Deutsche Forschungsgemeinschaft

Priority Program 1253

Optimization with Partial Differential Equations

T. Apel, T. G. Flaig and S. Nicaise

A priori error estimates for finite element methods for $H^{(2,1)}$-elliptic equations

December 2012

Preprint-Number SPP1253-139

http://www.am.uni-erlangen.de/home/spp1253

A priori error estimates for finite element methods for $H^{(2,1)}$-elliptic equations

Thomas Apel* Thomas G. Flaig† Serge Nicaise‡

December 7, 2012

Abstract

The convergence of finite element methods for linear elliptic boundary value problems of second and forth order is well understood. In this paper we introduce finite element approximations of some linear semi-elliptic boundary value problem of mixed order on a two dimensional rectangular domain Q. The equation is of second order in one direction and forth order in the other. We establish a regularity result and estimates for the finite element error of conforming approximations of this equation. The finite elements in use have a tensor product structure, in one dimension we use linear, quadratic or cubic Lagrange elements in the other dimension cubic Hermite elements. For these elements we prove the error bound $\mathcal{O}(h^2+\tau^k)$ in the energy norm and $\mathcal{O}\left((h^2+\tau^k)(h^2+\tau)\right)$ in the $L^2(Q)$-norm.

This type of equations appears in the optimal control of parabolic partial differential equations if one eliminates the control and the state (or the adjoint state) in the first order optimality conditions.

1 Introduction and general setting

Finite element error estimates are well known for second and forth order elliptic boundary value problems, but the situation changes, if the differential equation has different orders in different dimensions, e.g. if we discuss the differential equation

$$-y_{tt}+y_{xxxx}+y=f \qquad \text{in } Q=(0,X)\times(0,T), \tag{1}$$

with the conditions

$$\begin{aligned} y(x,0)&=0, & y_t(x,T)-y_{xx}(x,T)&=0,\\ y(0,t)&=0, & y_x(X,t)&=0.\\ y_{xx}(0,t)&=0, & y_{xxx}(X,t)&=0. \end{aligned}$$

*thomas.apel@unibw.de, Universität der Bundeswehr München, Institut für Mathematik und Bauinformatik, D-85579 Neubiberg, Germany

†thomas.flaig@unibw.de, Universität der Bundeswehr München, Institut für Mathematik und Bauinformatik, D-85579 Neubiberg, Germany

‡snicaise@univ-valenciennes.fr, LAMAV, Institut des Sciences et Techniques de Valenciennes, Université de Valenciennes et du Hainaut Cambrésis, B.P. 311, 59313 Valenciennes Cedex, France

It is easy to verify that the corresponding bilinear form is $V-$elliptic in the Sobolev space $V = H^{(2,1)}(Q)$, the Sobolev space of all $L^2(Q)$-functions whose second derivative with respect to x and first derivative with respect to t are square integrable. The equation itself does not satisfy the strong conditions of ellipticity, i.e. that all eigenvalues of the leading part of the operator have the same sign, but only the conditions of semi-ellipticity for which also zero eigenvalues in the leading part of the operator are allowed.

Function spaces with variable order of differentiation in different dimensions are discussed in [21, 22, 24, 30, 34]. A priori estimates in Besov spaces for the equation (1) on the unit circle are discussed by Triebel [34]. Estimates for general semi-elliptic equations can be found e.g. in [3, 4, 31, 20]. Eastham and Peterson discuss an isotropic finite element discretization of the Onsager pancake equation, which is another example for a semi-elliptic partial differential equation, in [14].

Such equations are not only of academic interest but appear in the optimal control of parabolic partial differential equation, see Section 2. To our knowledge in the context of parabolic optimal control problems such equations have been derived first by Büttner [10] and properties, such as ellipticity, have been discussed by Neitzel, Prüfert and Slawig [28, 29]. Gong, Hinze and Zhou [17] discuss a priori and a posteriori error estimates for mixed finite element discretizations.

We discuss the approximation error of a $H^{(2,1)}(Q)$ conforming finite element method for this equation and use interpolation error estimates based on the technique used by Rachowicz [32]. The main results are a regularity result, error estimates up to the order $h^2 + \tau^k$ in the energy norm and error estimates in the $L^2(Q)$-norm of order $(\tau^k + h^2)(\tau + h^2)$ for $k = 1, 2, 3$.

As announced we discuss in the following Section the connection of parabolic optimal control problems with semi-elliptic equations. In Section 3 we introduce the corresponding Sobolev spaces and prove an a priori regularity estimate for the exact solution. In Section 4 we present our finite element discretization and our main result, the error estimate for the Hermite-Lagrange tensor product finite elements, where the proof of the interpolation error is moved to the Section 5. Finally we present numerical examples in Section 6.

2 Optimal control problems

We consider an optimal control problem with a parabolic partial differential equation, i.e.

$$\begin{aligned} \min \int_0^T \frac{1}{2}(\, \|y - y_d\|^2_{L^2(\Omega)} + \nu \, \|u\|^2_{L^2(\Omega)} \,) \, \mathrm{d}\, t, & \\ \text{s.th.} \qquad y_t + Ay = u, \qquad & \text{in } Q = \Omega \times (0,T), \\ y = 0, \qquad & \text{in } \Omega \times \{0\}, \end{aligned}$$

where $\Omega = (a, b)$ with $a, b \in \mathbb{R}$ such that $a < b$ and A is a self-adjoint, second order elliptic operator with the boundary conditions

$$y = 0 \text{ on } \Sigma_1 = \Gamma_1 \times (0,T), \qquad \frac{\partial}{\partial n} y = 0 \text{ on } \Sigma_2 = \Gamma_2 \times (0,T).$$

with $\partial\Omega = \Gamma_1 \cup \Gamma_2$. The optimality conditions for this optimal control problem are well known and given by a system of partial differential equations (see e.g. [23, 35])

$$y_t + Ay = \frac{1}{\nu} p \text{ in } Q, \qquad y(x,0) = 0 \text{ in } \Omega, \qquad y = 0 \text{ on } \Sigma_1, \qquad \frac{\partial}{\partial n} y = 0 \text{ on } \Sigma_2, \tag{2}$$

$$p_t - Ap = y - y_d \text{ in } Q, \qquad p(x,T) = 0 \text{ in } \Omega, \qquad p = 0 \text{ on } \Sigma_1, \qquad \frac{\partial}{\partial n} p = 0 \text{ on } \Sigma_2. \tag{3}$$

We have chosen the sign of the adjoint state p so that $u = \frac{1}{\nu}p$, as in [8, 9, 19], the other choice with $u = -\frac{1}{\nu}p$ is also popular. The choice of the sign of the adjoint state only influences the right hand sides of the partial differential equations.

If we use the first equation (3) as definition for the state, $y = p_t - Ap + y_d$, and insert this definition into the equation (2), we end up with the equation

$$-p_{tt} + A^2 p + \frac{1}{\nu} p = y_{d,t} + A y_d. \tag{4}$$

The corresponding boundary conditions are

$$\begin{aligned} p(x,T) &= 0 \text{ in } \Omega, & p &= 0 \text{ on } \Sigma_1, & \frac{\partial}{\partial n} p &= 0 \text{ on } \Sigma_2, \\ p_t(x,0) - Ap(x,0) &= -y_d(x,0) \text{ in } \Omega, & Ap &= y_d \text{ on } \Sigma_1, & \frac{\partial}{\partial n} Ap &= y_d \text{ on } \Sigma_2. \end{aligned}$$

On the other hand if we use the first equation (2) as definition for $p = \nu y_t + \nu A y$ and insert this into the equation (3) then we end up with the equation

$$-y_{tt} + A^2 y + \frac{1}{\nu} y = \frac{1}{\nu} y_d. \tag{5}$$

with the boundary conditions

$$\left.\begin{aligned} y(x,0) &= 0 \text{ in } \Omega, & y &= 0 \text{ on } \Sigma_1, & \frac{\partial}{\partial n} y &= 0 \text{ on } \Sigma_2, \\ y_t(x,T) + Ay(x,T) &= 0 \text{ in } \Omega, & Ay &= 0 \text{ on } \Sigma_1, & \frac{\partial}{\partial n} Ay &= 0 \text{ on } \Sigma_2. \end{aligned}\right\} \tag{6}$$

For the equation (4) we need the time derivative and the application of the operator A to the desired state whereas for (5) we need only the desired state itself. Therefore (5) may also be used if y_d is less regular.

Remark 2.1. *This approach can be extended to optimal control problems with constraints for the control, i.e.*

$$c_1 \leq u = \frac{1}{\nu} p \leq c_2.$$

In this case we have for the adjoint state the nonlinear equation

$$-p_{tt} + A^2 p + \Pi_{c_1,c_2}\left(\frac{1}{\nu} p\right) = y_{d,t} + A y_d,$$

where Π_{c_1,c_2}*, given by*

$$\Pi_{c_1,c_2}\left(\frac{1}{\nu} p\right) = \max\left\{c_1, \min\left\{\frac{1}{\nu} p, c_2\right\}\right\},$$

is the projection to the set of admissible controls.

Remark 2.2. *A slightly different approach for the elimination of the state or the adjoint state can be found in [10] or [28]. They also prove that the corresponding bilinear form is* V*-elliptic in an appropriately chosen Hilbert space.*

Remark 2.3. *Beside parabolic optimal control problems the Onsager pancake equation*

$$
\begin{aligned}
\left(\mathrm{e}^{x}\left(e^{x}u_{xx}\right)_{xx}\right)_{xx}+bu_{yy}&=f(x,y), && \text{in } (0,1)^2,\\
u_x(0,y)=u_{xx}(0,y)=0,\ \left(\mathrm{e}^{x}\left(\mathrm{e}^{x}\,u_{xx}\right)_{xx}\right)_x(0,y)&=g(y),\\
u(1,y)=u_x(1,y)=0,\ \left(e^{x}u_{xx}\right)_x(1,y)&=0,\\
-bu_y(x,0)&=d\left(\mathrm{e}^{x/2}\,u_x\right)_x+\gamma_0(x),\\
bu_y(x,1)&=d\left(\mathrm{e}^{x/2}\,u_x\right)_x+\gamma_1(x),
\end{aligned}
$$

is another example of a semi-elliptic equation with different order of differentiation in different dimension. This equation models the flow in a gas centrifuge, a physical motivation can be found in [36]. First finite element discretizations of this equation are given in [6, 18]. In [14] a tensor product finite element with B-spline basis functions is used. In contrast to our technique based on an anisotropic interpolation result they only achieve second order of convergence in $L^2\left((0,1)^2\right)$. *It is likely that the transfer of our approach leads to better approximation rates and suggestions for coupling of the discretization parameters for* x *and* y.

3 Definitions and a priori estimates

3.1 Anisotropic Sobolev spaces with application to the model problem

For the multi-index $\alpha=(\alpha_1,\alpha_2)\in\mathbb{N}_0^2$ we define as usual the differential operator with respect to a multi-index as

$$D^\alpha=\left(\frac{\partial}{\partial t}\right)^{\alpha_2}\cdot\left(\frac{\partial}{\partial x}\right)^{\alpha_1}.$$

Theorem 3.1. *[24, Chapter 2.1] For* $r,s\in\mathbb{N}$, *the Sobolev spaces* $H^{(r,s)}(Q)$ *defined as*

$$H^{(r,s)}(Q)=L^2(0,T;H^r(\Omega))\cap H^s(0,T;L^2(\Omega))$$

is an Hilbert space with the norm

$$|||y|||^2_{H^{(r,s)}(Q)}=\|y\|^2_{L^2(0,T;H^r(\Omega))}+\|y\|^2_{H^s(0,T;L^2(\Omega))}\,.$$

Remark 3.2. *The Sobolev space* $H^{(2,1)}(Q)$ *was introduced by several authors. The definition above can be found in Lions and Magenes [24] or Ladyzhenskaya, Solonnikov and Ural'ceva [21]. They assume that lower order derivatives in every direction have to be in* $L^2(Q)$. *The definition of Nikol'skiĭ [30] and Triebel [34] is slightly different. They assume only that the function itself and its highest order derivative in every direction are* $L^2(Q)$*-functions. But Nikol'skiĭ also shows which lower order derivatives can be estimated by the norm of* y *and the norm of the highest order derivatives. We repeat a special case of Nikol'skiĭ's theorem in the next Theorem and prove the equivalence of two norms of* $H^{(2,1)}(Q)$ *in Theorem 3.4.*

Theorem 3.3. *Let* $Q=(a,b)\times(0,T)$. *Suppose that* y, $D^{(r,0)}y$, $D^{(0,s)}y\in L^2(Q)$ *and the multi-index* $l=(l_1,l_2)$ *fulfills*

$$1-\frac{l_1}{r}-\frac{l_2}{s}\geq 0,$$

then we have also

$$\left\| D^l y \right\|_{L^2(Q)} \lesssim \|y\|_{L^2(Q)} + \left\| D^{(r,0)} y \right\|_{L^2(Q)} + \left\| D^{(0,s)} y \right\|_{L^2(Q)} .$$

Proof. This is the anisotropic version of a similar result for the isotropic case which can be found in the book of Smirnow [33, Comment in Section IV.112.]. The isotropic case is proven for domains that are star shaped with respect to a non-empty ball in [33, Section IV.116.]. In [33, Section IV.118.] this is generalized in for domains which can be decomposed into finitely many subdomains which are star shaped with respect to a non-empty sphere.

For $Q = \mathbb{R}^n$ the anisotropic version of the result of Theorem 3.3 can be found in the book of Nikol'skiĭ [30, Theorem 9.2.2.].

For a class of special finite domains (including rectangular domains) a more general result, which includes the result of Theorem 3.3 as special case, is proven in [7, Theorem 13.6.1.]. □

Theorem 3.4. *For a one-dimensional domain* $\Omega = (a, b)$ *and* $Q = \Omega \times (0, T)$ *the norms*

$$\begin{aligned} |||y|||^2_{H^{(2,1)}(Q)} &= |y|^2_{H^{(2,1)}(Q)} + 2\,\|y\|^2_{L^2(Q)} + \|y_x\|^2_{L^2(Q)} , \\ \|y\|^2_{H^{(2,1)}(Q)} &= |y|^2_{H^{(2,1)}(Q)} + \|y\|^2_{L^2(Q)} , \end{aligned}$$

where the semi-norms are defined by

$$\begin{aligned} |y|^2_{H^{(r,s)}(Q)} &= |y|^2_{H^{(r,0)}(Q)} + |y|^2_{H^{(0,s)}(Q)} , \\ |y|^2_{H^{(r,0)}(Q)} &= \iint\limits_Q \left| D^{(r,0)} y \right|^2 \mathrm{d}\,x \,\mathrm{d}\,t, \qquad |y|^2_{H^{(0,s)}(Q)} = \iint\limits_Q \left| D^{(0,s)} y \right|^2 \mathrm{d}\,x \,\mathrm{d}\,t, \end{aligned}$$

are equivalent norms for the space $H^{(2,1)}(Q)$.

Proof. The inequality $\|y\|^2_{H^{(2,1)}(Q)} \leq |||y|||^2_{H^{(2,1)}(Q)}$ is clear. On the other hand if $y, y_{xx} \in L^2(Q)$ it follows by Theorem 3.3 that

$$\|y_x\|_{L^2(Q)} \lesssim \|y\|_{L^2(Q)} + \|y_{xx}\|_{L^2(Q)}$$

which proves $|||y|||_{H^{(2,1)}(Q)} \lesssim \|y\|_{H^{(2,1)}(Q)}$. □

Furthermore we introduce Sobolev spaces with respect to a multi-index set. Let the set A be a finite set of multi-indices, then we define the Sobolev space

$$H^A(Q) = \{u \in L^2(Q) : \ D^\alpha u \in L^2(Q), \ \forall \alpha \in A\}.$$

Theorem 3.5. *For the variational formulation of the model equation* (5) *with the boundary conditions* (6) *and* A *the Laplace operator, given by*

$$\left.\begin{aligned} a(y, \varphi) &= (f, \varphi) \qquad \forall \varphi \in V, \\ a(y, \varphi) &= \iint_Q y_t \varphi_t + y_{xx}\varphi_{xx} + \frac{1}{\nu} y\varphi \,\mathrm{d}\,x \,\mathrm{d}\,t + \int_\Omega y_x(x, T)\varphi_x(x, T) \,\mathrm{d}\,x, \\ (f, \varphi) &= \iint_Q \frac{1}{\nu} y_d \varphi \,\mathrm{d}\,x \,\mathrm{d}\,t, \\ V &= \left\{ v \in H^{(2,1)}(Q) : \ v(x, 0) = 0, \ v = 0 \ \text{on } \Sigma_1, \ v_x = 0 \ \text{on } \Sigma_2 \right\}, \end{aligned}\right\} \quad (7)$$

there exists a unique solution $y \in V$ *for* $y_d \in V^*$.

Proof. We prove this Theorem for the case $\nu = 1$. The modifications for arbitrary $\nu \in \mathbb{R}^+$ are obvious. The existence of a unique solution follows with the Lax-Milgram lemma, if we can prove the $V-$ellipticity and continuity of the bilinear form $a(\cdot,\cdot)$. The V-ellipticity follows directly as

$$\|y\|_{H^{(2,1)}(Q)} = a(y,y) - \int_\Omega (y_x(x,T))^2 \,\mathrm{d}\,x \le a(y,y).$$

For the continuity we use the Cauchy-Schwarz inequality

$$a(y,\varphi) \le \|y\|_{H^{(2,1)}(Q)} \|\varphi\|_{H^{(2,1)}(Q)} + \|y_x(x,T)\|_{L^2(\Omega)} \|\varphi_x(x,T)\|_{L^2(\Omega)}.$$

As $H^{(2,1)}(Q) \hookrightarrow C([0,T];H^1(\Omega))$ (see e.g. [12, (XVIII.1.61.iii)]) we have

$$\|y_x\|_{L^2(\Omega)} \le \|y\|_{H^1(\Omega)} \le \|y\|_{C([0,T],H^1(\Omega))} \lesssim \|y\|_{H^{(2,1)}(Q)}.$$

With this estimate we have proven the continuity of the bilinear form $a(\cdot,\cdot)$, and therefore the existence of the unique solution y follows. □

3.2 Regularity estimate for the $H^{(2,1)}(Q)$-elliptic equation

In this Subsection we provide an a priori estimate for semi-elliptic equations, which is needed for the proof of an $L^2(Q)$-error estimate with the Aubin-Nitsche trick.

Remark 3.6. *As the proof for several spatial dimensions is not more complicated than the spatially one dimensional case, we present the regularity estimate for the general case and a general elliptic operator A of second order.*

Theorem 3.7. *If $f \in L^2(Q)$ and A is a self adjoint second order elliptic operator, then the solution y of the problem*

$$\begin{aligned}
-y_{tt} + A^2 y + \frac{1}{\nu} y &= f && \text{in } Q,\\
y &= 0, && \text{on } \Sigma_1,\\
\frac{\partial}{\partial n_A} y &= 0, && \text{on } \Sigma_2,\\
Ay &= 0, && \text{on } \Sigma_1,\\
\frac{\partial}{\partial n_A} Ay &= 0, && \text{on } \Sigma_2,\\
y(x,0) &= 0, && \text{in } \Omega \times \{0\}\\
y_t(x,T) + Ay(x,T) &= 0, && \text{in } \Omega \times \{T\},
\end{aligned}$$

fulfills the estimate

$$\|y\|^2_{L^2(D(A^2))} + \|y\|^2_{H^1(D(A))} + \|y\|^2_{H^2(L^2(\Omega))} \lesssim \|f\|^2_{L^2(Q)}.$$

Remark 3.8. *If the domain Ω is smooth, we have $D(A) = H^2(\Omega)$ and $D(A^2) = H^4(\Omega)$ and therefore in this case the estimate of the Theorem is*

$$\|y\|_{H^{(4,2)}(Q)} \lesssim \|f\|_{L^2(Q)}.$$

Proof. We introduce the set $\{\varphi_k\}_{k=1}^{\infty}$ of orthonormal eigenfunctions of the operator A with the corresponding eigenvalues λ_k^2, which also fulfill the boundary conditions

$$\varphi_k = 0, \qquad \text{on } \Gamma_1, \qquad \frac{\partial}{\partial n}\varphi_k = 0, \qquad \text{on } \Gamma_2.$$

It is well known that the orthonormal eigenfunctions of a self-adjoint elliptic operator form an orthonormal basis of $L^2(\Omega)$ [15, Theorem 6.5.1.]. By the definition of the eigenfunctions we have $A\varphi_k = \lambda_k^2\varphi_k$ and therefore the boundary conditions

$$A\varphi_k = 0, \qquad \text{on } \Gamma_1, \qquad \frac{\partial}{\partial n}A\varphi_k = 0, \qquad \text{on } \Gamma_2.$$

are also fulfilled. So we write the solution of the equation as eigenfunction expansion

$$y = \sum_{k=1}^{\infty} y_k(t)\varphi_k$$

with time-dependent coefficients $y_k(t)$. When we insert this representation into the differential equation, this yields

$$-y_{k,tt} + \left(\lambda_k^4 + \frac{1}{\nu}\right) y_k = f_k \tag{8}$$

for every k with the (time-dependent) Fourier coefficients $f_k = \int_\Omega f\varphi_k \,\mathrm{d}\,\omega$ of the right hand side and initial and terminal conditions

$$\begin{aligned} y_k(0) &= 0,\\ y_{k,t}(T) + \lambda_k^2 y_k(T) &= 0. \end{aligned}$$

The weak form of this problem for every y_k is

$$\begin{aligned} \int_0^T f_k z \,\mathrm{d}\,t = \int_0^T y_{kt}z_t + \left(\lambda_k^4 + \frac{1}{\nu}\right) y_k z \,\mathrm{d}\,t + \lambda_k^2 y_k(T)z(T) =: a_k(y_k, z),\\ \forall z \in H^1(0,T) : z(0) = 0. \end{aligned}$$

If we use y_k as test function and the Cauchy-Schwarz inequality we have the estimate

$$\begin{aligned} a_k(y_k, y_k) = \|y_{k,t}\|^2_{L^2(Q)} + \left(\lambda_k^4 + \frac{1}{\nu}\right)\|y_k\|^2_{L^2(Q)} + \lambda_k^2 y_k^2(T) = \int_0^T f_k y_k \,\mathrm{d}\,t\\ \le \|f_k\|_{L^2(Q)}\|y_k\|_{L^2(Q)}. \end{aligned} \tag{9}$$

This yields directly

$$\left(\lambda_k^4 + \frac{1}{\nu}\right)\|y_k\|_{L^2(Q)} \lesssim \|f_k\|_{L^2(\Omega)}. \tag{10}$$

With (9) and (10) we can also estimate

$$\|y_{k,t}\|^2_{L^2(Q)} \le \|f_k\|_{L^2(Q)}\|y_k\|_{L^2(Q)} \le \|f_k\|^2_{L^2(Q)}\frac{1}{\lambda_k^4 + \frac{1}{\nu}}.$$

Taking the square root gives

$$\lambda_k^2 \|y_{k,t}\|_{L^2(Q)} \lesssim \|f_k\|_{L^2(Q)} .$$

Further we have an estimate for $y_{k,tt}$ with (8), the triangle inequality and (10)

$$\begin{aligned}\|y_{k,tt}\|_{L^2(Q)} &\leq \|f_k\|_{L^2(Q)} + \left(\lambda_k^4 + \frac{1}{\nu}\right) \|y_k\|_{L^2(Q)} \\ &\lesssim \|f_k\|_{L^2(Q)} .\end{aligned}$$

Altogether the estimate

$$\|y_{k,tt}\|_{L^2(Q)}^2 + \left\|\lambda_k^4 y_{k,t}\right\|_{L^2(Q)}^2 + \left\|\lambda_k^8 y_k\right\|_{L^2(Q)}^2 \lesssim \|f_k\|_{L^2(Q)}^2$$

is established.

Summing up over k implies $y \in H^2(0,T;L^2(\Omega)) \cap H^1(0,T;D(A)) \cap L^2(0,T;D(A^2))$ and the bound

$$\|y_{tt}\|_{L^2(Q)}^2 + \|Ay_t\|_{L^2(Q)}^2 + \left\|A^2 y\right\|_{L^2(Q)}^2 \lesssim \|f\|_{L^2(Q)}^2 ,$$

which is the desired estimate. □

4 $H^{(2,1)}(Q)$-conforming finite element methods

For the discretization we use a space-time finite element method with space-time finite element mesh $\mathcal{T}_{h,\tau}$ with rectangular elements θ. We use a tensor product ansatz with a continuously linear, quadratic or cubic Lagrangian ansatz in time and a continuous differentiable cubic Hermite ansatz in space for which the discretization parameter for the spatial discretization h and temporal discretization τ can be chosen independently. This approximation space is denoted by $V_{h\tau}$. We approximate the variational problem (7) in $V_{h\tau}$ by

$$a(y_{h\tau}, \varphi_{h\tau}) = (f, \varphi_{h\tau}), \qquad \forall \varphi_{h\tau} \in V_{h\tau}. \tag{11}$$

We define the interpolation operator on the finite element mesh with the corners (x_i, t_j) by

$$\begin{aligned} I_{h\tau}^k : H^{(3,2)}(Q) &\to \mathcal{C}^1(0,X) \otimes \mathcal{C}^0(0,T), \\ I_{h\tau}^k w\Big|_\theta &\in \mathbb{P}^3 \otimes \mathbb{P}^k, \\ I_{h\tau}^k w\left(x_i, t_j + \frac{m}{k}\tau\right) &= w\left(x_i, t_j + \frac{m}{k}\tau\right), \qquad \text{for } m = 0, \dots, k, \\ D^{(1,0)} I_{h\tau}^k w\left(x_i, t_j + \frac{m}{k}\tau\right) &= D^{(1,0)} w\left(x_i, t_j + \frac{m}{k}\tau\right), \qquad \text{for } m = 0, \dots, k, \end{aligned}$$

where $k = 1$, 2 or 3.

Theorem 4.1 (Interpolation error estimate)**.** *For a function $y \in H^A(Q) \cap H^{(3,2)}(Q)$ with the multi-index set $A = \{(0,0), (0,k+1), (i,1), (4,0), (2,j)\}$ with $i \in \{1,\dots,4\}$ and $j \in \{1,\dots,k\}$ the interpolation error can be estimated by*

$$\begin{aligned}\left\|y - I_{h\tau}^k y\right\|_{H^{(2,1)}(Q)} &\lesssim h^i \left\|D^{(i,1)} y\right\|_{L^2(Q)} + \tau^k \left\|D^{(0,k+1)} y\right\|_{L^2(Q)} \\ &\quad + \tau^j \left\|D^{(2,j)} y\right\|_{L^2(Q)} + h^2 \left\|D^{(4,0)} y\right\|_{L^2(Q)} .\end{aligned}$$

We use this interpolation error estimate for the estimates of the finite element error. As the proof is technical and rather long, we give it in the following section. Before, we apply it to the discretization error estimates for the finite element solution $y_{h\tau}$.

Theorem 4.2 (Error estimate in the energy norm)**.** *If for the exact solution of (7) $y \in H^A(Q) \cap H^{(3,2)}(Q)$ with the multi-index set $A = \{(0,0),(0,k+1),(i,1),(4,0),(2,j)\}$ with $i \in \{1,\dots,4\}$ and $j \in \{1,\dots,k\}$ holds, the approximation error for finite element solution $y_{h\tau}$ with an ansatz of polynomial degree k in time can be bounded by*

$$\|y - y_{h\tau}\|_{H^{(2,1)}(Q)} \lesssim h^i \left\| D^{(i,1)} y \right\|_{L^2(Q)} + \tau^k \left\| D^{(0,k+1)} y \right\|_{L^2(Q)} + \tau^j \left\| D^{(2,j)} y \right\|_{L^2(Q)} + h^2 \left\| D^{(4,0)} y \right\|_{L^2(Q)}.$$

Proof. As the bilinear form $a(\cdot,\cdot)$ is V-elliptic, $V_{h\tau} \subseteq V$ and the functions y and $y_{h\tau}$ are the solutions of the variational problems (7) and (11), we can apply the usual Céa Lemma (see e.g. [11, Theorem 2.4.1]) to get

$$\|y - y_{h\tau}\|_{H^{(2,1)}(Q)} \lesssim \inf_{v_h \in V_h} \|y - v_h\|_{H^{(2,1)}(Q)}.$$

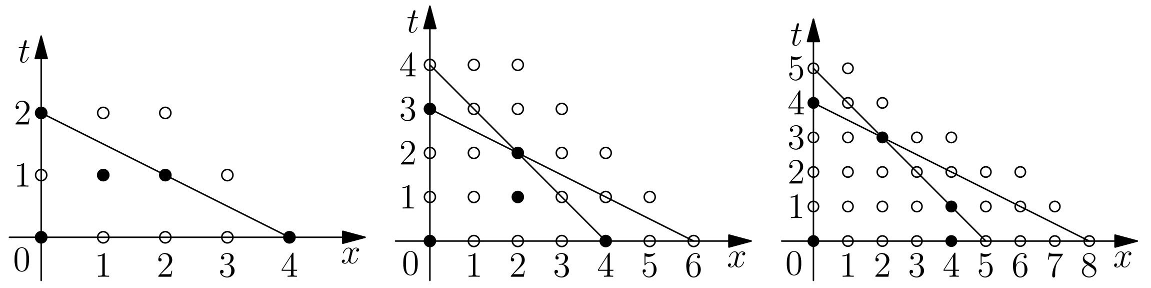
To bound the best approximation error we can use the interpolation error estimate of Theorem 4.1 and the proof is done. □

Remark 4.3. *In Theorem 4.2 the regularity assumption is given in terms of Sobolev spaces with respect to a multi-index set. We discuss now, for which Sobolev spaces $H^{(r,s)}(Q)$ the regularity assumptions are fulfilled in the most interesting case $i = 2$ and $j = k$, in which the Theorem provides an error estimate of order 2 with respect to the spatial discretization and of order k with respect to the temporal discretization. Our tool for this discussion is Theorem 3.3. In Figure 1 we have illustrated, which mixed derivatives are bounded for certain Sobolev spaces $H^{(r,s)}(Q)$:*

1. *For $k = 1$ the multi-index set is $A = \{(0,0),(0,2),(4,0),(1,1),(2,1)\}$. These derivatives exists for functions in the space $H^{(4,2)}(Q)$.*
2. *For $k = 2$ the multi-index set is $A = \{(0,0),(0,3),(4,0),(2,1),(2,2)\}$. These derivatives exist for functions in the spaces $H^{(4,4)}(Q)$ or $H^{(6,3)}(Q)$.*
3. *For $k = 3$ the multi-index set is $A = \{(0,0),(0,4),(4,0),(4,1),(2,3)\}$. These derivatives exist for functions in the spaces $H^{(5,5)}(Q)$ or $H^{(8,4)}(Q)$.*

Theorem 4.4 ($L^2(Q)$-error estimate with the Aubin-Nitsche trick)**.** *For a solution y, which fulfills the regularity assumptions of Theorem 4.2 for $i = 2$ and $j = k$, the error in the $L^2(Q)$-norm can be estimated by*

$$\|y - y_{h\tau}\|_{L^2(Q)} \lesssim (h^2 + \tau^k)(h^2 + \tau) \|y\|_{H^A(Q)}.$$



(a) Multi-indices for the temporal discretization with polynomials of degree $k = 1$.

(b) Multi-indices for the temporal discretization with polynomials of degree $k = 2$.

(c) Multi-indices for the temporal discretization with polynomials of degree $k = 3$.

Figure 1: With these figures we illustrate the multi-indices which are needed for the estimates in Theorem 4.2 for the case $i = 2$ and $j = k$ in black. For better overview additional multi-indices are added as circles. For the Sobolev space $H^{(r,s)}(Q)$ all the derivatives corresponding to the multi-indices below the line, which connects $(0, s)$ and $(r, 0)$ are $L^2(Q)$ functions, according to Theorem 3.3.

Proof. The adaption of the usual proof with the Aubin-Nitsche trick (see e.g. [11, Theorem 3.2.4 and Theorem 3.2.5]) is straight forward. The a priori estimate, which is needed for this proof, was established in Theorem 3.7. □

Remark 4.5. *The error estimates in Theorems 4.2 and 4.4 imply the following choice of discretization parameters:*

1. *For the linear ansatz in time the error estimates of Theorem 4.2 and Theorem 4.4 imply a choice of $\tau \sim h^2$ for balancing the discretization error in the energy-norm and the $L^2(Q)$-norm.*

   With this choice of discretization parameters the error in the energy norm behaves asymptotically like $\mathcal{O}(h^2)$ and the error in the $L^2(Q)$-norm behaves asymptotically like $\mathcal{O}(h^4)$. With respect to the number of unknowns N the errors are of order $N^{-2/3}$ and $N^{-4/3}$, respectively.

2. *For the quadratic ansatz in time the error estimate of Theorem 4.2 implies a choice of $\tau \sim h$ for balancing the discretization error in the energy-norm.*

   This choice of the discretization parameters leads to an asymptotic error behavior of $\mathcal{O}(h^2)$ in the energy norm and $\mathcal{O}(\tau h^2 + \tau^3 + \tau^2 h^2 + h^4) \sim \mathcal{O}(h^3)$ in the $L^2(Q)$-norm. With respect to the number of unknowns N the errors are of order N^{-1} and $N^{-3/2}$.

3. *For the quadratic ansatz in time the error estimate of Theorem 4.4 implies at least a choice of $\tau \sim h^2$ to get an error estimate of order h^4 in the $L^2(Q)$-norm. So the asymptotic error is like $\mathcal{O}(h^2)$ in the energy norm and $\mathcal{O}(h^4)$ in the $L^2(Q)$-norm. With respect to the number of unknowns N the errors are of order $N^{-2/3}$ and $N^{-4/3}$, i.e. worse in comparison with the choice $\tau \sim h$.*

4. *For the cubic ansatz in time the error estimates of Theorem 4.2 implies a choice of $\tau \sim h^{2/3}$ for second oder convergence in the energy norm. This choice of the discretization*

parameters leads to an asymptotic error behavior of $\mathcal{O}(h^2)$ in the energy norm and $\mathcal{O}(h^{8/3})$ in the $L^2(Q)$-norm. With respect to the number of unknowns N the errors are of order $N^{-6/5}$ and $N^{-24/15}$.

5 Proof of the interpolation error estimate

We split the proof of the Theorem 4.2 into three lemmas. We will prove an estimate on the reference element $R = (0,1)^2$ and get the convergence order by transformation to the world element.

Lemma 5.1. *Let $y \in H^A(Q) \cap H^{(3,2)}(Q)$ with the multi-index set $A = \{(0,k+1),\ (i,1)\}$ with $i \in \{1,\dots,4\}$. Then the time derivative of the interpolation error on one element θ can be bounded by*

$$\left\| D^{(0,1)}\left(y - I_{h\tau}^k y\right)\right\|_{L^2(\theta)} \lesssim \tau^k \left\| D^{(0,k+1)} y\right\|_{L^2(\theta)} + h^i \left\| D^{(i,1)} y\right\|_{L^2(\theta)}.$$

Proof. For the proof we use the standard transfer to the reference element $R = (0,1)^2$ and follow the ideas of [32, Section 2.1] On R, we denote all quantities by $\hat{\cdot}$. We start with

$$\left\| D^{(0,1)}\left(y - I_{h\tau}^k y\right)\right\|_{L^2(\theta)}^2 = \int_R \frac{\tau h}{\tau^2} \left(\hat{D}^{(0,1)}\left(\hat{y} - \hat{I}_{h\tau}^k \hat{y}\right)\right)^2 \mathrm{d}\,\hat{\omega}.$$

Next we introduce the temporal interpolation

$$\begin{aligned} \hat{I}_\tau^k : H^{(3,2)}(R) &\to H^3\left((0,1)\right) \otimes \mathcal{C}^0(0,T), \\ \hat{I}_\tau^k \hat{y} &\in H^3\left((0,1)\right) \otimes \mathbb{P}^k, \\ \hat{I}_\tau^k \hat{y}\left(\hat{x}, \frac{m}{k}\right) &= \hat{y}\left(\hat{x}, \frac{m}{k}\right), \qquad \text{for } m = 0,\dots,k, \end{aligned}$$

that is well-defined for almost all $\hat{x} \in (0,1)$. By adding and subtracting this function and the triangle inequality we have to estimate

$$\begin{aligned} \left\| D^{(0,1)}\left(y - I_{h\tau}^k y\right)\right\|_{L^2(\theta)} \le{}& \sqrt{\frac{h}{\tau}} \left(\int_R \left(\hat{D}^{(0,1)}\left(\hat{y} - \hat{I}_\tau^k \hat{y}\right)\right)^2 \mathrm{d}\,\hat{\omega}\right)^{1/2} \\ &+ \sqrt{\frac{h}{\tau}} \left(\int_R \left(\hat{D}^{(0,1)}\left(\hat{I}_\tau^k \hat{y} - \hat{I}_{h\tau}^k \hat{y}\right)\right)^2 \mathrm{d}\,\hat{\omega}\right)^{1/2}. \end{aligned} \tag{12}$$

For some fixed $\hat{x}^* \in (0,1)$ we can use the standard one dimensional interpolation result

$$\int_0^1 \left(\hat{D}^{(0,1)}\left(\hat{y} - \hat{I}_\tau^k \hat{y}\right)\left(\hat{x}^*, \hat{t}\right)\right)^2 \mathrm{d}\,\hat{t} \lesssim \int_0^1 \left(\hat{D}^{(0,k+1)} \hat{y}\left(\hat{x}^*, \hat{t}\right)\right)^2 \mathrm{d}\,\hat{t},$$

which yields

$$\int_R \left(\hat{D}^{(0,1)}\left(\hat{y} - \hat{I}_\tau^k \hat{y}\right)\right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(0,k+1)} \hat{y}\right)^2 \mathrm{d}\,\hat{\omega}.$$

The other integral in the estimate (12) is also an one-dimensional interpolation error as $\hat{I}^k_{h\tau}\hat{y}$ is an interpolant of $\hat{I}^k_\tau\hat{y}$. The application of the standard one dimensional interpolation result yields for $i = 1, \ldots, 4$ the estimate

$$\int_R \left(\hat{D}^{(0,1)}\left(\hat{I}^k_\tau\hat{y} - \hat{I}^k_{h\tau}\hat{y}\right)\right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(i,1)}\hat{I}^k_\tau\hat{y}\right)^2 \mathrm{d}\,\hat{\omega}.$$

To end the proof of this lemma we need finally to prove the estimate

$$\int_R \left(\hat{D}^{(i,1)}\hat{I}^k_\tau\hat{y}\right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(i,1)}\hat{y}\right)^2 \mathrm{d}\,\hat{\omega}. \tag{13}$$

With the nodal Lagrangian interpolation basis $\varphi_i(t) \in \mathbb{P}^k$, $i = 0, \ldots, k$ with

$$\varphi_i\left(\frac{i}{k}\right) = \delta_{ik}$$

the action of the temporal interpolation operator $\hat{I}^k_\tau$ can be described by

$$\hat{I}^k_\tau\hat{w}(\hat{x},\hat{t}) = \sum_{i=0}^{k} \hat{w}(\hat{x},\hat{t}_i)\varphi_i(\hat{t}).$$

With the basis $\chi_i = \sum_{j=0}^{i}\varphi_j$ (see also [1, Section 5]) the interpolation can be written as

$$\begin{aligned}\hat{I}^k_\tau\hat{w} &= \sum_{i=0}^{k-1}\left(\hat{w}(\hat{x},\hat{t}_i) - \hat{w}(\hat{x},\hat{t}_{i+1})\right)\chi_i(\hat{t}) + \hat{w}(x,\hat{t}_k)\\ &= -\sum_{i=0}^{k-1}\left(\int_{t_i}^{t_{i+1}} \hat{D}^{(0,1)}\hat{w}(\hat{x},\hat{s})\,\mathrm{d}\,\hat{s}\right)\chi_i(\hat{t}) + \hat{w}(\hat{x},\hat{t}_k).\end{aligned}$$

Therefore the first derivative of the interpolant is given by

$$\hat{D}^{(0,1)}\hat{I}^k_\tau\hat{w} = -\sum_{i=0}^{k-1}\left(\int_{t_i}^{t_{i+1}} \hat{D}^{(0,1)}\hat{w}\,\mathrm{d}\,\hat{s}\right)\chi_i'(\hat{t})$$

and the $L^2(R)$-norm of this derivative can be estimated as

$$\begin{aligned}\left\|\hat{D}^{(0,1)}\hat{I}^k_\tau\hat{w}\right\|_{L^2(R)} &\leq \sum_{i=0}^{k-1}\left\|\hat{D}^{(0,1)}\hat{w}\right\|_{L^1(t_i,t_{i+1};L^2(0,1))}\left\|\chi_i'(\hat{t})\right\|_{L^2((0,1))}\\ &\lesssim \left\|\hat{D}^{(0,1)}\hat{w}\right\|_{L^1(0,1;L^2(0,1))} \lesssim \left\|\hat{D}^{(0,1)}\hat{w}\right\|_{L^2(R)}.\end{aligned}$$

Choosing $\hat{w} = \hat{D}^{(i,0)}\hat{y}$ yields the estimate (13). Alltogether we have proven the estimate

$$\left\|D^{(0,1)}\left(y - I^k_{h\tau}y\right)\right\|^2_{L^2(\theta)} \leq \frac{\tau h}{\tau^2}\int_R \left(\hat{D}^{(0,k+1)}\hat{y}\right)^2 \mathrm{d}\,\hat{\omega} + \frac{\tau h}{\tau^2}\int_R \left(\hat{D}^{(i,1)}\hat{y}\right)^2 \mathrm{d}\,\hat{\omega}.$$

Transfering the integrals back on the element θ yields the result. □

Lemma 5.2. *Assume that $y \in H^A(Q) \cap H^{(3,2)}(Q)$ with the multi-index set $A = \{(4,0),(2,j)\}$ with $j \in \{1,\dots,k\}$. Then the second spatial derivative of the interpolation error on the element θ can be bounded by*

$$\left\| D^{(2,0)} \left(y - I_{h\tau}^k y \right) \right\|_{L^2(\theta)} \lesssim \tau^j \left\| D^{(2,j)} y \right\|_{L^2(\theta)} + h^2 \left\| D^{(4,0)} y \right\|_{L^2(\theta)}.$$

Proof. As in the proof of the previous lemma we follow the ideas of [32, Section 2.1] and transfer the integral onto the reference element, where we denote quantities on the reference element by $\hat{\cdot}$. This yields

$$\left\| D^{(2,0)} \left(y - I_{h\tau}^k y \right) \right\|_{L^2(\theta)}^2 = \int_R \frac{\tau h}{h^4} \left(\hat{D}^{(2,0)} \left(\hat{y} - \hat{I}_{h\tau}^k \hat{y} \right) \right)^2 \mathrm{d}\,\hat{\omega}.$$

Next we introduce the spatial interpolation

$$\begin{aligned} \hat{I}_h : H^{(3,2)}(R) &\to \mathcal{C}^1((0,1)) \otimes H^2(0,T), \\ \hat{I}_h \hat{y} &\in \mathbb{P}^3 \otimes H^2((0,T)), \\ D^{(i,0)} \hat{I}_h \hat{y}\left(m, \hat{t}\right) &= D^{(i,0)} \hat{y}\left(m, \hat{t}\right), \qquad \text{for } i = 0,1 \text{ and } m = 0,1. \end{aligned}$$

By adding and subtracting this interpolant and the triangle inequality we split the integral into

$$\begin{aligned} \left\| \hat{D}^{(2,0)} \left(y - I_{h\tau}^k y \right) \right\|_{L^2(\theta)}^2 \lesssim{}& \frac{\tau h}{h^4} \int_R \left(\hat{D}^{(2,0)} \left(\hat{y} - \hat{I}_h \hat{y} \right) \right)^2 \mathrm{d}\,\hat{\omega} \\ &+ \frac{\tau h}{h^4} \int_R \left(\hat{D}^{(2,0)} \left(\hat{I}_h \hat{y} - \hat{I}_{h\tau}^k \hat{y} \right) \right)^2 \mathrm{d}\,\hat{\omega}. \end{aligned} \tag{14}$$

As in the previous lemma the first integral can be estimated as an one dimensional interpolation error, which yields

$$\int_R \left(\hat{D}^{(2,0)} \left(\hat{y} - \hat{I}_h \hat{y} \right) \right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(4,0)} \hat{y} \right)^2 \mathrm{d}\,\hat{\omega}.$$

Again the other integral in the estimate (14) is also an one-dimensional interpolation error as $\hat{I}_{h\tau}^k \hat{y}$ is an interpolant of $\hat{I}_h \hat{y}$. The application of the standard one dimensional interpolation result yields with $j = 1,\dots,k$ the estimate

$$\int_R \left(\hat{D}^{(2,0)} \left(\hat{I}_h \hat{y} - \hat{I}_{h\tau}^k \hat{y} \right) \right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(2,j)} \hat{I}_h \hat{y} \right)^2 \mathrm{d}\,\hat{\omega}.$$

To end the proof of this Lemma we need finally to prove the estimate

$$\int_R \left(\hat{D}^{(2,j)} \hat{I}_h \hat{y} \right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(2,j)} \hat{y} \right)^2 \mathrm{d}\,\hat{\omega}.$$

To this end let

$$f(\hat{x}) = \hat{D}^{(0,j)} \hat{y} \Big|_{\hat{t}=\hat{t}^*}, \qquad g(\hat{x}) = \hat{D}^{(0,j)} \hat{I}_h \hat{y} \Big|_{\hat{t}=\hat{t}^*} = \hat{I}_h \hat{D}^{(0,j)} \hat{y} \Big|_{\hat{t}=\hat{t}^*},$$

for some fixed $\hat{t}^*$.

As the solution of the variational problem

$$\min_{p\in H^2(0,T)} \int_0^1 \left(\frac{\mathrm{d}^2}{\mathrm{d}\,\hat{x}^2} p(\hat{x})\right)^2 \mathrm{d}\,\hat{x}$$
$$\text{s. th. } p(0)=a,\ p(1)=b,\ \frac{\mathrm{d}}{\mathrm{d}\,x}p(0)=c,\ \frac{\mathrm{d}}{\mathrm{d}\,x}p(1)=d,$$

is given by the Hermite interpolant (using that the corresponding Euler-Lagrange equation is $p_{xxxx}=0$) we have

$$\int_0^1 \left(\frac{\mathrm{d}^2}{\mathrm{d}\,\hat{x}^2} g(\hat{x})\right)^2 \mathrm{d}\,\hat{x} \le \int_0^1 \left(\frac{\mathrm{d}^2}{\mathrm{d}\,\hat{x}^2} f(\hat{x})\right)^2 \mathrm{d}\,\hat{x}$$

Returning to the definition of the functions f and g and recalling that $\hat{t}^*$ was chosen arbitarily, the estimate holds for (almost) all $\hat{t}\in(0,1)$ and therefore we have

$$\int_R \left(\hat{D}^{(2,j)}\hat{I}_h\hat{y}\right)^2 \mathrm{d}\,\hat{\omega} \lesssim \int_R \left(\hat{D}^{(2,j)}\hat{y}\right)^2 \mathrm{d}\,\hat{\omega}.$$

Altogether we have proven the estimate

$$\left\|D^{(2,0)}\left(y - I^k_{h\tau}y\right)\right\|^2_{L^2(\theta)} \lesssim \frac{\tau h}{h^4}\int_R \left(\hat{D}^{(4,0)}\hat{y}\right)^2 \mathrm{d}\,\hat{\omega} + \frac{\tau h}{h^4}\int_R \left(\hat{D}^{(2,j)}\hat{y}\right)^2 \mathrm{d}\,\hat{\omega}.$$

Transfering the integrals back on the element θ yields the result. □

For the interpolation error estimate in the $L^2(\theta)$-norm we need a result about equivalent norms in $H^{(r,s)}(Q)$, which is provided in the next Lemma.

Lemma 5.3. *Assume that there exist $r\cdot s$ linear and bounded functionals $l_i, i=1,\cdots,r\cdot s$ in $H^{(r,s)}(Q)$ which do not vanish at the same time for any non-zero polynomial of degree at most $r-1$ in x and at most $s-1$ in t. Then the norm*

$$\|y\|^2_* = |y|^2_{H^{(r,s)}(Q)} + \sum_{i=1}^{r\cdot s} |l_i(y)|^2$$

is an equivalent norm on $H^{(r,s)}(Q)$.

Proof. As recommended in [31, Proof of Theorem 1] we follow the ideas of the proof of the isotropic case, which can be found e.g. in [27, Theorem 4.5.1] and [33, Theorem IV.114.3.]. The inequality

$$\|y\|_* \lesssim |||y|||_{H^{(r,s)}(Q)}$$

is clear by the definition of $\|\cdot\|_*$ as the functionals l_i are bounded in $H^{(r,s)}(Q)$.

We prove the inequality

$$|||y|||_{H^{(r,s)}(Q)} \lesssim \|y\|_* \tag{15}$$

by contradiction. Therefore we assume, that there is a sequence $\{v_n\}_{n=0}^\infty \in H^{(r,s)}(Q)$ with

$$|||v_n|||_{H^{(r,s)}(Q)} > n\,\|v_n\|_*\,. \tag{16}$$

Obviuously we have $v_n \neq 0$ and without loss of generality we can assume that the members of this sequence are normed, i.e. $|||v_n|||_{H^{(r,s)}(Q)} = 1$. The sequence v_n is bounded in $H^{(r,s)}(Q)$ and therefore compact in $L^2(Q)$, as $H^{(2,1)}(Q) \hookrightarrow H^1(Q) \hookrightarrow L^2(Q)$ and the embedding $H^1(Q) \hookrightarrow L^2(Q)$ is compact. So there is a convergent subsequence with limit v. We denote this subsequence again by v_n. Therefore we have

$$\|v_n - v\|_{L^2(Q)} \to 0.$$

With the assumption (16) we have

$$\|v_n\|_* < \frac{1}{n} \, |||v_n|||_{H^{(r,s)}(Q)}$$

and therefore $\|v_n\|_* \to 0$, which also implies, by the definition of the norm $\|\cdot\|_*$, that $|v_n|_{H^{(r,s)}(Q)} \to 0$. Therefore we have $D^{(r,0)}v = 0$ and $D^{(0,s)}v = 0$ in the sense of $L^2(Q)$. This implies also $D^{(r+n,0)}v = 0$ and $D^{(0,s+n)}v = 0$ for all $n \in \mathbb{N}$. By choosing n large enough and the Sobolev embedding theorem the function v is a (r,s)-times continuously differentiable function.

This implies that the limit v is a polynomial of degree $s-1$ in t and degreee $r-1$ in x. As the function v is the limit of the sequence $\{v_n\}$ it follows that $\|v\|_* = 0$ and therefore

$$\sum_{i=1}^{r \cdot s} |l_i(v)| = 0.$$

As the functionals l_i do not vanish for any non-zero polynomial of degree $s-1$ in t and degreee $r-1$ in x, this implies $v \equiv 0$ which is a contradiction to the assumption $\|v_n\|_{H^{(2,1)}(Q)} = 1$. $\square$

Lemma 5.4. *Assume that $y \in H^A(Q)$ with the multi-index set*

$$A = \{(0,k+1),\ (j,1),\ (4,0),\ (2,i)\} \qquad \text{with } j \in \{1,\dots,4\} \text{ and } i \in \{1,\dots,k\}.$$

Then the interpolation error on an element can be bounded by

$$\left\| y - I^k_{h\tau} y \right\|_{L^2(\theta)} \lesssim \tau^{k+1} \left\| D^{(0,k+1)} y \right\|_{L^2(\theta)} + h^4 \left\| D^{(4,0)} y \right\|_{L^2(\theta)}.$$

Proof. As in the Lemmas 5.1 and 5.2 we transfer the error to the reference element. On the reference element we can estimate the $L^2(R)$-norm by the stronger $H^{(4,k+1)}(R)$-norm, and by using Lemma 5.3, we get

$$\begin{aligned}
\left\| \hat{y} - \hat{I}^k_{h\tau} \hat{y} \right\|_{L^2(R)} &\leq \left\| \hat{y} - \hat{I}^k_{h\tau} \hat{y} \right\|_{H^{(4,k+1)}(R)} \\
&\lesssim \left| \hat{y} - \hat{I}^k_{h\tau} \hat{y} \right|_{H^{(4,k+1)}(R)} + \sum_{i=1}^{4\cdot(k+1)} \left| l_i(\hat{y} - \hat{I}^k_{h\tau} \hat{y}) \right| \\
&= |\hat{y}|_{H^{(4,k+1)}(R)} + \sum_{i=1}^{4\cdot(k+1)} \left| l_i(\hat{y} - \hat{I}^k_{h\tau} \hat{y}) \right|.
\end{aligned}$$

For the linear functionals l_i we choose

$$l_i(y) = y\left(0, \frac{i-1}{k}\right), \qquad \text{for } i = 1, \dots, k+1,$$

$$l_i(y) = y\left(1, \frac{i-(k+2)}{k}\right), \qquad \text{for } i = k+2, \dots, 2(k+1),$$

$$l_i(y) = D^{(1,0)}y\left(0, \frac{i-2(k+1)-1}{k}\right), \qquad \text{for } i = 2(k+1)+1, \dots, 3(k+1),$$

$$l_i(y) = D^{(1,0)}y\left(1, \frac{i-3(k+1)-1}{k}\right), \qquad \text{for } i = 3(k+1)+1, \dots, 4(k+1).$$

By the uniqueness of the polynomial interpolation it is clear, that the condition on the functionals of Lemma 5.3 is fulfuilled. With the properties of the interpolation operator $I_{h\tau}^k$ we see that

$$\sum_{i=1}^{4\cdot(k+1)} \left| l_i(\hat{y} - \hat{I}_{h\tau}^k \hat{y}) \right| = 0.$$

By transfering back to the element θ the proof is finished. □

So we have proven all results which we need to prove the interpolation error estimate of Theorem 4.1.

Proof of Theorem 4.1. The interpolation error on every element is bounded with Lemma 5.1, Lemma 5.2 and Lemma 5.4. For an interpolation error estimate on the whole domain we split the integration over the domain to the integration over the elements and sum up. □

6 Numerical example

For the observation of the convergence rates we consider the problem

$$\begin{aligned} -y_{tt} + y_{xxxx} + y &= f, && \text{in } (0,1)^2,\\ y &= 0, && \text{on } \{0\} \times (0,1),\\ y_{xx} &= 0, && \text{on } \{0\} \times (0,1),\\ y_x &= 0, && \text{on } \{1\} \times (0,1),\\ y_{xxx} &= 0, && \text{on } \{1\} \times (0,1),\\ y &= 0, && \text{on } (0,1) \times \{0\},\\ y_t - y_{xx} &= 0, && \text{on } (0,1) \times \{T\}, \end{aligned}$$

where f is chosen so that

$$y = (t-1)^2\, tx^3\, (x-1)^4 \tag{17}$$

is the exact solution of the boundary value problem. For the tensor product finite elements with the linear ansatz in time we choose the step size combination $\tau = h^2$. Therefore we observe second order convergence in h in the semi-norms of $H^{(2,0)}(Q)$ and $H^{(0,1)}(Q)$ and fourth order convergence in the $L^2(Q)$-norm, see Figure 2.

The numerical examples were implemented in Matlab. For the (faster) assemblation of the finite element matrices ideas of Davis [13] and Funken, Praetorius and Wissgott [16] were used.

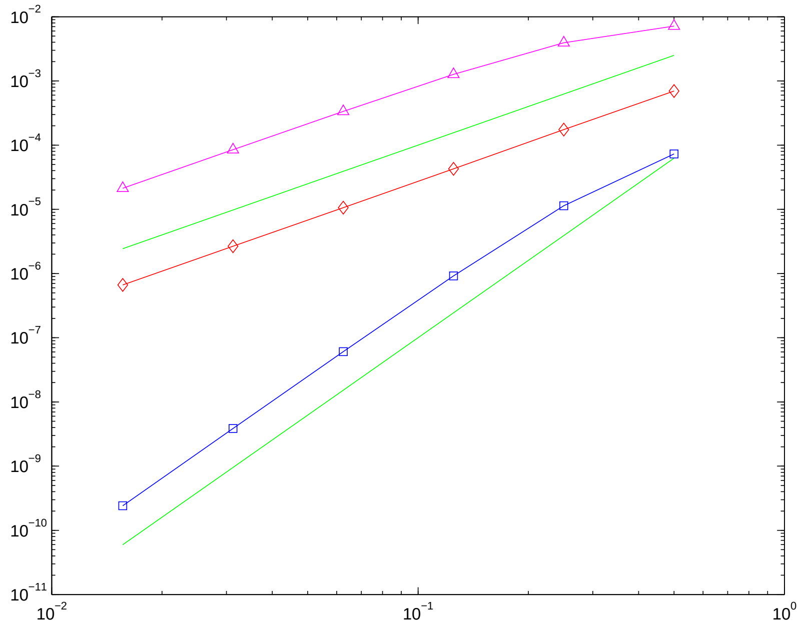


Figure 2: Convergence of the tensor product elements with linear Lagrange ansatz in time and cubic Hermite in space for the example with exact solution (17). The $L^2(Q)$-norm of the error is plotted in blue with squares, the $H^{(0,1)}(Q)$-semi-norm in red with diamonds and the $H^{(2,0)}(Q)$-semi-norm in magenta with triangles. The lines in green without any markers indicate h^2 and h^4.

7 Conclusions and outlook

In this paper we have established an a priori bound for a $H^{(2,1)}$-elliptic equation and a priori error bounds for tensor product finite element discretizations of semi-elliptic equations. A numerical example confirms the expected convergence rates.

An application of such equations is given by optimal control problems with parabolic partial differential equations. For this case the question is open whether discontinuous Galerkin time stepping schemes [5, 25] and continuous Galerkin time stepping [2, 5, 26] for the optimality system, which are commonly in use for these problems, can be understood as mixed finite element approximations of the $H^{(2,1)}(Q)$-elliptic equation.

Acknowledgements

The work was partially supported by the DFG priority program SPP 1253.

References

[1] Thomas Apel. interpolation of non-smooth functions on anisotropic finite element meshes. *Mathematical Modelling and Numerical Analysis*, 33(6):1149–1185, 1999.

[2] Thomas Apel and Thomas G. Flaig. Crank-Nicolson schemes for optimal control problems with evolution equations. *SIAM Journal on Numerical Analysis*, 50(3):1484–1512, 2012.

[3] Ralph A. Artino. On semielliptic boundary value problems. *Journal of Mathematical Analysis and Applications*, 42:610–626, 1973.

[4] Ralph A. Artino and J. Barros-Neto. Regular semielliptic boundary value problems. *Journal of Mathematical Analysis and Applications*, 61:40–57, 1977.

[5] Roland Becker, Dominik Meidner, and Boris Vexler. Efficient Numerical Solution of Parabolic Optimization Problems by Finite Element Methods. *Optimization Methods and Software*, 22(5):813 – 833, 2007.

[6] M. H. Berger. Finite element analysis of flow in a gas-filled rotating annulus. *International Journal for Numerical Methods in Fluids*, 7:215–231, 1987.

[7] Oleg Vladimirovich Besov, Valentin Petrovich Il'in, and Sergeĭ Michaĭlovič Nikol'skiĭ. *Integral representations of functions and imbedding Theorems - Volume I.* V. H. Winston & Sons, Washington, D.C.; Halsted Press [John Wiley & Sons], New York, 1978.

[8] Alfio Borzì and Volker Schulz. Multigrid methods for PDE optimization. *SIAM Review*, 51(2):361–395, 2009.

[9] Alfio Borzì and Volker Schulz. *Computational Optimization of Systems Governed by Partial Differential Equations.* SIAM, Philadelphia, 2012.

[10] Guido Büttner. *Ein Mehrgitterverfahren zur optimalen Steuerung parabolischer Probleme.* PhD thesis, TU Berlin, 2004.

[11] Philippe G. Ciarlet. *The finite element method for elliptic problems.* North-Holland Publishing Company, Amsterdam, 1979.

[12] Robert Dautray and Jacques-Louis Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution problems I.* Springer, Berlin, 1992.

[13] Tim Davis. Creating sparse finite-element matrices in MATLAB. Guest blog in: Loren Shure: Loren on the Art of MATLAB, `http://blogs.mathworks.com/loren/2007/03/01/creating-sparse-finite-element-matrices-in-matlab`, 2007.

[14] J. F. Eastham and J. S. Peterson. The finite element method in anisotropic Sobolev spaces. *Computers and Mathematics with Applications*, 47:1775–1786, 2004.

[15] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* AMS, Providence, Rhode Island, 1998.

[16] Stefan Funken, Dirk Praetorius, and Philipp Wissgott. Efficient implementation of adaptife P1-FEM in Matlab. *Computational Methods in Applied Mathematics*, 11(4):460–490, 2011.

[17] Wei Gong, Michael Hinze, and Zhaojie Zhou. Space-time finite element approximation of parabolic optimal control problems. DFG Priority Program 1253, Preprint-Number SPP1253-126, 2011.

[18] Max D. Gunzburger and Houston G. Wood III. A finite element method for the Onsager pancake equation. *Computer Methods in Applied Mechanics and Engineering*, 31:43–59, 1982.

[19] Max D. Gunzburger and Angela Kunoth. Space-time adaptive wavelet method for optimal control problems constrained by parabolic evolution equations. *SIAM J. Control Optim.*, 49(3):1150–1170, 2011.

[20] G. N. Hile, C. P. Mawata, and Chiping Zhou. A priori bounds for semielliptic operators. *Journal of Differential Equations*, 176:29–64, 2001.

[21] O. A. Ladyzhenskaya, V. A. Solonnikov, and N.N. Ural'ceva. *Linear and quasilinear equations of parabolic type*, volume 23 of *Translations of mathematical monographs*. AMS, Providence, Rhode Island, 1968.

[22] Hans-Gerd Leopold. On function spaces of variable order of differentiation. *Forum Mathematicum*, 3:1–21, 1991.

[23] Jacques Louis Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Sprigner, Berlin, 1971.

[24] Jacques Louis Lions and Enrico Magenes. *Non-Homogeneous Boundary Value Problems and Applications II*. Springer, Berlin, 1972.

[25] Dominik Meidner and Boris Vexler. A Priori Error Estimates for Space-Time Finite Element Discretization of Parabolic Optimal Control Problems. Part I: Problems without Control Constraints. *SIAM Journal on Control and Optimization*, 47(3):1150 – 1177, 2008.

[26] Dominik Meidner and Boris Vexler. A priori error analysis of the Petrov Galerkin Crank Nicolson scheme for parabolic optimal control problems. *SIAM Journal on Control and Optimization*, 49(5):2183 – 2211, 2011. Preprint-Nr.: SPP1253-109.

[27] Solomon G. Michlin. *Partielle Differentialgleichungen in der mathematischen Physik*. Verlag Harri Deutsch, Thun, 1978.

[28] Ira Neitzel, Uwe Prüfert, and Thomas Slawig. On solving parabolic optimal control problems by using space-time discretization. Technical Report 05-2009, TU Berlin, 2009.

[29] Ira Neitzel, Uwe Prüfert, and Thomas Slawig. A smooth regularization of the projection formula for constrained parabolic optimal control problems. *Numerical Functional Analysis and Optimization*, 32(12), 2011.

[30] Sergeĭ Michaĭlovič Nikol'skiĭ. *Approximation of Functions of Several Variables and Imbedding Theorems*. Springer, Berlin, 1975.

[31] N.V. Oganesyan. Solution of the first boundary value problem for model semielliptic equation by projection-grid method. *Soviet Journal of Contemporary Mathematical Analysis. Armenian Academy of Sciences.*, 4:82–90, 1989. translation from Izv. Akad. Nauk Arm. SSR. Matematika, Vol. 24, No.4, 393-402 (1989).

[32] Waldemar Rachowicz. An anisotropic h-type mesh-refinement strategy. *Computer Methods in Applied Mechanics and Engineering*, 109:169–181, 1993.

[33] Wladimir Iwanowitsch Smirnow. *Lehrgang der höheren Mathematik, Teil 5*. VEB Deutscher Verlag der Wissenschaften, Berlin, 10. edition, 1988.

[34] Hans Triebel. A priori estimates and boundary value problems for semi-elliptic differential equations: A model case. *Communications in partial differential equations*, 8(15):1621–1664, 1983.

[35] Fredi Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen - Theorie, Verfahren und Anwendungen.* Vieweg, Wiesbaden, 2005.

[36] Houston G. Wood and J. B. Morton. Onsager's pancake approximation for the fluid dynamics of a gas centrifuge. *Journal of Fluid Mechanics*, 101(1):1–31, 1980.