\big\|_{L^2(\Omega;\mathbb{R}^n)}, \tag{3.46}$$

simplifying by $\big\|u_h(\cdot,t)\big\|_{L^2(\Omega;\mathbb{R}^n)}$, observing that $\omega_h \le 1/t_\mathrm{c} \le 1/t_\star$, and applying Grönwall-Bellman's inequality yields the result. □

It is worth emphasising that the discrete quasi-dissipative property of the operator $(A_h, V_{h\star})$ plays a crucial role in the derivation of the stability estimate in Eq. (3.44). Similar, and, in fact, even sharper bounds may be deduced accounting for the contribution of the trace terms, as done in the next Chap. 4 concerning fully discrete DGMs.

The next result concerns the convergence of the error as the meshsize tends to zero.

Convergence of space semi-discrete DGMs

To simplify the notation, recalling the definition of the projection $\pi_h : L^2(\Omega;\mathbb{R}^n) \mapsto V_h$, the following auxiliary quantities are defined:

$$\xi_h(x,t) \triangleq u_h(x,t) - \pi_h u(x,t), \tag{3.47a}$$

$$\xi_\pi(x,t) \triangleq u(x,t) - \pi_h u(x,t), \tag{3.47b}$$

so that the approximation error may be decomposed as

$$u(x,t) - u_h(x,t) = \xi_\pi(x,t) - \xi_h(x,t). \tag{3.48}$$

Theorem 3.2.2 asserts the main convergence result concerning the space semi-discrete DGMs.

Theorem 3.2.2 (Convergence of space semi-discrete DGMs)**.** *Consider the IBVP* (2.7) *and the space semi-discrete problem* (3.29) *and assume* $u \in C^0([0,T];H^{k+1}(\Omega;\mathbb{R}^n))$ *for the exact solution. Then, the following estimate holds:*

$$\left\|u(\cdot,t) - u_h(\cdot,t)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left(\int_0^t \left|u\left(\cdot,t'\right) - u_h\left(\cdot,t'\right)\right|^2_{\frac{\varepsilon_h}{2}} \mathrm{d}t'\right)^{1/2} \lesssim \mathrm{e}^{C_{\mathrm{sta}}t/t_\star} \chi h^{k+1/2}, \quad \text{for } t \in [0,T], \tag{3.49}$$

where

$$\chi \triangleq \sqrt{\eta_{\mathrm{c}} T} \left\|u(\cdot)\right\|_{C^0([0,T];H^{k+1}(\Omega;\mathbb{R}^n))}, \tag{3.50}$$

and the constant C_{sta} *is independent of* h *and the data* $f(x,t)$*,* $C(t)$*, and* $a(x,t)$*.*

Proof. Using the definitions (3.47) combined with Eqs. (2.7) and (3.29), and recalling the consistency property (3.33) proved in Corollary 3.2.1, the following relationship for the error equation may be deduced according to the abstract formulation:

$$\frac{\mathrm{d}\xi_h(t)}{\mathrm{d}t} = A_h(t)\xi_h(t) - A_h(t)\xi_\pi(t), \qquad \text{for } t \in (0,T). \tag{3.51}$$

Hence, taking the inner product of Eq. (3.51) with $\xi_h(x,t)$ on $L^2(\Omega;\mathbb{R}^n)$ and invoking the quasi-dissipativity property (3.37) of the discrete operators $(A_h, V_{h\star})$ proved in Lemma (3.2.2) yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left|\xi_h(\cdot,t)\right|^2_{\varepsilon_h} \leq \omega_h \left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} - \left\langle A_h(t)\xi_\pi(t), \xi_h(t)\right\rangle, \quad \text{for } t \in (0,T). \tag{3.52}$$

By noticing that $\xi_\pi(t) \in V_{h\star}$, recalling the boundedness on orthogonal subscales (3.42) provides

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left|\xi_h(\cdot,t)\right|^2_{\varepsilon_h} &\leq \omega_h\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left\|\xi_\pi(\cdot,t)\right\|_{h\star}\left\|\xi_h(\cdot,t)\right\|_h \\ &\leq \left\|\xi_\pi(\cdot,t)\right\|_{h\star}\left(\frac{1}{\sqrt{t_\mathrm{c}}}\left\|\xi_h(\cdot,t)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left|\xi_h(\cdot,t)\right|_\eta\right) \\ &\quad + \omega_h\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}, \quad \text{for } t \in (0,T).\end{aligned} \tag{3.53}$$

Applying the generalised version of Young's inequality for products to the last term, observing that $\omega_h \leq 1/t_\mathrm{c} \leq 1/t_\star$, and recalling the definition of the seminorm $|\cdot|_{\frac{\varepsilon_h}{2}}$ as in Eq. (3.26d) gives then

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \frac{1}{2}\left|\xi_h(\cdot,t)\right|^2_{\frac{\varepsilon_h}{2}} \lesssim \frac{C_\mathrm{sta}}{t_\star}\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left\|\xi_\pi(\cdot,t)\right\|^2_{h\star}, \quad \text{for } t \in (0,T). \tag{3.54}$$

Thus, using Grönwall–Bellman's inequality and noticing that $\xi_h(x,0) = 0$, it may be inferred that

$$\left\|\xi_h(\cdot,t)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \int_0^t \left|\xi_h\left(\cdot,t'\right)\right|^2_{\frac{\varepsilon_h}{2}} \mathrm{d}t' \lesssim \mathrm{e}^{C_\mathrm{sta}t/t_\star} t_\star \max_{t\in[0,T]}\left\|\xi_\pi(\cdot,t)\right\|^2_{h\star}, \quad \text{for } t \in [0,T]. \tag{3.55}$$

Therefore, resorting to the triangle inequality and observing that $|\cdot|^2_{\frac{\varepsilon_h}{2}} \leq \|\cdot\|^2_{h\star}$ and that $t_\star \leq T$ yields

$$\begin{aligned}\left\|u(\cdot,t) - u_h(\cdot,t)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left(\int_0^t \left|u\left(\cdot,t'\right) - u_h\left(\cdot,t'\right)\right|^2_{\frac{\varepsilon_h}{2}} \mathrm{d}t'\right)^{1/2} &\lesssim \mathrm{e}^{C_\mathrm{sta}t/t_\star}\sqrt{T}\max_{t\in[0,T]}\left\|\xi_\pi(\cdot,t)\right\|_{h\star} \\ &\quad + \left\|\xi_\pi(\cdot,t)\right\|_{L^2(\Omega;\mathbb{R}^n)}, \\ &\quad \text{for } t \in [0,T].\end{aligned} \tag{3.56}$$

Using the polynomial approximation properties asserted by Lemmata 3.1.4 and 3.1.5, and defining χ according to Eq. (3.50), also provides

$$\left\|\xi_\pi(\cdot,t)\right\|^2_{h\star} \lesssim \eta_\mathrm{c} h^{2k+1}\left\|u(\cdot,t)\right\|^2_{H^{k+1}(\Omega;\mathbb{R}^n)}, \tag{3.57a}$$

$$\left\|\xi_\pi(\cdot,t)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim h^{k+1}\left\|u(\cdot,t)\right\|_{H^{k+1}(\Omega;\mathbb{R}^n)} \lesssim \chi h^{k+1/2}, \tag{3.57b}$$

owing to Assumption 3.2.1 on the meshsize. Combining Eqs. (3.56) and (3.57), the desired result follows. □

Chapter 4

Time discretisation and explicit RK schemes

By selecting an opportune basis spanning the functional space V_h, the semi-discrete problem derived in Chap. 3 may be converted into a linear system of ODEs, for which existence and uniqueness follow from standard arguments. In practice, these systems often need to be solved with the aid of a computer, which motivates the need for analysing numerical schemes in time. In particular, explicit Runge-Kutta (RK) algorithms are widely employed for the solution of linear problems. In this context, the aim of the present Chapter is that of analysing the stability and convergence properties of RK1 and RK2 schemes in combination with the DGMs discussed in Chap. 3. More specifically, the forward Euler method is adduced as an example of a low-order RK1 scheme, whereas the higher-order RK2 algorithms explored in this thesis cover the explicit midpoint method, Heun's second-order method, and Ralston's method.

The remainder of the Chapter is organised as follows. Section 4.1 discusses the main RK schemes considered in the thesis, providing also an alternative representation of the discrete governing equations that is more conveniently adopted for the error analysis conducted then in Sect. 4.3. In this context, the main convergence results are asserted in Sect. 4.2 concerning both the forward Euler and the RK2 algorithms. Finally, Sect. 4.3 guides the reader through the proof of the main results.

4.1 Time discretisation

Typically, the semi-discrete problem formulated according to Eqs. (3.29) needs to be discretised also in time in order to be solved numerically, leading to fully discrete DGMs. In the following, a fixed time step δt is considered such that $T = N\delta t$, with $N \in \mathbb{N}$. For $n \in \{0,\dots,N\}$, the discrete time is defined as $t^n = n\delta t$, and more generally the superscript is used to indicate functions evaluated at the discrete time $t = t^n$. For example, the solution to the IBVP evaluated at $t = t^n$ is denoted by $u(t^n) = u^n$, and similarly for the forcing term $f^n = f(t^n)$. In the same spirit, given a real number $\rho \in [0,1]$, the solution evaluated at $t^n + \rho\delta t$ is indicated with $u^{n+\rho} = u(t^n + \rho\delta t)$, and so on for other time-dependent quantities.

The mild assumption

$$\delta t \leq t_\star, \tag{4.1}$$

with $t_\star$ as in Eq. (3.23), is also introduced in the following to facilitate the error analysis conducted in Sect. 4.3. A very common way of approximating the problem described by Eqs.

(3.29) consists in resorting to an explicit RK algorithm. A possible choice for formulating a general RK scheme of order s for the semi-discrete problem (3.29) is

$$k_i = A_h\Big(t^n + c_i\delta t\Big)\left(u_h^n + \delta t\sum_{j=1}^{s} a_{ij}k_j\right) + f_h\Big(t^n + c_i\delta t\Big), \qquad \text{for } i \in \{1,\dots,s\}, \tag{4.2a}$$

$$u_h^{n+1} = u_h^n + \sum_{i=1}^{s} b_i k_i. \tag{4.2b}$$

In Eqs. (4.2), $(a_{ij})_{1\leq i,j\leq s}$ are real numbers, $(b_i)_{1\leq i\leq s}$ are real numbers satisfying $\sum_{i=1}^{s} b_i = 1$, and $(c_i)_{1\leq i\leq s}$ are real numbers in $[0,1]$ such that $c_i = \sum_{j=1}^{s} a_{ij}$ for all $i \in \{1,\dots,s\}$. These quantities are conventionally collected in the so-called *Butcher's tableau*

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}. \tag{4.3}$$

In particular, RK schemes are explicit whenever $a_{ji} = 0$ for all $i \geq j$.

More specifically, two examples of explicit RK schemes are discussed in the following: the forward Euler algorithm, detailed in the next Sect. 4.1.1, and the two-stage RK2 algorithms, presented in Sect. 4.1.2.

4.1.1 Forward Euler scheme

The simplest, and perhaps most intuitive approximation to the semi-discrete problem (3.29) takes the form

$$\frac{u^{n+1} - u^n}{\delta t} = A_h^n u^n + f_h^n, \tag{4.4}$$

leading to the alternative representation

$$u^{n+1} = u^n + \delta t A_h^n u^n + \delta t f_h^n, \tag{4.5}$$

which corresponds to an explicit RK1 scheme with Butcher's tableau

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}. \tag{4.6}$$

The discrete Eq. (4.5) is already in a suitable form for the error analysis conducted in Sect. 4.3.

4.1.2 Explicit RK2 schemes

Concerning the semi-discrete problem (3.29), in the most general form, RK2 schems admit a general representation as in Eqs. (4.2) according to

$$k_1 = A_h^n u_h^n + f_h^n, \tag{4.7a}$$

$$k_2 = A_h^{n+\lambda}(u_h^n + \lambda\delta t k_1) + f_h^{n+\lambda}, \tag{4.7b}$$

$$u_h^{n+1} = u_h^n + \left(1 - \frac{1}{2\lambda}\right) k_1 + \frac{1}{\lambda} k_2, \tag{4.7c}$$

with Butcher's tableau reading

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \lambda & \lambda & 0 \\ \hline & 1 - \frac{1}{2\lambda} & \frac{1}{\lambda} \end{array}, \tag{4.8}$$

where $\lambda \in [1/2, 1]$. In particular, the Butcher's tableau in Eq. (4.8) corresponds to the explicit midpoint method for $\lambda = 1/2$, to Heun's second-order method for $\lambda = 1$, and to Ralston's method for $\lambda = 2/3$.

Following the approach outlined by Shu and Osher [55], the representation of Eqs. (4.7) may be converted into the following one by introducing intermediate steps on the solution:

$$w_h^n = u_h^n + \delta t A_h^n u_h^n + \delta t f_h^n, \tag{4.9a}$$

$$u_h^{n+1} = \frac{u_h^n}{2\lambda} + \left(1 - \frac{1}{2\lambda}\right) w_h^n + \frac{\delta t}{2} A_h^{n+\lambda} \left[\left(\frac{1}{\lambda} - 1\right) u_h^n + w_h^n \right] + \frac{\delta t}{2\lambda} f_h^{n+\lambda}, \tag{4.9b}$$

which is more convenient to proceed with the error analysis performed in Sect. 4.1.2. It is also worth observing that, whenever the operators $(A, D(A))$ and $(A_h, V_{h\star})$ are time-independent, yielding thus $A_h^n = A_h^{n+\lambda} = A_h$ in Eqs. (4.7) and (4.9), all the above-mentioned methods admit a representation in the form (4.9) with $\lambda = 1$ (see, e.g., [39]).

4.2 Main convergence results

The present Section states the main convergence results for the considered DGMs with RK discretisation algorithms for the time variable. In particular, the following Courant-Friedrichs-Lewy (CFL) conditions are introduced, namely, the usual CFL condition

$$\delta t \leq \rho \frac{h}{\eta_c}, \tag{4.10}$$

the refined $4/3$-CFL condition

$$\delta t \leq \rho t_\star^{-1/3} \left(\frac{h}{\eta_c}\right)^{4/3}, \tag{4.11}$$

for some $\rho \in \mathbb{R}_{>0}$ (not necessarily the same), and the 2-CFL condition

$$\delta t \leq \rho' \frac{1}{t_\star} \left(\frac{h}{\eta_c}\right)^2, \tag{4.12}$$

for some $\rho' \in \mathbb{R}_{>0}$.

It is worth observing that the standard CFL condition (4.10) is implied by the $4/3$-one (4.11), which is in turn implied by the 2-CFL condition (4.12).

4.2.1 Forward Euler and finite volume schemes

Inspired by [39], and continuing the discussion initiated in Sect. 4.1.1, the following scheme is considered concerning the forward Euler method:

$$u_h^{n+1} = u_h^n + \delta t A_h^n u_h^n + \delta t f_h^n, \tag{4.13a}$$
$$u_h^0 = \pi_h u_0, \tag{4.13b}$$

with $V_h = \mathbb{P}_d^0(\mathscr{T}_h;\mathbb{R}^n)$. The corresponding main convergence result is asserted by Theorem 4.2.1 below.

Theorem 4.2.1 (Convergence for the forward Euler scheme)**.** *Assume $u \in C^2([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];H^1(\Omega;\mathbb{R}^n))$ for the exact solution and set $V_h \triangleq \mathbb{P}_d^0(\mathscr{T}_h;\mathbb{R}^n)$. Moreover, assume that the CFL condition in Eq. (4.10) is satisfied with a suitable threshold ρ independent of h, δt, and the data $f(x,t)$, $C(t)$ and $a(x,t)$. Then, the following estimate holds:*

$$\left\| u^N(\cdot) - u_h^N(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)} + \left(\sum_{n=0}^{N-1} \delta t \left| u^n(\cdot) - u_h^n(\cdot) \right|_{\frac{\varepsilon_h}{2}}^2 \right)^{1/2} \lesssim \mathrm{e}^{C_\star T/t_\star} \left(\chi_1 \delta t + \chi_2 \sqrt{h} \right), \tag{4.14}$$

where

$$\chi_1 \triangleq \sqrt{t_\star T} \left\| \frac{\partial^2 u(\cdot,\cdot)}{\partial t^2} \right\|_{C^0([0,T];L^2(\Omega;\mathbb{R}^n))}, \tag{4.15a}$$
$$\chi_2 \triangleq \sqrt{\eta_c T} \left\| \frac{\partial^2 u(\cdot,\cdot)}{\partial t^2} \right\|_{C^0([0,T];H^1(\Omega;\mathbb{R}^n))}, \tag{4.15b}$$

the seminorm $|\cdot|_{\frac{\varepsilon_h}{2}}$ is defined according to Eq. (3.26d), and the constant $C_\star$ is independent of h, δt, and the data $f(x,t)$, $C(t)$, and $a(x,t)$.

Proof. See Sect. 4.3.1. □

The proof of Theorem 4.2.1 is almost identical to that of Theorem 3.7 in [39], but is anyway detailed in Sect. 4.3.1, since it relies on the introduction of the seminorms $|\cdot|_{\varepsilon_h}$ and $|\cdot|_{\frac{\varepsilon_h}{2}}$ which are not defined in [39]. It is also worth clarifying that the specific choice of the polynomial space $V_h \triangleq \mathbb{P}_d^0(\mathscr{T}_h;\mathbb{R}^n)$ in Theorem 4.2.1 is motivated by the fact that higher polynomial degrees require enforcing the 2-CFL condition (4.12) in order to provide optimal convergence. Since the forward Euler approximation is less than first-order accurate in space, such a requirement is too stringent, and is instead explored in the context of RK2 schemes. A more elaborated discussion concerning this aspect is reported in [39].

4.2.2 Explicit RK2 schemes

Inspired by [39], and continuing the discussion initiated in Sect. 4.1.2, the following explicit RK2 schemes are considered:

$$w_h^n = u_h^n + \delta t A_h^n u_h^n + \delta t f_h^n, \tag{4.16a}$$

$$u_h^{n+1} = \frac{u_h^n}{2\lambda} + \left(1 - \frac{1}{2\lambda}\right) w_h^n + \frac{\delta t}{2} A_h^{n+\lambda} \left[\left(\frac{1}{\lambda} - 1\right) u_h^n + w_h^n\right] + \frac{\delta t}{2\lambda} f_h^{n+\lambda}, \tag{4.16b}$$

$$u_h^0 = \pi_h u_0, \tag{4.16c}$$

with $V_h = \mathbb{P}_d^k(\mathscr{T}_h; \mathbb{R}^n)$, $k \geq 1$. The corresponding main convergence result is asserted by Theorem 4.2.2 below.

Theorem 4.2.2 (Convergence for the explicit RK2 schemes)**.** *Assume $u \in C^3([0,T]; L^2(\Omega; \mathbb{R}^n)) \cap C^s([0,T]; H^{k+1-s}(\Omega; \mathbb{R}^n))$, $s \in \{0,1\}$, for the exact solution, $f \in C^2([0,T]; L^2(\Omega; \mathbb{R}^n))$, $A \in C^2([0,T]; \mathscr{L}(H^1(\Omega; \mathbb{R}^n); L^2(\Omega; \mathbb{R}^n)))$, and set $V_h \triangleq \mathbb{P}_d^k(\mathscr{T}_h; \mathbb{R}^n)$ for $k \geq 1$. Moreover, assume that the 2-CFL condition in Eq. (4.12) is satisfied with a suitable threshold ρ' independent of h, δt, and the data $f(x,t)$, $C(t)$ and $a(x,t)$. Then, the following estimate holds:*

$$\begin{aligned}
&\left\| u^N(\cdot) - u_h^N(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)} \\
&\quad + \left(\sum_{n=0}^{N-1} \frac{2\lambda - 1}{\lambda} \delta t \left| u^n(\cdot) - u_h^n(\cdot) \right|_{\frac{\varepsilon_h}{2}}^2 + \lambda \delta t \left| \frac{1-\lambda}{\lambda} \left(u^n(\cdot) - u_h^n(\cdot) \right) + w^n(\cdot) - w_h^n(\cdot) \right|_{\frac{\varepsilon_h}{2}}^2 \right)^{1/2} \\
&\lesssim \mathrm{e}^{C_\star T / t_\star} \left(\chi_1 \delta t^2 + \chi_2 \delta t^3 + \chi_3 h^{k+1/2} \right),
\end{aligned} \tag{4.17}$$

where

$$\chi_1 \triangleq \sqrt{t_\star T} C_{fu}, \tag{4.18a}$$

$$\chi_2 \triangleq \sqrt{t_\star T} C_u, \tag{4.18b}$$

$$\chi_3 \triangleq \sqrt{\eta_c T} \sum_{s=0}^{1} \eta_c^{-s} \left\| \frac{\partial^s u(\cdot,\cdot)}{\partial t^s} \right\|_{C^0([0,T]; H^{k+1-s}(\Omega;\mathbb{R}^n))}, \tag{4.18c}$$

the seminorm $|\cdot|_{\frac{\varepsilon_h}{2}}$ is defined according to Eq. (3.26d), the constant $C_\star$ is independent of h, δt, and the data $f(x,t)$, $C(t)$, and $a(x,t)$, and C_{fu} and C_u are given by

$$\begin{aligned}
C_{fu} \triangleq & \left\| \frac{\partial^3 u(\cdot,\cdot)}{\partial t^3} \right\|_{C^0([0,T]; L^2(\Omega;\mathbb{R}^n))} + \eta_2 \left\| u(\cdot,\cdot) \right\|_{C^0([0,T]; H^1(\Omega;\mathbb{R}^n))} \\
& + \eta_1 \left\| \frac{\partial u(\cdot,\cdot)}{\partial t} \right\|_{C^0([0,T]; H^1(\Omega;\mathbb{R}^n))} + \left\| \frac{\partial^2 f(\cdot,t)}{\partial t^2} \right\|_{C^0([0,T]; L^2(\Omega;\mathbb{R}^n))},
\end{aligned} \tag{4.19a}$$

$$C_u \triangleq \eta_2 \left\| \frac{\partial u(\cdot,\cdot)}{\partial t} \right\|_{C^0([0,T]; H^1(\Omega;\mathbb{R}^n))}, \tag{4.19b}$$

with

$$\eta_1 \triangleq \max\left\{ \sup_{t\in[0,T]} \left\| \frac{\partial a(\cdot,t)}{\partial t} \right\|_\infty, \sup_{t\in[0,T]} \left\| \frac{\mathrm{d}C(t)}{\mathrm{d}t} \right\| \right\}, \tag{4.20a}$$

$$\eta_2 \triangleq \max\left\{ \sup_{t\in[0,T]} \left\| \frac{\partial^2 a(\cdot,t)}{\partial t^2} \right\|_\infty, \sup_{t\in[0,T]} \left\| \frac{\mathrm{d}^2 C(t)}{\mathrm{d}t^2} \right\| \right\}. \tag{4.20b}$$

Proof. See Sect. 4.3.2. □

The proof of Theorem 4.2.2 follows similar steps as those needed to prove Theorem 3.10 in [39]. In this context, it is however worth observing that the RK2 schemes considered in the present thesis are slightly more general than those analysed in [39]. The main difference resides in the fact that the operators $(A, D(A))$ and $(A_h, V_{h\star})$ are time-varying. As a result, except for Heun's second-order method (corresponding to $\lambda = 1$), the discrete Eqs. (4.16) cannot be recast in the same form as that studied in [39], and necessitate a dedicated analysis. An interesting conclusion is that, owing to the presence of time-varying operator $(A, D(A))$ and $(A_h, V_{h\star})$, optimal estimates may be achieved with the aid of the same techniques outlined in [39] owing to a more stringent 2-CFL condition, whereas only the 4/3-CFL condition (4.11) is required in [39].

To the best of the author's knowledge, the results advocated in this thesis concerning the analysis of the considered RK2 schemes are novel.

4.3 Error analysis

The present Section is dedicated to the error analysis of the fully discrete DGMs introduced in Sect. 4.2, covering the explicit forward Euler and RK2 schemes.

The *iter* is analogous to that reported in [39], with some major modifications concerning the RK2 schemes, which are required to accommodate the time-dependency of the continuous and discrete operators $(A, D(A))$ and $(A_h, V_{h\star})$, respectively.

4.3.1 Forward Euler scheme

The error analysis for the forward Euler schemes is articulated in different steps. First, an equation is derived governing the discrete time evolution of the error between the exact and approximated solutions to the considered IBVP. After deducing a suitable energy estimate and bound on the discrete operator $(A_h, V_{h\star})$, the stability of the scheme is proved, which then allows to show uniform convergence as the meshsize and time step tend to zero.

Error equation

The first step in the error analysis involves the derivation of the error equation. In particular, by defining

$$\xi_h^n \triangleq u_h^n - \pi_h u^n, \tag{4.21a}$$

$$\xi_\pi^n \triangleq u^n - \pi_h u^n, \tag{4.21b}$$

the approximation error may be decomposed as

$$u^n - u_h^n = \xi_\pi^n - \xi_h^n. \tag{4.22}$$

Since a bound on ξ_π^n may be inferred from the polynomial approximation properties and the smoothness of u^n, a suitable upper bound to the error in Eq. (4.22) is deduced by first finding an upper bound for ξ_h^n in terms of ξ_π^n, and then resorting to the triangle inequality.

More specifically, the error equation is derived according to Lemma (4.3.1).

Lemma 4.3.1 (Error equation (Di Pietro and Ern [39]))**.** *Assume* $u \in C^2([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];H^1(\Omega;\mathbb{R}^n))$ *for the exact solution. Then, the error equation satisfies*

$$\xi_h^{n+1} = \xi_h^n + \delta t A_h^n \xi_h^n - \delta t \alpha_h^n, \tag{4.23}$$

where

$$\alpha_h^n = A_h^n \xi_\pi^n + \pi_h \theta^n, \tag{4.24a}$$

$$\theta^n = \frac{1}{\delta t} \int_{t^n}^{t^{n+1}} \left(t^{n+1} - t \right) \frac{\mathrm{d}^2 u}{\mathrm{d}t^2} \, \mathrm{d}t. \tag{4.24b}$$

Proof. A second-order Taylor expansion in time with integral remainder for the exact solution u yields

$$u^{n+1} = u^n + \delta t \frac{\mathrm{d}u^n}{\mathrm{d}t} + \delta t \theta^n. \tag{4.25}$$

Projecting onto V_h and recalling the consistency property (3.33) proved in Corollary 3.2.1 at discrete time t^n provides

$$\pi_h u^{n+1} = \pi_h u^n + \delta t \pi_h \frac{\mathrm{d}u^n}{\mathrm{d}t} + \delta t \pi_h \theta^n = \pi_h u^n + \delta t (A_h^n u^n + f_h^n) + \delta t \pi_h \theta^n. \tag{4.26}$$

Subtracting the latter Eq. (4.26) from (4.13) gives

$$\xi_h^{n+1} = \xi_h^n + \delta t A_h^n (\xi_h^n - \xi_\pi^n) - \delta t \pi_h \theta^n, \tag{4.27}$$

from which the assertion follows. □

Energy estimate

The next step involves obtaining an energy estimate for the term ξ_h^{n+1} appearing in Eq. (4.23). The result is formalised in Lemma 4.3.2 below.

Lemma 4.3.2 (Energy estimate)**.** *The error Eq.* (4.23) *satisfies the following energy estimate:*

$$\begin{aligned} \frac{1}{2} \left\| \xi_h^{n+1}(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2 - \frac{1}{2} \left\| \xi_h^n(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \delta t \left| \xi_h^n(\cdot) \right|_{\varepsilon_h}^2 \le \frac{1}{2} & \left\| \xi_h^{n+1}(\cdot) - \xi_h^n(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2 \\ & - \delta t \langle \alpha_h^n, \xi_h^n \rangle_{L^2(\Omega;\mathbb{R}^n)} \\ & + \delta t \omega_h \left\| \xi_h^n(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2, \end{aligned} \tag{4.28}$$

where the constant ω_h *is given as in Eq.* (3.36).

Proof. The proof is almost identical to that of Lemma 3.17 in [39]. Taking the inner product of Eq. (4.23) on $L^2(\Omega;\mathbb{R}^n)$ and recalling the quasi-dissipativity property (3.37) of the discrete operator $(A_h, V_{h\star})$ proved in Lemma 3.2.2 yields

$$\begin{aligned} \left\langle \xi_h^{n+1}, \xi_h^n \right\rangle_{L^2(\Omega;\mathbb{R}^n)} &= \left\| \xi_h^n(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \delta t \langle A_h^n \xi_h^n, \xi_h^n \rangle_{L^2(\Omega;\mathbb{R}^n)} - \delta t \langle \alpha_h^n, \xi_h^n \rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &\le \left\| \xi_h^n(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2 - \delta t \left| \xi_h^n(\cdot) \right|_{\varepsilon_h}^2 + \delta t \omega_h \left\| \xi_h^n(\cdot) \right\|_{L^2(\Omega;\mathbb{R}^n)}^2 - \delta t \langle \alpha_h^n, \xi_h^n \rangle_{L^2(\Omega;\mathbb{R}^n)}. \end{aligned} \tag{4.29}$$

Using the identity $2ab = a^2 + b^2 - (a-b)^2$ on the left-hand side of Eq. (4.29), it may be concluded that

$$\begin{aligned}\frac{1}{2}\left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \frac{1}{2}\left\|\xi_h^{n}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \leq \frac{1}{2}\left\|\xi_h^{n+1}(\cdot) - \xi_h^{n}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left\|\xi_h^{n}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\ - \delta t\left|\xi_h^{n}(\cdot)\right|^2_{\varepsilon_h} + \delta t\,\omega_h\left\|\xi_h^{n}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\ - \delta t\langle\alpha_h^n, \xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)},\end{aligned} \tag{4.30}$$

which is the desired result. □

Bound on the discrete operator

The next step consists in deducing a suitable upper bound for the discrete operator $(A_h, V_{h\star})$, for which purpose the norms introduced according to Eqs. (3.27a) and (3.27b) are crucial, as asserted by Lemma (4.3.3).

Lemma 4.3.3 (Boundedness of the discrete operator (Di Pietro and Ern [39])). *The following inequalities hold:*

$$\left\langle A_h(t)v, w_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)} \leq C_{1\star}\left\|v(\cdot)\right\|_{h\star}\left\|w_h(\cdot)\right\|_h, \qquad \text{for } (v, w_h) \in V_{h\star}\times V_h, \tag{4.31a}$$

$$\left\|A_h(t)v_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \leq C_{2\star}\sqrt{\frac{\eta_{\mathrm{c}}}{h}}\left\|v_h(\cdot)\right\|_h, \qquad \text{for } v_h \in V_h, \tag{4.31b}$$

where the constants $C_{1\star}$ and $C_{2\star}$ are independent of h, δt, and the data $f(x,t)$, $C(t)$, and $a(x,t)$.

Proof. The proof is almost identical to that of Lemma 3.18 in [39]. To prove (4.31a), it is first convenient to integrate by part the advective term in Eq. (3.28), which, recalling that $w_h(x)$ is piecewise constant, yields

$$\begin{aligned}\left\langle A_h(t)v, w_h\right\rangle &= \int_\Omega \Big(\nabla\cdot a(x,t)\Big)v^{\mathrm{T}}(x)w_h(x)\,\mathrm{d}x - \int_\Omega \Big(a(x,t)\cdot\nu(x)\Big)^{\oplus} v^{\mathrm{T}}(x)w_h(x)\,\mathrm{d}s \\ &- \sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F a(x,t)\cdot\nu_F(x)\Big\{\Big\{v(x)\Big\}\Big\}^{\mathrm{T}}[\![w_h(x)]\!]\,\mathrm{d}s \\ &- \frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F \left|a(x,t)\cdot\nu_F(x)\right|[\![v(x)]\!]^{\mathrm{T}}[\![w_h(x)]\!]\,\mathrm{d}s \\ &+ \int_\Gamma v^{\mathrm{T}}(x)C^{\mathrm{T}}(t)\,\mathrm{d}s\int_\Omega w_h(x)\,\mathrm{d}x, \quad \text{for } (v, w_h) \in V_{h\star}\times V_h.\end{aligned} \tag{4.32}$$

Concerning the first two terms on the right-hand side of Eq. (4.32), applying Cauchy-Schwarz' inequality gives

$$\int_\Omega \Big(\nabla\cdot a(x,t)\Big)v^{\mathrm{T}}(x)w_h(x)\,\mathrm{d}x - \int_\Omega \Big(a(x,t)\cdot\nu(x)\Big)^{\oplus} v^{\mathrm{T}}(x)w_h(x)\,\mathrm{d}s \lesssim \left\|v(\cdot)\right\|_h\left\|w_h(\cdot)\right\|_h. \tag{4.33}$$

Similarly, for the fourth term:

$$\frac{1}{2}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F \left|a(x,t)\cdot\nu_F(x)\right|[\![v(x)]\!]^{\mathrm{T}}[\![w_h(x)]\!]\,\mathrm{d}s \leq \left\|v(\cdot)\right\|_h\left\|w_h(\cdot)\right\|_h. \tag{4.34}$$

Considering instead the third term, it may be deduced that

$$\begin{aligned}\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F a(x,t)\cdot\nu_F(x)\left\{\!\!\left\{v(x)\right\}\!\!\right\}^{\mathrm{T}}[\![w_h(x)]\!]\,\mathrm{d}s &\leq \left(\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F\left|a(x,t)\cdot\nu_F(x)\right|\left\|\left\{\!\!\left\{v(x)\right\}\!\!\right\}\right\|_2^2\mathrm{d}s\right)^{1/2}\\ &\quad\times\left\|w_h(\cdot)\right\|_h\\ &\leq\left(\sum_{T\in\mathscr{T}_h}\eta_{\mathrm{c}}\left\|v(\cdot)\right\|_{L^2(\partial T;\mathbb{R}^n)}^2\right)^{1/2}\left\|w_h(\cdot)\right\|_h.\end{aligned}\tag{4.35}$$

Finally, the last term may be bounded similarly as in Eq. (3.43):

$$\begin{aligned}\int_\Gamma v^{\mathrm{T}}(x)C^{\mathrm{T}}(t)\,\mathrm{d}s\int_\Omega w_h(x)\,\mathrm{d}x &\leq\sqrt{\sup_{t\in[0,T]}\left\|C(t)\right\|^2}\left\|v(\cdot)\right\|_{L^2(\Gamma;\mathbb{R}^n)}\left\|w_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\\ &=\sqrt{\inf_{(x,t)\in\Gamma\times[0,T]}\left|a(x,t)\cdot\nu(x)\right|}\left\|v(\cdot)\right\|_{L^2(\Gamma;\mathbb{R}^n)}\\ &\quad\times\sqrt{\frac{\sup_{t\in[0,T]}\left\|C(t)\right\|^2}{\inf_{(x,t)\in\Gamma\times[0,T]}\left|a(x,t)\cdot\nu(x)\right|}}\left\|w_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\\ &\lesssim\left|v(\cdot)\right|_\eta\sqrt{\varepsilon_h}\left\|w_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\lesssim\left|v(\cdot)\right|_\eta\left\|w_h(\cdot)\right\|_h.\end{aligned}\tag{4.36}$$

Combining the estimates in Eqs. (4.33), (4.34), (4.35) and (4.36), the result (4.31a) follows.

To obtain (4.31b), concerning the original definition of the operator $(A_h, V_{h\star})$ as in Eq. (3.28), observing that $(a(x,t)\cdot\nabla_h)v_h(x)=0$ since the $v(x)$ is piecewise constant, the first and third terms may be bounded together as

$$\int_\Omega\left[\left(a(x,t)\cdot\nabla_h\right)v_h(x)\right]^{\mathrm{T}}w_h(x)\,\mathrm{d}x+\int_\Gamma\left(a(x,t)\cdot\nu(x)\right)^{\ominus}v_h^{\mathrm{T}}(x)w_h(x)\,\mathrm{d}s\leq 2\left\|v_h(\cdot)\right\|_h\left\|w_h(\cdot)\right\|_h.\tag{4.37}$$

Moreover, using the Cauchy-Schwarz' and the discrete trace inequalities to bound the fourth term provides

$$\sum_{F\in\mathscr{F}_h^{\mathrm{i}}}\int_F a(x,t)\cdot\nu_F(x)\,[\![v(x)]\!]^{\mathrm{T}}\left\{\!\!\left\{w_h(x)\right\}\!\!\right\}\mathrm{d}s\leq\sqrt{\frac{\eta_{\mathrm{c}}}{h}}\left\|v_h(\cdot)\right\|_h\left\|w_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}.\tag{4.38}$$

Combining the estimates in Eqs. (4.37) and (4.38) with those already derived in (4.34) and (4.36), and recalling the fact that $\|w_h\|_h\lesssim\sqrt{\eta_{\mathrm{c}}/h}\|w_h\|_{L^2(\Omega;\mathbb{R}^n)}$ since $h\leq\eta_{\mathrm{c}}t_\star$ together with

$$\left\|A_h(t)v_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}=\sup_{w_h\in V_h\setminus\{0\}}\frac{\left\langle A_h(t)v_h,w_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)}}{\left\|w_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}},\tag{4.39}$$

leads to the desired result. □

Stability

The penultimate step involves the derivation of upper bounds for terms appearing on the right-hand side of the energy estimate in Eq. (4.28). In particular, the first quantity relates to the anti-dissipative nature of the forward Euler scheme, whereas the second one contains the contribution of the space approximation error. The stability of the considered scheme is proved in Lemma 4.3.4.

Lemma 4.3.4 (Stability of the forward Euler scheme)**.** *Assume that $u \in C^2([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^0([0,T];H^1(\Omega;\mathbb{R}^n))$ and that the CFL condition in Eq. (4.10) holds with $\rho \leq (2C_{2\star})^{-2}$, i.e.,*

$$\delta t \leq \frac{h}{4C_{2\star}^2\eta_{\mathrm{c}}}. \tag{4.40}$$

Then, there exists $C_\star$ independent of h, δt, $f(x,t)$, $C(t)$, and $a(x,t)$, such that

$$\left\|\xi_h^{n+1}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \delta t\left|\xi_h^n(\cdot)\right|_{\frac{\varepsilon_h}{2}}^2 \leq \left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \delta t C_\star\left(\left\|\xi_\pi^n(\cdot)\right\|_{h\star}^2 + \frac{1}{t_\star}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \delta t^2 t_\star C_u^2\right), \tag{4.41}$$

where the seminorm $|\cdot|_{\frac{\varepsilon_h}{2}}$ reads as in Eq. (3.26d).

Proof. The proof is almost identical to that of Lemma 3.20 in [39]. In particular, the three terms on the right-hand side of Eq. (4.28), renamed $\mathscr{T}_1$, $\mathscr{T}_2$, and $\mathscr{T}_3$ for convenience, need to be bounded. Starting with the third term, it holds that

$$|\mathscr{T}_3| \triangleq \delta t\,\omega_h\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 \lesssim \frac{\delta t}{t_{\mathrm{c}}}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2. \tag{4.42}$$

To bound $\mathscr{T}_2$, the inequality (4.31a) is crucial. Recalling the definition provides

$$|\mathscr{T}_2| = \delta t\left|\langle\alpha_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}\right| \leq \delta t\left|\langle A_h^n\xi_\pi^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}\right| + \delta t\left|\langle\pi_h\theta^n\xi_\pi^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}\right|. \tag{4.43}$$

Resorting to the bound (4.31a) and applying the generalised verison of Young's inequality to the first term on the right-hand side yields

$$\delta t\left|\langle A_h^n\xi_\pi^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}\right| \leq \delta t C_{1\star}\left\|\xi_\pi^n(\cdot)\right\|_{h\star}\left\|\xi_h^n(\cdot)\right\|_h \leq \frac{1}{4}\delta t\left|\xi_h^n(\cdot)\right|_\eta^2 + \delta t C\left(\left\|\xi_\pi^n(\cdot)\right\|_{h\star}^2 + \frac{1}{t_{\mathrm{c}}}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2\right). \tag{4.44}$$

Concerning instead the second term on the right-hand side of Eq. (4.43), it may be deduced that

$$\left\|\theta^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 \leq \frac{1}{3}\delta t^2 \max_{t\in[t^k,t^{k+1}]}\left\|\frac{\partial^2 u(\cdot,t)}{\partial t^2}\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 \lesssim \delta t^2 C_u^2. \tag{4.45}$$

Consequently

$$\delta t\left|\langle \pi_h\theta^n\xi_\pi^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}\right| = \delta t\left|\langle \theta^n\xi_\pi^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}\right| \leq \delta t\left\|\theta^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}$$
$$\leq \delta t^3 t_\star C_u^2 + \frac{\delta t}{t_\star}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}. \tag{4.46}$$

Collecting the two bounds and observing that $t_\star \leq t_c$ gives finally

$$|\mathscr{T}_2| \leq \frac{1}{4}\delta t\left|\xi_h^n(\cdot)\right|^2_\eta + \delta t C\left(\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \frac{1}{t_\star}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t^2 t_\star C_u^2\right). \tag{4.47}$$

To bound the first term, the estimate (4.31b) is essential. First, from the error Eq. (4.23) it follows that

$$|\mathscr{T}_1| \triangleq \frac{1}{2}\left\|\xi_h^{n+1}(\cdot)-\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} = \frac{1}{2}\delta t^2\left\|A_h^n\xi_h^n(\cdot)-\alpha_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}$$
$$\leq \delta t^2\left\|A_h^n\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t^2\left\|\alpha_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}. \tag{4.48}$$

Concerning the first term on the right-hand side of Eq. (4.48), resorting to the bound (4.31b) and specifying the CFL condition (4.10) gives

$$\delta t^2\left\|A_h^n\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \leq \delta t^2 C_{2\star}^2\frac{\eta_c}{h}\left\|\xi_h^n(\cdot)\right\|^2_h \leq \frac{1}{4}\left\|\xi_h^n(\cdot)\right\|^2_h = \frac{1}{4}\delta t\left|\xi_h^n(\cdot)\right|^2_\eta + \frac{1}{4}\frac{\delta t}{t_c}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}. \tag{4.49}$$

Considering the second term, using the inequality (4.31a), the definition (4.39) and the estimate (4.45) in combination with the CFL condition (4.10) leads to

$$\delta t^2\left\|\alpha_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \lesssim \delta t^2\left\|A_h^n\xi_\pi^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t^2\left\|\pi_h\theta^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}$$
$$\lesssim \delta t^2\frac{\eta_c}{h}\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \delta t^2\left\|\theta^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \lesssim \delta t\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \delta t^4 C_u^2. \tag{4.50}$$

Combining Eqs. (4.49) and (4.50) and recalling that $\delta t \leq t_\star$ provides

$$|\mathscr{T}_1| \leq \frac{1}{4}\delta t\left|\xi_h^n(\cdot)\right|^2_\eta + \delta t C\left(\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \frac{1}{t_c}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t^2 t_\star C_u^2\right). \tag{4.51}$$

Hence, collecting the estimates derived according to Eqs. (4.42), (4.47), and (4.51) yields

$$|\mathscr{T}_1|+|\mathscr{T}_2|+|\mathscr{T}_3| \leq \frac{1}{2}\delta t\left|\xi_h^n(\cdot)\right|^2_\eta + \delta t C\left(\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \frac{1}{t_\star}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t^2 t_\star C_u^2\right). \tag{4.52}$$

□

Proof of Theorem 4.2.1

Based on the results advocated above, it is finally possible to establish the proof Theorem 4.2.1, asserting the convergence of the proposed forward Euler schemes.

Proof of Theorem 4.2.1. The proof is almost identical to that of Theorem 3.7 in [39]. For all $n \in \mathbb{N}_0$, set

$$a^n \triangleq \left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}, \qquad b^n \triangleq \delta t \left|\xi_h^n(\cdot)\right|^2_{\frac{\varepsilon_h}{2}}, \qquad d^n \triangleq \delta t C_\star \left(\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \delta t^2 t_\star C_u^2 \right), \tag{4.53}$$

so that the stability estimate may be recast more conveniently as

$$a^{n+1} + b^n \le (1+\gamma) a^n + d^n, \qquad n \in \mathbb{N}_0, \tag{4.54}$$

where $\gamma \triangleq \delta t C_\star / t_\star$. Then, it may be easily proved by induction that

$$a^{n+1} + \sum_{m=0}^{n} (1+\gamma)^{n-m} b^m \le (1+\gamma)^{n+1} a^0 + \sum_{m=0}^{n} (1+\gamma)^{n-m} d^m. \tag{4.55}$$

Specifying $n = N-1$ and observing that $1 \le (1+\gamma)^{n-m} \le (1+\gamma)^N \le \mathrm{e}^{N\gamma} = \mathrm{e}^{C_\star T/t_\star}$ and that $a^0 = 0$ since $u_h^0 = \pi_h u_0$ and $\xi_h^0 = 0$ yields

$$a^N + \sum_{n=0}^{N-1} b^n \le \mathrm{e}^{C_\star T/t_\star} \sum_{n=0}^{N-1} d^n. \tag{4.56}$$

From an application of the triangle inequality, and recalling that $\xi_\pi^n = u^n - \pi_h u^n$ and that $\left|\xi_\pi^n(\cdot)\right|_{\frac{\varepsilon_h}{2}} \le \left\|\xi_\pi^n(\cdot)\right\|_{h\star}$, it follows that

$$\begin{aligned} \left\|u^N(\cdot) - u_h^N(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left(\sum_{n=0}^{N-1} \delta t \left|u^n(\cdot) - u_h^n(\cdot)\right|^2_{\frac{\varepsilon_h}{2}} \right)^{1/2} \lesssim & \left(\mathrm{e}^{C_\star T/t_\star} \sum_{n=0}^{N-1} d^n \right)^{1/2} \\ & + \left\|\xi_\pi^N(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}. \end{aligned} \tag{4.57}$$

Resorting to the definition of the norm $\|\cdot\|_{h\star}$ as in Eq. (3.27b), the polynomial approximation properties asserted by Lemmata 3.1.4 and 3.1.5, and to Assumption 3.2.1, it may be inferred that

$$\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} \lesssim \eta_c h \left\|u^n(\cdot)\right\|^2_{H^1(\Omega;\mathbb{R}^n)}, \tag{4.58}$$

and hence

$$\begin{aligned} \sum_{n=0}^{N-1} d^n \lesssim \sum_{n=0}^{N-1} \delta t C_\star \left(\left\|\xi_\pi^n(\cdot)\right\|^2_{h\star} + \delta t^2 t_\star C_u^2 \right) &\le \delta t^2 t_\star T C_u^2 + h \eta_c T \left\|u(\cdot,\cdot)\right\|^2_{C^0([0,T];H^1(\Omega;\mathbb{R}^n))} \\ &= \delta t^2 \chi_1^2 + h \chi_2^2, \end{aligned} \tag{4.59}$$

with χ_1 and χ_2 defined according to Eqs. (4.15). Finally, invoking again Assumption 3.2.1 provides

$$\left\|\xi_\pi^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim h \left\|u^N(\cdot)\right\|_{H^1(\Omega;\mathbb{R}^n)} \le \sqrt{\frac{h^2}{\eta_c T}} \chi_2 \le \sqrt{h} \chi_2. \tag{4.60}$$

Combining the bounds derived above, the desired result follows. □

4.3.2 Explicit RK2 schemes

The error analysis for the forward RK2 schemes follows a similar *iter* to that already outlined for the forward Euler approximation. The main difference resides in the fact that bounds on the discrete operator $(A_h, V_{h\star})$ are more difficult to infer, since terms involving the broken gradient do not identically vanish for $k \geq 1$. The first step in the derivation of the main result consists in deducing of a discrete time equation governing the evolution of the approximation error.

Error equation

The present Section is devoted to deriving the error equation. In particular, by defining the quantities ξ_h^n and ξ_π^n as in Eqs. (4.21) and introducing

$$\zeta_h^n \triangleq w_h^n - \pi_h w^n, \tag{4.61a}$$
$$\zeta_\pi^n \triangleq w^n - \pi_h w^n, \tag{4.61b}$$

the errors may be decomposed as

$$u^n - u_h^n = \xi_\pi^n - \xi_h^n, \tag{4.62a}$$
$$w^n - w_h^n = \zeta_\pi^n - \zeta_h^n, \tag{4.62b}$$

with

$$w \triangleq u + \delta t \frac{\mathrm{d}u}{\mathrm{d}t}. \tag{4.63}$$

Departing from Eqs. (4.21), (4.61), and (4.63), it is possible to derive the error equations as in the following Lemma 4.3.5.

Lemma 4.3.5. *Assume $u \in C^3([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^1([0,T];H^1(\Omega;\mathbb{R}^n))$ for the exact solution, $f \in C^2([0,T];L^2(\Omega;\mathbb{R}^n))$, and $A \in C^2([0,T];\mathscr{L}(H^1(\Omega;\mathbb{R}^n);L^2(\Omega;\mathbb{R}^n)))$. Then, the error equation satisfies*

$$\zeta_h^n = \xi_h^n + \delta t A_h^n \xi_h^n - \delta t \alpha_h^n, \tag{4.64a}$$
$$\xi_h^{n+1} = \frac{\xi_h^n}{2\lambda} + \left(1 - \frac{1}{2\lambda}\right)\zeta_h^n + \frac{\delta t}{2} A_h^{n+\lambda}\left[\left(\frac{1}{\lambda} - 1\right)\xi_h^n + \zeta_h^n\right] - \frac{\delta t}{2}\beta_h^n, \tag{4.64b}$$

where

$$\alpha_h^n \triangleq A_h^n \xi_\pi^n, \tag{4.65a}$$
$$\beta_h^n \triangleq A_h^{n+\lambda}\left[\left(\frac{1}{\lambda} - 1\right)\xi_\pi^n + \zeta_\pi^n\right] - \delta t(\pi_h \Lambda^n u^n + \pi_h F^n) - \lambda\,\delta t^2 \pi_h\left(\frac{\mathrm{d}A^n}{\mathrm{d}t} + \Lambda^n\right)\frac{\mathrm{d}u^n}{\mathrm{d}t} + \pi_h \theta^n, \tag{4.65b}$$

with

$$\theta^n \triangleq \frac{1}{\delta t}\int_{t^n}^{t^{n+1}} \left(t^{n+1} - t\right)^2 \frac{\mathrm{d}^3 u(t)}{\mathrm{d}t^3}\,\mathrm{d}t, \tag{4.66a}$$
$$\Lambda^n \triangleq \frac{1}{\lambda\,\delta t}\int_{t^n}^{t^{n+\lambda}} \left(t^{n+\lambda} - t\right) \frac{\mathrm{d}^2 A(t)}{\mathrm{d}t^2}\,\mathrm{d}t, \tag{4.66b}$$
$$F^n \triangleq \frac{1}{\lambda\,\delta t}\int_{t^n}^{t^{n+\lambda}} \left(t^{n+\lambda} - t\right) \frac{\mathrm{d}^2 f(t)}{\mathrm{d}t^2}\,\mathrm{d}t. \tag{4.66c}$$

Proof. From Corollary 3.2.1, consistency at discrete time t^n yields

$$\pi_h w^n = \pi_h u^n + \delta t \pi_h \frac{\mathrm{d}u^n}{\mathrm{d}t} = \pi_h u^n + \delta t A_h^n u^n + \delta t f_h^n. \tag{4.67}$$

Subtracting the above Eq. (4.67) from (4.16a) and defining α_h^n according to (4.65a) provides (4.64a). Moreover, a second-order Taylor expansion with integral remainder gives

$$u^{n+1} = u^n + \delta t \frac{\mathrm{d}u^n}{\mathrm{d}t} + \frac{\delta t^2}{2}\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} + \frac{\delta t}{2}\theta^n = w^n + \frac{\delta t}{2}A^n(w^n - u^n) + \frac{\delta t^2}{2}\frac{\mathrm{d}A^n}{\mathrm{d}t}u^n + \frac{\delta t^2}{2}\frac{\mathrm{d}f^n}{\mathrm{d}t} + \frac{\delta t}{2}\theta^n, \tag{4.68}$$

with θ^n reading as in Eq. (4.66a). Analogously, a first-order Taylor expansion with integral remainder of $A^{n+\lambda}$ gives

$$A^{n+\lambda} = A^n + \lambda\,\delta t \frac{\mathrm{d}A^n}{\mathrm{d}t} + \lambda\,\delta t \Lambda^n, \tag{4.69}$$

with Λ^n defined according to Eq. (4.66b). Substituting the latter expression into Eq. (4.68) and performing a similar first-order expansion for $f^{n+\lambda}$ provides, after some manipulations,

$$u^{n+1} = \frac{u^n}{2\lambda} + \left(1 - \frac{1}{2\lambda}\right)w^n + \frac{\delta t}{2}A^{n+\lambda}\left[\left(\frac{1}{\lambda} - 1\right)u^n + w^n\right] + \frac{\delta t}{2}\delta^n, \tag{4.70}$$

with

$$\begin{aligned}
\delta^n &\triangleq -\delta t(\Lambda^n u^n + F^n) + \left(A^{n+\lambda} - A^n\right)(u^n - w^n) + \frac{f^{n+\lambda}}{\lambda} + \theta^n \\
&= -\delta t(\Lambda^n u^n + F^n) - \lambda\,\delta t^2\left(\frac{\mathrm{d}A^n}{\mathrm{d}t} + \Lambda^n\right)\frac{\mathrm{d}u^n}{\mathrm{d}t} + \frac{f^{n+\lambda}}{\lambda} + \theta^n,
\end{aligned} \tag{4.71}$$

where F^n reads as in Eq. (4.66c).

Projecting Eq. (4.70) onto V_h and invoking the time-shifted consistency property proved in Lemma 3.2.1 for both $A_h^{n+\lambda}u^n$ and $A_h^{n+\lambda}w^n$ therefore yields

$$\pi_h u^{n+1} = \frac{\pi_h u^n}{2\lambda} + \left(1 - \frac{1}{2\lambda}\right)\pi_h w^n + \frac{\delta t}{2}A_h^{n+\lambda}\left[\left(\frac{1}{\lambda} - 1\right)u^n + w^n\right] + \frac{\delta t}{2}\pi_h\delta^n, \tag{4.72}$$

in which

$$\pi_h\delta^n \triangleq -\delta t(\pi_h\Lambda^n u^n + \pi_h F^n) - \lambda\,\delta t^2 \pi_h\left(\frac{\mathrm{d}A^n}{\mathrm{d}t} + \Lambda^n\right)\frac{\mathrm{d}u^n}{\mathrm{d}t} + \frac{f_h^{n+\lambda}}{\lambda} + \pi_h\theta^n. \tag{4.73}$$

Subtracting Eq. (4.72) from (4.16b) and defining β_h^n as in (4.65b) gives the desired result. □

Energy estimate

The next step involves obtaining an energy estimate for the term ξ_h^{n+1} appearing in Eq. (4.64b). The result is formalised in Lemma 4.3.6 below.

Lemma 4.3.6 (Energy estimate)**.** *The error Eqs.* (4.64) *satisfy the following energy estimate:*

$$
\begin{aligned}
\left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} &+ \frac{2\lambda-1}{\lambda}\delta t\left|\xi_h^n(\cdot)\right|^2_{\varepsilon_h} + \lambda\delta t\left|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right|^2_{\varepsilon_h} \\
&\le (1-\lambda)\left\|\xi_h^{n+1}(\cdot)-\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \lambda\left\|\xi_h^{n+1}(\cdot)-\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\
&\quad + \left(1+\frac{2\lambda-1}{\lambda}\delta t\omega_h\right)\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\
&\quad + \lambda\delta t\omega_h\left\|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\
&\quad - \frac{2\lambda-1}{\lambda}\delta t\langle\alpha_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} - \lambda\delta t\left\langle\beta_h^n,\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)},
\end{aligned}
\tag{4.74}
$$

where the constant ω_h *is given as in Eq.* (3.36)*.*

Proof. Taking the inner product of Eq. (4.64a) with $(2\lambda-1)/\lambda\,\xi_h^n(x)$ on $L^2(\Omega;\mathbb{R}^n)$ provides

$$
\langle\xi_h^n,\zeta_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} = \left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t\langle A_h^n\xi_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} - \delta t\langle\alpha_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}. \tag{4.75}
$$

Similarly, taking the inner product of Eq. (4.64b) with $2(1-\lambda)\xi_h^k(x)+2\lambda\zeta_h^k(x)$ on $L^2(\Omega;\mathbb{R}^n)$ yields

$$
\begin{aligned}
2(1-\lambda)\left\langle\xi_h^{n+1},\xi_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} &+ 2\lambda\left\langle\xi_h^{n+1},\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} = \frac{1-\lambda}{\lambda}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\
&+ (2\lambda-1)\left\|\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \frac{4\lambda-2\lambda^2-1}{\lambda}\langle\xi_h^n,\zeta_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} \\
&+ \lambda\delta t\left\langle A_h^{n+\lambda}\left(\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right),\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} \\
&- \lambda\delta t\left\langle\beta_h^n,\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)}.
\end{aligned}
\tag{4.76}
$$

Observing that

$$
2\left\langle\xi_h^{n+1},\xi_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} = \left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} - \left\|\xi_h^{n+1}(\cdot)-\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}, \tag{4.77a}
$$

$$
2\left\langle\xi_h^{n+1},\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} = \left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \left\|\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} - \left\|\xi_h^{n+1}(\cdot)-\zeta_h^k(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}, \tag{4.77b}
$$

gives

$$\begin{aligned}\left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} &= (1-\lambda)\left\|\xi_h^{n+1}(\cdot)-\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \lambda\left\|\xi_h^{n+1}(\cdot)-\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \frac{(1-\lambda)^2}{\lambda}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + (\lambda-1)\left\|\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \frac{4\lambda-2\lambda^2-1}{\lambda}\langle\xi_h^n,\zeta_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \lambda\,\delta t\left\langle A_h^{n+\lambda}\left(\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right), \frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &- \lambda\,\delta t\left\langle \beta_h^n, \frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)}.\end{aligned} \tag{4.78}$$

Hence, multipying Eq. (4.75) by $(2\lambda-1)/\lambda$ and adding the resulting expression to the above (4.78) provides

$$\begin{aligned}\left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} &= (1-\lambda)\left\|\xi_h^{n+1}(\cdot)-\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \lambda\left\|\xi_h^{n+1}(\cdot)-\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \lambda\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + (\lambda-1)\left\|\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + 2(1-\lambda)\langle\xi_h^n,\zeta_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \frac{2\lambda-1}{\lambda}\delta t\langle A_h^n\xi_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} - \frac{2\lambda-1}{\lambda}\delta t\langle\alpha_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \lambda\,\delta t\left\langle A_h^{n+\lambda}\left(\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right), \frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &- \lambda\,\delta t\left\langle \beta_h^n, \frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)}.\end{aligned} \tag{4.79}$$

Resorting to Young's inequality for products to bound the term $\langle\xi_h^n,\zeta_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)}$ yields then

$$\begin{aligned}\left\|\xi_h^{n+1}(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} &= (1-\lambda)\left\|\xi_h^{n+1}(\cdot)-\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \lambda\left\|\xi_h^{n+1}(\cdot)-\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \frac{2\lambda-1}{\lambda}\delta t\langle A_h^n\xi_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &- \frac{2\lambda-1}{\lambda}\delta t\langle\alpha_h^n,\xi_h^n\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &+ \lambda\,\delta t\left\langle A_h^{n+\lambda}\left(\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right), \frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} \\ &- \lambda\,\delta t\left\langle \beta_h^n, \frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)}.\end{aligned} \tag{4.80}$$

Finally, recalling the quasi-disspativity property (3.37) of the discrete operator $(A_h, V_{h\star})$ proved in Lemma 3.2.2 leads to the desired result. □

Preliminary stability bounds

The next result, formalised in Lemma 4.3.7, delivers some preliminary stability bounds that are necessary to ensure stability and convergence of the considered RK2 schemes.

For what follows, the additional norm is introduced on $V_{h\star}$:

$$\|v(\cdot)\|_{\star\star}^2 \triangleq \|v(\cdot)\|_{h\star}^2 + \eta_c h \sum_{i=1}^{n} \|\nabla_h v_i(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)}^2. \tag{4.81}$$

Furthermore, to allow for ease of notation, the following energy-like quantity, collecting the contributions of the space and time approximation errors, is defined:

$$\mathscr{E}_h^n \triangleq \|\xi_\pi^n(\cdot)\|_{\star\star} + \|\zeta_\pi^n(\cdot)\|_{\star\star} + \sqrt{t_\star}\left(C_{fu}\delta t^2 + C_u \delta t^3\right) + \frac{1}{\sqrt{t_\star}}\|\xi_h^n(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)}. \tag{4.82}$$

where C_{fu} and C_u read as in Eqs. (4.19).

Lemma 4.3.7 (Preliminary stability bounds)**.** *Assume $u \in C^3([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^1([0,T];H^1(\Omega;\mathbb{R}^n))$ for the exact solution, $f \in C^2([0,T];L^2(\Omega;\mathbb{R}^n))$, and $A \in C^2([0,T];\mathscr{L}(H^1(\Omega;\mathbb{R}^n);L^2(\Omega;\mathbb{R}^n)))$. Then, if the CFL condition in Eq. (4.10) holds, there exists $C_\star$ independent of h, δt, and the data $f(x,t)$, $C(t)$, and $a(x,t)$ such that*

$$\begin{aligned}
&\left\|\xi_h^{n+1}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \frac{2\lambda-1}{2\lambda}\delta t\left|\xi_h^n(\cdot)\right|_{\frac{\varepsilon_h}{2}}^2 + \frac{\lambda}{2}\delta t\left|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot) + \zeta_h^n(\cdot)\right|_{\frac{\varepsilon_h}{2}}^2 \\
&\quad\lesssim \left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + (1-\lambda)\left\|\xi_h^{n+1}(\cdot) - \xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \lambda\left\|\xi_h^{n+1}(\cdot) - \zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \\
&\quad + C_\star \delta t (\mathscr{E}_h^n)^2.
\end{aligned} \tag{4.83}$$

Proof. The proof involves four different steps. First, it is shown that

$$\|A_h(t)v(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \sqrt{\frac{\eta_c}{h}}\|v(\cdot)\|_{\star\star}, \qquad \text{for } v \in V_{h\star}. \tag{4.84}$$

To this end, from the definition of the norm $\|\cdot\|_{\star\star}$ according to Eq. (4.81) and Assumption 3.2.1, it may be deduced that

$$\begin{aligned}
\left\langle A_h(t)v, w_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)} &\lesssim \left(\frac{1}{t_c}\|v(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} + \eta_c\|\nabla_h v(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} + \sqrt{\varepsilon_h}|v(\cdot)|_\eta\right)\|w_h(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} \\
&\quad + |v(\cdot)|_\eta |w_h(\cdot)|_\eta + |v(\cdot)|_\eta \left(\sum_{F\in\mathscr{F}_h^i}\int_F |a(x,t)\cdot \nu_F(x)| \left\{\!\!\left\{w_h(x)\right\}\!\!\right\} \mathrm{d}x\right)^{1/2} \\
&\lesssim \left(\frac{1}{t_c}\|v(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} + \eta_c\|\nabla_h v(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} + \sqrt{\frac{\eta_c}{h}}|v(\cdot)|_\eta\right)\|w_h(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)} \\
&\lesssim \sqrt{\frac{\eta_c}{h}}\|v(\cdot)\|_{\star\star}\|w_h(\cdot)\|_{L^2(\Omega;\mathbb{R}^n)}, \quad \text{for } (v,w_h) \in V_{h\star}\times V_h.
\end{aligned} \tag{4.85}$$

Since

$$\left\|A_h(t)v(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} = \sup_{w_h\in V_h\setminus\{0\}} \frac{\left\langle A_h(t)v, w_h\right\rangle_{L^2(\Omega;\mathbb{R}^n)}}{\left\|w_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}}, \tag{4.86}$$

Eq. (4.85) provides (4.84). Moreover, applying the inverse and trace inequalities, yields

$$\left\|v_h(\cdot)\right\|_{\star\star} \lesssim \sqrt{\frac{\eta_c}{h}}\left\|v_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}, \qquad \text{for } v\in V_h, \tag{4.87}$$

which, combined with Eq. (4.84), leads to

$$\left\|A_h(t)v_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \frac{\eta_c}{h}\left\|v_h(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}, \qquad \text{for } v\in V_h. \tag{4.88}$$

The next step consists in deriving upper bounds for the terms $\alpha_h^n(x)$ and $\beta_h^n(x)$. Concerning the first quantity, the bound in Eq. (4.84) and the usual CFL condition (4.10) imply

$$\sqrt{\delta t}\left\|\alpha_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \left\|\xi_\pi^n(\cdot)\right\|_{\star\star} \leq \mathscr{E}_h^n. \tag{4.89}$$

Moving to the analysis of the term $\beta_h^n(x)$, using the triangle inequality gives

$$\begin{aligned}\left\|\beta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \leq{}& \left(\frac{1}{\lambda}-1\right)\left\|A_h^{n+\lambda}\xi_\pi^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left\|A_h^{n+\lambda}\zeta_\pi^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\\ &+\delta t\left\|\pi_h\Lambda^n u^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \delta t\left\|\pi_h F^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left\|\pi_h\theta^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\\ &+\lambda\delta t^2\left\|\pi_h\frac{\partial A^n}{\partial t}\frac{\partial u^n}{\partial t}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \lambda\delta t^2\left\|\Lambda^n\frac{\partial u^n}{\partial t}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}.\end{aligned} \tag{4.90}$$

The first two terms appearing in Eq. (4.90) may be bounded as

$$\sqrt{\delta t}\left\|A_h^{n+\lambda}\xi_\pi^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \left\|\xi_\pi^n(\cdot)\right\|_{\star\star}, \tag{4.91a}$$

$$\sqrt{\delta t}\left\|A_h^{n+\lambda}\zeta_\pi^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \left\|\zeta_\pi^n(\cdot)\right\|_{\star\star}. \tag{4.91b}$$

Moreover, proceeding similarly as in the derivation of Eq. (4.46), it may be deduced that

$$\left\|\pi_h F^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \leq \left\|F^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \lambda\delta t\left\|\frac{\partial^2 f(\cdot,t)}{\partial t^2}\right\|_{C^0([0,T];L^2(\Omega;\mathbb{R}^n))}, \tag{4.92a}$$

$$\left\|\pi_h\theta^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \leq \left\|\theta^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \delta t^2\left\|\frac{\partial^3 u(\cdot,t)}{\partial t^3}\right\|_{C^0([0,T];L^2(\Omega;\mathbb{R}^n))}. \tag{4.92b}$$

Finally, the third, second last and last terms may instead be bounded as

$$\left\|\Lambda^n u^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \lambda\delta t\eta_2\left\|u(\cdot,\cdot)\right\|_{C^0([0,T];H^1(\Omega;\mathbb{R}^n))}, \tag{4.93a}$$

$$\left\|\Lambda^n\frac{\partial u^n}{\partial t}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \lambda\delta t\eta_2\left\|\frac{\partial u(\cdot,\cdot)}{\partial t}\right\|_{C^0([0,T];H^1(\Omega;\mathbb{R}^n))}, \tag{4.93b}$$

$$\left\|\pi_h \frac{\partial A^n}{\partial t}\frac{\partial u^n}{\partial t}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \eta_1 \left\|\frac{\partial u(\cdot,\cdot)}{\partial t}\right\|_{C^0([0,T];H^1(\Omega;\mathbb{R}^n))}. \tag{4.93c}$$

Combining all the above estimates and recalling that $t \leq t_\star$ therefore yields

$$\sqrt{\delta t}\left\|\beta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \left\|\xi_\pi^n(\cdot)\right\|_{\star\star} + \left\|\zeta_\pi^n(\cdot)\right\|_{\star\star} + \sqrt{t}\left(C_{fu}\delta t^2 + C_u\delta t^3\right) \leq \mathscr{E}_h^n, \tag{4.94}$$

with the constants C_{fu} and C_u defined according to Eqs. (4.19). Next, it is necessary to bound the term $\zeta_h^n(x)$. Departing from the error equation (4.64a), and resorting to the triangle inequality, the bounds (4.88) and (4.89) and the usual CFL condition (4.10) provides

$$\begin{aligned}\left\|\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} &\leq \left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \delta t\left\|A_h^n\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \delta t\left\|\alpha_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\\ &\leq \left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \delta t\frac{\eta_c}{h}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \sqrt{\delta t}\left\|\xi_\pi^n(\cdot)\right\|_{\star\star}\\ &\lesssim \left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \sqrt{\delta t}\left\|\xi_\pi^n(\cdot)\right\|_{\star\star}.\end{aligned} \tag{4.95}$$

Since $\delta t \leq t_\star \leq t_c$, the above expression gives finally

$$\frac{\delta t}{t_c}\left\|\zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \leq \frac{\delta t}{t_c}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} + \delta t\left\|\xi_\pi^n(\cdot)\right\|^2_{\star\star}. \tag{4.96}$$

Lastly, bounds on the quantities $\delta t\langle\alpha_h^n,\xi_h^n\rangle$ and $\delta t\langle\beta_h^n,(1-\lambda)\xi_h^n/\lambda+\zeta_h^n\rangle$ should be deduced. Using the boundedness on orthogonal subscales (3.42) stated in Proposition 3.2.1, it is first possible to infer that

$$\begin{aligned}\delta t\langle\alpha_h^n,\xi_h^n\rangle &\triangleq \delta t\langle A_h^n\xi_\pi^n,\xi_h^n\rangle \lesssim \delta t\left\|\xi_\pi^n(\cdot)\right\|_{\star\star}\left\|\xi_h^n(\cdot)\right\|_h\\ &\lesssim \delta t\left\|\xi_\pi^n(\cdot)\right\|_{\star\star}\left(\left|\xi_h^n(\cdot)\right|_\eta + \frac{1}{\sqrt{t_c}}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\right).\end{aligned} \tag{4.97}$$

Hence, an application of Young's inequality for product yields

$$\delta t\langle\alpha_h^n,\xi_h^n\rangle - \frac{\delta t}{2}\left|\xi_h^n(\cdot)\right|^2_\eta \lesssim \delta t\left\|\xi_\pi^n(\cdot)\right\|^2_{\star\star} + \frac{\delta t}{t_c}\left\|\xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \leq \delta t(\mathscr{E}_h^n)^2. \tag{4.98}$$

Similarly, resorting again to Eq. (3.42), the bound previously deduced according to (4.94), and Cauchy-Schwarz' inequality provides

$$\begin{aligned}\delta t\left\langle\beta_h^n,\frac{1-\lambda}{\lambda}\xi_h^n+\zeta_h^n\right\rangle_{L^2(\Omega;\mathbb{R}^n)} &\lesssim \delta t\left\|\frac{1-\lambda}{\lambda}\xi_\pi^n(\cdot)+\zeta_\pi^n(\cdot)\right\|_{\star\star}\left\|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right\|_h\\ &\quad + \delta t\left(C_{fu}\delta t^2 + C_u\delta t^3\right)\left\|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\\ &\lesssim \delta t\left(\frac{1-\lambda}{\lambda}\left\|\xi_\pi^n(\cdot)\right\|_{\star\star} + \left\|\zeta_\pi^n(\cdot)\right\|_{\star\star}\right)\\ &\quad\times\left(\left|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right|_\eta + \frac{1}{\sqrt{t_c}}\left\|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\right)\\ &\quad + \delta t\left(C_{fu}\delta t^2 + C_u\delta t^3\right)\left\|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot)+\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}.\end{aligned} \tag{4.99}$$

Using again Young's inequality for product therefore gives

$$\delta t \left\langle \beta_h^n, \frac{1-\lambda}{\lambda}\xi_h^n + \zeta_h^n \right\rangle_{L^2(\Omega;\mathbb{R}^n)} - \frac{\delta t}{2}\left|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot) + \zeta_h^n(\cdot)\right|_\eta^2 \lesssim \delta t\left(\left\|\xi_\pi^n(\cdot)\right\|_{\star\star}^2 + \left\|\zeta_\pi^n(\cdot)\right\|_{\star\star}^2\right)$$
$$+ \frac{\delta t}{t_c}\left(\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \left\|\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2\right)$$
$$+ \delta t\left(C_{fu}\delta t^2 + C_u \delta t^3\right)\left(\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \left\|\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}\right). \tag{4.100}$$

Invoking the estimate (4.95) to bound the terms involving $\left\|\zeta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}$, it may be finally deduced that

$$\delta t \left\langle \beta_h^n, \frac{1-\lambda}{\lambda}\xi_h^n + \zeta_h^n \right\rangle_{L^2(\Omega;\mathbb{R}^n)} - \frac{\delta t}{2}\left|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot) + \zeta_h^n(\cdot)\right|_\eta^2 \lesssim \delta t (\mathscr{E}_h^n)^2. \tag{4.101}$$

Recalling that $\omega_h \le 1/t_c \le 1/t_\star$, collecting the above bounds and inserting them into Eq. (4.74) leads to the desired result. □

Stability

Stability for the RK2 schemes is finally proved by inferring opportune bounds on the anti-dissipative terms appearing on the right-hand side of energy estimate in Eq. (4.74). To this end, the 2-CFL condition (4.12) is invoked. Lemma 4.3.8 asserts the result.

Lemma 4.3.8 (Stability of RK2 schemes)**.** *Assume $u \in C^3([0,T];L^2(\Omega;\mathbb{R}^n)) \cap C^1([0,T];H^1(\Omega;\mathbb{R}^n))$ for the exact solution, $f \in C^2([0,T];L^2(\Omega;\mathbb{R}^n))$, and $A \in C^2([0,T];\mathscr{L}(H^1(\Omega;\mathbb{R}^n);L^2(\Omega;\mathbb{R}^n)))$. Then, if the 2-CFL condition in Eq. (4.12) holds for some $\rho' \in \mathbb{R}_{>0}$, there exists $C_\star$ independent of h, δt, and the data $f(x,t)$, $C(t)$, and $a(x,t)$ such that*

$$\left\|\xi_h^{n+1}(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2 + \frac{2\lambda-1}{2\lambda}\delta t\left|\xi_h^n(\cdot)\right|_{\frac{\varepsilon_h}{2}}^2 + \frac{\lambda}{2}\delta t\left|\frac{1-\lambda}{\lambda}\xi_h^n(\cdot) + \zeta_h^n(\cdot)\right|_{\frac{\varepsilon_h}{2}}^2 \le \left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)}^2$$
$$+ C_\star \delta t (\mathscr{E}_h^n)^2, \tag{4.102}$$

where the seminorm $|\cdot|_{\frac{\varepsilon_h}{2}}$ reads as in Eq. (3.26d).

Proof. From the error Eq. (4.64b), it is possible to deduce that

$$\xi_h^{n+1} - \zeta_h^n = \frac{\delta t}{2} A_h^{n+\lambda}\left(\frac{1-\lambda}{\lambda}\xi_h^n + \zeta_h^n\right) - \frac{\delta t}{2\lambda}A_h^n \xi_h^n + \frac{\delta t}{2\lambda}\alpha_h^n - \frac{\delta t}{2}\beta_h^n. \tag{4.103}$$

Substituting the expression for $\zeta_h^n(x)$ from Eq. (4.64a) also yields

$$\xi_h^{n+1} - \zeta_h^n = \frac{\delta t^2}{2} A_h^{n+\lambda} A_h^n \xi_h^n + \frac{\delta t}{2\lambda}\left(A_h^{n+\lambda} - A_h^n\right)\xi_h^n + \frac{\delta t}{2\lambda}\left(\alpha_h^n - \lambda\,\delta t A_h^{n+\lambda}\alpha_h^n - \lambda\beta_h^n\right). \tag{4.104}$$

Recalling that $\delta t \le t_\star$, the last quantity appearing on the right-hand side may be bounded with the aid of the estimates in Eqs. (4.88), (4.89), and (4.94) as

$$\begin{aligned} |\mathscr{T}_3| &\triangleq \frac{\delta t}{2\lambda}\left\|\alpha_h^n(\cdot) - \lambda\,\delta t A_h^{n+\lambda}\alpha_h^n(\cdot) - \lambda\beta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \\ &\le \frac{\delta t}{2\lambda}\left\|\alpha_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \frac{\delta t^2}{2}\left\|A_h^{n+\lambda}\alpha_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \frac{\delta t}{2}\left\|\beta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \\ &\lesssim \delta t\left\|\alpha_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \delta t\left\|\beta_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \sqrt{\delta t}\,\mathscr{E}_h^n. \end{aligned} \tag{4.105}$$

Concerning the second term, combining the bound in Eq. (4.88) with the 2-CFL condition (4.12), it may be immediately deduced that

$$\begin{aligned} |\mathscr{T}_2| &\triangleq \frac{\delta t}{2\lambda}\left\|\left(A_h^{n+\lambda} - A_h^n\right)\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \le \frac{\delta t}{2\lambda}\left\|A_h^{n+\lambda}\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} + \frac{\delta t}{2\lambda}\left\|A_h^n\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \\ &\lesssim \delta t\frac{\eta_c}{h}\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \sqrt{\delta t}\,\mathscr{E}_h^n. \end{aligned} \tag{4.106}$$

Finally, the first quantity may be bounded by applying two times the estimate (4.12) and invoking the 4/3-CFL condition (4.11) (which is implied by the 2-CFL conditon in Eq. (4.12)). This yields

$$|\mathscr{T}_1| \triangleq \frac{\delta t^2}{2}\left\|A_h^{n+\lambda}A_h^n\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \delta t^2\left(\frac{\eta_c}{h}\right)^2\left\|\xi_h^n(\cdot)\right\|_{L^2(\Omega;\mathbb{R}^n)} \lesssim \sqrt{\delta t}\,\mathscr{E}_h^n. \tag{4.107}$$

Therefore, combining Eqs. (4.105), (4.106), and (4.107) provides

$$\left\|\xi_h^{n+1}(\cdot) - \zeta_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \lesssim |\mathscr{T}_1|^2 + |\mathscr{T}_2|^2 + |\mathscr{T}_3|^2 \lesssim \delta t(\mathscr{E}_h^n)^2. \tag{4.108}$$

Similarly, also from the error Eq. (4.64b), it may be inferred that

$$\begin{aligned} \xi_h^{n+1} - \xi_h^n = &\left(1 - \frac{1}{2\lambda}\right)\delta t A_h^n\xi_h^n + \frac{\delta t}{2\lambda}\left(A_h^{n+\lambda} - A_h^n\right)\xi_h^n \\ &- \frac{\delta t}{2}\left[\left(1 - \frac{1}{2\lambda}\right)\alpha_h^n + \delta t A_h^{n+\lambda}\alpha_h^n + \beta_h^n\right]. \end{aligned} \tag{4.109}$$

Therefore, following a similar rationale as in the derivation of Eq. (4.108), it may be concluded that

$$\left\|\xi_h^{n+1}(\cdot) - \xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)} \lesssim \delta t(\mathscr{E}_h^n)^2. \tag{4.110}$$

Adding Eqs. (4.108) and (4.110) together yields the result. □

Remark 4.3.1. *It is worth observing that, in the proof of Lemma 4.3.8 above, concerning the derivation of Eq. (4.108), the 2-CFL condition (4.12) was invoked only to provide an upper bound on the term $\mathscr{T}_2$. If the operators $(A, D(A))$ and $(A_h, V_{h\star})$ do not depend upon the time variable, the quantity $\mathscr{T}_2$ vanishes. Moreover, all the RK2 schemes analysed in the thesis reduce to the form considered in [39], which may be obtained directly from Eqs. (4.16) be specifying $\lambda = 1$. Accordingly, the contribution relating to $\left\|\xi_h^{n+1}(\cdot) - \xi_h^n(\cdot)\right\|^2_{L^2(\Omega;\mathbb{R}^n)}$ disappears from the energy estimate deduced as in Eq. (4.74). In this case, the less stringent 4/3-CFL condition (4.11) might be invoked.*

Proof of Theorem 4.2.2

The results derived previously permit to assert the main result, namely Theorem 4.2.2. The proof is only sketched below.

Sketch of the proof of Theorem 4.2.2. The proof is almost analogous to that of Theorem 4.2.1, with the difference that the norm $\|\cdot\|_{h\star}$ should be replaced by $\|\cdot\|_{\star\star}$, and the term $\left\|\zeta_\pi^n(\cdot)\right\|_{\star\star}$ need to be accounted for, in addition to $\left\|\xi_\pi^n(\cdot)\right\|_{\star\star}$. The reader is redirected to [39] for further details. □

Chapter 5

Numerical experiments

The present Chapter is dedicated to the numerical approximation of the IBVPs introduced in Chap. 2, using the techniques developed in 3 and 4. The emphasis is mainly on problems in one space dimension that include boundary and trace terms, as those analysed in Sect. 2.2.1. More specifically, Sect. 5.1 focuses on the error analysis concerning sufficiently regular solutions, whereas Sect. 5.2 deals with the main rolling contact problems found in the literature. In the following, the discussion is mainly restricted to RK2 schemes, since the result advocated in Theorem 4.2.1 constitutes a minor modification of that already enounced in [39, 40].

5.1 Analysis for smooth solutions

The hyperbolic problems typically encountered in rolling contact mechanics do not enjoy sufficient regularity to satisfy the conditions required by Theorem 4.2.2. Therefore, to verify numerically the theoretical bound derived according to Theorem 4.2.2, simplified IBVPs are first considered. The effect of boundary terms and time-varying operators are investigated separately in Sects. 5.1.1 and 5.1.2, respectively.

5.1.1 Effect of boundary terms

In order to investigate numerically the rate of convergence for the total error predicted by Theorem 4.2.2, it may first be beneficial to consider Eqs. (2.12) in the scalar case, i.e., $u(x,t) \in \mathbb{R}$, with constant data $a(x,t) = a = 1$, $B(t) = B = 0$, $C(t) = C = 1/2$, and $f(x,t) = f = 0$. With this choice for the transport velocity $a(x,t)$, the matrix $C(t)$, and the forcing term $f(x,t)$, Assumption 2.2.2 holds for all $t \in [0,T]$ and consequently the problem admits a unique strict solution, as asserted by Theorem 2.2.2. In fact, since the transport velocity is constant, the corresponding IBVP problem admits a closed-form solution, consisting of an integral expression combined with a delay-differential equation (DDE) for the boundary term $u(1,t)$ [32]. From the discussion initiated in Sect. 2.3.2, it follows that the solution is sufficiently smooth owing to opportune assumptions made on the regularity of the IC. In particular, the IC

$$u_0(x) = x^3 + \frac{3}{2}x^2 + \frac{5}{2}x, \qquad \text{for } x \in [0,1], \tag{5.1}$$

satisfies the compatibility condition up to the second order both in time and space, thus ensuring the existence of sufficiently smooth solutions as those required by Theorem 4.2.2.

The total error, calculated using the expression appearing on the left-hand side of Eq. (4.17) by specifying $\psi_h = 3/2$, is reported in Table 5.1 for decreasing values of the meshsize h and

Meshsize h	**Total error** $(k=1)$	**Total error** $(k=2)$
0.025	0.0015	$5.59\cdot10^{-6}$
0.017	$8.59\cdot10^{-4}$	$2.00\cdot10^{-6}$
0.0125	$5.63\cdot10^{-4}$	$9.73\cdot10^{-7}$
0.01	$3.92\cdot10^{-4}$	$5.56\cdot10^{-7}$

Table 5.1: Error convergence for polynomial degrees $k=1$ and 2 predicted using Heun's second-order method ($\lambda=1$) for the IBVP described by Eqs. (2.12) and (5.1).

polynomial degrees $k=1$ and 2, considering explicit RK2 schemes with $\lambda=1$ (corresponding to Heun's second-order method) owing to the refined $4/3$-CFL condition (4.11) with $\rho=0.2$. In fact, it is worth emphasising that, when the problem data are constant, the operators $(A,D(A))$ and $(A_h,V_{h\star})$ are also time-independent, and, according to Remark 4.3.1, the $4/3$-CFL condition may be more conveniently invoked in place of the 2-CFL one to derive quasi-optimal error convergence for the complete discrete RK2 schemes analysed in Sect. 4.3.2. Therefore, according to Theorem 4.2.2, under the refined $4/3$-CFL condition (4.11), the RK2 scheme with polynomial degree $k=1$ produces an error convegence in the order of $O(h^{8/3}+h^{3/2})$, whereas the polynomial degree $k=2$ yields $O(h^{8/3}+h^{5/2})$ accuracy. In both case, the time error, relating to the contribution $h^{8/3}$, is dominated by that produced by the space discretisation, proportional to either $h^{3/2}$ or $h^{5/2}$. The values reported in Table 5.1 and the trends illustrated in Fig. 5.1 seem to confirm the bound derived according to Theorem 4.2.2 to be sharp.

The maximum simulation times, corresponding to a meshsize of $h=0.01$ with 100 elements simulated in MATLAB/Simulink® on a personal computer, amounted to 1.94 and 4.48 s for $k=1$ and $k=2$, respectively.

5.1.2 Effect of time-varying operators

To investigate numerically the effect connected with the presence of a time-varying operator, Eqs. (2.12) are again considered in the scalar case, i.e., $u(x,t)\in\mathbb{R}$, with $B(t)=B=0$, $C(t)=C=0$, $f(x,t)=f=0$, and transport velocity and IC assigned as

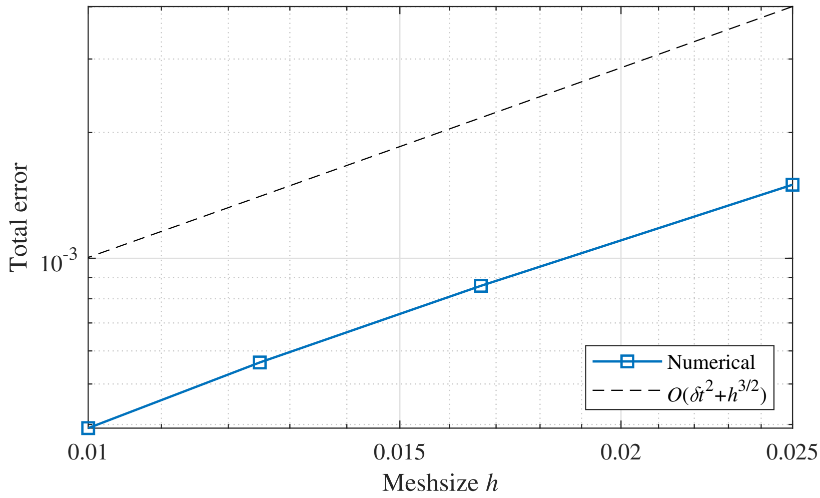
$$a(x,t)=a(t)=1+t, \qquad \text{for } t\in[0,T], \tag{5.2a}$$

$$u_0(x)=x^3, \qquad \text{for } x\in[0,1]. \tag{5.2b}$$

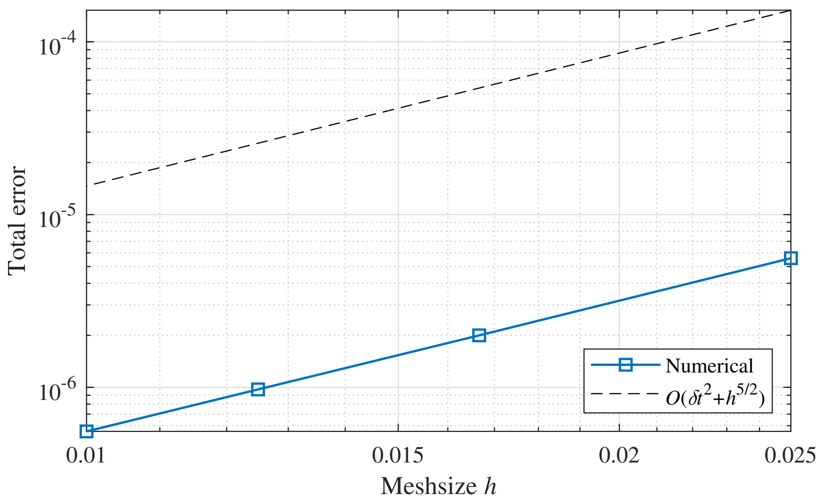
The above expressions for the transport velocity and IC ensure the existence and uniqueness of sufficiently smooth solutions satisfying the assumptions of Theorem 4.2.2.

The total error, calculated as in the right-hand side of Eq. (4.17), is reported in Table 5.2 for decreasing values of the meshsize h and polynomial degrees $k=1$ and 2, using Heun's method for time discretisation (RK2 schemes with $\lambda=1$), with timestep δt obeying the 2-CFL condition of Eq. (4.12) with $\rho'=5$ and 10 for $k=1$ and 2, respectively. Figure 5.2 seems to numerically corroborate that the bound derived in Theorem 4.2.2 is sharp. In particular, it is evident that, especially concerning the RK scheme with polynomial degree $k=2$, the refined $4/3$-CFL condition (4.11) cannot ensure the optimal rate of convergence predicted by Theorem 4.2.2.

The maximum simulation times, corresponding to a meshsize of $h=0.0125$ with 80 elements, amounted to 1.57 and 3.15 s for $k=1$ and $k=2$, respectively.



(a) Convergence of the total error predicted according to Theorem 4.2.2 (polynomial degree $k = 1$).

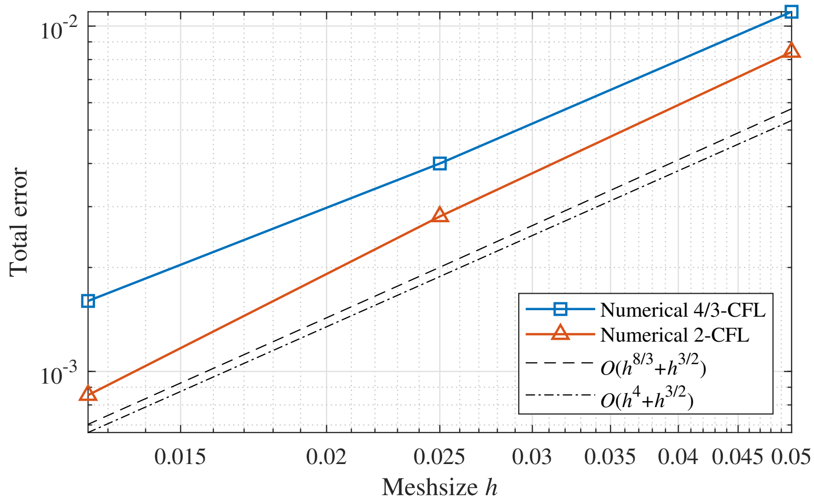


(b) Convergence of the total error predicted according to Theorem 4.2.2 (polynomial degree $k = 2$).

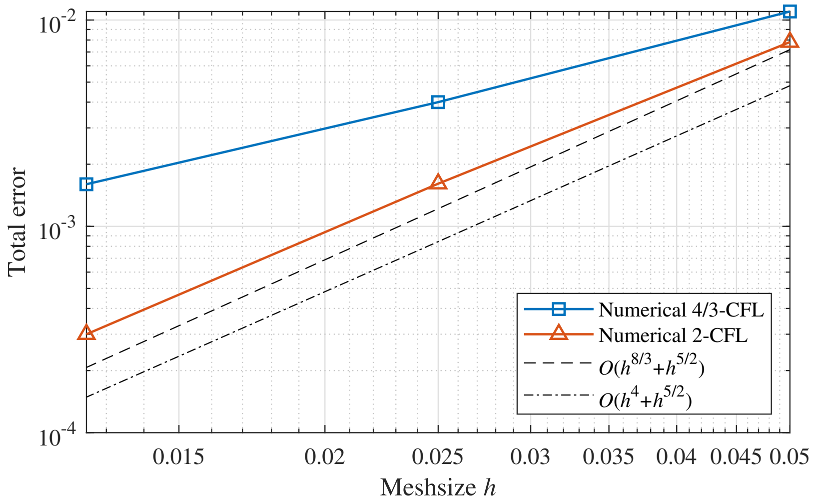
Figure 5.1: Convergence of the total error for different polynomial degrees $k = 1$ and 2 for the IBVP described by Eqs. (2.12) and (5.1).

Meshsize h	**Total error** $(k = 1)$	**Total error** $(k = 2)$
0.05	$8.41 \cdot 10^{-3}$	$7.82 \cdot 10^{-3}$
0.025	$2.81 \cdot 10^{-3}$	$2.58 \cdot 1.61^{-3}$
0.0125	$8.54 \cdot 10^{-4}$	$4.17 \cdot 2.99^{-4}$

Table 5.2: Error convergence for polynomial degrees $k = 1$ and 2 predicted using Heun's second-order method ($\lambda = 1$) for the IBVP described by Eqs. (2.12) and (5.2).



(a) Convergence of the total error predicted according to Theorem 4.2.2 (polynomial degree $k = 1$).



(b) Convergence of the total error predicted according to Theorem 4.2.2 (polynomial degree $k = 2$).

Figure 5.2: Convergence of the total error for different polynomial degrees $k = 1$ and 2 for the IBVP described by Eqs. (2.12) and (5.2).

5.2 Application to linear hyperbolic rolling contact problems

The present Section is dedicated to the numerical approximation of some typical rolling contact problems encountered in the literature.

5.2.1 Systems in one space dimension

The systems in one space dimension considered in this thesis are those governing the brush and LuGre-brush models on time-varying domains. Similar formulations find frequently application in the study of instability phenomena connected with oscillating normal and tangential forces [56–58].

Brush models on a time-varying domain

After opportunely performing a change of variables [33], the brush models on a time-varying domain may be recast in the form of Eqs. (2.12) with $u(x,t) \in \mathbb{R}^2$, $B(t) = 0$, and

$$a(x,t) \triangleq \frac{1}{2\alpha(t)}\Big(1+(1-2x)\dot{\alpha}(t)\Big), \tag{5.3a}$$

$$C(t) \triangleq \Big(1-\dot{\alpha}(t)\Big)\Big(I_2+2\alpha(t)M\Big)^{-1}M, \tag{5.3b}$$

$$f(x,t) \triangleq \Big(I_2+\alpha(t)M\Big)^{-1}\sigma(t)+\alpha(t)\begin{bmatrix}0\\1-2x\end{bmatrix}\varphi(t), \tag{5.3c}$$

where $\alpha \in C^1([0,T];[\alpha_{\min},\alpha_{\max}])$ with $\alpha_{\min} > 0$, $\dot{\alpha} \in C^1([0,T];[\dot{\alpha}_{\min},\dot{\alpha}_{\max}])$, $\max\{|\dot{\alpha}_{\min}|,|\dot{\alpha}_{\max}|\} < 1$, $(\sigma,\varphi) \in C^1([0,T];\mathbb{R}^3)$, and $M \in \mathbf{GL}_2(\mathbb{R})$ is a positive definite, diagonal matrix [33]. Note that the assumptions on $\alpha(t)$ and $\dot{\alpha}(t)$ imply also that $C \in C^1([0,T];\mathbf{M}_{2\times 2}(\mathbb{R}))$.

When $\dot{\alpha}(t) = 0$ for all $t \in [0,T]$, Assumption 2.2.2 is automatically satisfied and the corresponding IBVP even admits a closed-form solution [32]. When $\dot{\alpha}(t) \neq 0$, the condition in Assumption 2.2.2 is satisfied if the matrix

$$\begin{aligned}\Sigma(t) &\triangleq I_2-\Big(1-\dot{\alpha}(t)\Big)\int_0^1 \frac{2\alpha(t)}{1+(1-2x)\dot{\alpha}(t)}\,\mathrm{d}x\Big(I_2+2\alpha(t)M\Big)^{-1}M \\ &= I_2-\Big(1-\dot{\alpha}(t)\Big)\frac{\tanh^{-1}\Big(\dot{\alpha}(t)\Big)}{\dot{\alpha}(t)}\Big(I_2+2\alpha(t)M\Big)^{-1}2\alpha(t)M\end{aligned} \tag{5.4}$$

is invertible, i.e., $\Sigma \in C^1([0,T];\mathbf{GL}_2(\mathbb{R}))$. Such a criterion is always verified provided that

$$(1-\dot{\alpha}_{\min})\frac{\tanh^{-1}(\dot{\alpha}_{\min})}{\dot{\alpha}_{\min}}\Big(1+2\alpha_{\max}\lambda_{\max}(M)\Big)^{-1}2\alpha_{\max}\lambda_{\max}(M) < 1, \tag{5.5}$$

where $\lambda_{\max}(M)$ denotes the largest eigenvalue of M. The formulation presented above has been introduced in [7] in the context of transient tyre modelling, limited to the case of a fixed contact patch ($\dot{\alpha}(t) = 0$ for all $t \in [0,T]$), and then further developed in [16, 32]. Concerning applications in railway dynamics, the same PDEs have been obtained for a time-varying contact patch in [33], with $C(t) = 0$ for all $t \in [0,T]$. In both the automotive and railway fields, the variable $u(x,t) \in \mathbb{R}^2$ collects the tangential deformations of the material particles travelling inside the contact patch, relatively to the road or rail surface, respectively. The transport velocity

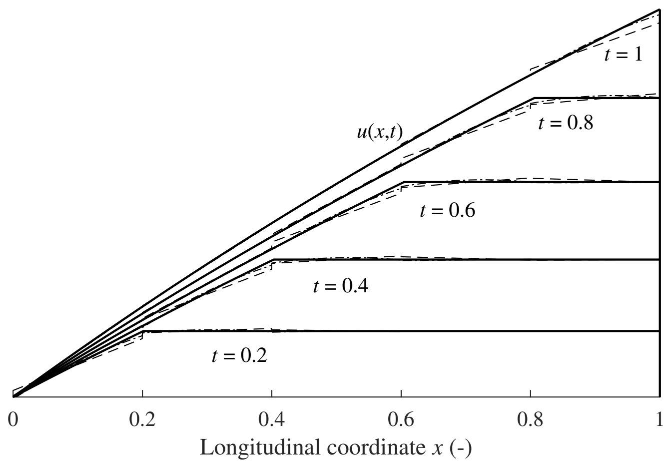


Figure 5.3: Analytical solution and DGM approximations to the rolling contact problem described by Eqs. (2.12) and (5.3), with constant transport velocity $a(x,t) = a = 1/(2\alpha)$. Solid line: analytical solution [32]; dashed line: RK2 with polynomial degree $k = 1$ and 5 mesh elements; dash-dotted line: RK2 with polynomial degree $k = 2$ and 5 mesh elements. Model parameters taken from [32]: $\alpha = 0.075$, $M = 4.44$, $\sigma = 0.7$, $\varphi = 0$.

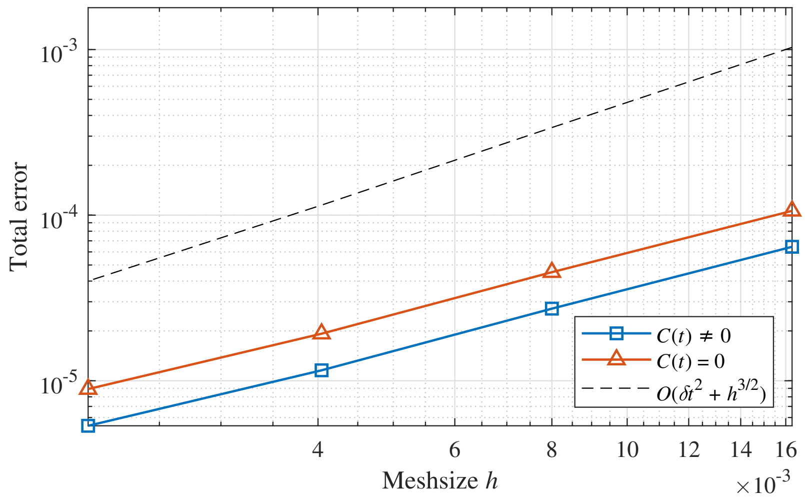
$a(x,t) \in \mathbb{R}$, representing instead the rolling speed of the tyre or railway wheel, is clearly the same for all the components of $u(x,t)$.

In order to investigate numerically the rate of convergence for the total error predicted by Theorem 4.2.2, it may first be beneficial to consider Eqs. (5.3) with constant data $\dot{\alpha}(t) = 0$ and $C(t) = C$. In this case, the transport velocity also becomes constant, and, as in Sect. 5.1.1, the corresponding IBVP problem admits a closed-form solution. In particular, when the matrix $C(t) = C \neq 0$, the analytical solution consists of an integral expression combined with a delay-differential equation (DDE) for the boundary term $u(1,t)$ [32], whereas the case $C(t) = 0$ may be studied by resorting to the classic method of the characteristic lines. More specifically, by relying on simple fixed-point arguments, the exact solution may be inferred to be even continuous for any $f \in C^0([0,1] \times [0,T]; \mathbb{R}^n)$ [32].

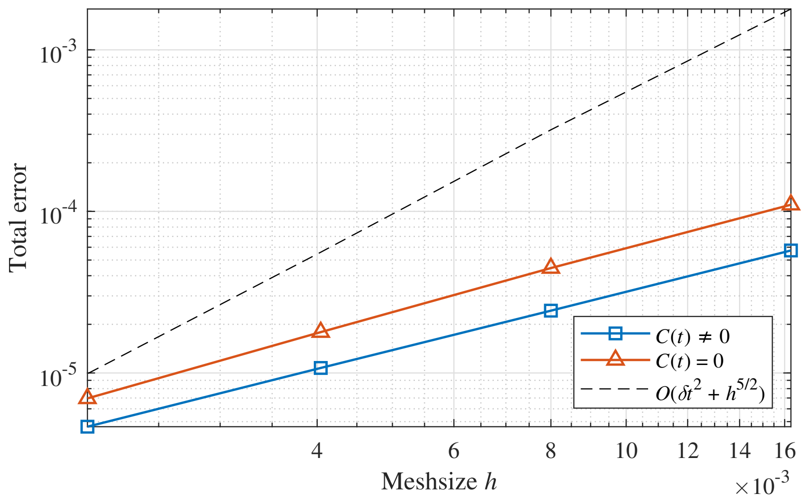
With respect to the scalar case, Fig. 5.3 illustrates two approximations of the solution to the rolling contact problem described by Eqs. (2.12) and (5.3), obtained using Heun's second order method (RK2 schemes with $\lambda = 1$), with $\rho' = 0.2$ and different polynomial degrees $k = 1$ and 2 (dashed and dash-dotted lines, respectively). The corresponding closed-form expression derived according to [32] is also reported (solid line). Both the complete discrete DGMs used to generate Fig. 5.3 employed 5 mesh elements ($h = 0.2$), already yielding a satisfactory agreement with the exact solution. The numerical simulation required a total time of 0.79 s on a personal computer.

The error convergence is illustrated instead in Fig. 5.4 for both polynomial degrees $k = 1$ and 2. In particular, it is worth observing that, since the IC only satisfies the zeroth-order compatibility condition, Theorem 2.2.2 yields $C^0([0,T]; H^1((0,1); \mathbb{R}^n))$-regularity, and therefore the error converges at a slower rate than that predicted by Theorem 4.2.2. More specifically, both the RK2 schemes are characterised by $O(h^{1/2})$ accuracy, independently of the value assumed by the matrix of coefficients $C(t)$.

A time-varying transport velocity appears when oscillating normal load and tangential forces are considered. Concerning the scalar case, typical expressions for the time-dependent transport velocity $a(x,t)$, matrix $C(t)$ and forcing term $f(x,t)$ read according to Eqs. (5.4), with $\varphi(t) = 0$



(a) Convergence of the total error (polynomial degree $k = 1$).



(b) Convergence of the total error (polynomial degree $k = 2$).

Figure 5.4: Convergence of the total error for different polynomial degrees $k = 1$ and 2 for the IBVP described by Eqs. (2.12) and (5.3).

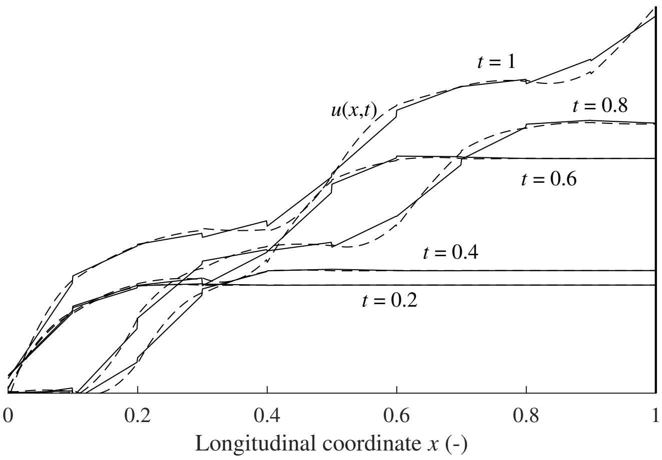


Figure 5.5: DGM approximations to the rolling contact problem described by Eqs. (2.12) and (5.3), with time varying-data according to Eqs. (5.6). Solid line: RK2 with polynomial degree $k = 1$ and 10 mesh elements; dashed line: RK2 with polynomial degree $k = 2$ and 10 mesh elements. Model parameters: $\alpha_0 = 0.075$, $\alpha_{\mathrm{I}} = 0.1 \cdot \alpha_0$, $M = 4.44$, $\sigma_0 = 0.7$, $\sigma_{\mathrm{I}} = 0.1 \cdot \sigma_0$, $\varphi = 0$, $\omega = 100$. Total simulation time $T = 2(\alpha_0 + \alpha_{\mathrm{I}})$.

and

$$\alpha(t) = \alpha_0 + \alpha_{\mathrm{I}} \sin(\omega t), \tag{5.6a}$$
$$\sigma(t) = \sigma_0 + \sigma_{\mathrm{I}} \sin(\omega t), \tag{5.6b}$$

where both signals are characterised by the same frequency ω.

In this case, a closed-form solution is unfortunately not available, but it is still worth comparing the approximations obtained by employing DGMs with different polynomial degrees. Figure 5.5 illustrates the numerical solution obtained using first and second-order polynomial functions in conjunction with Heun's second order method in time. In both cases, the plotted solutions refer to a mesh with 10 elements, with $\rho' = 0.2$ in Eq. (4.12). It may be observed that the trend predicted by the DGM with $k = 2$ is much smoother than that yielded by the lower-order polynomial degree. The total simulation time amounted to 14.92 and 25.46 s for $k = 1$ and 2, respectively. In this case, the heavier computational cost should be ascribed to the more stringent 2-CFL condition in Eq. (4.12) than the $4/3$-one invoked previously concerning the problem with constant transport velocity.

For the IBVP described by Eqs. (2.12) and (5.3), the error convergence is illustrated in Fig. 5.6 for the usual polynomial degrees $k = 1$ and 2, respectively. As observed previously, the low regularity of the exact solution does not fulfil the criteria required by Theorem 4.2.2, and hence the rate of convergence is not optimal. In fact, the accuracy is approximately in the order of $O(h^{1/2})$. Concerning the problem under investigation, the simulation time was also prohibitive, amounting at more than 420 s for the smallest meshsize $h = 0.033$ in combination with the 2-CFL condition of Eq. (4.12).

LuGre-brush models on a time-varying domain

After opportunely performing a change of variables [33], the LuGre-brush models with a spatially constant pressure distribution on a time-varying domain may be recast in the form of Eqs. (2.12)

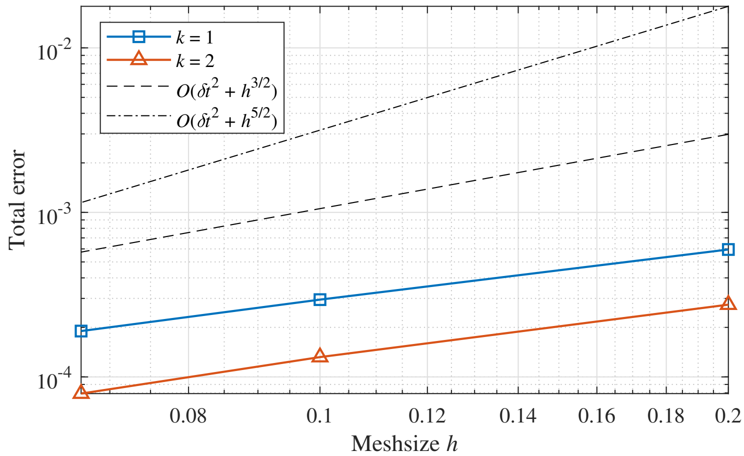


Figure 5.6: Convergence of the total error for different polynomial degrees $k = 1$ and 2 for the rolling contact problem described by Eqs. (2.12) and (5.3), with time varying-data according to Eqs. (5.6).

with $u(x,t) \in \mathbb{R}^2$, $a(x,t)$ and $f(x,t)$ reading as in Eqs. (5.3a) and (5.3c), respectively, and

$$C(t) \triangleq \Big(1-\dot{\alpha}(t)\Big)\Big(I_2+2\alpha(t)M_1M_2p(t)\Big)^{-1}M_1M_2p(t), \tag{5.7a}$$

$$\tilde{B}(t) \triangleq -\beta(t)M_1, \tag{5.7b}$$

$$K(x,t) \equiv K(t) \triangleq 2\alpha(t)\Big(1-\dot{\alpha}(t)\Big)\Big(I_2+2\alpha(t)M_1M_2p(t)\Big)^{-1}M_1M_2\left(\beta(t)M_1p(t)-\frac{\partial p(t)}{\partial t}\right), \tag{5.7c}$$

where $\alpha(t)$ and $\dot{\alpha}(t)$ satisfy the same assumptions as previously, $\beta \in C^1([0,T];\mathbb{R}_{\geq 0})$, $p \in C^1([0,T];[p_{\min}, p_{\max}])$, with $p_{\min} > 0$, and M_1, $M_2 \in \mathbf{GL}_2(\mathbb{R})$ are positive definite diagonal matrices [33].

Again, if $\dot{\alpha}(t) = 0$ for all $t \in [0,T]$, Assumption 2.2.2 is identically verified. When $\dot{\alpha}(t) \neq 0$, the analogous condition of Eq. (5.4) which satisfies Assumption 2.2.2 reads

$$\begin{aligned}\Sigma(t) &\triangleq I_2-\Big(1-\dot{\alpha}(t)\Big)\int_0^1 \frac{2\alpha(t)}{1+(1-2x)\dot{\alpha}(t)}\,\mathrm{d}x\Big(I_2+2\alpha(t)M_1M_2p(t)\Big)^{-1}M_1M_2p(t)\\ &= I_2-\Big(1-\dot{\alpha}(t)\Big)\frac{\tanh^{-1}\Big(\dot{\alpha}(t)\Big)}{\dot{\alpha}(t)}\Big(I_2+2\alpha(t)M_1M_2p(t)\Big)^{-1}2\alpha(t)M_1M_2p(t).\end{aligned} \tag{5.8}$$

Clearly, $\Sigma \in C^1([0,T];\mathbf{GL}_2(\mathbb{R}))$ if the following criterion is fulfilled:

$$(1-\dot{\alpha}_{\min})\frac{\tanh^{-1}(\dot{\alpha}_{\min})}{\dot{\alpha}_{\min}}\Big(1+2\alpha_{\max}\lambda_{\max}(M_1M_2)p_{\max}\Big)^{-1}2\alpha_{\max}\lambda_{\max}(M_1M_2)p_{\max} < 1, \tag{5.9}$$

where this time $\lambda_{\max}(M_1M_2)$ denotes the largest eigenvalue of M_1M_2. In the context of tyre dynamics, and limited to the case $\dot{\alpha}(t) = 0$ and $C(t) = 0$ for all $t \in [0,T]$, the LuGre-brush models were derived in [17–20] and studied extensively also in [21–23]. The first formulation accounting for the presence of boundary terms has recently appeared in [32]. According to such a model, $u(x,t) \in \mathbb{R}^2$ is interpreted either as a tangential deformation or as an internal frictional variable.

The IBVP described by Eqs. (2.12), (5.3a), (5.3c), and (5.7) may be studied within similar techniques as in Sect. 5.2.1, and are not dealt with explicitly in this thesis.

5.2.2 System in two space dimensions

The systems in two space dimensions considered in this thesis are those governing the equations of the brush and LuGre-brush models on fixed domains.

In particular, when accounting for large spin slips, both formulations are in the form of Eqs. (2.44), with $u(x,t) \in \mathbb{R}^2$, $\Omega \subset \mathbb{R}^2$, and

$$a(x,t) \equiv a(x) = \begin{bmatrix} -\varepsilon_1 + \gamma x_2 \\ \varepsilon_2 - \gamma x_1 \end{bmatrix}, \tag{5.10a}$$

$$B(t) \equiv B = \begin{bmatrix} -\kappa_1 & -\psi \\ \psi & -\kappa_2 \end{bmatrix}, \tag{5.10b}$$

$$f(x,t) \equiv f(x) = \sigma + \begin{bmatrix} -\varphi x_2 \\ \varphi x_1 \end{bmatrix}, \tag{5.10c}$$

where $\kappa_1 = \kappa_2 = 0$ identically for the standard brush model and $\kappa_1,\ \kappa_2 \geq 0$ for the LuGre-brush models. For $B(t)$ constant, or commuting with its integral in the time-varying case, the corresponding evolution operator introduced in Sect. 2.2.2 reads evidently $U_B(t,\tilde{t}) = \exp(\int_{\tilde{t}}^{t} B(t')\,\mathrm{d}t')$. Concerning the standard brush models, the matrix $B(t)$ does not only commute even in the time-varying case, but it is also skew-symmetric, i.e., $B(t) \in \mathbf{Skew}_2(\mathbb{R})$, and therefore the evolution operator is unitary ($\kappa_1 = \kappa_2 = 0$ implies more specifically $U_B(t,\tilde{t}) \in \mathbf{SO}_2(\mathbb{R})$). The problem may, in principle, be solved analytically using the method of the characteristic lines even when $B = B(t)$ and $f(x) = f(x,t)$ are time-varying. This may be accomplished either directly or by converting the original IBVP (2.44) into the equivalent one (2.49). Analytical solutions are reported, for example, in [14, 15, 24] concerning rectangular, circular, and elliptical domains.

The numerical treatment of the two-dimensional problem is not discussed explicitly in this thesis, whereas a summary of salient results may be found again in [39, 40] for exact solutions enjoying sufficiently smooth regularity.

Chapter 6

Conclusions

The present thesis addressed the problem of recovering numerical solutions to linear hyperbolic IBVPs encountered in rolling contact mechanics. All the equations considered in this work assumed a brush-like representation for the friction model, as customary in vehicle dynamics and automotive applications. Given the peculiar structure of the hyperbolic PDEs under investigation, which included integral and boundary terms in the one-dimensional case, the first part of the thesis was dedicated to establishing the well-posedness of the corresponding IBVPs. Existence and uniqueness were proved for the continuous problem within the mathematical framework provided by the semigroup theory. This allowed, in most cases, to derive strict solutions enjoying sufficient regularity properties to satisfy the assumptions required for the subsequent numerical analyses. With respect to problems involving several space dimensions, a classic result by Bardos [50] was conveniently recalled which applies to the governing equations of spinning and rolling bodies. In this context, it is also worth emphasising that, limited to the one-dimensional problems, the results advocated in Chap. 2 concerning the well-posedness of the considered equations represent already a novelty in the context of the mathematical analysis of rolling contact models.

The second part of the work was then devoted to the development of numerical schemes to approximate the exact solutions of the rolling contact problems described in Chap. 2. In particular, this was accomplished by combining discontinuous Galerkin finite element methods (DGMs) with explicit Runge-Kutta (RK) algorithms of first and second-order. Whilst the semi-discrete problem and the discrete operator were introduced and analysed in Chap. 3, the complete discrete formulation was fully developed in Chap. 4, where the two main results were also asserted regarding the convergence of the proposed schemes. More specifically, departing from the analyses initiated in [39, 40], analogous convergence results were established in the case of time-independent continuous and discrete operators $(A, D(A))$ and $(A_h, V_{h\star})$ when accounting for the presence of integral and boundary terms. In particular, under the assumption of sufficiently smooth exact solutions and owing to a refined $4/3$-CFL condition, accuracy in the order of $O(h^{4/3} + h^{1/2})$ and $O(h^{8/3} + h^{k+1/2})$ was proved respectively for the finite-volume approximation in combination with RK1 algorithms, and for the DGMs in conjunction with RK2 schemes. The more involved case of time-dependent operators $(A, D(A))$ and $(A_h, V_{h\star})$ yielded similar results in terms of convergence rate, but required instead a stricter 2-CFL condition to hold. Applications to IBVPs typically encountered in the study of rolling contact phenomena were discussed extensively in Chap. 5, with particular emphasis on the one-dimensional problems that account for the presence of boundary and integral terms. The developed algorithms showed an overall promising agreement with the exact solution derived in [32] for the case of constant transport velocity, requiring also relatively small simulation times. On the other hand, the

2-CFL condition needed to ensure convergence in the time-dependent case seems to pose heavy limitations concerning real-time applications.

Future research efforts may be directed to the analysis of higher-order RK schemes, which are already available in virtual environments like MATLAB/Simulink®, and possibly able to overcome the drawbacks connected with too stringent CFL conditions. Moreover, different applications from those considered in this thesis could be explored in forthcoming studies.

Bibliography

[1] Evans LC. Partial differential equations. 2nd ed. American Mathematical Society; 1996.

[2] Larsson S, Thomee V. Partial Differential Equations with Numerical Methods. 1st ed. Springer, Berlin, Heidelberg. 2009. Available from: https://doi.org/10.1007/978-3-540-88706-5.

[3] Johnson KL. Contact Mechanics. Cambridge University Press; 1985. Available from: https://doi.org/10.1017/CBO9781139171731.

[4] Goryacheva IG. Contact Mechanics in Tribology. Springer Science & Business Media. 1998.

[5] Kalker JJ. On the rolling contact of two elastic bodies in the presence of dry friction. Dissertation, TH Delft, Delft (1967).

[6] Kalker JJ. Three-Dimensional Elastic Bodies in Rolling Contact. Springer: Dordrecht; 1990. Available from: https://doi.org/10.1007/978-94-015-7889-9.

[7] Guiggiani M. The Science of Vehicle Dynamics, 2nd ed. Cham(Switzerland): Springer International; 2018.

[8] Pacejka HB. Tire and vehicle dynamics. 3rd ed. Amsterdam: Elsevier/BH; 2012.

[9] Frendo F, Bucchi F. ”Brush model” for the analysis of flat belt transmissions in steady-state conditions. Mechanism and Machine Theory. 2020. 143. Available from https://doi.org/10.1016/j.mechmachtheory.2019.103653.

[10] Frendo F, Bucchi F. Enhanced brush model for the mechanics of power transmission in flat belt drives under steady–state conditions: Effect of belt elasticity. Mechanism and Machine Theory. 2020. 153. Available from: https://doi.org/10.1016/j.mechmachtheory.2020.103998.

[11] Bucchi F, Frendo F. Validation of the brush model for the analysis of flat belt transmissions in steady-state conditions by finite element simulation. Mechanism and Machine Theory. 2022. 167. Available from: https://doi.org/10.1016/j.mechmachtheory.2021.104556.

[12] Marques F, Flores P, Pimenta Claro JC et al. A survey and comparison of several friction force models for dynamic analysis of multibody mechanical systems. Nonlinear Dyn 2016;86:1407–1443. Available from: https://doi.org/10.1007/s11071-016-2999-3.

[13] Marques F, Flores P, Claro JCP et al. Modeling and analysis of friction including rolling effects in multibody dynamics: a review. Multibody Syst Dyn 45, 223–244 (2019). Available from: https://doi.org/10.1007/s11044-018-09640-6.

[14] Romano L, Timpone F, Bruzelius F, Jacobson B. Analytical results in transient brush tyre models: theory for large camber angles and classic solutions with limited friction, Meccanica 57:165-191 (2022). Available from: http://dx.doi.org/10.1007/s11012-021-01422-3.

[15] Romano L, Timpone F, Bruzelius F, Jacobson B. Rolling, tilting and spinning spherical wheels: analytical results using the brush theory, Mechanism and Machine Theory

173 (2022). Available from: https://doi.org/10.1016/j.mechmachtheory.2022.104836.

[16] Romano L. Advanced Brush Tyre Modelling. SpringerBriefs in Applied Sciences. Springer, Cham (2022). Available from: https://doi.org/10.1007/978-3-030-98435-9.

[17] Deur J. Modeling and Analysis of Longitudinal Tire Dynamics Based on the Lugre Friction Model. IFAC Proceedings Volumes. 2001;34(1):91-96. Available from: https://doi.org/10.1016/S1474-6670(17)34383-5.

[18] Deur J, Asgari J, Hrovat D. A dynamic tire friction model for combined longitudinal and lateral motion. InProceedings of the ASME-IMECE World Conference. 2001.

[19] Deur J, Asgari J, Hrovat D. A 3D Brush-type Dynamic Tire Friction Model. Vehicle Syst. Dyn. 2004;42(3):133-173. Available from: https://doi.org/10.1080/00423110412331282887.

[20] Deur J, Ivanovic V, Troulis M, Miano C, Hrovat D, Asgari J. Extensions of the LuGre tyre friction model related to variable slip speed along the contact patch length. Vehicle Syst. Dyn. 2005;43(supp):508-524. Available from: https://doi.org/10.1080/00423110500229808.

[21] Canudas-de-Wit C, Tsiotras P, Velenis E, Basset M, Gissinger G. Dynamic friction models for road/tire longitudinal interaction. Vehicle Syst. Dyn. 2003;39(3)189-226. Available from: https://doi.org/10.1076/vesd.39.3.189.14152.

[22] Tsiotras P, Velenis E, Sorine M. A LuGre Tire Friction Model With Exact Aggregate Dynamics. Vehicle Syst. Dyn. 2004;42(3):195-210. Available from: https://doi.org/10.1080/00423110412331289835.

[23] Velenis E, Tsiotras P, Canudas-de-Wit C, Sorine M. Dynamic tyre friction models for combined longitudinal and lateral vehicle motion. Vehicle Syst. Dyn. 2005;43(1):3-29. Available from: https://doi.org/10.1080/00423110412331290464.

[24] Romano L, Bruzelius F, Jacobson B. An extended LuGre-brush tyre model for large camber angles and turning speeds, Vehicle Syst. Dyn. (2022). Available from: https://doi.org/10.1080/00423114.2022.2086887.

[25] Marques F, Woliński Ł, Wojtyra M, Flores P, Lankarani HM. An investigation of a novel LuGre-based friction force model. Mechanism and Machine Theory. 2021;166. Available from: https://doi.org/10.1016/j.mechmachtheory.2021.104493.

[26] Kalker JJ. Transient Phenomena in Two Elastic Cylinders Rolling Over Each Other With Dry Friction. Journal of Applied Mechanics. 1970, 37(3): 677-688. Available from: https://doi.org/10.1115/1.3408597.

[27] Kalker JJ. Transient rolling contact phenomena. ASLE Trans. 1971;14(3):177-184. Available from: https://doi.org/10.1080/05698197108983240.

[28] Vollebregt EAH. User guide for CONTACT, Vollebregt & Kalker's rolling and sliding contact model. Techical report TR09-03, version 13.

[29] Vollebregt EAH, Wilders P. FASTSIM2: a second-order accurate frictional rolling contact algorithm. Computational Mechanics 47 (1), 105-116; 2011. Available from: https://doi.org/10.1007/s00466-010-0536-7.

[30] Vollebregt EAH, Iwnicki SD, Xie G, Shackelton P. Assessing the accuracy of different simplified frictional rolling contact algorithms. Vehicle System Dynamics 50 (1), 1-17; 2012. Available from: https://doi.org/10.1080/00423114.2011.552618.

[31] Vollebregt EAH. Numerical modeling of measured railway creep versus creep-force curves with CONTACT. Wear 314(1-2), 87-95; 2014. Available from: https://doi.org/10.1016/j.wear.2013.11.030.

[32] Romano L, Bruzelius F, Jacobson B. Transient tyre models with a flexible carcass, Vehicle Syst. Dyn. (2023).

[33] Romano L, Maglio M, Bruni S. Transient wheel-rail rolling contact theories, Tribology International (2023).

[34] Romano L, Bruzelius F, Hjort M, Jacobson B. Development and analysis of the two-regime transient tyre model for combined slip. Vehicle Syst. Dyn. 2022.

[35] Cockburn B, Hou S, Shu CW. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case, Math. Comp. 54(190):545–581 (1990).

[36] Cockburn B, Karniadakis GE, Shu CW. Discontinuous Galerkin Methods - Theory, Computation and Applications, volume 11 of Lecture Notes in Computer Science and Engineering. Springer, 2000.

[37] Cockburn B, Lin S, Shu CW. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems. J. Comput. Phys. 84(1):90–113 (1989).

[38] Cockburn B, Shu CW. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. Math. Comp. 52(186):411–435 (1989).

[39] Di Pietro DA, Ern A. Mathematical Aspects of Discontinuous Galerkin Methods, 1st. ed. Springer Berlin, Heidelberg (2011). Available from: https://doi.org/10.1007/978-3-642-22980-0.

[40] Burman E, Ern E, Fernández MA. Explicit Runge–Kutta Schemes and Finite Elements with Symmetric Stabilization for First-Order Linear PDE Systems, SIAM Journal on Numerical Analysis 48(6):2019-2042 (2010). Available from: https://doi.org/10.1137/090757940.

[41] Massey FJ. Abstract evolution equations and the mixed problem for symmetric hyperbolic systems, Trans. Amer. Math. Soc. 168:165-188 (1972). Available from: https://www.ams.org/journals/tran/1972-168-00/S0002-9947-1972-0298231-4.

[42] Tanabe H. Equations of Evolution, 1st ed. Pitman (1979).

[43] Tanabe H. Functional Analytic Methods for Partial Differential Equations, 1st ed. CRC Press (1997).

[44] Pazy A. Semigroups of Linear Operators and Applications to Partial Differential Equations, 1st. ed. Springer New York, NY (1983). Available from: https://doi.org/10.1007/978-1-4612-5561-1.

[45] Kato T. Linear evolution equations of 'hyperbolic' type, J. Fac. Sci. Univ. Tokyo 17:241–258 (1970).

[46] Kato T. Perturbation Theory for Linear Operators, 2nd ed. Springer Berlin, Heidelberg (1995). Available from: https://doi.org/10.1007/978-3-642-66282-9.

[47] Curtain RF, Zwart H. An Introduction to Infinite-Dimensional Linear Systems Theory, 1st. ed. Springer New York, NY (2012). Available from: https://doi.org/10.1007/978-1-4612-4224-6.

[48] Curtain RF, Zwart H. Introduction to Infinite-Dimensional Systems Theory: A State-Space Approach, 1st. ed. Springer New York, NY (2020). Available from: https://doi.org/10.1007/978-1-0716-0590-5.

[49] Taylor AE. Introduction to Functional Analysis, Wiley, New York, (1958).

[50] Bardos C. Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels, théorèmes d'approximation application à l'équation de transport,

Ann. Sci. Éc. Norm. Super. 4(3):185-233 (1970).

[51] Kreiss HO, Lorenz J. Initial-Boundary Value Problems and the Navier-Stokes Equations, Society for Industrial and Applied Mathematics (2004). Available from: https://epubs.siam.org/doi/abs/10.1137/1.9780898719130.

[52] Gustafsson B, Kreiss HO, Oliger J. Time-Dependent Problems and Difference Methods, 2nd ed. Wiley, Pure and Applied Mathematics (2013).

[53] Crippa G, Donadello C, Spinolo LV. Initial–boundary value problems for continuity equations with BV coefficients, Journal de Mathématiques Pures et Appliquées 102(1) (2022). Available from: https://doi.org/10.1016/j.matpur.2013.11.002.

[54] Benzoni-Gavage S, Serre D. Multi-dimensional hyperbolic partial differential equations: First-order systems and applications, 1st. ed. Oxford Academic (2006). Available from: https://doi.org/10.1093/acprof:oso/9780199211234.001.0001.

[55] Shu CW, Osher S. Efficient implementation of essentially nonoscillatory shock-capturing schemes, J. Comput. Phys. 77(2):439-471 (1988).

[56] Ciavarella M, Barber J. Influence of longitudinal creepage and wheel inertia on short-pitch corrugation: A resonance-free mechanism to explain the roaring rail phenomenon. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology 222(3); 2008. Available from: https://doi.org/10.1243/13506501JET373.

[57] Afferrante L, Ciavarella M. Short-pitch rail corrugation: A possible resonance-free regime as a step forward to explain the "enigma"? Wear 266(9-10); 2008. Available from: https://doi.org/10.1016/j.wear.2008.12.003.

[58] Afferrante L, Ciavarella M. Short pitch corrugation of railway tracks with wooden or concrete sleepers: An enigma solved? Tribology International 43(3); 2010. Available from: https://doi.org/10.1016/j.triboint.2009.09.010.