Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Time Discretization of Parabolic Problems by the hp-Version of the Discontinuous Galerkin Finite Element Method

D. Schötzau and C. Schwab

Research Report No. 99-04
February 1999

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Time Discretization of Parabolic Problems by the hp-Version of the Discontinuous Galerkin Finite Element Method

D. Schötzau and C. Schwab

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 99-04 February 1999

Abstract

The Discontinuous Galerkin Finite Element Method (DGFEM) for the time discretization of parabolic problems is analyzed in a *hp*-version context. Error bounds which are explicit in the time step as well as the approximation order are derived and it is shown that the *hp*-DGFEM gives spectral convergence in problems with smooth time dependence. In conjunction with geometric time partitions it is proved that the *hp*-DGFEM results in exponential rates of convergence for piecewise analytic solutions exhibiting singularities induced by incompatible initial data or piecewise analytic forcing terms. For the *h*-version DGFEM algebraically graded time partitions are determined that give the optimal algebraic convergence rates. A fully discrete *hp* scheme is discussed exemplarily for the heat equation. The use of certain mesh-design principles for spatial discretizations yields exponential rates of convergence in time and space. Numerical examples confirm the theoretical results.

Keywords: Abstract Parabolic Problems, Discontinuous Galerkin Methods, *hp*-Version of the Finite Element Method

AMS Subject Classification: 65M60, 65J10

1. Introduction

The Discontinuous Galerkin Finite Element Method (DGFEM) was originally introduced as a time-stepping method for the numerical solution of ordinary differential equations (see [7, 10, 19, 23] and the references there for recent developments). In a series of important papers the DGFEM time discretization technique has been applied to parabolic partial differential equations by Eriksson, Johnson, Larsson, Thomée, Wahlbin and their co-workers (cf. [11, 12, 13, 14, 15, 16, 17, 18, 24]). We refer also to the recent monograph [38] and the references there. Issues such as optimal order error estimates, a posteriori error analysis, adaptivity and nonlinearities have been addressed and the method has reached a certain maturity by now. We mention also Makridakis and Babuška who establish in [26] the quasioptimality of the DGFEM in certain mesh-dependent norms.

However, all these works considered the "h-version" DGFEM where the convergence is achieved by decreasing the time steps Δt at a fixed (mostly low) approximation order r. As $\Delta t \to 0$, this results typically in algebraic error estimates of the order $O(\Delta t^r)$ for solutions depending smoothly on t. Error bounds that are explicit in r can usually not be found in the literature.

In the 1980ies, the p- and hp-versions of the Finite Element Method (FEM) were introduced by Babuška, Szabó and their co-workers (see the survey articles [3, 4], the monograph [35] and the references there). In particular, it was shown for elliptic problems with piecewise analytic solutions that the hp-FEM can achieve exponential rates of convergence.

Solutions of parabolic problems exhibit similar analyticity properties: After a start-up singularity induced by incompatibilities of the initial data the solutions are smoothened and are analytic in time. This behaviour suggests that p- and hp-version concepts can be applied to discretization methods for parabolic problems. The p-version approach is to achieve convergence by increasing the polynomial approximation order r on fixed time steps Δt, whereas the hp-version combines judiciously h- and p-refinement techniques. An attempt in that direction has been made by Babuška and Janik in [2]. There, the p- and hp-version of a Petrov-Galerkin method in time are analyzed. However, severe restrictions on the space discretizations were required and no exponential convergence results were derived.

In this work we introduce the hp-version of the DGFEM for the temporal discretization of parabolic equations. We show for this hp-version approach that spectral and exponential convergence rates can be obtained.

Complementary to the usual Method of Lines, the hp-DGFEM reduces the original transient equation to a sequence of stationary problems which have still to be solved numerically in order to get a fully discrete scheme. Arbitrary variations in the temporal approximation order r as well as in the time step Δt are allowed and for linear parabolic problems the hp-DGFEM is unconditionally stable independently of the spatial discretization. This is crucial, since hp methods in space require highly anisotropic meshes for the efficient resolution of layers and fronts which tend to produce very stringent CFL limitations in explicit time-stepping schemes. The underlying variational structure of the hp-DGFEM is moreover well suited for a-posteriori analysis and adaptivity.

Whereas we employ in the present work a DG approach for the time discretization, it should be remarked here that DG methods are also successfully applied in spatial discretizations, particularly in the context of convection-diffusion problems. Recent results in that direction can be found in [5, 9, 30] and the references there. We refer also to the survey article [8].

The outline of this paper is as follows: In Section 2 we present our parabolic setting based on a Gelfand triple of Hilbert spaces and we formulate the DG Finite Element Method. In Section 3 error estimates that are explicit in r and Δt are derived and it is shown that p-version (or spectral) accuracy can be achieved in transient problems with smooth time dependence.

Often, smooth time dependence may be an unrealistic assumption, since solutions of parabolic problems can exhibit singular behaviour induced e.g. by incompatible initial data or by piecewise analytic forcing terms. In Section 4 we analyze the structure of such singularities and establish bounds on the growth of the solutions' time derivatives which are explicit in the regularity order and the time. We confine ourselves to self-adjoint operators and refer to the forthcoming work [34] where these regularity results are extendend to the non-selfadjoint case by the use of semigroup theory.

These estimates are crucial in Section 5 where it is proved that the hp-DGFEM based on geometric time partitions and linearly increasing time approximation orders results in exponential rates of convergence for solutions of parabolic problems that are piecewise analytic in time. For fixed approximation orders algebraically graded time meshes can be employed and for each approximation order we determine in Section 5 the graded mesh that yields the optimal algebraic convergence rate.

A complete hp discretization in time and space is discussed exemplarily for the heat equation in Section 6. At each time step a system of possibly singularly perturbed reaction-diffusion equations is solved. If the hp-Finite Element Methods for these spatial problems take into account certain mesh design principles using anisotropic and geometric mesh refinement techniques, exponential rates of convergence in time and space can be achieved. In Section 6.4 the theoretical results are confirmed in numerical examples.

Throughout, standard notations and conventions are followed. For two Banach spaces X, Y we denote by $\mathcal{L}(X, Y)$ the Banach space of all linear and bounded operators $X \to Y$ equipped with the operator norm.

To describe time discretizations we use Bochner spaces of functions which map a (time) interval $I = (a, b)$ into X: We denote by $L^p(I; X)$, $1 \leq p \leq \infty$, and $H^k(I; X)$, $0 \leq k \in \mathbb{R}$, the corresponding Lebesgue and Sobolev spaces. $C^k(\overline{I}; X)$ are the functions that are k times continuously differentiable and $\mathcal{D}(I; X)$ are the C^∞-functions with compact support in I. $\mathcal{P}^r(I; X)$ denotes the set of all polynomials of degree $\leq r$ with coefficients in X, i.e. $p(t) \in \mathcal{P}^r(I; X)$ if and only if $p(t) = \sum_{j=0}^r x_j t^j$ for some $x_j \in X$ and $t \in I$. If $X = \mathbb{R}$, the dependence on X is omitted.

2. DGFEM for Abstract Parabolic Problems

We present in Section 2.1 our parabolic setting and formulate in Section 2.2 the Discontinuous Galerkin Finite Element Methods (DGFEM).

2.1. Abstract Parabolic Problems

Let X and H be complex, separable Hilbert spaces, with dense injection $X \overset{d}{\hookrightarrow} H$ and norms $\|\cdot\|_X$ and $\|\cdot\|_H$, respectively. Denoting by $(\cdot,\cdot)_H$ the scalar product on H and identifying H with H^*, the antidual of H, we get the Gelfand triple

$$X \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} X^*. \tag{2.1}$$

We denote by $(\cdot,\cdot)_{X^*\times X}$ the duality pairing in $X^* \times X$ and by $\|\cdot\|_{X^*}$ the norm in X^*. For $y \in H$, $x \in X$ we have $(y,x)_{X^*\times X} = (y,x)_H$. We assume in addition that

$$X \overset{d}{\hookrightarrow} H \text{ is compactly imbedded.} \tag{2.2}$$

In the triple (2.1) we consider the abstract (linear) parabolic problem

$$u'(t) + Lu(t) = g(t), \qquad t \in J = (0,T),\ 0 < T < \infty, \tag{2.3}$$

$$u(0) = u_0. \tag{2.4}$$

We assume $L \in \mathcal{L}(X, X^*)$ to be an elliptic "spatial differential" operator given as $(Lu,v)_{X^*\times X} = a(u,v)$ where $a : X\times X \to \mathbb{C}$ is a continuous, coercive and self-adjoint sesquilinear form with

$$|a(u,v)| \le \alpha\|u\|_X\|v\|_X, \qquad u,v \in X, \tag{2.5}$$

$$\text{Re } a(u,u) \ge \beta\|u\|_X^2, \qquad u \in X, \tag{2.6}$$

$$a(u,v) = \overline{a(v,u)}, \qquad \forall u,v \in X. \tag{2.7}$$

L is then an isomorphism in $\mathcal{L}(X,X^*)$ with $\|L\|_{\mathcal{L}(X,X^*)} \le \alpha$ and $\|L^{-1}\|_{\mathcal{L}(X^*,X)} \le \frac{1}{\beta}$. The standard weak form of (2.3), (2.4) is: Find $u \in L^2(J;X) \cap H^1(J;X^*)$ (which implies $u \in C([0,T];H)$) such that $u(0) = u_0$ and

$$-\int_J (u(t),v)_H \varphi'(t)dt + \int_J a(u,v)\varphi(t)dt = \int_J (g(t),v)_{X^*\times X}\varphi(t)dt \tag{2.8}$$

for all $v \in X$ and $\varphi \in \mathcal{D}(J)$.
Examples of parabolic problems that fit into that framework are the Stokes problem or the standard heat equation [33].
Concerning the data we assume always at least that $u_0 \in H$ and $g \in L^2(J;H)$. In that case (2.3), (2.4) has a unique weak solution $u \in L^2(J;X)\cap H^1(J;X^*)$ and there holds the a-priori estimate [25]

$$\|u\|_{L^2(J;X)} + \|u'\|_{L^2(J;X^*)} \le C(\|g\|_{L^2(J;H)} + \|u_0\|_H). \tag{2.9}$$

To describe the smoothness of the initial values we introduce spaces between H and X defined by the K-method of interpolation [25, 39] as

$$H_\theta = (H,X)_{\theta,2}, \qquad 0 \le \theta \le 1$$

(with the usual convention that $H_0 = H$ and $H_1 = X$). In Section 4 we investigate the time analyticity of solutions of (2.3), (2.4) where

$$u_0 \in H_\theta \text{ for some } 0 \le \theta \le 1 \tag{2.10}$$

and where the function $g \in L^2(J;H)$ is analytic as a function on $[0,T]$ with values in H, i.e. the time derivatives of g can be controlled by

$$\|g^{(l)}(t)\|_H \leq Cl!d^l, \qquad t \in [0,T],\ l \in \mathbb{N}_0 \tag{2.11}$$

with C and d independent of l and t.
In the error analysis ahead we need sometimes the assumption

$$u \in C([\varepsilon,T];X), \qquad \forall \varepsilon > 0. \tag{2.12}$$

Condition (2.12) holds true if u belongs to $H^1(J;X)$ or if assumption (2.11) on the right hand side is fulfilled.

Remark 2.1 The compactness assumption (2.2) and the self-adjointness assumption (2.7) are not essential. They allow us in Section 4 to study the time regularity by convenient Fourier series techniques. However, by means of classical semigroup theory all these regularity results can be generalized to the non-selfadjoint case, where (2.2) and (2.7) are not valid anymore. These rather technical issues are presented in the forthcoming work [34].

2.2. DGFEM Discretization

Let $\mathcal{M}$ be a partition of $J = (0,T)$ into $M(\mathcal{M})$ subintervals $\{I_m = (t_{m-1}, t_m)\}_{m=1}^M$. The time step k_m is $k_m := t_m - t_{m-1}$. We define the one-sided limits in H (or X) of a function u as

$$u_m^+ = \lim_{s\to 0,\ s>0} u(t_m + s),\ 0 \leq m \leq M-1, \qquad u_m^- = \lim_{s\to 0,\ s>0} u(t_m - s),\ 1 \leq m \leq M$$

and set $[u]_m = u_m^+ - u_m^-$, $1 \leq m \leq M-1$.
For $u, v \in L^2(I_m;X)$ with $u', v' \in L^2(I_m;X^*)$ we have $u, v \in C(\overline{I}_m;H)$ and the one-sided limits exist in H. On the mesh $\mathcal{M} = \{I_m\}_{m=1}^M$ we introduce the space $C_b(\mathcal{M};X) = \{u : J \to X : u|_{I_m} \in C_b(I_m;X)\}$ of X-valued functions which are bounded and piecewise continuous. ($C_b(I_m;X)$ denotes the bounded continuous functions on I_m.) We define

$$\begin{aligned} B_{DG}(u,v) &:= \sum_{m=1}^M \int_{I_m} \{(u',v)_{X^*\times X} + a(u,v)\}dt \\ &+ \sum_{m=2}^M ([u]_{m-1}, v_{m-1}^+)_H + (u_0^+, v_0^+)_H, \end{aligned} \tag{2.13}$$

$$F_{DG}(v) := \sum_{m=1}^M \int_{I_m} (g(t),v)_{X^*\times X}dt + (u_0, v_0^+)_H. \tag{2.14}$$

It can easily be seen by integration by parts that:

Lemma 2.2 *Let $u \in L^2(J;X) \cap H^1(J;X^*)$ be a weak solution of (2.3), (2.4) in the sense of (2.8). Then it satisfies $B_{DG}(u,v) = F_{DG}(v)$ for all $v \in C_b(\mathcal{M};X)$.*

We associate with each time interval I_m an approximation order $r_m \geq 0$ and store these temporal orders in the vector $\underline{r} := \{r_m\}_{m=1}^M$. The semidiscrete space in which we want to discretize (2.3), (2.4) in time is

$$\mathcal{V}^{\underline{r}}(\mathcal{M}; X) = \{u : J \to X : u|_{I_m} \in \mathcal{P}^{r_m}(I_m; X),\ 1 \leq m \leq M\}. \tag{2.15}$$

$\mathcal{V}^{\underline{r}}(\mathcal{M}; X)$ is a linear space consisting of piecewise polynomials with coefficients in X. If $\underline{r}$ is constant on each time interval, i.e. $r_m = r$ for all $1 \leq m \leq M$, we write simply $\mathcal{V}^r(\mathcal{M}; X)$. We set $\mathrm{NRDOF}(\mathcal{V}^{\underline{r}}(\mathcal{M}; X)) := \sum_{m=1}^{M(\mathcal{M})}(r_m+1)$ for the number of degrees of freedom of the time discretization. We note in passing that $\mathrm{NRDOF}(\mathcal{V}^{\underline{r}}(\mathcal{M}; X))$ spatial problems must still be discretized to obtain a fully discrete scheme (see Section 6 ahead). Therefore, $\mathrm{NRDOF}(\mathcal{V}^{\underline{r}}(\mathcal{M}; X))$ can be viewed as a crude measure for the cost of a given time-stepping scheme.

DGFEM 2.3 *Let $\mathcal{M} = \{I_m\}$ be a partition of $J = (0, T)$ and $\underline{r}$ an approximation order distribution on $\mathcal{M}$. The DGFEM for (2.3), (2.4) is to find $U \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$ such that*

$$B_{DG}(U, V) = F_{DG}(V)$$

for all $V \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$.

Remark 2.4 Note that in the form B_{DG} we have $(U', V)_{X^* \times X} = (U', V)_H$ for all $U, V \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$.

Remark 2.5 Owing to the discontinuity of the trial and test space the DGFEM 2.3 can be interpreted as implicit time marching scheme, where U is obtained by solving successively evolution problems on I_m for $m = 1, \ldots, M$ with initial values U_{m-1}^-. More precisely: If U is already given on the time intervals I_k, $1 \leq k \leq m-1$, we determine U on I_m by solving:
Find $U \in \mathcal{P}^{r_m}(I_m; X)$ such that

$$\int_{I_m} \{(U', V)_H + a(U, V)\}dt + (U_{m-1}^+, V_{m-1}^+)_H = \int_{I_m} (g, V)_{X^* \times X} dt + (U_{m-1}^-, V_{m-1}^+)_H \tag{2.16}$$

for all $V \in \mathcal{P}^{r_m}(I_m; X)$. Here we set $U_0^- = u_0$.

Proposition 2.6 *The DGFEM 2.3 has a unique solution $U \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$. If u is the solution of (2.3), (2.4), we have the Galerkin orthogonality*

$$B_{DG}(u - U, V) = 0 \qquad \forall V \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X).$$

Proof: This is shown using similar arguments as in [38]. □

Lemma 2.7 *For all $V, W \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$ there holds*

$$B_{DG}(V, W) = \sum_{m=1}^{M} \int_{I_m} (-V, W')_H + a(V, W)dt$$

$$- \sum_{m=1}^{M-1} (V_m^-, [W]_m)_H + (V_M^-, W_M^-)_H, \tag{2.17}$$

$$\text{Re } B_{DG}(V-W, V-W) = \sum_{m=1}^{M} \int_{I_m} \text{Re } a(V-W, V-W)dt$$
$$+ \frac{1}{2}\|(V-W)_0^+\|_H^2 + \frac{1}{2}\sum_{m=1}^{M-1} \|[V-W]_m\|_H^2 + \frac{1}{2}\|(V-W)_M^-\|_H^2. \tag{2.18}$$

Proof: Integration by parts yields

$$B_{DG}(V,W) = \sum_{m=1}^{M} \int_{I_m} \{-(V, W')_H + a(V,W)\}dt$$
$$+ \sum_{m=1}^{M} (V_m^-, W_m^-)_H - (V_{m-1}^+, W_{m-1}^+)_H + \sum_{m=2}^{M} ([V]_{m-1}, W_{m-1}^+)_H + (V_0^+, W_0^+)_H.$$

Rearranging the nodal contributions gives the first claim (2.17). To prove (2.18), we write

$$B_{DG}(V-W, V-W) = \frac{1}{2}B_{DG}(V-W, V-W) + \frac{1}{2}B_{DG}(V-W, V-W) =: T_1 + T_2.$$

Evaluating T_1 with (2.13) and T_2 with (2.17) shows the assertion. □

3. hp Error Analysis

In Sections 3.1, 3.2 and 3.3 we derive error estimates for the DGFEM which are explicit in the time steps k_m and in the polynomial orders r_m. As a consequence, we establish in Section 3.4 optimal convergence rates for the h- and p-version DGFEM on quasiuniform temporal partitions. In particular, the obtained bounds show that the DGFEM gives spectral accuracy for solutions with smooth time dependence.

3.1. Abstract Error Analysis of the DGFEM

First, we introduce a projector and show that it is well defined (we refer also to [38, p. 185] where the same projector is introduced, but with values in H).

Definition 3.1 *Let $I = (-1, 1)$. For a function $u \in L^2(I; X)$ which is continuous at $t = 1$ we define $\Pi^r u \in \mathcal{P}^r(I; X)$, $r \in \mathbb{N}_0$, via the $r + 1$ conditions*

$$\int_I (\Pi^r u - u, q)_H dt = 0 \ \forall q \in \mathcal{P}^{r-1}(I; X), \qquad \Pi^r u(+1) = u(+1) \in X. \tag{3.1}$$

Lemma 3.2 *Π^r in Definition 3.1 is well defined.*

Proof: Assume that u_1 and u_2 are two polynomials in $\mathcal{P}^r(I; X)$ which satisfy (3.1). Especially, we have $u_1(1) = u_2(1)$. Denote by L_i, $i \geq 0$, the Legendre polynomial of

degree i in $\mathcal{P}^i(I)$. The difference $u_1 - u_2$ can be developed into the series $u_1 - u_2 = \sum_{i=0}^{r} v_i L_i$ with $v_i = \int_I (u_1 - u_2) L_i dt \in X$. Fix now $k \in \{0, \ldots, r-1\}$. From (3.1) follows that $\int_I (u_1 - u_2, v L_k)_H dt = 0$ for all $v \in X$. Using the orthogonality properties of the Legendre polynomials we get $(v_k, v)_H = 0$ for all $v \in X$. Since X is dense in H, we conclude that $v_k = 0$ in H and thus $v_k = 0$ in X. The difference $u_1 - u_2$ is therefore given by $u_1 - u_2 = v_r L_r$, $v_r \in X$. Because of $u_1(1) = u_2(1)$ we have $v_r = 0$, which proves the uniqueness of a polynomial satisfying the conditions in Definition 3.1. The existence follows similarly by setting

$$\Pi^r u = \sum_{i=0}^{r-1} u_i L_i + (u(1) - \sum_{i=0}^{r-1} u_i) L_r = \sum_{i=0}^{r-1} u_i L_i + (\sum_{i=r}^{\infty} u_i) L_r, \tag{3.2}$$

where $u = \sum_{i=0}^{\infty} u_i L_i$ is the Legendre expansion of u. □

On an arbitrary interval (a, b) with $k := b - a > 0$ we define $\Pi^r_{(a,b)}$ via the linear map $Q : (-1, 1) \to (a, b)$, $\xi \mapsto x = \frac{1}{2}(a + b + \xi k)$ as $\Pi^r_{(a,b)} u = [\Pi^r (u \circ Q)] \circ Q^{-1}$.

Proposition 3.3 *Let u be the exact solution of (2.3), (2.4) and U the semidiscrete solution of the DGFEM 2.3 in $\mathcal{V}^{\underline{r}}(\mathcal{M}; X)$. Assume (2.12). Let $\mathcal{I}u \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$ be the interpolant of u which is defined on each time interval I_m as $\mathcal{I}u|_{I_m} = \Pi^{r_m}_{I_m}(u|_{I_m})$. Then there holds*

$$\|u - U\|_{L^2(J;X)} \leq C(1 + \frac{\alpha}{\beta}) \|u - \mathcal{I}u\|_{L^2(J;X)}.$$

The constant C is in particular independent of T.

Remark 3.4 Due to (2.12) the point values in X used for the intervalwise definition of $\mathcal{I}$ are well defined.

Proof: By Lemma 2.7 and the coercivity condition (2.6) we have

$$\begin{aligned} &\text{Re } B_{DG}(V - W, V - W) \geq \beta \int_J \|V - W\|_X^2 dt \\ &+ \frac{1}{2} \|(V - W)_0^+\|_H^2 + \frac{1}{2} \sum_{m=1}^{M-1} \|[V - W]_m\|_H^2 + \frac{1}{2} \|(V - W)_M^-\|_H^2 \end{aligned}$$

for all $V, W \in \mathcal{V}^{\underline{r}}(\mathcal{M}; X)$. Hence, we get

$$\beta \int_J \|U - \mathcal{I}u\|_X^2 dt \leq \text{Re } B_{DG}(U - \mathcal{I}u, U - \mathcal{I}u) \leq |B_{DG}(u - \mathcal{I}u, U - \mathcal{I}u)|,$$

where we used in the last step Proposition 2.6. Writing Θ for $U - \mathcal{I}u$ we get with Lemma 2.7 and the definition of $\mathcal{I}$ that

$$\begin{aligned} \beta \int_J \|U - \mathcal{I}u\|_X^2 dt &\leq \int_J |\{-(u - \mathcal{I}u, \Theta')_H\} + a(u - \mathcal{I}u, \Theta)| dt \\ &\quad + \sum_{m=1}^{M-1} |((u - \mathcal{I}u)_m^-, [\Theta]_m)_H| + |((u - \mathcal{I}u)_M^-, \Theta_M^-)_H| \\ &= \int_J |a(u - \mathcal{I}u, \Theta)| dt \leq \alpha \int_J \|u - \mathcal{I}u\|_X \|\Theta\|_X dt. \end{aligned}$$

We conclude now with the inequality of Cauchy-Schwarz that $\int_J \|U - \mathcal{I}u\|_X^2 dt \leq \frac{\alpha^2}{\beta^2} \int_J \|u - \mathcal{I}u\|_X^2 dt$. The assertion follows with the triangle inequality. $\square$

3.2. Properties of the Projector Π^r

We analyze the projector Π^r in Definition 3.1. Let $I = (-1, 1)$ and denote by $\{L_i\}_{i\geq 0}$, $L_i \in \mathcal{P}^i(I)$, the Legendre polynomials on I.

Lemma 3.5 *Let $u \in L^2(I;X)$ be continuous at $t = 1$ and let $u = \sum_{i=0}^{\infty} u_i L_i$ be the Legendre expansion of u with coefficients $u_i = \int_I u L_i(t) dt \in X$. For $r \in \mathbb{N}_0$ we denote by P^r the $L^2(I;X)$-projection onto $\mathcal{P}^r(I;X)$. There holds:*

$$\|u - \Pi^r u\|^2_{L^2(I;X)} \leq C\|u - P^r u\|^2_{L^2(I;X)} + C \max(1, r)^{-1} (\sum_{i=r+1}^{\infty} \|u_i\|_X)^2.$$

Proof: From (3.2) we have $\Pi^r u = \sum_{i=0}^{r-1} u_i L_i + (\sum_{i=r}^{\infty} u_i) L_r$. Therefore,

$$u - \Pi^r u = \sum_{i=r}^{\infty} u_i L_i - (\sum_{i=r}^{\infty} u_i) L_r = \sum_{i=r+1}^{\infty} u_i L_i - (\sum_{i=r+1}^{\infty} u_i) L_r.$$

Due to the orthogonality properties of the Legendre polynomials this is $u - \Pi^r u = (u - P^r u) - (\sum_{i=r+1}^{\infty} u_i) L_r$ and the assertion follows by the triangle inequality. $\square$

Lemma 3.6 *For $r \in \mathbb{N}_0$ and $u \in H^1(I;X)$ we have*

$$\|u - \Pi^r u\|^2_{L^2(I;X)} \leq C\{\|u - P^r u\|^2_{L^2(I;X)} + \frac{\|u'\|^2_{L^2(I;X)}}{\max(1, r^2)}\}, \tag{3.3}$$

$$\|u - \Pi^r u\|^2_{L^2(I;X)} \leq C \inf_{q \in \mathcal{P}^r(I;X)} \{\|u - q\|^2_{L^2(I;X)} + \frac{\|u' - q'\|^2_{L^2(I;X)}}{\max(1, r^2)}\}. \tag{3.4}$$

Proof: Taking into account Lemma 3.5 the first assertion in (3.3) follows if we prove

$$\sum_{i=r+1}^{\infty} \|u_i\|_X \leq \frac{C}{\max(1, r)^{\frac{1}{2}}} \|u'\|_{L^2(I;X)} \tag{3.5}$$

where $u = \sum_{i=0}^{\infty} u_i L_i$. We develop u' into the Legendre series $u' = \sum_{i=0}^{\infty} b_i L_i$ with coefficients $b_i \in X$. Then u can be written as $u(t) = \sum_{i=0}^{\infty} b_i \int_{-1}^{t} L_i(s) ds + u(-1)$. Recall that $\int_{-1}^{t} L_i(s) ds = \frac{1}{2i+1}(L_{i+1}(t) - L_{i-1}(t))$ for $i \geq 1$. Hence,

$$\begin{aligned} u &= (b_0 + \frac{u(-1)}{2}) L_0 + \sum_{i=1}^{\infty} \frac{b_i}{2i+1} L_{i+1} - \sum_{i=1}^{\infty} \frac{b_i}{2i+1} L_{i-1} \\ &= (b_0 + \frac{u(-1)}{2}) L_0 + \sum_{i=2}^{\infty} \frac{b_{i-1}}{2i-1} L_i - \sum_{i=0}^{\infty} \frac{b_{i+1}}{2i+3} L_i. \end{aligned}$$

Comparing equal coefficients in the Legendre expansions we get

$$\begin{aligned} u_0 &= b_0 + \frac{u(-1)}{2} - \frac{b_1}{3}, \qquad u_1 = -\frac{b_2}{5}, \\ u_i &= \frac{b_{i-1}}{2i-1} - \frac{b_{i+1}}{2i+3}, \qquad i \geq 2. \end{aligned}$$

Consequently, we have for $r \geq 1$

$$\begin{aligned} \sum_{i=r+1}^{\infty} u_i &= \sum_{i=r+1}^{\infty} \frac{b_{i-1}}{2i-1} - \frac{b_{i+1}}{2i+3} = \sum_{i=r}^{\infty} \frac{b_i}{2i+1} - \sum_{i=r+2}^{\infty} \frac{b_i}{2i+1} \\ &= \frac{b_r}{2r+1} + \frac{b_{r+1}}{2r+3}. \end{aligned}$$

This yields

$$\sum_{i=r+1}^{\infty} \|u_i\|_X \leq C\left(\frac{\|b_r\|_X^2}{(2r+1)^2} + \frac{\|b_{r+1}\|_X^2}{(2r+3)^2}\right)^{\frac{1}{2}} \leq C r^{-\frac{1}{2}} \|u'\|_{L^2(I;X)},$$

which proves (3.5) and thus (3.3). For $r = 0$ the inequality (3.5) is established similarly.
Let now $q \in \mathcal{P}^r(I;X)$ be arbitrary. Insert $u - q$ into (3.3): The assertion (3.4) follows from the reproducing properties of the projectors Π^r and P^r and from the fact that $\|u - P^r u\|_{L^2(I;X)} \leq \|u - q\|_{L^2(I;X)}$. □

Lemma 3.7 *For $u \in L^2(I;X)$ which is continuous at $t = 1$ and $r \in \mathbb{N}_0$ we have*

$$\|\Pi^r u\|_{L^2(I;X)}^2 \leq C \max(1,r) \|u\|_{L^2(I;X)}^2 + \frac{C}{\max(1,r)} \|u(+1)\|_X^2.$$

Proof: Assume first $r \geq 1$. We develop u into the Legendre series $u = \sum_{i=0}^{\infty} u_i L_i$. Due to (3.2) we have $\Pi^r u = \sum_{i=0}^{r-1} u_i L_i + (u(1) - \sum_{i=0}^{r-1} u_i) L_r$. Then by the triangle inequality

$$\begin{aligned} \|\Pi^r u\|_{L^2(I;X)}^2 &\leq \sum_{i=0}^{r-1} \|u_i\|_X^2 \frac{2}{2i+1} + \frac{C}{\max(1,r)} \|u(1) - \sum_{i=0}^{r-1} u_i\|_X^2 \\ &\leq \|u\|_{L^2(I;X)}^2 + \frac{C}{\max(1,r)} \|u(1)\|_X^2 + \frac{C}{\max(1,r)} \Big(\sum_{i=0}^{r-1} \|u_i\|_X\Big)^2. \end{aligned}$$

The last sum can be bounded by

$$\Big(\sum_{i=0}^{r-1} \|u_i\|_X\Big)^2 \leq \Big(\sum_{i=0}^{r-1} \|u_i\|_X^2 \frac{2}{2i+1}\Big)\Big(\sum_{i=0}^{r-1} \frac{2i+1}{2}\Big) \leq \frac{2r+1}{2} r \|u\|_{L^2(I;X)}^2,$$

which proves the assertion. The modifications for $r = 0$ are obvious. □

Lemma 3.8 *Consider two polynomial degrees $0 \le r' \le r$. Then we have for functions $u \in L^2(I;X)$ which are continuous at $t = 1$:*

$$\|u - \Pi^r u\|^2_{L^2(I;X)} \le C \max(1,r) \|u - \Pi^{r'} u\|^2_{L^2(I;X)}.$$

Proof: Applying Lemma 3.7 for $u - \Pi^{r'} u$ yields

$$\|\Pi^r u - \Pi^{r'} u\|_{L^2(I;X)} \le C \max(1,r) \|u - \Pi^{r'} u\|_{L^2(I;X)}.$$

The assertion follows with the triangle inequality. □

3.3. hp Approximation Results

We recall the following *hp* approximation result from [35]. There, the proof is presented for real-valued functions, but the extension to the Bochner spaces considered here is straightforward.

Proposition 3.9 *Let $I = (-1,1)$ and let $u \in H^{s_0+1}(I;X)$ for some integer $s_0 \in \mathbb{N}_0$. Then there exists $q \in \mathcal{P}^r(I;X)$, $r \in \mathbb{N}_0$, such that*

$$\begin{aligned} \|u' - q'\|^2_{L^2(I;X)} &\le C \frac{(r-s)!}{(r+s)!} |u'|^2_{H^s(I;X)}, \\ \|u - q\|^2_{L^2(I;X)} &\le C \frac{1}{\max(1,r^2)} \frac{(r-t)!}{(r+t)!} |u'|^2_{H^t(I;X)} \end{aligned}$$

for any $0 \le s,t \le \min(r, s_0)$. Additionally, $q(\pm 1) = u(\pm 1)$ if $r \ge 1$.

The application of Proposition 3.9 in Lemma 3.6 and scaling gives estimates for Π^r_I. By interpolation these estimates can be extended straightforwardly to noninteger Sobolev spaces.

Theorem 3.10 *Let $I = (a,b)$, $k = b - a$, $r \in \mathbb{N}_0$ and $u \in H^{s_0+1}(I;X)$ for some $s_0 \in \mathbb{N}_0$. Then*

$$\|u - \Pi^r_I u\|^2_{L^2(I;X)} \le C (\frac{k}{2})^{2(s+1)} \frac{1}{\max(1,r)^2} \frac{(r-s)!}{(r+s)!} \|u^{(s+1)}\|^2_{L^2(I;X)}$$

for any integer $0 \le s \le \min(r, s_0)$. Moreover, we get by interpolation

$$\|u - \Pi^r_I u\|^2_{L^2(I;X)} \le \frac{C}{\max(1,r)^2} \frac{\Gamma(r+1-s)}{\Gamma(r+1+s)} (\frac{k}{2})^{2(s+1)} \|u\|^2_{H^{s+1}(I;X)}$$

for any real $0 \le s \le \min(r, s_0)$.

Proposition 3.3 and Theorem 3.10 result in *hp* error estimates for the DGFEM 2.3 which are valid if the exact solution is at least in $H^1(J;X)$:

Theorem 3.11 *Let u be the exact solution of (2.3), (2.4) and U the semidiscrete solution of the DGFEM 2.3 in $\mathcal{V}^{\underline{r}}(\mathcal{M};X)$. Assume that $u|_{I_m} \in H^{s_{0,m}+1}(I_m;X)$ for $1 \le m \le M$ and $s_{0,m} \in \mathbb{N}_0$. Then we have*

$$\|u - U\|^2_{L^2(J;X)} \le C \sum_{m=1}^{M} (\frac{k_m}{2})^{2(s_m+1)} \max(1,r_m)^{-2} \frac{\Gamma(r_m+1-s_m)}{\Gamma(r_m+1+s_m)} \|u\|^2_{H^{s_m+1}(I_m;X)}$$

for any $0 \le s_m \le \min(r_m, s_{0,m})$.

3.4. h- and p-Version DGFEM on Quasiuniform Time Partitions

In the h-version of the DGFEM the approximation order is kept fixed, $r_m = r$, where r is typically low. The convergence of the semidiscrete DGFEM solutions U in $\mathcal{V}^r(\mathcal{M}; X)$ to the exact one is obtained by decreasing the length of the time steps in a quasi- or non-quasiuniform way. The p-version of the DGFEM on the other hand uses a fixed temporal partition $\mathcal{M}$ which is typically quasiuniform. The convergence is achieved by letting the time approximation order $r \to \infty$. From Theorem 3.11 we can easily establish the following convergence rates for the h- and p-version on quasiuniform time partitions. Remark that $\frac{\Gamma(r+1-s)}{\Gamma(r+1+s)} \sim r^{-2s}$ for $r \to \infty$, as can be seen from Stirling's formula [31].

Corollary 3.12 *Let $\mathcal{M}$ be a time partition with quasiuniform steps where $k := \max\{k_m\}/k_m \leq \tau$ for a constant $\tau > 0$. Let the approximation order be constant, $r_m = r \in \mathbb{N}_0$. Let $u \in H^{s_0+1}(J; X)$, $s_0 \in \mathbb{N}_0$, be the exact solution of (2.3), (2.4) and U the semidiscrete solution of the DGFEM 2.3 in $\mathcal{V}^r(\mathcal{M}; X)$. Then there holds*

$$\|u - U\|^2_{L^2(J;X)} \leq C \frac{k^{2\{\min(s_0,r)+1\}}}{r^{2(s_0+1)}} \|u\|^2_{H^{s_0+1}(J;X)}, \tag{3.6}$$

where C depends on s_0, but is independent of k and r.

Remark 3.13 The error estimates in Corollary 3.12 are uniform in k and r. They show that the DGFEM converges either as the time steps are reduced ($k \to 0$) or as the temporal order r is increased ($r \to \infty$). They also show that for smooth solutions where s_0 is large it is more advantageous to increase r than to reduce k at fixed, low r. The estimates in Corollary 3.12 give algebraic convergence rates in $N = \mathrm{NRDOF}(\mathcal{V}^r(\mathcal{M}; X))$, namely $\|u - U\|_{L^2(J;X)} \leq C N^{-\min(r,s_0)-1}$, if the h-version is used with a fixed approximation order r (in this case $N \sim 1/k$). If the p-version is employed, we get accordingly $\|u - U\|_{L^2(J;X)} \leq C N^{-s_0-1}$ (here $N \sim r$).

Remark 3.14 For solutions u which are analytic in $\overline{J} = \overline{(0, T)}$ the p-version results in exponential rates of convergence. Namely, by standard approximation theory for analytic functions there holds as a consequence of Proposition 3.3 and Lemma 3.6 that $\|u - U\|^2_{L^2(J;X)} \leq C \exp(-br)$. Therefore, the p-version of the DGFEM is especially recommended for problems where the exact solution is analytic in time.

4. Time Regularity

In this section we study the analyticity properties of solutions u of (2.3), (2.4). In Section 4.1 we use Duhamel's formula to express the evolution operator and we derive in Section 4.2 bounds on the derivatives of u that are explicit in l and t.

4.1. The Evolution Operator

Due to (2.7) the form $(u, v) \mapsto a(u, v)$ is an inner product on X and we can equip X with the energy norm $\|u\|^2_E = a(u, u)$. We have (cf. (2.5), (2.6)) $\beta\|u\|^2_X \leq \|u\|^2_E \leq$

$\alpha\|u\|_X^2$. The assumption (2.2) implies the existence of sequences $\{\lambda_i\}_{i\in\mathbb{N}}$ and $\{\varphi_i\}_{i\in\mathbb{N}}$ of real eigenvalues $\lambda_i > 0$ and eigenfunctions $\varphi_i \in X$ of the eigenvalue problem

$$u \in X: \ a(u,v) = \lambda(u,v)_H, \qquad \forall v \in X.$$

We can assume that the eigenvalues λ_i are non-decreasing and that the eigenfunctions $\{\varphi_i\}$ form an orthonormal basis of H. We have

$$\|u\|_H^2 = \sum_{i=1}^{\infty} |(u,\varphi_i)_H|^2, \qquad \|u\|_E^2 = \sum_{i=1}^{\infty} \lambda_i |(u,\varphi_i)_H|^2 \tag{4.1}$$

for functions $u \in H$ and $u \in X$, respectively. We define for $t \geq 0$ the linear operator $T(t)$ on H as

$$T(t)u = \sum_{i=1}^{\infty} e^{-\lambda_i t}(u,\varphi_i)_H \varphi_i, \qquad u \in H, \ t \geq 0. \tag{4.2}$$

Obviously, we have the semigroup properties

$$T(t+s) = T(t)T(s) \quad \text{for } t,s \geq 0, \quad T(0) = I. \tag{4.3}$$

In terms of the semigroup $T(t)$ the solution u of (2.3), (2.4) with data $u_0 \in H$ and $g \in L^2(J;H)$ can be expressed by Duhamel's formula

$$u(t) = T(t)u_0 + \int_0^t T(t-s)g(s)ds, \qquad 0 \leq t \leq T. \tag{4.4}$$

Proposition 4.1 *There exist constants $C_1, C_2, d_1, d_2 > 0$ such that*

$$\|T^{(l)}(t)\|_{\mathcal{L}(H,X)}^2 \leq C_1 d_1^{2l+1}\Gamma(2l+2)t^{-(2l+1)}, \qquad \|T^{(l)}(t)\|_{\mathcal{L}(X,X)}^2 \leq C_2 d_2^{2l}\Gamma(2l+1)t^{-2l}$$

for all $t > 0$ and $l \in \mathbb{N}_0$.

Proof: Let first u be in H and set $u_i = (u,\varphi_i)_H$. We have for $l \in \mathbb{N}_0$ that $T^{(l)}(t)u(x) = \sum_{i=1}^{\infty}(-\lambda_i)^l e^{-\lambda_i t}u_i\varphi_i(x)$ and get with (4.1)

$$\|T^{(l)}(t)u\|_E^2 = \sum_{i=1}^{\infty} \lambda_i^{2l} e^{-2\lambda_i t}|u_i|^2\lambda_i = \sum_{i=1}^{\infty} \lambda_i^{2l+1} e^{-2\lambda_i t}|u_i|^2.$$

The function $f(\lambda) = \lambda^{2l+1}e^{-2\lambda t}$ takes its maximal value on $\mathbb{R}^+$ at $\lambda_{max} = \frac{2l+1}{2t}$. This gives

$$\|T^{(l)}(t)u\|_E^2 \leq f(\lambda_{max})\|u\|_H^2 \leq (2l+1)^{2l+1}(\frac{1}{2})^{2l+1}t^{-(2l+1)}e^{-(2l+1)}\|u\|_H^2.$$

An application of the formula of Stirling [31] yields

$$(2l+1)^{2l+1}e^{-(2l+1)} \leq C\Gamma(2l+2)\frac{1}{\sqrt{2l+1}} \leq C\Gamma(2l+2),$$

which allows us to conclude that

$$\beta\|T^{(l)}(t)u\|_X^2 \le \|T^{(l)}(t)u\|_E^2 \le C(\frac{1}{2})^{2l+1}\Gamma(2l+2)t^{-(2l+1)}\|u\|_H^2.$$

This proves the differentiability of T as a function of $t>0$ with values in $\mathcal{L}(H,X)$ and the first assertion.
Fix now $u \in X$ and set, as before, $u_i = (u,\varphi_i)_H$. We consider only $l \ge 1$, the case $l=0$ being completely analogous. We get $\|T^{(l)}(t)u\|_E^2 = \sum_{i=1}^\infty \lambda_i^{2l} e^{-2\lambda_i t}|u_i|^2\lambda_i$. Define $g(\lambda) = \lambda^{2l}e^{-2\lambda t}$. Here, the maximal value is at $\lambda_{max} = \frac{2l}{2t}$, which gives

$$\|T^{(l)}(t)u\|_E^2 \le g(\lambda_{max})\|u\|_E^2 \le (2l)^{2l}(\frac{1}{2})^{2l}t^{-2l}e^{-2l}\|u\|_E^2.$$

The second assertion follows again with the formula of Stirling. □

Interpolating the estimates in Proposition 4.1 gives:

Proposition 4.2 *There are constants $C,d>0$ such that*

$$\|T^{(l)}(t)\|_{\mathcal{L}(H_\theta,X)}^2 \le Cd^{2l+1-\theta}\Gamma(2l+2-\theta)t^{-(2l+1)+\theta}$$

for $t>0$, $l \in \mathbb{N}_0$, $0 \le \theta \le 1$.

4.2. Analyticity

We split the solution u of (2.3), (2.4) into $u = u_1 + u_2$, where u_1 and u_2 solve

$$u_1' + Lu_1 = 0,\ u_1(0) = u_0, \qquad u_2' + Lu_2 = g,\ u_2(0) = 0. \tag{4.5}$$

Proposition 4.2 and (4.4) allow us to control u_1. It remains to derive analogous bounds for u_2:

Lemma 4.3 *Under the analyticity assumption (2.11) we have for u_2 in (4.5) the solution formula $u_2(t) = \int_0^t T(s)g(t-s)ds$ and there holds*

$$u_2^{(l)}(t) = \sum_{i=0}^{l-1} T^{(i)}(t)g^{(l-1-i)}(0) + \int_0^t T(s)g^{(l)}(t-s)ds, \qquad l \ge 1.$$

Proof: u_2 can be represented according to (4.4). The first claim is then a simple change of variables, the second one is obtained from the first one with an induction argument. □

Lemma 4.4 *Assume (2.11). Then there exist constants $C,d>0$ such that*

$$\|u_2^{(l)}(t)\|_X \le Cd^l\Gamma(l+1)(t^{1/2} + \sum_{i=0}^{l-1} t^{-i-1/2})$$

for $l \in \mathbb{N}_0$, $t>0$. (For $l=0$ the last sum is empty.)

Proof: Because of Lemma 4.3 we have

$$\|u_2^{(l)}(t)\|_X \leq \sum_{i=0}^{l-1} \|T^{(i)}(t)\|_{\mathcal{L}(H,X)} \|g^{(l-1-i)}(0)\|_H + \int_0^t \|T(s)\|_{\mathcal{L}(H,X)} \|g^{(l)}(t-s)\|_H ds \tag{4.6}$$

We first bound the sum. Recall from Proposition 4.2 that for any $t > 0$, $i \in \mathbb{N}_0$

$$\|T^{(i)}(t)\|_{\mathcal{L}(H,X)} \leq C d_1^{i+1/2} \Gamma(2i+2)^{1/2} t^{-i-1/2}.$$

Employing standard properties of the Gamma function [31] we get

$$\Gamma(2i+2)^{1/2} \leq C 2^{i+1/2} \Gamma(i+1)^{1/2} \Gamma(i+3/2)^{1/2} \leq C 2^{i+1/2} (i+1)!$$

and consequently

$$\|T^{(i)}(t)\|_{\mathcal{L}(H,X)} \leq C d_2^{i+1/2} (i+1)! t^{-i-1/2}, \qquad i \in \mathbb{N}_0. \tag{4.7}$$

Using (2.11) and (4.7) we conclude

$$\begin{aligned}
&\sum_{i=0}^{l-1} \|T^{(i)}(t)\|_{\mathcal{L}(H,X)} \|g^{(l-1-i)}(0)\|_H \leq C d_3^{l-1+1/2} \sum_{i=0}^{l-1} (i+1)!(l-1-i)! t^{-i-1/2} \\
&\leq C d_3^{l-1/2} l! \sum_{i=0}^{l-1} \binom{l}{i+1}^{-1} t^{-i-1/2} \leq C d_3^{l-1/2} \Gamma(l+1) \sum_{i=0}^{l-1} t^{-i-1/2}.
\end{aligned}$$

Using again (2.11) and (4.7) the integral I in (4.6) can be bounded by

$$I \leq C d_4^l l! \int_0^t s^{-1/2} ds = C d_4^l \Gamma(l+1) t^{1/2}.$$

This proves the assertion. □

Proposition 4.5 *Assume (2.11). Then there exist constants $C, d > 0$ such that $\|u_2^{(l)}(t)\|_X^2 \leq C d^{2l} \Gamma(2l+2) t^{-2l+1}$ for $l \in \mathbb{N}_0$ and $0 < t \leq \min(1, T)$.*

Proof: For $0 < t \leq \min(1, T)$ we have from Lemma 4.4

$$\|u_2^{(l)}(t)\|_X \leq C d_1^l \Gamma(l+1)(l+1) t^{-l+1-1/2} \leq C d_2^l \Gamma(l+1) t^{-l+1/2}.$$

There holds $\Gamma(l+1)^2 \leq \Gamma(l+1)\Gamma(l+3/2) = C\Gamma(2l+2) 2^{-2(l+1)}$, which finishes the proof. □

Combining the estimates for u_1 and u_2 obtained in Proposition 4.2 and Proposition 4.5 results in:

Proposition 4.6 *Let $u_0 \in H_\theta$ for $0 \leq \theta \leq 1$ and let $g \in L^2(J; H)$ fulfill (2.11). Then the solution u of (2.3), (2.4) satisfies*

$$\|u^{(l)}(t)\|_X^2 \leq C d^{2l} \Gamma(2l+2) t^{-(2l+1)+\theta}$$

for $l \in \mathbb{N}_0$ and $0 < t \leq \min(1, T)$. The constants C and d depend on u_0 and g, but are independent of t and l. Further, if $0 < a < b \leq \min(1, T)$, then

$$\int_a^b \|u^{(l)}(t)\|_X^2 dt \leq C d^{2l} \Gamma(2l+2) a^{-2l+\theta}$$

for all $l \in \mathbb{N}$. By interpolation, we get immediately

$$\|u\|^2_{H^s((a,b);X)} \leq C d^{2s} \Gamma(2s+3) a^{-2s+\theta} \tag{4.8}$$

for all real numbers $s \geq 1$ and $0 < a < b \leq \min(1, T)$.

Proposition 4.6 describes the singular temporal structure of the solution at $t = 0$. The regularity is given explicitly in terms of t and l (respectively s). However, the farther away from the origin we move the better the solution is smoothened and the better it behaves analytically. This is expressed (qualitatively) in the following proposition. It is a direct consequence of Proposition 4.2 and Lemma 4.4.

Proposition 4.7 *Let $u_0 \in H$ and let $g \in L^2(J; H)$ satisfy (2.11). Fix $0 < t_0 < T$. Then we have*

$$\|u^{(l)}(t)\|_X \leq C d^l l!, \qquad l \in \mathbb{N}_0,\ t_0 \leq t \leq T$$

with constants C and d just depending on t_0, T, u_0 and g. In particular, the solution u satisfies (2.12).

On the first time interval I_1 near $t = 0$ we approximate in Section 5 the solution u by a constant polynomial and use the following Hardy-type inequality:

Proposition 4.8 *Let $u_0 \in H_\theta$ for some $0 \leq \theta \leq 1$ and let g satisfy (2.11). Then there holds $\int_0^k \|u(t) - u(k)\|_X^2 dt \leq C k^\theta$ for any $0 < k \leq \min(1, T)$ with a constant C independent of k. (The point value $u(k) \in X$ is well defined due to (2.12).)*

Proof: We split the solution u into $u_1 + u_2$ as in (4.5). Then:

$$\int_0^k \|u - u(k)\|_X^2 dt \leq C \int_0^k \|u_1(t)\|_X^2 dt + C \int_0^k \|u_1(k)\|_X^2 dt$$
$$+C \int_0^k \|u_2(t)\|_X^2 dt + C \int_0^k \|u_2(k)\|_X^2 dt =: T_1 + T_2 + T_3 + T_4.$$

We bound first T_1: To do so, let S be the operator $u_0 \mapsto S u_0 = u_1$ which maps the initial condition u_0 to the solution u_1 of the homogeneous equation in (4.5). By (4.4) and Proposition 4.2 we see that $\|S u_0\|^2_{L^2((0,k);X)} \leq C k \|u_0\|_X^2$. Using a-priori estimates as in (2.9) it can moreover be seen that $\|S u_0\|^2_{L^2((0,k);X)} \leq \|S u_0\|^2_{L^2(J;X)} \leq C \|u_0\|_H^2$ with a constant C independent of k. Interpolating these estimates results in

$$\|S u_0\|^2_{L^2((0,k);X)} \leq C k^\theta \|u_0\|^2_{H_\theta} \qquad \text{and} \qquad T_1 \leq C k^\theta. \tag{4.9}$$

To bound T_2 we use again (4.4) and Proposition 4.2. We get

$$T_2 \leq C \int_0^k k^{-1+\theta}\, dt \leq C k^\theta. \tag{4.10}$$

T_3 and T_4 are bounded by Proposition 4.5:

$$T_3 \le C \int_0^k t \; dt \le Ck^2, \qquad T_4 \le C \int_0^k k \; dt \le Ck^2. \tag{4.11}$$

Combining (4.9), (4.10) and (4.11) proves the assertion. □

5. DGFEM on Geometric and Graded Temporal Meshes

In order to resolve the start-up singularity at $t = 0$ due to incompatibilities of the initial data we use time partitions that are refined towards $t = 0$. In the *hp*-version of the DGFEM geometrically refined time meshes and linearly increasing approximation orders are employed. We prove in Section 5.1 that this combination of *h*- and *p*- refinement leads to exponential rates of convergence.
We show in Section 5.2 that the use of graded meshes in the *h*-version DGFEM yields the optimal algebraic convergence rates, independently of the compatibility of the initial data.

5.1. Exponential Convergence on Geometric Time Meshes

Approximation on basic geometric partitions: We consider (2.3), (2.4) on $J = (0, 1)$ discretized on geometric temporal partitions.

Definition 5.1 *The (basic) geometric partition $\mathcal{M}_{n,\sigma} = \{I_m\}_{m=1}^{n+1}$ of $J = (0, 1)$ with grading factor $\sigma \in (0, 1)$ and $n + 1$ time intervals I_m is given by the nodes*

$$t_0 = 0, \qquad t_m = \sigma^{n-m+1}, \qquad 1 \le m \le n + 1. \tag{5.1}$$

For $2 \le m \le n+1$ the time steps $k_m = t_m - t_{m-1}$ satisfy $k_m = \lambda t_{m-1}$ with $\lambda = \frac{1-\sigma}{\sigma}$. We adress first the approximation on the intervals $\{I_m\}_{m=2}^{n+1}$ away from $t = 0$. We choose on each of these intervals I_m a polynomial order $r_m \ge 1$. As before, the regularity of the solution on I_m is measured by the parameter $s_m \ge 0$.

Lemma 5.2 *Fix an interval $I_m \in \mathcal{M}_{n,\sigma}$ for $2 \le m \le n + 1$ and set $s_m = \alpha_m r_m$ with $\alpha_m \in (0, 1)$. Then there exist constants $C, d > 0$ such that*

$$\|u - \Pi_{I_m}^{r_m} u\|^2_{L^2(I_m;X)} \le C\sigma^{(n-m+2)\theta}((\gamma d)^{2\alpha_m}[\frac{(1-\alpha_m)^{1-\alpha_m}}{(1+\alpha_m)^{1+\alpha_m}}])^{r_m}$$

with C, d only depending on $u_0 \in H_\theta$, $0 \le \theta \le 1$, and $g \in L^2(J; H)$ satisfying (2.11).

Proof: Since we consider a fixed interval $I = I_m$, $2 \le m \le n + 1$, we omit for simplicity the subscript m and write r, α, k and s in the calculations. We set also $t = t_{m-1}$. Proposition 4.6 and Theorem 3.10 yield

$$\|u - \Pi_I^r\|^2_{L^2(I;X)} \;\le\; C\frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)}(\frac{k}{2})^{2(s+1)}\|u\|^2_{H^{s+1}(I;X)}$$

$$
\begin{aligned}
&\leq C\frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)}(\frac{\lambda}{2})^{2(s+1)}t^{2s+2}\|u\|^2_{H^{s+1}(I;X)}\\
&\leq C(\frac{\gamma d_1}{2})^{2(s+1)}\frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)}\Gamma(2(s+1)+3)t^{2s+2}t^{-2s-2+\theta}\\
&\leq C(\frac{\gamma d_2}{2})^{2s}\frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)}\Gamma(2s+1)t^{\theta}.
\end{aligned}
$$

Inserting now $s=\alpha r$ we get with Stirling's formula

$$
\begin{aligned}
&\frac{\Gamma(r+1-s)}{\Gamma(r+1+s)}\Gamma(2s+1)\leq C\frac{(r-s)^{r-s}e^{-(r-s)}\sqrt{r-s}}{(r+s)^{r+s}e^{-(r+s)}\sqrt{r+s}}(2s)^{2s}e^{-2s}\sqrt{s}\\
&\leq \frac{r^{r(1-\alpha)}}{r^{r(1+\alpha)}}(\frac{(1-\alpha)^{1-\alpha}}{(1+\alpha)^{1+\alpha}})^r e^{2s}2^{2s}\alpha^{2s}r^{2\alpha r}e^{-2s}\sqrt{r}\leq r^{1/2}2^{2s}[\frac{(1-\alpha)^{1-\alpha}}{(1+\alpha)^{1+\alpha}}]^r.
\end{aligned}
$$

The claim follows. □

To prove exponential convergence in the *hp*-DGFEM, we start at Proposition 3.3 with the interpolant $\mathcal{I}u$ (which is defined intervalwise in Definition 3.1) and use Lemma 3.8 to get the error bound

$$
\|u-U\|^2_{L^2(J;X)}\leq C\max(1,r_1)\|u-\Pi^0_{I_1}u\|^2_{L^2(I_1;X)}+\sum_{m=2}^{n+1}\|u-\Pi^{r_m}_{I_m}u\|^2_{L^2(I_m;X)} \quad (5.2)
$$

for all $U\in\mathcal{V}^{\underline{r}}(\mathcal{M}_{n,\sigma};X)$ and polynomial vectors $\underline{r}$.
We consider orders which increase linearly in time:

Definition 5.3 *A polynomial degree vector $\underline{r}=\{r_m\}_{m=1}^{n+1}$ is called linear with slope $\mu>0$ on the geometric partition $\mathcal{M}_{n,\sigma}$ if $r_1=0$ and $r_m=\lfloor\mu m\rfloor$ for $2\leq m\leq n+1$.*

Theorem 5.4 *Consider the parabolic problem (2.3), (2.4) on $J=(0,1)$ with initial value $u_0\in H_\theta$ for some $0<\theta\leq 1$ and right-hand side g satisfying (2.11). This equation is discretized in t using the DGFEM 2.3 on a geometric partition $\mathcal{M}_{n,\sigma}$. Then there exists $\mu_0>0$ such that for all linear polynomial degree vectors $\underline{r}=\{r_m\}_{m=1}^{n+1}$ with slope $\mu\geq\mu_0$ the semidiscrete DGFEM solution U obtained in $\mathcal{V}^{\underline{r}}(\mathcal{M}_{n,\sigma};X)$ satisfies*

$$
\|u-U\|^2_{L^2(J;X)}\leq C\exp(-bN^{\frac{1}{2}})
$$

with constants C and b independent of $N=\mathrm{NRDOF}(\mathcal{V}^{\underline{r}}(\mathcal{M}_{n,\sigma};X))$.

Proof: Let

$$
\mu>\max\{1,\frac{\theta\ln(\sigma)}{\ln(f_{min})}\}, \quad (5.3)
$$

where $0<f_{min}<1$ is defined below. Set $r_1=0$ and $r_m=\lfloor\mu m\rfloor\geq 1$ for $2\leq m\leq n+1$. As before, $s_m=\alpha_m\rho_m$, $2\leq m\leq n+1$, for $\alpha_m\in(0,1)$ to be selected. We

start with (5.2): In the first interval I_1 near the origin we use the estimate derived in Proposition 4.8 and for $I_2, \ldots, I_{n+1}$ the one in Lemma 5.2. This yields

$$\begin{aligned} \|u-U\|^2_{L^2(J;X)} &\leq C\sigma^{n\theta} + C\sum_{m=2}^{n+1} \sigma^{(n-m+2)\theta} f_{\gamma,d}(\alpha_m)^{r_m} \\ &\leq C\sigma^{n\theta}\{1 + \sum_{m=2}^{n+1} \sigma^{(2-m)\theta} f_{\gamma,d}(\alpha_m)^{r_m}\}, \end{aligned}$$

where $f_{\gamma,d}(\alpha) = (\gamma d)^{2\alpha}[\frac{(1-\alpha)^{1-\alpha}}{(1+\alpha)^{1+\alpha}}]$. $f_{\gamma,d}(\alpha)$ satisfies

$$0 < \inf_{0<\alpha<1} f_{\gamma,d}(\alpha) = f_{\gamma,d}(\alpha_{min}) < 1 \qquad \text{with } \alpha_{min} = \frac{1}{\sqrt{1+\gamma^2 d^2}}.$$

Set $f_{min} = f_{min}(\gamma, d) = f_{\gamma,d}(\alpha_{min})$ and select $\alpha_m = \alpha_{min}$ for $2 \leq m \leq n+1$. Hence,

$$\|u-U\|^2_{L^2(J;X)} \leq C\sigma^{n\theta}\{1 + \sum_{m=2}^{n+1} \sigma^{(2-m)\theta} f_{min}^{r_m}\}. \tag{5.4}$$

Since

$$\sigma^{(2-m)\theta} f_{min}^{r_m} = \sigma^{2\theta} \frac{f_{min}^{r_m}}{\sigma^{m\theta}} = \sigma^{2\theta} \frac{f_{min}^{\lfloor \mu m \rfloor}}{\sigma^{m\theta}} \leq C\sigma^{2\theta} (\frac{f_{min}^{\mu}}{\sigma^{\theta}})^m$$

and $f_{min}^{\mu} < \sigma^{\theta}$ by (5.3), we conclude that the sum in (5.4) can be bounded by

$$\sum_{m=2}^{n+1} \sigma^{(2-m)\theta} f_{min}^{r_m} \leq C\sigma^{2\theta} \sum_{m=2}^{n+1} q^m$$

with $q = f_{min}^{\mu}/\sigma^{\theta} < 1$. We have $\sum_{m=2}^{\infty} q^m < \infty$ and therefore

$$\|u-U\|^2_{L^2(J;X)} \leq C_\sigma \sigma^{n\theta}.$$

Observing that $N = \mathrm{NRDOF}(\mathcal{V}^{\underline{r}}(\mathcal{M}_{n,\sigma}; X)) \leq C\mu n^2$ completes the proof. □

Remark 5.5 If the orders are constant on $\mathcal{M}_{n,\sigma}$, i.e. $r_m = r$ for all m, and if r is proportional to the number of layers, i.e. $r = \lfloor \mu(n+1) \rfloor$, then exponential convergence results for all $\sigma \in (0,1)$,

$$\|u-U\|_{L^2(J;X)} \leq C\exp(-br).$$

In this case condition (5.3) on the slope $\mu > 0$ is not necessary.

The hp-version of DGFEM: Consider now (2.3), (2.4) on $J = (0,T)$ with $u_0 \in H_\theta$ for some $0 < \theta \leq 1$ and with right-hand side g satisfying (2.11). The time discretization by the *hp*-version of the DGFEM is now the following: In a fixed interval $(0, t_0)$ $(0 < t_0 \leq 1)$ near the origin $t = 0$ where the exact solution u exhibits singular behaviour induced by incompatibilities of the initial conditions, we use a geometrically refined partition $\mathcal{M}_{n,\sigma}$ as in Definition 5.1. If we select on $\mathcal{M}_{n,\sigma}$ a

linearly increasing polynomial approximation order according to Definition 5.3, then u is approximated on $(0, t_0)$ at an exponential rate of convergence (cf. Theorem 5.4). On the remaining time interval (t_0, T) away from $t = 0$ the solution u is analytic due to the smoothing property of parabolic evolution operators (see Proposition 4.7). Therefore, on (t_0, T) we can use the p-version of the DGFEM to get exponential rates of convergence, that is we increase the approximation order r on a fixed quasiuniform partition $\mathcal{M}_q$ of (t_0, T). This results in exponential convergence rates on (t_0, T) as in Remark 3.14: If we choose r on $\mathcal{M}_q$ as $\max(1, \lfloor \mu(n+1) \rfloor)$, where $n+1$ is the number of time steps in $\mathcal{M}_{n,\sigma}$, we get $\|u - U\|^2_{L^2((t_0,T);X)} \leq C \exp(-bn)$ with C, b independent of n. On the left-hand side in Figure 1 the partition of J in the hp-version is shown schematically. Together we get exponential rates of convergence, that is

$$\|u - U\|^2_{L^2(J;X)} \leq C \exp(-bN^{\frac{1}{2}}) \tag{5.5}$$

with constants C and b independent of the $N = \mathrm{NRDOF}(\mathcal{V}^{\underline{r}}(\mathcal{M}; X))$.

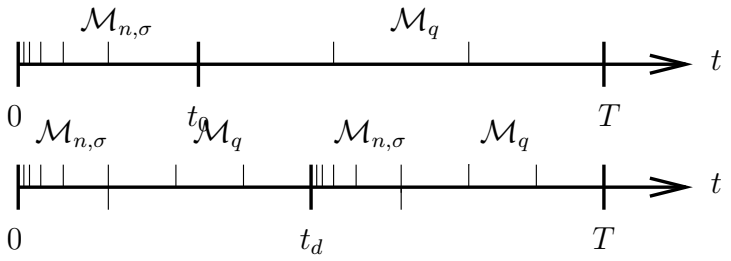

Figure 1: Left: The partition of $(0, T)$ in the hp-version DGFEM: Geometric refinement $\mathcal{M}_{n,\sigma}$ near $t = 0$ and a quasiuniform refinement $\mathcal{M}_q$ after t_0. Right: The partition of $(0, T)$ for piecewise analytic g: Geometric refinement near $t = 0$ and $t = t_d$.

Remark 5.6 The exponential convergence results in Theorem 5.4 and (5.5) are derived under the assumption that the right-hand side g is analytic on $J = [0, T]$ (see (2.11)). However, if g is only piecewise analytic, we still get exponential rates of convergence if we employ time steps that are additionally geometrically refined from above towards "discontinuities" of g: To illustrate this assume that g is analytic on $J_1 = [0, t_d)$ and on $J_2 = [t_d, T]$ (as a function of t with values in H). The point t_d separates the analyticity intervals and at t_d the right-hand side g can be discontinuous. From the semigroup properties (4.3) and Duhamel's formula (4.4) it follows that the solution on the second time interval J_2 is precisely the solution of a parabolic problem on J_2 started at t_d with initial value $u(t_d)$ and analytic right-hand side $g|_{J_2}$. Therefore, we can apply the regularity results of Section 4 also for the second time interval J_2.
Based on this observation the corresponding strategy in the hp-DGFEM is the following: Firstly, we discretize (2.3), (2.4) on J_1 with initial value u_0 with a geometrically refined partition near $t = 0$ and a linearly increasing degree vector $\underline{r}$. Secondly, at

t_d we restart again with a geometric partition in $J_2 = (t_d, t_d + T)$ as shown in the right-hand side of Figure 1 to resolve the singular solution behaviour caused by the discontinuity of the data at $t = t_d$. Away from $t = 0$ and $t = t_d$ quasiuniform time steps are employed.

5.2. The h-Version on Graded Time Partitions

It can be seen from the estimates in Remark 3.13 that in the h-version DGFEM on quasiuniform partitions the best possible convergence rate of N^{-r-1} is lost if the solution u is not smooth enough in time, i.e. if $s_0 < r$ in (3.6).
In the Finite Element Method one uses graded meshes to compensate this loss of convergence. We show that the same mechanisms also work for the DGFEM. Assume for simplicity that $T = 1$.
A graded temporal mesh $\mathcal{M}$ is defined by a grading function $h : [0, 1] \to [0, 1]$ which is strictly increasing and satisfies

$$h \in C^0([0,1]) \cap C^1((0,1)), \qquad h(0) = 0, \qquad h(1) = 1. \tag{5.6}$$

The nodes in $\mathcal{M}$ are given by

$$t_m = h(\frac{m}{M}), \qquad m = 0, \dots, M(\mathcal{M}). \tag{5.7}$$

Again, M is the number of time steps and $N = \mathrm{NRDOF}(\mathcal{V}^r(\mathcal{M}; X))$ is then proportional to M, i.e. $N = (r + 1)M$. We have:

Proposition 5.7 *Let $u_0 \in H_\theta$ for some $0 < \theta \leq 1$ and let the right-hand side g satisfy (2.11). Let $\mathcal{M}$ be given by a grading function h where additionally $h(t) = o(t^{\frac{2r+2}{\theta}})$ and $h'(t)^{2r+3}h(t)^{-(2r+3)+\theta}$ is Riemann integrable on $J = (0, 1)$. Then we have asymptotically, as $M \to \infty$, for the DGFEM solution U in $\mathcal{V}^r(\mathcal{M}; X)$*

$$\|u - U\|^2_{L^2(J;X)} \leq CM^{-(2r+2)} \int_0^1 h'(t)^{2r+3}h(t)^{-(2r+3)+\theta}dt.$$

Proof: From (5.2) and Theorem 3.10 follows

$$\|u - U\|^2_{L^2(J;X)} \leq C\|u - \Pi^0_{I_1}\|^2_{L^2(I_1;X)} + C\sum_{m=2}^{M} k_m^{2(r+1)} \int_{t_{m-1}}^{t_m} \|u^{(r+1)}\|^2_X dt \tag{5.8}$$

with $C = C(r)$. We bound the error contribution on the first element I_1 by Proposition 4.8, i.e. $\|u - \Pi^0_{I_1}\|_{L^2(I_1;X)} \leq Ck_1^\theta$. On the elements away from $t = 0$ there holds $\int_{t_{m-1}}^{t_m} \|u^{(r+1)}\|^2_X dt = k_m\|u^{(r+1)}(\xi_m)\|^2_X$ for some $\xi_m = h(x_m) \in (t_{m-1}, t_m)$. Together with the regularity statements in Proposition 4.6 we conclude that (5.8) can be estimated by

$$\|u - U\|^2_{L^2(J;X)} \leq Ck_1^\theta + C\sum_{m=2}^{M} k_m^{2r+3}\xi_m^{-(2r+3)+\theta}. \tag{5.9}$$

Since

$$k_m = h(\frac{m}{M}) - h(\frac{m-1}{M}) = \frac{h'(\tilde{x}_m)}{M}, \qquad \tilde{x}_m \in (\frac{m-1}{M}, \frac{m}{M}),$$

we get from (5.9)

$$\|u - U\|^2_{L^2(J;X)} \leq CM^{-2r-2}(M^{2r+2}h(1/M)^\theta + \sum_{m=2}^{M} h'(\tilde{x}_m)^{2r+3}h(x_m)^{-(2r+3)+\theta}k_m). \tag{5.10}$$

As $M \to \infty$, we have $\lim_{M\to\infty} M^{2r+2}h(1/M)^\theta = \lim_{t\to 0}(t^{-\frac{2r+2}{\theta}}h(t))^\theta = 0$. The sum over all elements away from $t = 0$ is a Riemann sum. As $M \to \infty$, it approaches therefore $\int_0^1 h'(t)^{2r+3}h(t)^{-(2r+3)+\theta}dt$, which finishes the proof. □

Remark 5.8 To prove the assertions in Proposition 5.7 only finite regularity of the right-hand side $g(t)$ is needed and assumption (2.11) can thus be relaxed.

We cite from [21, Part II, Lemma 2.4] the following result:

Proposition 5.9 *Let I be the functional $I[h] = \int_0^1 h(t)^{\sigma-n}h'(t)^n dt$ with $\sigma > 0$ and $n \in \mathbb{N}$ defined for grading functions h as in (5.6). Then $I[h]$ has a unique minimizer $h(t) = t^{\frac{n}{\sigma}}$ with $I[h] = (\frac{n}{\sigma})^n$.*

Apply Proposition 5.9 with $n = 2r + 3$ and $\sigma = \theta$. Then the function $h(t) = t^{\frac{2r+3}{\theta}}$ minimizes the integral $\int_0^1 h'(t)^{2r+3}h(t)^{-(2r+3)+\theta}dt$ arising in Proposition 5.7. It can easily be seen that h satisfies the additional assumptions in Proposition 5.7. Consequently:

Theorem 5.10 *Let u be the solution of (2.3), (2.4) with $u_0 \in H_\theta$ for some $0 < \theta \leq 1$ and right-hand side g satisfying (2.11). Consider the h-version DGFEM at a fixed approximation order r on the graded time mesh $\mathcal{M}$ with M intervals given by the grading function $h(t) = t^{\frac{2r+3}{\theta}}$. Let $N = \mathrm{NRDOF}(\mathcal{V}^r(\mathcal{M}; X))$. Then, as $M \to \infty$ or $N \to \infty$, there holds for the DGFEM solution $U \in \mathcal{V}^r(\mathcal{M}; X)$*

$$\|u - U\|_{L^2(J;X)} \leq CM^{-(r+1)} \qquad or \qquad \|u - U\|_{L^2(J;X)} \leq CN^{-(r+1)}$$

with C depending only on u_0, g and r. The use of this grading function recovers the optimal h-version convergence rate and minimizes the functional $I[h]$ in Proposition 5.9.

Remark 5.11 The use of graded meshes for non-smooth initial data is also discussed in [32, Chapter 10]. Depending on the degree of roughness several grid topologies were introduced which are all refined towards $t = 0$. The application of these partitions permitted to recover the optimal h-convergence rates for the Euler and Trapezoidal Method where $r = 0$ and $r = 1$. However, by Proposition 5.9, $h(t) = t^{\frac{2r+3}{\theta}}$ is optimal for each approximation order r and Theorem 5.10 generalizes in this sense the results of [32].

6. hp Discretization in Time and Space

The DGFEM reduces the parabolic equation (2.3), (2.4) in each time step I_m to a coupled elliptic reaction-diffusion system of $r_m + 1$ equations (see Section 6.1). In order to obtain a fully discrete solution this system has to be solved numerically which is very costly, particularly for large r_m. To overcome these difficulties we propose in Section 6.2 two decoupling methods that transform the system into r_m+1 independent reaction-diffusion equations of the type arising in the backward Euler scheme. However, the equations are singularly perturbed for small time steps k_m or large approximation orders r_m and require careful spatial discretizations.
For the case of the heat equation on a domain $\Omega \subset \mathbb{R}^2$ we present in Section 6.3 an *hp* Finite Element Method (FEM) for the solution of these singularly perturbed reaction-diffusion equations based on the recent work in [27, 28, 29, 36, 37, 40]. Provided that certain mesh-design principles are observed, the *hp*-FEM features exponential rates of convergence that are robust, i.e. independent of the perturbation parameter. Similar arguments can be applied to general elliptic operators and our analysis provides sufficient insight for an efficient *hp* FE-approximation in space. Finally, we show in conjunction with the *hp*-version of DGFEM that exponential rates of convergence in time and space can be obtained.
In Section 6.4 we confirm the theoretical results of the present work in numerical examples.

6.1. The Spatial Problems

On a generic time step $I = (t_0, t_1)$ with length $k = t_1 - t_0 > 0$ and approximation order r the DGFEM semiapproximation $U|_I = U$ is found by solving the problem in (2.16). The right-hand side $g(t)$ and the initial condition $u_{initial}$ are the known data on the time step.
Let $\{\widehat{\varphi}_i\}_{i=0}^r$ and $\{\widehat{\psi}_i\}_{i=0}^r$ be two bases of the reference polynomial space $\mathcal{P}^r((-1,1))$ chosen as normalized Legendre polynomials, i.e.

$$\widehat{\varphi}_i = \widehat{\psi}_i := \sqrt{i+1/2} \cdot L_i, \qquad i = 0, \dots, r. \tag{6.1}$$

Here, L_i is the usual Legendre polynomial of degree i on $(-1,1)$. These bases define transported variants $\{\varphi_i\}_{i=0}^r$ and $\{\psi_i\}_{i=0}^r$ on $\mathcal{P}^r((t_0,t_1))$ given by $\varphi_i \circ Q(\hat{t}) = \widehat{\varphi}_i(\hat{t})$ and $\psi_i \circ Q(\hat{t}) = \widehat{\psi}_i(\hat{t})$, where Q is the transformation $t = Q(\hat{t}) = \frac{1}{2}(t_0 + t_1 + \hat{t}k)$ from $(-1,1)$ onto (t_0,t_1).
In (2.16), the trial polynomial $U \in \mathcal{P}^r(I;X)$ and the test polynomial $V \in \mathcal{P}^r(I;X)$ can uniquely be written as $U = \sum_{j=0}^r u_j \varphi_j$ and $V = \sum_{i=0}^r v_i \psi_i$ with coefficients $u_j, v_i \in X$. Problem (2.16) is then equivalent to the elliptic system with the following variational formulation:
Find $\{u_j\}_{j=0}^r \subset X$ such that

$$\sum_{i,j=0}^{r} \{[\int_I \varphi_j' \psi_i dt + \varphi_j^+(t_0)\psi_i^+(t_0)](u_j, v_i)_H + [\int_I \varphi_j \psi_i dt]\ a(u_j, v_i)\} \tag{6.2}$$
$$= \sum_{i=0}^{r} \{(\int_I g\psi_i dt, v_i)_H + (u_{initial}, v_i)_H \psi_i^+(t_0)\}$$

for all $\{v_i\}_{i=0}^r \subset X$.
We introduce the matrices

$$\widehat{A}_{ij} := \widehat{A}^1_{ij} + \widehat{A}^2_{ij} = \int_{-1}^1 \widehat{\varphi}'_j \widehat{\psi}_i d\hat{t} + \widehat{\varphi}^+_j(-1)\widehat{\psi}^+_i(-1), \qquad \widehat{B}_{ij} := \int_{-1}^1 \widehat{\varphi}_j \widehat{\psi}_i d\hat{t}, \tag{6.3}$$

which are expressed in terms of the bases $\{\widehat{\varphi}_i\}$ and $\{\widehat{\psi}_i\}$ on $(-1,1)$ and which are therefore independent of k. Then (6.2) is to find $\{u_j\}_{j=0}^r \subset X$ such that

$$\sum_{i,j=0}^r \widehat{A}_{ij}(u_j, v_i)_H + \frac{k}{2}\widehat{B}_{ij} a(u_j, v_i) = \sum_{i=0}^r \frac{k}{2}(\widehat{f}^1_i, v_i)_H + (\widehat{f}^2_i, v_i)_H \qquad \forall \{v_i\}_{i=0}^r \subset X.$$

Here, the right-hand sides $\widehat{f}^1_i$ and $\widehat{f}^2_i$ are defined by

$$\widehat{f}^1_i(v) = (\int_{-1}^1 [g \circ Q]\widehat{\psi}_i d\hat{t}, v)_H, \qquad \widehat{f}^2_i(v) = (u_{initial}, v)_H \widehat{\psi}^+_i(-1), \qquad v \in X.$$

The strong form (observing that $\widehat{B}_{ij} = \delta_{ij}$ due to the choice (6.1)) reads

$$\sum_{j=0}^r \{\widehat{A}_{ij} u_j + \frac{k}{2}\delta_{ij} L u_j\} = \frac{k}{2}\widehat{f}^1_i + \widehat{f}^2_i, \qquad i = 0, \ldots, r. \tag{6.4}$$

To obtain a fully discrete approximation of (2.3), (2.4) the system (6.4) has to be solved numerically by a Finite Element Method. If $\{U^F_j\}_{j=0}^r$ is a FE solution of (6.4) in X, then $U^F = \sum_{j=0}^r U^F_j \varphi_j$ approximates $U = \sum_{j=0}^r u_j \varphi_j$ on the time step I. We get for the error

$$\begin{aligned} \|U - U^F\|^2_{L^2(I;X)} &= \int_I \| \sum_{j=0}^r (u_j - U^F_j)\varphi_j \|^2_X dt \\ &= \sum_{j=0}^r \|u_j - U^F_j\|^2_X (j + 1/2) \int_I [L_j \circ Q^{-1}]^2 dt = \frac{k}{2} \sum_{j=0}^r \|u_j - U^F_j\|^2_X, \end{aligned}$$

where we used the orthogonality properties of the Legendre polynomials. Thus:

Proposition 6.1 *Let u be the exact solution of (2.3), (2.4) on $J = (0,T)$ and let U be the time discretization of u obtained in $\mathcal{V}^r(\mathcal{M}; X)$ with the DGFEM 2.3. On each time interval I_m we develop $U|_{I_m}$ into $U|_{I_m} = \sum_{j=0}^{r_m} u_{m,j}\varphi_{m,j}$ where $\{\varphi_{m,j}\}$ are the basis functions (6.1) scaled to I_m and where the coefficients $\{u_{m,j}\}$ solve the system (6.4). Let $\{U^F_{m,j}\}$ be a Finite Element approximation of (6.4) and let $U^F \in L^2(J;X)$ be the fully discrete solution given by $U^F|_{I_m} = \sum_{j=0}^{r_m} U^F_{m,j}\varphi_{m,j}$. Then we have the error estimate*

$$\|u - U^F\|^2_{L^2(J;X)} \le C\|u - U\|^2_{L^2(J;X)} + C \sum_{m=1}^{M(\mathcal{M})} k_m \sum_{j=0}^{r_m} \|u_{m,j} - U^F_{m,j}\|^2_X. \tag{6.5}$$

The first term in the error estimate (6.5) is the error of the time discretization. This error can be made exponentially small when the *hp*-version of the DGFEM is applied. The second error contribution stems from the spatial discretizations and will be discussed in more details in the next subsections.

Remark 6.2 In (6.1) we chose normalized Legendre polynomials as basis functions. However, other choices are possible. The optimal choice of the basis functions would be the one where the matrices $\widehat{A}$ and $\widehat{B}$ in (6.3) diagonalize simultaneously. Then the equations in (6.2) decouple completely and could be solved independently. This can be done in parallel if each equation is assigned to one processor. Unfortunately, it does not seem to be possible to find such diagonalizations in $\mathbb{R}$. In Section 6.2 we present decoupling methods in $\mathbb{C}$.

If the time step k in (6.4) approaches zero, the coefficient of the principal part tends to zero as well and the system (6.4) becomes singularly perturbed. The same may happen for large r. In the following lemma we analyze the dependence of the coefficients in the matrix $\widehat{A}$ on r.

Lemma 6.3 *Let $\lambda \in \mathbb{C}$ be an eigenvalue of the matrix $\widehat{A}$ in (6.3). Then $0 <$ Re $\lambda \le C_1 \max(1, r^2)$ and $0 < C_2 \le |\lambda| \le C_3 \max(1, r^2)$ with constants independent of $r \in \mathbb{N}_0$.*

Proof: Eigenvalues of $\widehat{A}$ correspond in a one-to-one relation to eigenvalues of the following variational problem:
Find $p \in \mathcal{P}^r((-1,1);\mathbb{C})$, $p \neq 0$, such that

$$\mathcal{B}(p,q) := \int_{-1}^{1} p'\overline{q}dt + p(-1)\overline{q(-1)} = \lambda \int_{-1}^{1} p\overline{q}dt \tag{6.6}$$

for all $q \in \mathcal{P}^r((-1,1);\mathbb{C})$.
We proceed now in several steps and assume that $r \ge 1$ (the modifications for $r = 0$ are obvious).

Step (i): There holds

$$\int_{-1}^{1} |\pi(t)|^2 dt \le \frac{1}{k}|\int_{-1}^{1} \pi(t)dt|^2 + \frac{1}{2}\int_{-1}^{1}(1-t^2)|\pi'(t)|^2 dt \tag{6.7}$$

for all $\pi(t) \in \mathcal{P}^r((a,b);\mathbb{C})$ and $r \ge 0$.
To prove (6.7), we denote by $\{L_k\}_{k\ge 0}$ the Legendre polynomials of degree k on $I = (-1,1)$. We can develop π and π' into the series $\pi(t) = \sum_{k=0}^{r} a_k L_k(t)$ and $\pi'(t) = \sum_{k=1}^{r} a_k L_k'(t)$. There holds $2a_0 = \int_I \pi(t)dt$. Since

$$\int_I L_k'(t)L_l'(t)(1-t^2)dt = l(l+1)\int_I L_k(t)L_l(t)dt$$

for $k, l \ge 1$, the derivatives $\{L_k'\}$ are orthogonal with respect to $(1-t^2)dt$. We get

$$\int_I |\pi'(t)|^2(1-t^2)dt + |\int_I \pi(t)dt|^2 = \sum_{k=1}^{r} |a_k|^2 \int_I L_k'(t)^2(1-t^2)dt + 4|a_0|^2$$

$$= \sum_{k=1}^{r} |a_k|^2 k(k+1)\int_I L_k(t)^2 dt + 4|a_0|^2 \ge 2\sum_{k=0}^{r} |a_k|^2 \int_I L_k(t)^2 dt = 2\int_I |\pi(t)|^2 dt.$$

This is (6.7).

Step (ii): We show that $0 \leq \text{Re } \lambda \leq Cr^2$ and $0 \leq |\lambda| \leq Cr^2$: Selecting $q = p$ in (6.6) yields

$$\text{Re } (\int_{-1}^{1} p'\overline{p}dt + p(-1)\overline{p(-1)}) = \text{Re } \lambda \cdot \int_{-1}^{1} |p|^2 dt.$$

Integrating by parts the left-hand side gives

$$\frac{1}{2}(|p(+1)|^2 + |p(-1)|^2) = \text{Re } \lambda \cdot \int_{-1}^{1} |p|^2 dt. \tag{6.8}$$

By inverse estimates,

$$\frac{1}{2}(|p(+1)|^2 + |p(-1)|^2) \leq \|p\|^2_{L^\infty((-1,1);\mathbb{C})} \leq Cr^2 \|p\|^2_{L^2((-1,1);\mathbb{C})}.$$

Therefore, we have $0 \leq \text{Re } \lambda \leq Cr^2$. Analogously, we get

$$|\lambda| \leq \frac{\int_{-1}^{1} |p'p|dt + |p(-1)|^2}{\int_{-1}^{1} |p|^2 dt}. \tag{6.9}$$

By Cauchy-Schwarz and Schmidt's inverse inequality we have

$$\int_{-1}^{1} |p'p|dt \leq (\int_{-1}^{1} |p'|^2 dt)^{\frac{1}{2}} (\int_{-1}^{1} |p|^2 dt)^{\frac{1}{2}} \leq Cr^2 \int_{-1}^{1} |p|^2 dt \tag{6.10}$$

and, as before, $|p(-1)|^2 \leq Cr^2 \|p\|^2_{L^2((-1,1);\mathbb{C})}$. This implies $|\lambda| \leq Cr^2$.

Step (iii): We prove by contradiction that $\text{Re } \lambda > 0$: If $\text{Re } \lambda = 0$, we conclude from (6.8) that an eigenfunction $p \neq 0$ satisfies $p(+1) = p(-1) = 0$. We show that then also $p' \equiv 0$ and consequently $p \equiv 0$, which is not possible:
Selecting $q = (1 + t)p'$ in (6.6) yields

$$\int_{-1}^{1} |p'|^2 (1 + t)dt = \lambda \int_{-1}^{1} p\overline{p}'(1 + t)dt.$$

We may assume $\int_{-1}^{1} p\overline{p}'(1 + t)dt \neq 0$ (since otherwise $p' \equiv 0$ trivially). Then

$$0 = \text{Re } \lambda = \frac{\int_{-1}^{1} |p'|^2 (1 + t)dt \cdot \text{Re } \int_{-1}^{1} p\overline{p}'(1 + t)dt}{|\int_{-1}^{1} p\overline{p}'(1 + t)dt|}. \tag{6.11}$$

Integrating by parts gives $\text{Re } \int_{-1}^{1} p\overline{p}'(1 + t)dt = -2 \int_{-1}^{1} |p|^2 dt$. From (6.11) we see that $\int_{-1}^{1} |p'|^2 (1 + t)dt = 0$ and therefore $p' \equiv 0$.

Step (iv): It remains to establish the lower bound $|\lambda| \geq C > 0$. To do so, we introduce the norm $\|\cdot\|_E$ on $\mathcal{P}^r((-1, 1); \mathbb{C})$ by

$$\|p\|^2_E = \int_{-1}^{1} |p'|^2 (1 + t)dt + |p(+1)|^2 + |\int_{-1}^{1} pdt|^2.$$

From (6.7) in Step (i) follows that

$$\int_{-1}^{1} |p|^2 \, dt \le C| \int_{-1}^{1} p \, dt|^2 + C \int_{-1}^{1} (1-t^2)|p'|^2 dt$$

for all polynomials $p \in \mathcal{P}^r((-1,1);\mathbb{C})$ with C independent of r. Therefore,

$$\|p\|_{L^2} \le C\|p\|_E, \qquad C \text{ independent of } r. \tag{6.12}$$

We claim that

$$\inf_{0\neq p} \sup_{0\neq q} \frac{\text{Re } \mathcal{B}(p,q)}{\|p\|_E \|q\|_{L^2}} \ge \gamma > 0 \tag{6.13}$$

with a constant γ independent of r. (The infimum and supremum are taken over all polynomials in $\mathcal{P}^r((-1,1);\mathbb{C})$.) To prove (6.13), fix $0 \neq p \in \mathcal{P}^r((-1,1);\mathbb{C})$. Choosing in (6.6) $q = (1+t)p' + p(1) + (1-t)\int_{-1}^{1} p \, dt$ gives

$$\mathcal{B}(p,q) = \int_{-1}^{1} |p'|^2(1+t) \, dt + \overline{p(1)} \int_{-1}^{1} p' \, dt$$
$$+ \overline{(\int_{-1}^{1} p \, dt)} \int_{-1}^{1} p'(1-t) \, dt + p(-1)\overline{p(1)} + 2p(-1)\overline{(\int_{-1}^{1} p \, dt)} = \|p\|_E^2.$$

Moreover, since $(1+t)^2 \le C(1+t)$ for $-1 \le t \le 1$, we have $\|q\|_{L^2}^2 \le C\|p\|_E^2$. This shows (6.13).
For an eigenfunction $p \neq 0$ we conclude using (6.6), (6.13) and (6.12) that

$$0 < \gamma \le \sup_{0\neq q} \frac{\text{Re } \mathcal{B}(p,q)}{\|p\|_E \|q\|_{L^2}} \le \sup_{0\neq q} \frac{|\lambda| |\int_{-1}^{1} p\overline{q} dt|}{\|p\|_E \|q\|_{L^2}} \le |\lambda| \frac{\|p\|_{L^2}}{\|p\|_E} \le C|\lambda|.$$

This is the desired lower bound. □

In the left-hand side of Figure 1 we plot $\lambda_{min} = \min\{|\lambda| : \lambda \text{ eigenvalue of } \hat{A}\}$ and $\lambda_{max} = \max\{|\lambda| : \lambda \text{ eigenvalue of } \hat{A}\}$ for varying r. The graph is in agreement with the statements in Lemma 6.3.

6.2. Decoupling

A coupled system of the form (6.2) or (6.4) is costly to solve numerically and can not be further decoupled in $\mathbb{R}$. However, in $\mathbb{C}$ it is possible to transform (6.4) into upper triangular or diagonal form. This additional gain may be worth to switch over to complex arithmetic:

Method (i): The Schur Decomposition Theorem [20] guarantees the existence of a unitary matrix Q in $\mathbb{C}^{(r+1)\times(r+1)}$ such that $\widehat{A}$ in (6.3) can be transformed into an upper triangular matrix $T \in \mathbb{C}^{(r+1)\times(r+1)}$ given by $T = Q^h \widehat{A} Q$. The system (6.4) transforms into

$$\sum_{j=0}^{r} T_{ij} w_j + \frac{k}{2} \delta_{ij} L w_j = \text{r.h.s.}, \qquad i = 0, \ldots, r,$$

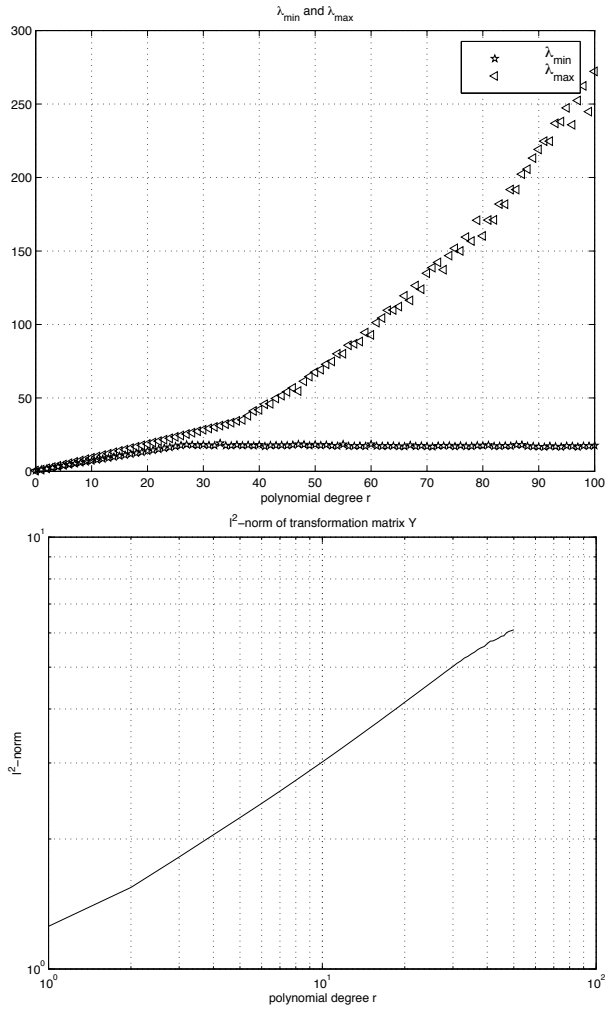

Figure 1: Left: The values of λ_{min} and λ_{max} against the approximation order r. Right: The norm of the transformation matrix Y in dependence on r.

where (in vector notation) the new unknowns $\vec{w}$ are given by $\vec{w} = Q^h \vec{u}$. Since T is upper triangular, a backward-solve process essentially amounts to the successive solution of $r+1$ complex scalar equations, namely $T_{jj} w_j + \frac{k}{2} L w_j = \text{r.h.s.}$, $j = 0, \ldots, r$. The coefficient T_{jj} is an eigenvalue $\lambda_j \in \mathbb{C}$ of $\widehat{A}$. After rescaling the coefficients, we get $r+1$ scalar reaction-diffusion problems of the form

$$\frac{k}{2\lambda_j} L w + w = f, \qquad j = 0, \ldots r. \tag{6.14}$$

Equation (6.14) is singularly perturbed for large r or small k (cf. Lemma 6.3).
If $\{W_j^F\}$ are FE-solutions to (6.14), we have with $\vec{U}^F = Q\vec{W}^F$ that the discretization error in (6.5) satisfies $\sum_{j=0}^r \|u_j - U_j^F\|_X^2 = \sum_{j=0}^r \|w_j - W_j^F\|_X^2$ due to the unitarity of the transformation matrix Q.

Method (ii): *Numerical* experiments show that the matrix $\widehat{A}$ in (6.3) is diagonalizable in $\mathbb{C}$ at least for $0 \le r \le 100$ (a theoretical proof of this fact is lacking): There

exists a matrix $Y \in \mathbb{C}^{(r+1)\times(r+1)}$ such that $Y^{-1}\widehat{A}Y = \mathrm{diag}(\lambda_1, \ldots, \lambda_{r+1})$ with pairwise complex conjugate eigenvalues $\{\lambda_j\}$. Unfortunately, the transformation matrix Y is not unitary anymore and becomes ill-conditioned for large r. Since we expect r to vary only in a relatively small range in practice, say $0 \le r \le 10$, this seems not to be a big disadvantage at the moment. The diagonalization process transforms the system (6.4) into the decoupled system

$$\lambda_j w_j + \frac{k}{2} L w_j = \text{r.h.s.}, \qquad j = 0, \ldots, r$$

with $\vec{w} = Y^{-1}\vec{u}$. As in Method (i), this are $r+1$ reaction-diffusion equations of the form (6.14). However, the matrix l^2-norm of Y is not 1 and now depends on r. In the right-hand side of Figure 1 we plot this norm for increasing r. We calculated Y numerically using the diagonalization commands in MATLAB 5.2.0[1]. The graph indicates that the norm of Y depends at most algebraically on r, that is we have in (6.5) $\sum_{j=0}^{r} \|u_j - U_j^F\|_X^2 \le C(r) \sum_{j=0}^{r} \|w_j - W_j^F\|_X^2$ with $C(r) \le Cr^\alpha$ for some $0 < \alpha \le 1$. In fact, Figure 1 suggests $\alpha = 1/2$.

6.3. hp Discretization in Time and Space

The Finite Element Method for the singular perturbation problems: Both transformation processes in Section 6.2 require in time step I_m the solution of $r_m + 1$ reaction-diffusion equations of the form (6.14). To discuss their discretizations consider the singularly perturbed model equation

$$\varepsilon^2 L w + w = f, \tag{6.15}$$

where $\varepsilon \in \mathbb{C}$ is a parameter with $\mathrm{Re}\ \varepsilon > 0$ and $\mathrm{Re}\ (\varepsilon^2) > 0$ whose modulus $|\varepsilon| \in (0, 1]$ can approach zero. The perturbation parameter in (6.14) is $\varepsilon = \sqrt{k_m/(2\lambda)}$, $\sqrt{\cdot}$ being the usual principal branch of the square root taken to be positive on $(0, \infty)$. λ is an eigenvalue of the matrix $\widehat{A}$ in (6.3). Thus, due to Lemma 6.3 the absolute value of ε is in the range $Ck_m \ge |\varepsilon|^2 \ge Ck_m/\max(1, r_m^2)$, where the lower bound is achieved (see Figure 1).
The weak formulation of (6.15) is

$$\text{Find } w \in X \text{ such that } b_\varepsilon(w, v) := \varepsilon^2 a(w, v) + (w, v)_H = f(v) \text{ for all } v \in X. \tag{6.16}$$

We have

$$\begin{aligned} \mathrm{Re}\ b_\varepsilon(w, w) &= \mathrm{Re}\ (\varepsilon^2)\mathrm{Re}\ a(w, w) - \mathrm{Im}\ (\varepsilon^2)\mathrm{Im}\ a(w, w) + \|w\|_H^2, & (6.17)\\ \mathrm{Im}\ b_\varepsilon(w, w) &= \mathrm{Im}\ (\varepsilon^2)\mathrm{Re}\ a(w, w) + \mathrm{Re}\ (\varepsilon^2)\mathrm{Im}\ a(w, w). & (6.18) \end{aligned}$$

The sesquilinear form b_ε is coercive in the energy norm $\|w\|_\varepsilon^2 := |\varepsilon|^2\|w\|_X^2 + \|w\|_H^2$.

Lemma 6.4 *For $\varepsilon \in \mathbb{C}$ with $\mathrm{Re}\ (\varepsilon^2) > 0$ we have $|b_\varepsilon(w, w)| \ge \sqrt{2}\min(1, \beta)\|w\|_\varepsilon^2$.*

[1]MATLAB is a trademark of The MathWorks Inc.

Proof: From (6.17), (6.18) we get

$$\begin{aligned}|b_\varepsilon(w,w)|^2 \geq\ & [(\text{Re }\varepsilon^2)^2 + (\text{Im }\varepsilon^2)^2](\text{Re }a(w,w))^2 \\ & +[(\text{Re }\varepsilon^2)^2 + (\text{Im }\varepsilon^2)^2](\text{Im }a(w,w))^2 \\ & +2\text{Re }(\varepsilon^2)\text{Re }a(w,w)\|w\|_H^2 - 2\text{Im }(\varepsilon^2)\text{Im }a(w,w)\|w\|_H^2 + \|w\|_H^4.\end{aligned}$$

Due to (2.7) we have Im $a(w,w) = 0$ and get with (2.6)

$$\begin{aligned}|b_\varepsilon(w,w)|^2 &\geq |\varepsilon|^4(\text{Re }a(w,w))^2 + 2\text{Re }(\varepsilon^2)\text{Re }a(w,w)\|w\|_H^2 + \|w\|_H^4 \\ &\geq |\varepsilon|^4\beta^2\|w\|_X^4 + \|w\|_H^4 \geq 2\min(1,\beta^2)(|\varepsilon|^2\|w\|_X^2 + \|w\|_H^2)^2.\end{aligned}$$

This yields the assertion. □

In the Finite Element Method a finite dimensional subspace $V_N \subset X$ of dimension $N = \dim(V_N)$ is chosen, and the finite element solution $W^F \in V_N$ of (6.15) is given by

$$b_\varepsilon(W^F, V) = f(V) \qquad \forall V \in V_N. \tag{6.19}$$

Because of Lemma 6.4 problem (6.19) has a unique solution and we have quasioptimality in the energy norm,

$$\|w - W^F\|_\varepsilon \leq C \inf_{V\in V_N} \|w - V\|_\varepsilon. \tag{6.20}$$

The question is then to choose the spaces V_N appropriately.

Robust exponential convergence in the hp-FEM: Exemplarily, we discuss in the following the numerical solution of the spatial problems (6.15) in the case of the heat equation on a Lipschitz domain $\Omega \subset \mathbb{R}^d$ where $L = -\Delta$, $H = L^2(\Omega)$ and $X = H_0^1(\Omega)$. The singular perturbation problems (6.15) read then

$$-\varepsilon^2\Delta w + w = f \text{ in } \Omega, \qquad w = 0 \text{ on } \partial\Omega. \tag{6.21}$$

The corresponding weak formulation in (6.16) is in this case (recall that the spatial problems are decoupled in $\mathbb{C}$): Find $w \in H_0^1(\Omega;\mathbb{C})$ such that

$$b_\varepsilon(w,v) := \varepsilon^2 \int_\Omega \{\nabla w\overline{\nabla v} + u\overline{v}\}dx = \int_\Omega f\overline{v}dx =: f(v) \qquad \forall v \in H_0^1(\Omega;\mathbb{C}). \tag{6.22}$$

The small parameter ε in (6.21) causes difficulties in the convergence of discretizations due to boundary layers that downgrade the approximation properties of the standard FEM. Boundary layers are solution components that show a rapid variation normal to the boundary and a smooth behaviour tangentially to it. In boundary fitted coordinates they are of the form $w_{BL}(\rho,s) = C(s)\exp(-\rho/\varepsilon)$ with $C(s)$ (piecewise) analytic, ρ denoting the normal distance to the wall and s being the arclength on $\partial\Omega$. Moreover, in polygonal domains with corners there arise corner singularities. These are solution components which are in polar coordinates (r,φ) near a corner basically of the form $w_c(r,\varphi) = r^\alpha\Phi(\varphi)$ for some $\alpha \in (0,1)$ and some analytic function Φ.
The efficient resolution of corner singularity or boundary layer phenomena in problems of the form (6.21) requires properly designed FE spaces V_N in (6.20). In the

hp-FEM context, the combination of anisotropic and geometric mesh refinement towards the boundaries and the corners with judiciously increased polynomial degrees allows one to approximate these solution components at a (robust) exponential rate of convergence. We mention [1, 22, 35] and the references there for the *hp*-approximation of corner singularities and [27, 28, 29, 36, 37, 40] for corresponding boundary layer approximation results.

We discuss the 2D case for (6.21): Let $\mathcal{T}$ be a partition of Ω into quadrilateral and triangular elements. Assume that for each $K \in \mathcal{T}$ there is a differentiable and bijective element mapping F_K from the generic reference element $\hat{K}$ which is either the unit square $(0,1)^2$ or the triangle $\{(x,y) : 0 < y < 1-x\}$ onto K. The FE space of piecewise mapped polynomials is then defined in the usual way:

$$V_N = S_0^{p,1}(\mathcal{T}) = \{v \in H_0^1(\Omega;\mathbb{C}) : v|_K = \pi_p \circ F_K^{-1} \text{ for some } \pi_p \text{ in } \mathcal{S}^p(\hat{K};\mathbb{C}),\ K \in \mathcal{T}\}. \tag{6.23}$$

The polynomial space $\mathcal{S}^p(\hat{K};\mathbb{C})$ is to be understood as the set of all polynomials of total degree $\leq p$ if $\hat{K}$ is the reference triangle and as the set of all polynomials of degree $\leq p$ in each variable if $\hat{K}$ is the unit square.

Assume first that $\partial\Omega$ is smooth and no corner singularities are present. In this case, we use "boundary layer meshes" where needle elements of size $O(p|\varepsilon|)$ are inserted near the boundary (we refer to [28] for the exact definition of admissible meshes). The interior of the domain is partitioned in a quasiuniform way. On the left-hand side of Figure 2 we show such a "boundary layer mesh" with the corresponding anisotropic refinement towards $\partial\Omega$. Note that the needle elements become fatter as p increases. The subsequent theorem can also be found in [28]:

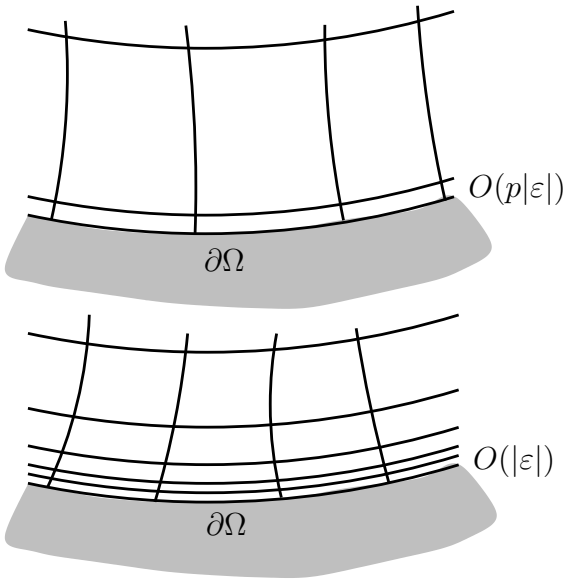

Figure 2: Left: Needle element of size $O(p|\varepsilon|)$ are inserted near $\partial\Omega$. Right: Geometric refinement towards the boundary, the smallest layer has width $O(|\varepsilon|)$.

Theorem 6.5 *Consider (6.21) in a domain $\Omega \subset \mathbb{R}^2$ with analytic boundary curve $\partial\Omega$. Let the right-hand side f be analytic on $\overline{\Omega}$. Let W^F be the Finite Element solution of (6.21) in $V_N = S_0^{p,1}(\mathcal{T})$, where $\mathcal{T}$ is a boundary layer mesh with needle elements of width $O(p|\varepsilon|)$. Then we have robust exponential convergence for the error, i.e., $\|w - W^F\|_\varepsilon \leq C \exp(-bp)$ with C, b independent of ε and p.*

Remark 6.6 In [28], Theorem 6.5 is proved for real $\varepsilon \in \mathbb{R}$ with $\varepsilon > 0$. Nevertheless, one can check that all the results there hold true verbatim with ε replaced by $|\varepsilon|$.

Remark 6.7 The width of the needle elements depends on ε as well as on p. In practice, it may be more convenient to fix a mesh and then increase the polynomial degree p until the desired accuracy is reached. This can be obtained by the use of meshes that are refined geometrically (anisotropically) towards the boundary in such a way that the smallest element has width $O(|\varepsilon|)$. If we increase p on such a fixed mesh, Theorem 6.5 still holds true [28]. In the right-hand side of Figure 2 the geometric refinement towards $\partial\Omega$ is illustrated.

Remark 6.8 The results in Theorem 6.5 are established by means of asymptotic expansion techniques. If the data f satisfy certain compatibility conditions, the leading order terms in these expansion series vanish and the strength of the actual layers in the solutions can then be considerably weaker. In that case the use of the boundary layer meshes above can possibly lead to an overrefinement in the spatial discretizations.

Remark 6.9 If the domain $\Omega \subset \mathbb{R}^2$ has corners, the regularity and approximation theory of equations of the form (6.21) gets more complicated due to the interaction of corner singularities and boundary layers. At present, a rigorous proof of robust exponential convergence as in Theorem 6.5 is lacking in the presence of corners. Nevertheless, in simple model situations [33] it is known that the use of tensor products of geometrically refined meshes near corners leads the robust exponential rates of convergence, namely

$$\|u - U_{FEM}\|_\varepsilon \leq C \exp(-bN_x^{1/3}), \tag{6.24}$$

where $N_x = \dim(S_0^{p,1}(\mathcal{T}))$ and b, C are independent of ε. The number of layers in the geometric refinement must be related linearly to p and the smallest layer has again to be of width $O(|\varepsilon|)$. In Figure 3 such geometric boundary layer meshes near convex and reentrant corners are shown (where we admitted also hanging nodes).

Remark 6.10 If the right-hand side f is piecewise analytic, Theorem 6.5 still holds true provided that the spatial mesh is correspondingly adapted in the interior of the domain to resolve interior layers.

Convergence properties of the fully discrete scheme: We consider now the *hp*-discretization of (2.3), (2.4) in time and space. To do so, assume that the right-hand side g is (piecewise) analytic in time and space and that the initial condition u_0 is (piecewise) analytic in space. We emphasize, however, that u_0 need not be

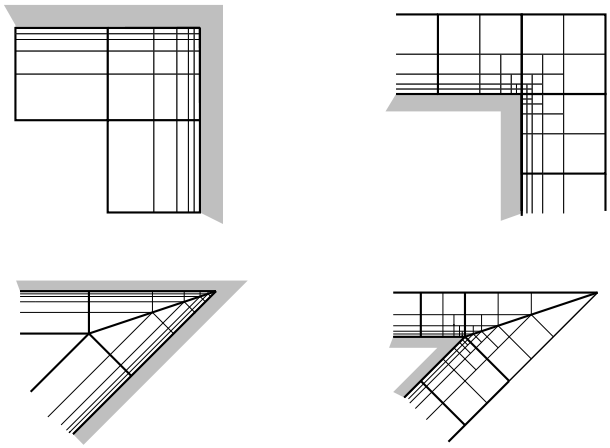

Figure 3: Geometric boundary layer meshes near convex and reentrant corners.

in $H_0^1(\Omega)$ and incompatibilities of the initial data with the zero Dirichlet boundary conditions are admitted. Condition (2.10) is then fulfilled for some $\theta > 0$.
We discretize (2.3), (2.4) in time by the *hp*-DGFEM on a geometric temporal partition $\mathcal{M}_{n,\sigma}$. The smallest singular perturbation parameter ε_m in the spatial problems (6.21) on time step I_m is of the order

$$|\varepsilon_m|^2 \sim k_m / \max(1, r_m^2) = C \frac{\sigma^{n-m}}{m^2}.$$

To solve these equations we use for simplicity a fixed geometric spatial boundary layer mesh $\mathcal{T}$ constructed to resolve $\varepsilon_{min} = \min_{m=1}^{n+1} |\varepsilon_m|$. That is: $\mathcal{T}$ is geometrically refined towards corners and boundaries as in Figure 3. The width of the smallest mesh layer is $O(\varepsilon_{min})$. The space problems are solved in $V_N = S_0^{p,1}(\mathcal{T})$ according to (6.19). Then all boundary layer and corner singularity phenomena present in the solutions of the reaction-diffusion equations (6.21) are captured and due to the analyticity assumption we can expect that all the spatial equations (6.21) are approximated at a (uniform) exponential rate of convergence (in the energy norm) as in (6.24), i.e. there holds on each time step I_m

$$\frac{k_m}{2r_m^2} \|\nabla(w_{j,m} - W_{j,m}^F)\|_{L^2(\Omega)}^2 + \|w_{j,m} - W_{j,m}^F\|_{L^2(\Omega)}^2 \leq C \exp(-bN_x^{1/3}) \qquad (6.25)$$

with $N_x = \dim(S_0^{p,1}(\mathcal{T}))$, $j = 0, \ldots, r_m$, $m = 1, \ldots, M$. (Recall that here $\{w_{j,m}\}$ are the exact solutions and $\{W_{j,m}^F\}$ the FE approximations of the $r_m + 1$ decoupled equations on time step I_m after transforming the system (6.4) as proposed in (i) or (ii) of Section 6.2.)
Using (6.25), the properties of the transformations in Section 6.2 and Theorem 5.4, the full discretization error in (6.5) of Proposition 6.1 becomes

$$\|u - U^F\|_{L^2(J;H_0^1(\Omega))}^2 \leq C \exp(-bN_t^{\frac{1}{2}}) + C \sum_{m=1}^{M} r_m^3 \exp(-bN_x^{\frac{1}{3}}),$$

where $N_t = \mathrm{nrdof}(\mathcal{V}^{\underline{r}}(\mathcal{M}; H_0^1(\Omega)))$, $N_x = \dim(S_0^{p,1}(\mathcal{T}))$, and $C \leq C r_m^\alpha$ if Method (ii) is employed. Since $\sum_{m=1}^M r_m \leq C N_t$ and $r_{max} = \max_{i=1}^M r_m \leq N_t$, we get

$$\|u - U^F\|^2_{L^2(J;H_0^1(\Omega))} \leq C \exp(-b N_t^{\frac{1}{2}}) + C N_t^3 \exp(-b N_x^{1/3}).$$

Since the coefficient in front of $\exp(-bN_x^{1/3})$ grows algebraically in N_t, it can be absorbed by the exponential decay in N_x if N_x and N_t are related algebraically, i.e. $N_t \leq C N_x^\beta$ for some $\beta \geq 0$. In that case we get

$$\|u - U^F\|_{L^2(J;H_0^1(\Omega))} \leq C \exp(-b N_t^{\frac{1}{2}}) + C \exp(-b N_x^{\frac{1}{3}}).$$

If the domain is in $\mathbb{R}^d$ with $d \geq 3$, analogous estimates hold true as soon as the spatial problems can be approximated exponentially as in Theorem 6.5 or (6.24).

Remark 6.11 With time the solutions of (2.3), (2.4) are strongly smoothened and after a few time steps the boundary layer phemonena actually encountered in the spatial problems may be considerably weaker than in the discussions above due to right-hand sides satisfying compatibility conditions as in Remark 6.8. In addition, for the geometric time meshes $\mathcal{M}_{n,\sigma}$ the singular parameter $|\varepsilon_m|^2 = k_m / \max(1, r_m^2)$ is of relatively moderate size after a few steps. Therefore, it is conceivable that the use of fixed boundary layer meshes leads to an overrefinement in the spatial discretizations. This can be overcome by coarsening the spatial meshes dynamically in time which compensates then to a certain extent the additional costs that arise for $r >> 1$. The development of such adaptive strategies in a *hp*-context requires a rigorous analytic regularity theory in time and space for problems in (2.3), (2.4) which does not lie within the scope of the present work, and, therefore, we do not consider such strategies here.
However, the following heuristic argument indicates that for highly incompatible initial conditions the geometric boundary layer meshes in space are needed in the first time steps: Consider the heat equation on a domain $\Omega \subset \mathbb{R}^d$ with right-hand side $g = 0$ and initial condition $u_0 = 1$. The first DGFEM time step with temporal approximation order $r = 0$ amounts in (2.16) to the solution of

$$\text{Find } U \in H_0^1(\Omega) \text{ such that } -k_m \Delta U + U = 1 \text{ in } \Omega, \qquad U = 0 \text{ on } \partial\Omega. \tag{6.26}$$

For small k_m problem (6.26) is singularly perturbed and clearly exhibits the previously discussed boundary layer phenomena due to the incompatibilities of the initial data u_0 with the zero Dirichlet boundary conditions.

6.4. Numerical Experiments

The model problems: In this section we present numerical results for the homogeneous heat equation where $L = -\Delta$, $H = L^2(\Omega)$ and $X = H_0^1(\Omega)$ obtained for the one dimensional domain $\Omega = (0,1)$ and the time interval $J = (0,1)$:

$$\begin{aligned} &\frac{\partial}{\partial t} u(x,t) - \frac{\partial^2}{\partial x^2} u(x,t) = 0 \text{ on } (0,1) \times (0,1), \\ &u(0,t) = u(1,t) = 0 \text{ on } (0,1), \qquad u(x,0) = u_0(x) \text{ on } (0,1). \end{aligned} \tag{6.27}$$

We investigate the performance of the DGFEM for the following three initial conditions

$$u_0^1(x) = \sin(\pi x), \qquad u_0^2(x) = x(1-x), \qquad u_0^3(x) = 1.$$

The first two initial conditions are compatible with being in $H_0^1(\Omega)$, while the third one is incompatible with the zero boundary conditions in (6.27). u_0^1 is an eigenfunction of the Laplacian and the corresponding solution u^1 given by $u^1(x,t) = \sin(\pi x)\exp(-\pi^2 t)$ is arbitrarily smooth in x and t (in fact, u^1 analytic on $\overline{J} \times \overline{\Omega}$). For the other two initial values the solutions can be represented by Fourier series of eigenfunctions of the Laplacian, that is

$$u^2(x,t) = 4\sum_{l=1}^{\infty} \frac{1-\cos(l\pi)}{l^3\pi^3} \exp(-l^2\pi^2 t)\sin(l\pi x), \tag{6.28}$$

$$u^3(x,t) = 2\sum_{l=1}^{\infty} \frac{1-\cos(l\pi)}{l\pi} \exp(-l^2\pi^2 t)\sin(l\pi x). \tag{6.29}$$

To determine the time regularity of the solutions u^2 and u^3 we write $a_{2,l} = \frac{1}{l^3}$ and $a_{3,l} = \frac{1}{l}$ for the size of the Fourier coefficients in the series. We have

$$\|\frac{\partial^s}{\partial t^s} u^i(x,t)\|^2_{H_0^1(\Omega)} \leq C(s) \sum_{l=1}^{\infty} l^2 a_{i,l}^2 \pi^{4s} l^{4s} \exp(-2l^2\pi^2 t)$$

and get

$$\|(u^i)^{(s)}\|^2_{L^2(I;H_0^1(\Omega))} \leq C \sum_{l=1}^{\infty} a_{i,l}^2 l^{4s}. \tag{6.30}$$

The above sum is finite for $s = 5/4 - \delta$ in the case $i = 2$ and for $s = 1/4 - \delta$ in the case $i = 3$, respectively, for any $\delta > 0$.

We discretize (6.27) in time by the DGFEM 2.3. In each time step I_m we obtain the system (6.4) which we decouple by the transformation Method (ii) in Section 6.2. The resulting $r_m + 1$ scalar reaction-diffusion equations of the form (6.21) are solved by the *hp*-FEM in (6.19) on meshes which are geometrically refined towards $x = 0$ and $x = 1$. The grading factor on these spatial meshes is 0.15 and the number of layers is chosen in such a way that the smallest scale ε_{min} is resolved, i.e. the first layer near $x = 0$ and $x = 1$ is of width $O(\varepsilon_{min})$. The polynomial degree in space is selected to be $p = 10$ such that the spatial problems are approximated very accurately and the overall error is dominated by the error of the time discretization.

h-DGFEM: We consider first the *h*-version of the DGFEM on an equidistant temporal partition $\mathcal{M}$ with a constant approximation order r. $\mathcal{M}$ consists of 2^i time steps, $i = 0, \ldots, 12$. The length of each time interval is then $k = 2^{-i}$. On the left-hand side in Figure 4 we plot the relative errors in $L^2(I, H_0^1(\Omega))$ against $N = \mathrm{NRDOF}(\mathcal{V}^r(\mathcal{M}; H_0^1(\Omega)))$ for the DGFEM solution of (6.27) with the initial condition $u_0 = u_0^1$ and $0 \leq r \leq 3$. The corresponding slopes (-1, -2, -3 and -4) predicted by Corollary 3.12 and Remark 3.13 can clearly be seen in the error graphs. The solutions u^2 and u^3 are not arbitrarily smooth in time anymore and the *h*-version convergence rates (for large r) are in that case determined by the

maximal regularity of the solution (cf. Corollary 3.12). Better results are expected with graded meshes (see Theorem 5.10). In Figure 5 we show the h-version DGFEM for the solutions u^2 and u^3 on quasiuniform and graded meshes. For u^2 we choose the grading function $h(t) = t^{2r+3}$ and employ the polynomial approximation order $r = 3$. The partitions consist still of 2^i intervals. In agreement with Corollary 3.12 and Remark 3.13 we get the slope $-5/4$ for equidistant time steps, whereas the slope -4 is recovered on the graded time mesh. For u^3 we take the grading function $h(t) = t^{3(2r+3)}$ and depict the performance for $r = 0$ and $r = 1$. On equidistant time meshes we observe the convergence rate $N^{-1/4}$, both for $r = 0$ and $r = 1$, according to Corollary 3.12. Again, the use of graded meshes yields the best possible rates with slopes -1 and -2, respectively.

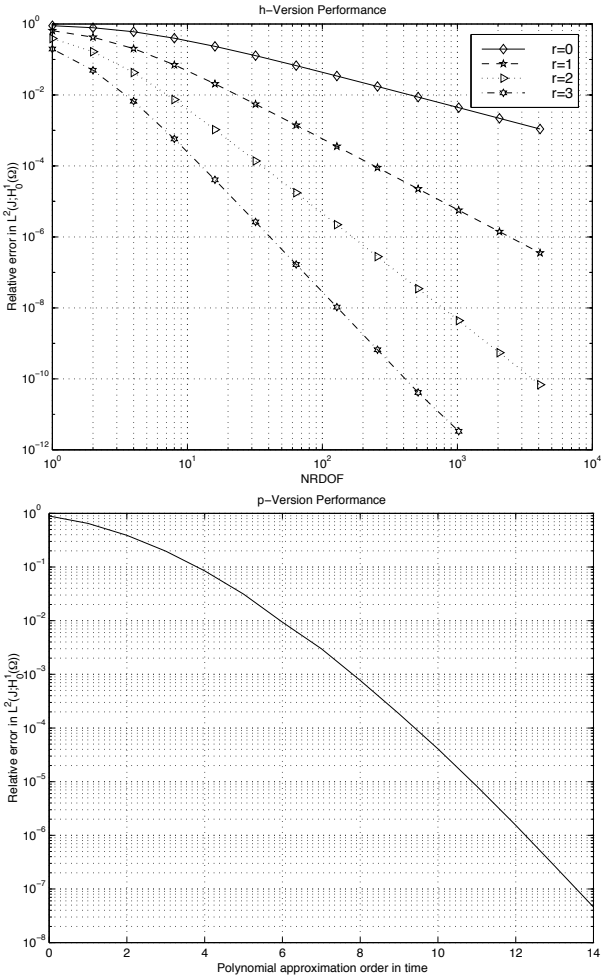

Figure 4: Convergence rates for the smooth solution u^1. Left: h-version DGFEM. Right: p-version DGFEM (NRDOF $= r + 1$).

p-DGFEM: In the p-version of the DGFEM the convergence is obtained by increasing the time approximation order r on fixed time intervals. If the solution is analytic, this results in exponential rates of convergence as in Remark 3.14 and is shown in the right-hand side of Figure 4 for $u_0 = u_0^1$. There, only one time step

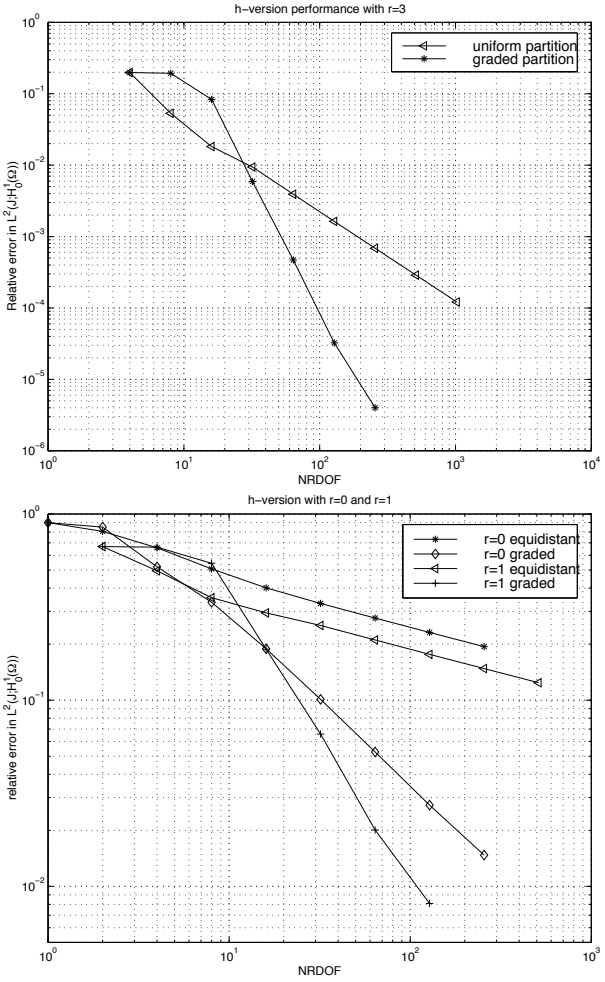

Figure 5: h-version DGFEM on quasiuniform and graded time steps. Left: Results for u^2. Right: Results for u^3.

$I_1 = (0, 1)$ was used. Such exponential convergence rates can not be expected anymore for u^2 and u^3. The performance of the p-version DGFEM for u^2 and u^3 is depicted in Figure 6. On the left we used one time step of length 1 and on the right four time steps of length 0.25. Due to Corollary 3.12 and (6.30) convergence rates of -1.25 and -0.25 are expected. However, we can see the slope -2.5 for u^2 and -0.5 for u^3. This doubling of the convergence rates is well known in the p-version (cf. [3, 4, 35]) and can be explained if the regularity of the solution is measured in certain weighted spaces. We refer to [33] for results in this direction.

hp-DGFEM: The hp-DGFEM combines judiciously h- and p-refinement. The time intervals $\{I_m\}$ are geometrically refined towards the origin on a partition $\mathcal{M}_{n,\sigma}$ as in Definition 5.1 and the polynomial degrees $\{r_m\}$ are linearly increasing from layer to layer (cf. Definition 5.3). In Figure 7 we consider the hp-version for initial data $u_0 = u_0^2$ and $u_0 = u_0^3$. We employ a geometrical grading factor $\sigma = 0.2$ and set the approximation order on layer m to $r_m = \lfloor \mu m \rfloor$ with a slope $\mu > 0$. The error graphs clearly show exponential rates of convergence as predicted by Theorem 5.4.

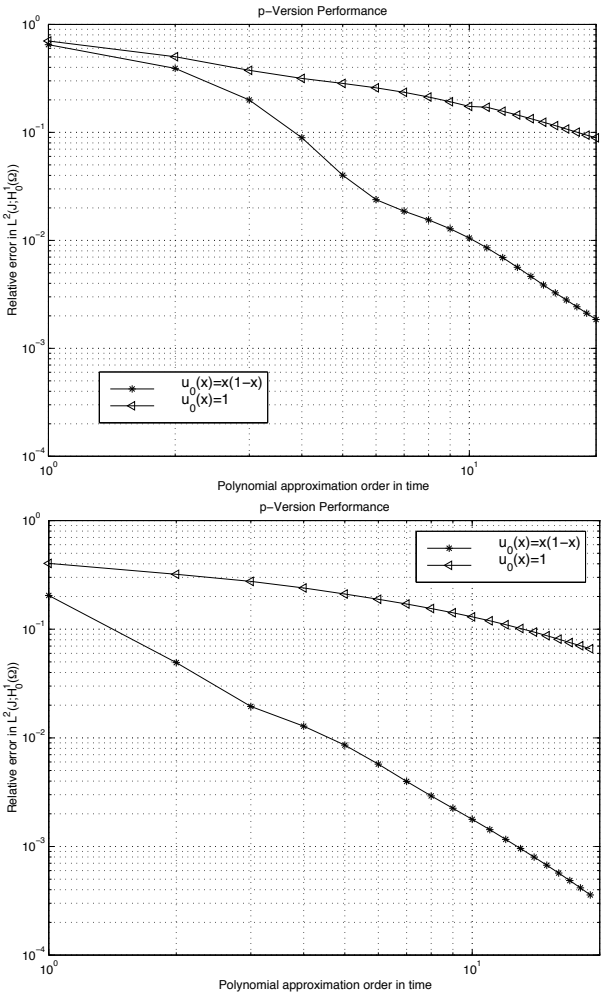

Figure 6: p-version DGFEM for u^2 and u^3. Left: Performance with one time step (NRDOF $= r+1$). Right: Performance with four time steps (NRDOF $= 4(r+1)$).

The best slope is ≈ 1.5 for u_2 and ≈ 0.5 for u_3. We conclude from these results that the optimal slope μ depends on θ of $u_0 \in H_\theta$. For u^3 we obtain a relative error of about 10^{-2} with ≈ 50 degrees of freedom. In the h-version approach on a graded mesh with $r = 1$ the same accuracy is obtained with ≈ 100 degrees of freedom (see Figure 5).

In the hp-version of the DGFEM the error can be orders of magnitude smaller when the grading factor σ is optimally chosen. We address this question numerically in Figure 8 where we vary σ on $\mathcal{M}_{n,\sigma}$ for $u_0 = u_0^2$ and $u_0 = u_0^3$ with $\mu = 1.5$ and $\mu = 0.5$, correspondingly. All the curves show exponential rates of convergence. It can be seen that $\sigma \approx 0.15$ gives the best results for u_0^2 and for u_0^3. This is in agreement to [21] where the optimal grading factor to resolve r^α-singularities is shown to be $\sigma \approx 0.17$, independently of α.

As indicated in Remark 6.11, it may not be necessary to approximate the spatial problems on boundary layer meshes. To investigate this numerically, we solve them in Figure 9 with a p-version FEM on a uniform spatial mesh consisting of four

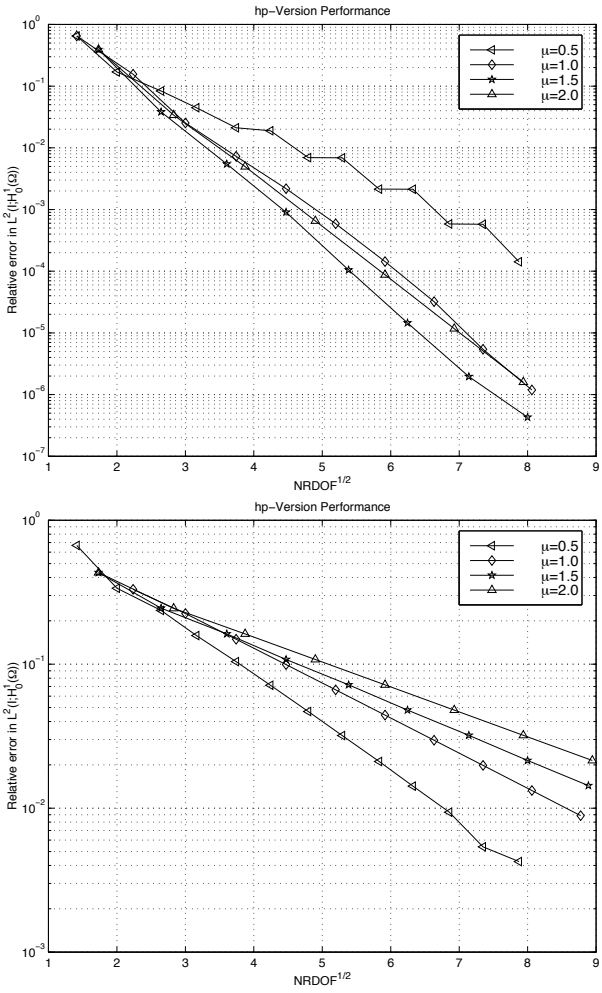

Figure 7: *hp*-version DGFEM: Exponential rates of convergence with $\sigma = 0.2$. Left: Results for u^2. Right: Results for u^3.

elements. The approximation order in space is selected as $p = 10$ and $p = 20$. The error curves show the performance of the *hp*-DGFEM with $\sigma = 0.2$ and with $\mu = 1.5$ for u^2 and $\mu = 0.5$ for u^3, respectively, in accordance to Figure 7. The methods first converge exponentially and then level off as soon as the error in the space discretization becomes dominant. Whereas the results for u^2 are more or less comparable with the ones in Figure 7 obtained with geometric boundary layer meshes in space, the error graphs for u^3 are clearly better if the spatial equations are solved on geometric meshes. This indicates that the strength of the layer phenomena actually present in the spatial problems depends on the compatibility of the initial data in agreement to the discussion in Remark 6.11. However, the performance on geometric boundary layer meshes in space seems to be more robust than on uniform meshes, at the disadvantage of a possible overrefinement.

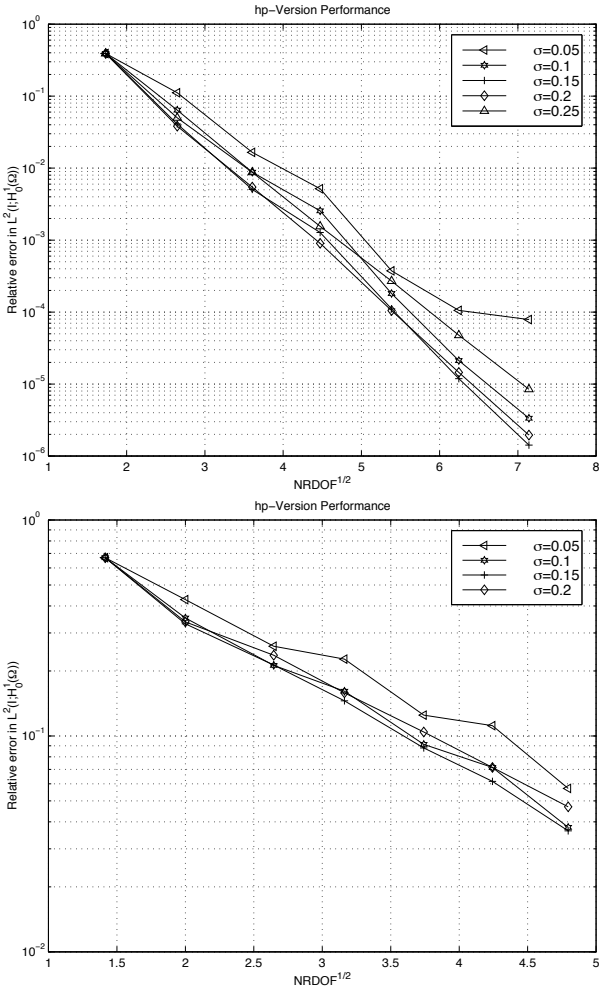

Figure 8: *hp*-version of DGFEM: Varying geometrical grading factors. Left: Results for u^2 ($\mu = 1.5$). Right: Results for u^3 ($\mu = 0.5$).

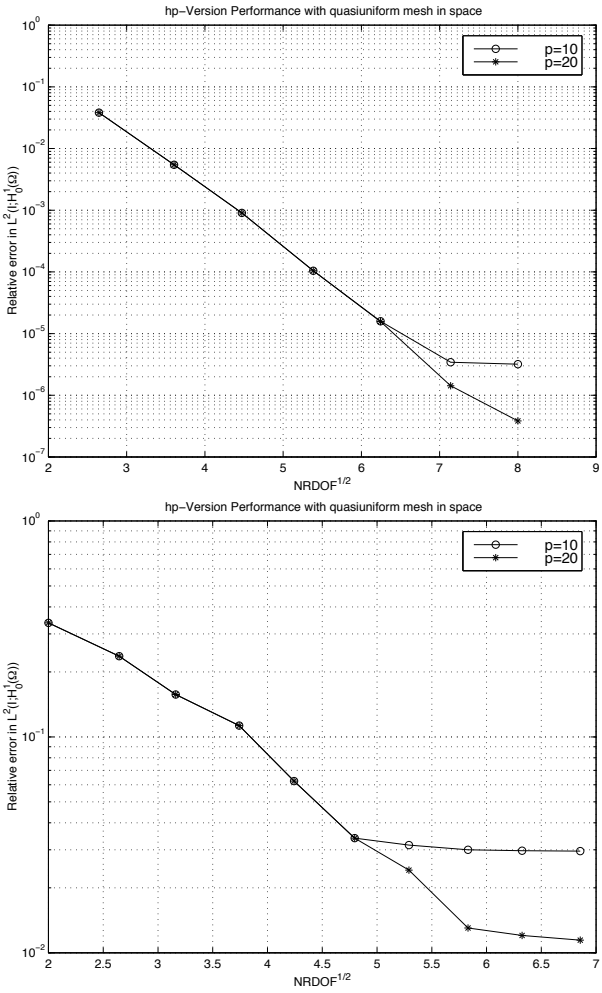

Figure 9: *hp*-version of DGFEM: *p*-version FEM in space on a quasiuniform mesh with four element. Left: Results for u^2. Right: Results for u^3.

References

[1] I. Babuška and B.Q. Guo: *The hp version of the Finite Element Method for domains with curved boundaries*, SIAM J. Numer. Anal. **25** (1988), 837-861.

[2] I. Babuška and T. Janik: *The hp-version of the Finite Element Method for parabolic equations, I: The p-version in time, II: The hp-version in time*, Numerical Methods for Partial Differential Equations **5** (1989), 363-399, and **6** (1990), 343-369.

[3] I. Babuška and M. Suri: *The p- and hp-versions of the Finite Element Method: An overview*, Computer Methods in Applied Mechanics and Engineering **80** (1990), 5-26.

[4] I. Babuška and M. Suri: *The p- and hp-versions of the Finite Element Method, basic principles and properties*, SIAM Review **36 (4)** (1994), 578-632.

[5] C.E. Baumann and J.T. Oden: *A discontinuous hp finite element method for convection diffusion problems*, Computer Methods in Applied Mechanics and Engineering (1999), to appear.

[6] J. Bergh and J. Löfström: **Interpolation Spaces**, Springer Verlag, New York, 1976.

[7] K. Böttcher and R. Rannacher: *Adaptive error control in solving ordinary differential equations by the Discontinuous Galerkin Method*, Preprint 96-53, IWR, Universität Heidelberg, 1996.

[8] B. Cockburn: *Discontinuous Galerkin methods for convection dominated problems*, Lecture Notes for the VKI/NASA-Ames special course on high order methods in CFD, to appear in Springer Lecture Notes, eds:. T.J. Barth and H. Deconinck, 1999, to appear.

[9] B. Cockburn and C.W. Shu: *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal. **35** (1998), 2440-2463.

[10] M. Delfour, W. Hager and F. Trochu: *Discontinuous Galerkin Methods for ordinary differential equations*, Math. Comp. **36** (1981), 455-473.

[11] K. Eriksson and C. Johnson: *Error estimates and automatic time step control for nonlinear parabolic problems I*, SIAM J. Numer. Anal. **24** (1987), 12-23.

[12] K. Eriksson and C. Johnson: *Adaptive Finite Element Methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal. **28** (1991), 43-77.

[13] K. Eriksson and C. Johnson: *Adaptive Finite Element Methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal. **32** (1995), 706-740.

[14] K. Eriksson and C. Johnson: *Adaptive Finite Element Methods for parabolic problems III: Time steps variable in space*, in preparation.

[15] K. Eriksson and C. Johnson: *Adaptive Finite Element Methods for parabolic problems IV: Nonlinear problems*, SIAM J. Numer. Anal. **32** (1995), 1729-1749.

[16] K. Eriksson and C. Johnson: *Adaptive Finite Element Methods for parabolic problems V: Long-time integration*, SIAM J. Numer. Anal. **32** (1995), 1750-1763.

[17] K. Eriksson, C. Johnson and S. Larsson: *Adaptive Finite Element Methods for parabolic problems VI: Analytic semigroups*, SIAM J. Numer. Anal. **35** (1998), 1315-1325.

[18] K. Eriksson, C. Johnson and V. Thomée: *Time discretization of parabolic problems by the Discontinuous Galerkin Method*, RAIRO Modél. Math. Anal. Numér. **19** (1985), 611-643.

[19] D. Estep: *A posteriori error bounds and global error control for approximation of ordinary differential equations*, SIAM J. Numer. Anal. **32** (1995), 1-48.

[20] G.H. Golub and C.F. van Loan: **Matrix computations**, 3rd Ed., The Johns Hopkins University Press, Maryland, 1996.

[21] W. Gui and I. Babuška: *The h-, p- and hp-versions of the Finite Element Method in one dimension, I: The error analysis of the p-version, II: The error analysis of the h- and hp-versions, III: The adaptive hp-version*, Numerische Mathematik **49** (1986), 577-612, 613-657 and 659-683.

[22] B.Q. Guo and I. Babuška: *The hp-version of the finite element method, Part I: The basic approximation results, Part II: General results and applications*, Comp. Mech. **1** (1986), 21-41 and 203-226.

[23] C. Johnson: *Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations*, SIAM J. Numer. Anal. **25** (1988), 908-926.

[24] S. Larsson, V. Thomée and L.B. Wahlbin: *Numerical solution of parabolic integro-differential equations by the Discontinuous Galerkin Method*, Math. Comp. **67 (221)** (1998), 45-71.

[25] J.L. Lions and E. Magenes: **Non-homogeneous Boundary Value Problems and Applications**, Volume I, Springer Verlag, New York, 1972.

[26] Ch.G. Makridakis and I. Babuška: *On the stability of the Discontinuous Galerkin Method for the heat equation*, SIAM J. Numer. Anal. **34 (1)** (1997), 389-401.

[27] J.M. Melenk: *On the robust exponential convergence of hp finite element methods for problems with boundary layers*, IMA J. Numer. Anal. **17** (1997), 577-601.

[28] J.M. Melenk and C. Schwab: *hp-FEM for reaction-diffusion equations, I: Robust exponential convergence*, SIAM J. Numer. Anal. **35 (4)** (1998), 1520-1557.

[29] J.M. Melenk and C. Schwab: *Analytic regularity for a singularly perturbed problem*, SIAM J. Math. Anal., in press.

[30] J.T. Oden, I. Babuška and C.E. Baumann: *A discontinuous hp finite element method for diffusion problems*, J. Comput. Phys. **146** (1998), 495-519.

[31] F.W.J. Olver: **Asymptotics and Special Functions**, Academic Press, San Diego, 1974.

[32] A. Prohl: **Projection and Quasi-Compressibility Methods for Solving the Incompressible Navier-Stokes Equations**, Advances in Numerical Mathematics, Teubner, Stuttgart, 1997.

[33] D. Schötzau: *hp-DGFEM for parabolic equations*, Doctoral Dissertation, ETH Zürich, to appear.

[34] D. Schötzau and C. Schwab: *Analytic Time Regularity of Non-Selfadjoint Parabolic Problems and Applications in the hp-Version of the Discontinous Galerkin Method*, in preparation.

[35] C. Schwab: **p- and hp-Finite Element Methods**, Oxford University Press, New York, 1998.

[36] C. Schwab and M. Suri: *The p- and hp-versions of the finite element method for problems with boundary layers*, Math. Comp. **65** (1996), 1403-1429.

[37] C. Schwab, M. Suri and C.A. Xenophontos: *The hp-version of the FEM for problems in mechanics with boundary layers*, Computer Methods in Applied Mechanics and Engineering **157** (1998), 311-333.

[38] V. Thomée: **Galerkin Finite Element Methods for Parabolic Problems**, Springer Verlag, New York, 1997.

[39] H. Triebel: **Interpolation Theory, Function Spaces, Differential Operators**, 2nd Ed., J.A. Barth Publishers, Heidelberg-Leipzig, 1995.

[40] C.A. Xenophontos: *The hp Finite Element Method for singularly perturbed problems in smooth domains*, Math. Models and Methods in Applied Sciences **8 (2)** (1988), 299-326.

Research Reports

No.	Authors	Title
99-04	D. Schötzau, C. Schwab	Time Discretization of Parabolic Problems by the hp-Version of the Discontinuous Galerkin Finite Element Method
99-03	S.A. Zimmermann	The Method of Transport for the Euler Equations Written as a Kinetic Scheme
99-02	M.J. Grote, A.J. Majda	Crude Closure for Flow with Topography Through Large Scale Statistical Theory
99-01	A.M. Matache, I. Babuška, C. Schwab	Generalized p-FEM in Homogenization
98-10	J.M. Melenk, C. Schwab	The *hp* Streamline Diffusion Finite Element Method for Convection Dominated Problems in one Space Dimension
98-09	M.J. Grote	Nonreflecting Boundary Conditions For Electromagnetic Scattering
98-08	M.J. Grote, J.B. Keller	Exact Nonreflecting Boundary Condition For Elastic Waves
98-07	C. Lage	Concept Oriented Design of Numerical Software
98-06	N.P. Hancke, J.M. Melenk, C. Schwab	A Spectral Galerkin Method for Hydrodynamic Stability Problems
98-05	J. Waldvogel	Long-Term Evolution of Coorbital Motion
98-04	R. Sperb	An alternative to Ewald sums, Part 2: The Coulomb potential in a periodic system
98-03	R. Sperb	The Coulomb energy for dense periodic systems
98-02	J.M. Melenk	On n-widths for Elliptic Problems
98-01	M. Feistauer, C. Schwab	Coupling of an Interior Navier–Stokes Problem with an Exterior Oseen Problem
97-20	R.L. Actis, B.A. Szabo, C. Schwab	Hierarchic Models for Laminated Plates and Shells
97-19	C. Schwab, M. Suri	Mixed *hp* Finite Element Methods for Stokes and Non-Newtonian Flow
97-18	K. Gerdes, D. Schötzau	*hp* FEM for incompressible fluid flow - stable and stabilized
97-17	L. Demkowicz, K. Gerdes, C. Schwab, A. Bajer, T. Walsh	HP90: A general & flexible Fortran 90 *hp*-FE code
97-16	R. Jeltsch, P. Klingenstein	Error Estimators for the Position of Discontinuities in Hyperbolic Conservation Laws with Source Terms which are solved using Operator Splitting