MATHEMATICS OF COMPUTATION
Volume 74, Number 249, Pages 85–102
S 0025-5718(04)01654-0
Article electronically published on April 20, 2004

CONVERGENCE OF A CONTINUOUS GALERKIN METHOD WITH MESH MODIFICATION FOR NONLINEAR WAVE EQUATIONS

OHANNES KARAKASHIAN AND CHARALAMBOS MAKRIDAKIS

Abstract. We consider space-time continuous Galerkin methods with mesh modification in time for semilinear second order hyperbolic equations. We show a priori estimates in the energy norm without mesh conditions. Under reasonable assumptions on the choice of the spatial mesh in each time step we show optimal order convergence rates. Estimates of the jump in the Riesz projection in two successive time steps are also derived.

1. Introduction

We consider space-time continuous Galerkin methods with mesh modification in time for the model problem

$$
\begin{aligned}
&u_{tt} - \Delta u = f(u), \quad \text{in } \Omega \times [0,T],\\
&u = 0, \quad \text{on } \partial\Omega \times [0,T],\\
&u(\cdot,0) = u^0, \quad \text{in } \Omega,\\
&u_t(\cdot,0) = u^1, \quad \text{in } \Omega,
\end{aligned}
\tag{1.1}
$$

where Ω is a bounded domain in $\mathbb{R}^2$, u is a real-valued function defined on $\overline{\Omega}\times[0,T]$, and f is a given Lipschitz function.

The continuous Galerkin method [AM], [BL], [FP] is a space-time finite element method whose Runge-Kutta counterpart is the class of Gauss-Legendre methods and thus is particularly useful for hyperbolic problems where the conservation of certain quantities is important. This method has been analyzed for linear wave equations in [BL], [FP]. For nonlinear problems with possible singular behavior it is important to consider methods that allow adaptive mesh refinement. In this direction a continuous Galerkin method with mesh modification in time was proposed and analyzed in [KM2] for the nonlinear Schrödinger equation; see also [D]. In [Y] a first order in time fully discrete method with mesh modification for linear hyperbolic problems is considered and error estimates in the energy norm are derived. In this paper we continue our investigation by considering continuous Galerkin methods with mesh modification for nonlinear wave equations of the form (1.1). The method proposed reduces to the classical one in the linear case, [BL], [FP], if we have the same spatial mesh for all times. For other space-time finite element methods for the wave equation cf. [F], [HH], [J].

Received by the editor November 20, 2001 and, in revised form, May 7, 2003.
2000 *Mathematics Subject Classification.* Primary 65M60, 65M12.

Mesh modification can have negative effects on the a priori analysis. In fact, in extreme cases it might cause the divergence of the numerical method, as was pointed out by Dupont [D]. Therefore the order of the method should not be taken for granted even when "reasonable" mesh modification takes place. In the error analysis we have to either impose restrictions in the selection of the mesh as was done, e.g., in [EJ] or allow new terms, known as "jump terms", to make an appearance in the error estimates as in [KM1], [KM2]. These terms, albeit small, are multiplied by the number of time levels where a spatial mesh modification has occurred. In this paper, we have devoted special attention to this important issue. In one approach, we show how to eliminate such terms (cf. Section 4) while the second consists in obtaining superconvergence results for these terms (cf. Section 5).

The continuous Galerkin method. To introduce the method, we need some notation. Let $H^\ell = H^\ell(\Omega)$ be the standard L^2-based Sobolev spaces of order ℓ with norm $\|\cdot\|_\ell$. Also $(\cdot,\cdot)$ denotes the inner product, and $\|\cdot\|$, the corresponding norm on $L^2(\Omega)$. $\|\cdot\|_\infty$, denotes the norm of $L^\infty(\Omega)$ and $\|\cdot\|_{1,\infty}$ the norm of $W^{1,\infty}(\Omega)$. We shall also use the "energy" norm on $H_0^1 \times L^2$ given by $|||(u,v)^T||| = \{\|\nabla u\|^2 + \|v\|^2\}^{1/2}$. For a space-time domain $\Omega \times I$, $L^p(I;H^\ell(\Omega))$ will denote the usual spaces of functions defined on I and with values in $H^\ell(\Omega)$ with norm denoted by $\|\cdot\|_{L^p(I_n;H^\ell)}$.

We consider a partition of $[0,T]$, $0 = t^0 < t^1 < \cdots < t^N = T$, and we let $I_n = (t^n, t^{n+1}]$, $\quad k_n = t^{n+1} - t^n$. We associate a partition $\mathcal{T}_{hn}$ of Ω and a finite element space S_h^n to each interval I_n :

$$S_h^n = \{\chi \in H_0^1(\Omega) : \chi|_K \in \mathbb{P}_{r-1}(K), K \in \mathcal{T}_{hn}\},$$

where $\mathbb{P}_p(S)$ is the space of polynomials of degree p. (We associate S_h^{-1} to $\{t^0\}$ but for simplicity we take $S_h^{-1} = S_h^0$.) In the sequel we shall denote by K an element of the partition $\mathcal{T}_{hn}$. Also h_K stands for the diameter of the element K, and $h_n = \max_{K\in\mathcal{T}_{hn}} h_K$.

For a positive integer q, let $V_q = V_{hk}(q)$ be the space of piecewise polynomial functions $\varphi : \Omega \times (0,T] \to \mathbb{R}$ of the form: $\varphi|_{\Omega\times I_n} = \sum_{j=0}^q t^j \chi_j(x), \quad \chi_j \in S_h^n$. Hence, the functions of V_q are for each $t \in I_n$ elements of S_h^n, but they may be discontinuous (in t) at the nodes t^n, $n = 0,\ldots,J-1$. For this reason, we introduce the notation $v^{n+} = \lim_{t\to t^{n}+} v(t)$. Let also $V_q^n = \{\varphi|_{\Omega\times I_n} : \varphi \in V_q\}$. Let $A_h^n : H_0^1(\Omega) \to S_h^n$ be the discrete operator defined by

$$(A_h^n v, \chi) = (\nabla v, \nabla \chi), \quad \forall \chi \in S_h^n\,, \tag{1.2}$$

and let $\mathcal{L}_h^n : H_0^1(\Omega) \times L^2(\Omega) \to S_h^n \times S_h^n$ be defined by

$$\mathcal{L}_h^n = \begin{pmatrix} 0 & -I \\ A_h^n & 0 \end{pmatrix}.$$

Henceforth, $((\cdot,\cdot))$ will denote the usual L^2 inner product on the product space $L^2 \times L^2$.

We will consider approximations of (1.1) written in the usual first order (in time) system form. In particular we will seek $U \in V_q \times V_q$, $U = (U_1, U_2)^T \approx (u, u_t)^T$ such that

$$\int_{I_n} \{((U_t,\Phi)) + ((\mathcal{L}_h^n U,\Phi))\}\, dt = \int_{I_n} ((F_h(U),\Phi))dt,\ \forall \Phi \in V_{q-1}^n \times V_{q-1}^n, \tag{1.3}$$
$$U^{n+} = \Pi^n U^n, \quad n = 0,\dots,N-1,$$

where $U^0 = (u^0,u^1)^T$ and $F_h(U) = (0, P^n f(U_1))^T$, P^n is the L^2 projection into S_h^n and Π^n denotes an appropriate projection operator into $S_h^n \times S_h^n$. The analysis in this paper requires the choice $\Pi^n = \begin{pmatrix} P_E^n & 0 \\ 0 & P^n \end{pmatrix}$, where P_E^n is the elliptic projection operator into S_h^n. Note however that if $S_h^{n-1} \subset S_h^n$, which would result if one refines the mesh locally, then $U^{n+} = U^n$. Also, in practice, one could presumably use a Lagrangian type interpolation operator as a less expensive alternative to Π^n. This will involve calculations only on the altered part of the mesh; cf. Remark 3.1 for a comment on the influence of this choice in the error analysis.

Main results. The convergence properties of this method are discussed in the next sections. To simplify the presentation, we assume here that f is globally Lipschitz; the general case of locally Lipschitz f can be treated as in [KM2]. Although we follow the basic ideas in [KM2], the approach taken here is different and does not follow with straightforward adaptations of the analysis in [KM2] for the nonlinear Schrödinger equation. One of the main difficulties presented here is the need to directly work with the energy norm. This requires special choices and more involved consistency analysis. On the other hand the analysis in this paper is more condensed and based rather on finite element type techniques compared with those of [KM1], [KM2].

In Theorem 3.1 we show that

$$\begin{aligned} &\max_{t\in[0,T]} \Big(\|u(t)-U_1(t)\| + \|u_t(t)-U_2(t)\| \Big) \\ &\quad \le C\Big\{ \max_n k_n^{q+1} C_t(u) + \max_n\ h_n^r C_x(u) + \sqrt{\mathcal{N}_C}\, \max_n \lVert\!| J^n |\!\rVert \Big\}, \end{aligned} \tag{1.4}$$

where $C_t(u)$ and $C_x(u)$ are quantities depending on various temporal and spatial derivatives of u, $\mathcal{N}_C$ denotes the number of times where $S_h^j \neq S_h^{j-1}$, $j = 1,\dots N-1$, and $J^n = (\omega^{n+} - \omega^n, \omega_t^{n+} - \omega_t^n)^T$ is the jump in the elliptic projection of $(u,u_t)^T$ at time t^n (cf. Theorem 3.1 for the details). Theorem 3.2 provides the "local" spatial mesh version of the estimate proved in Theorem 3.1. In addition a direct consequence of the estimates in the energy norm in Theorem 3.1 is the L^∞ estimate

$$\max_{t\in[0,T]} \|u(t)-U_1(t)\|_\infty \le C L_h \Big\{ \max_n k_n^{q+1} C_t(u) + \max_n\ h_n^r C_x(u) + \sqrt{\mathcal{N}_C}\, \max_n \lVert\!| J^n |\!\rVert \Big\}, \tag{1.5}$$

where L_h is a factor that grows logarithmically with h. As in [KM2], an interesting feature in the proofs is the right choice of two time interpolating operators. These are the interpolants at the Gauss-Legendre (stability analysis) and at the Lobatto points of each I_n (consistency analysis). Note in addition that due to the mesh modification with n we have chosen to work with the energy norm, rather than with the weaker norm introduced in [BB] and used in [BL], [FP]. We show that we can retain the optimal order of convergence in the L^2 and L^∞ norms. See also [M] where this was done in a different context.

Note, however, the presence of the gradient in the jump terms due to the mesh modification (compare with [KM2]). Indeed, $\|\!|J^n|\!\| = O(h^{r-1})$. In Section 4 we show that it is possible to eliminate the jump terms under the reasonable assumption that all meshes contain a reference mesh $\mathcal{T}_h$. Indeed, we show the optimal order estimate:

$$\max_{t\in[0,T]} \Big(\|u(t) - U_1(t)\| + \|u_t(t) - U_2(t)\| \Big) \le C \Big\{ \max_n k_n^{q+1} C_t(u) + h^r C_x(u) \Big\}. \tag{1.6}$$

In particular this shows that optimal convergence rates can be retained when mesh modification is performed in many practical applications.

Finally, in Section 5 we show how to obtain increased accuracy for the jump terms under the assumption that the meshes differ only in a region of small area. In the two-dimensional case, and assuming that this region is of area $O(h^2)$, we show that $\|\!|J^n|\!\| = O(h^r)$. Thus in this case also the convergence rates are optimal provided $\mathcal{N}_C$ remains bounded.

Plan of the analysis. The analysis in the forthcoming Sections 2 and 3 is based on several steps that we briefly describe. One of the main ingredients of our approach is the use of the connection of the continuous Galerkin method to the Runge-Kutta–Gauss-Legendre family. This connection is used as a motivation for the representation of U in terms of its values at the Gauss points of each I_n (cf. (2.3))

$$U(x,t) = \sum_{j=0}^{q} \hat{\ell}_{n,j}(t) U^{n,j}(x), \quad (x,t) \in \Omega \times I_n, \quad t^{n,0} = t^n.$$

Here $t^{n,j}$, $j = 1, \dots, q$, are the Gauss Legendre points of I_n and $\{\hat{\ell}_i\}_{i=0}^q$ are the Lagrange polynomials of degree q associated with the $q+1$ points $t^{n,j}$, $j = 1, \dots, q$, plus the point $t^{n,0} = t^n$. Here, of course, $U^{n,0} = U^{n+} = \Pi^n U^n$ is the given starting value. This representation is crucial throughout the paper. Its first use is in the uniqueness proof, Lemma 2.1 and Theorem 2.1: Indeed in view of (2.7) and Lemma 2.1 we can gain control of $\sum_{j=0}^q \|U^{n,j}\|^2$ by appropriate selection of the test function in (2.7). Then the existence-uniqueness follows by applying known arguments; cf., e.g., [KM2]. Essentially the same argument is used later in the stability analysis for the control of the $L^2(I_n; L^2)$ norm of the error by using (2.4); cf. Lemma 3.4.

Consistency analysis—the basic error equation. The error is decomposed as $U-u = (U-W) + (W-u)$, where $W = (W_1, W_2)^T \in V_q \times V_q$. The choice of W is essential since it should be chosen such that $W - u$ has the right order and in addition W has desirable consistency properties. This is achieved by the definition of the components of W through (3.6), (3.7). Note first that the definitions are based on the interpolation operator $\mathcal{I}_{Lo}^{n,q}$ at the $q+1$ Lobatto points of I_n and the elliptic projection operator, (3.1). Lobatto interpolation is important since it preserves continuity at *both* endpoints of the interval and its corresponding quadrature has the same accuracy as that of the Gauss rule with q points. A more natural choice for W_1 would be just $W_1(x,t)\big|_{I_n} = \mathcal{I}_{Lo}^{n,q}\omega(x,t)$, but then (3.8) is not valid. It turns out that (3.8) is essential in the sequel in order to avoid "spatial" error terms in the first component of the right-hand side of (3.9), and therefore W_1 given by (3.7) has all the desirable properties. The approximation properties of (W_1, W_2) are established in part (i) of Lemma 3.3. It remains to estimate $E = U - W$. In view of the definition of the scheme and the properties of W we conclude in Lemma 3.2

that E satisfies the basic error equation (3.9). Then the consistency terms that appear in its right-hand side are estimated in Lemma 3.3(ii).

Stability analysis—estimate of E. Next, the basic error equation (3.9) will be the starting point to prove the final estimate for E. Notice that we choose to work with the energy norm $\|\!|E|\!\| = \left(\|\nabla E_1\|^2 + \|E_2\|^2\right)^{1/2}$. This norm is a natural choice for the wave equation; in [BL], [FP] the weaker $L^2 \times H^{-1}$-like norm introduced in [BB] was used. The choice of the energy norm is important since it allows us to handle the mesh modification in a proper way. Next, by selecting $\Phi = \mathcal{P}_t^{n,q-1}\mathcal{A}_h^n E$ in (3.9), we conclude the first estimate (3.23). It is evident now that an additional bound of $\|\!|E|\!\|_{L^2(I_n;L^2)}$ is needed. This is obtained in Lemma 3.4. Indeed as in Section 2 we can gain control of $\sum_{j=0}^{q} \|\!|E^{n,j}|\!\|^2$ and thus of $\|\!|E|\!\|_{L^2(I_n;L^2)}$ in view of (2.4) by choosing appropriate test functions in (3.9) and using the stability Lemma 2.1. Using then Lemma 3.4 in (3.23), the only term that remains to be handled is $\|\!|E^{n+}|\!\|$. This is exactly the term where mesh modification will give rise to the coarsening terms; cf. Lemma 3.5. The proof is then completed in Theorem 3.1 by a refined Grownwall type argument. "Local" mesh size versions of these estimates rely on the validity of corresponding bounds for the elliptic finite element case. Theorem 3.2 provides the error estimates in this case.

2. Notation—preliminaries

As in [KM2], we will consider for $q \geq 1$, the Gauss-Legendre integration rule,

$$\int_0^1 g(\tau)d\tau \cong \sum_{j=1}^{q} w_j g(\tau_j), \quad 0 < \tau_1 < \cdots < \tau_q < 1, \tag{2.1}$$

which is *exact for all polynomials of degree* $\leq 2q-1$. Let $\{\ell_i\}_{i=1}^q$ be the Lagrange polynomials of degree $q-1$ associated with the abscissas $\tau_1, \ldots, \tau_q$.

Using the linear transformation $t = t^n + \tau k_n$ that maps [0,1] onto $\overline{I}_n$, we adapt the quadrature rule (2.1) to the interval $\overline{I}_n$ by defining its abscissas and weights as follows:

$$\begin{aligned} &t^{n,i} = t^n + \tau_i k_n, \\ &\ell_{n,i}(t) = \ell_i(\tau), \quad t = t^n + \tau k_n, \\ &w_{n,i} = \int_{t^n}^{t^{n+1}} \ell_{n,i}(t)dt = k_n \int_0^1 \ell_i(\tau)d\tau = k_n w_i, \qquad i = 1, \ldots, q. \end{aligned} \tag{2.2}$$

We shall also use the Lagrange polynomials $\{\hat{\ell}_i\}_{i=0}^q$ of degree q associated with the $q+1$ points $0 = \tau_0 < \tau_1 < \cdots < \tau_q$. In particular, $U|_{I_n}$ is determined by the functions $U^{n,j} \in S_h^n \times S_h^n$ ($U^{n,j} = U(x, t^{n,j})$) such that

$$U(x,t) = \sum_{j=0}^{q} \hat{\ell}_{n,j}(t) U^{n,j}(x), \quad (x,t) \in \Omega \times I_n, \quad t^{n,0} = t^n, \tag{2.3}$$

where $U^{n,0} = U^{n+} = \Pi^n U^n$ is given.

In the sequel, the following equivalence of norms will be useful:

$$C_1 \Big\{ k_n \sum_{j=0}^{q} \|v^j\|^2 \Big\}^{1/2} \leq \|v\|_{L^2(I_n;L^2)} \leq C_2 \Big\{ k_n \sum_{j=0}^{q} \|v^j\|^2 \Big\}^{1/2}, \quad v \in V_q \tag{2.4}$$

where $v = \sum_{j=0}^{q} \hat{\ell}_{n,j} v^j \in V_q^n$. This is a consequence of the $L^\infty - L^2$ inverse property

$$\max_{I_n} |y(t)| \le C_I k_n^{-1/2} \left(\int_{I_n} |y(t)|^2 dt \right)^{1/2}, \quad \forall y \in \mathbb{P}_q(I_n), \tag{2.5}$$

and of the bound $\int_{I_n} |\hat{\ell}_{n,j}(t)| dt \le c k_n$.

We consider the L^2-projection operator $P_t^{n,q-1} : \mathbb{P}_q[t^n, t^{n+1}] \to \mathbb{P}_{q-1}[t^n, t^{n+1}]$. Then (cf. [KM2])

$$P_t^{n,q-1} = \mathcal{I}_{GL}^{n,q-1}, \tag{2.6}$$

where $\mathcal{I}_{GL}^{n,q-1}$ is the Lagrange interpolation operator corresponding to the q Gauss-Legendre points $t^{n,1} < \cdots < t^{n,q}$. This is indeed the case since for $v \in \mathbb{P}_q[t^n, t^{n+1}]$,

$$\int_{I_n} (\mathcal{I}_{GL}^{n,q-1} v) \phi dt = \sum_{j=1}^{q} w_{n,j} v(t^{n,j}) \phi(t^{n,j}) = \int_{I_n} v \phi dt, \quad \forall \phi \in \mathbb{P}_{q-1}^n[t^n, t^{n+1}].$$

Now, for $\Phi \in V_{q-1}^n \times V_{q-1}^n$,

$$\int_{I_n} ((U_t, \Phi)) = \sum_{i,j=1}^{q} m_{ij} ((U^{n,j}, \Phi^i)) + \sum_{i=1}^{q} m_{i0} ((U^{n+}, \Phi^i)), \tag{2.7}$$

with

$$m_{ij} = \int_{I_n} \hat{\ell}'_{n,j}(t)\, \ell_{n,i}(t) dt, \quad i = 1, \ldots, q, \; j = 0, \ldots, q,$$

and $v^i = v(t^{n,i})$. The stability of the method relies on the positivity of the matrix $\mathcal{M}$

$$\mathcal{M}_{ij} = m_{ij}, \quad i, j = 1, \ldots, q.$$

In fact it is shown in [KM2] that the array $\widetilde{\mathcal{M}} = D^{-1/2} \mathcal{M} D^{1/2}$, where $D = \text{diag}\{\tau_1, \ldots, \tau_q\}$ is positive definite:

Lemma 2.1 ([KM2])**.** *Let* $\alpha := \frac{1}{2} \min_j \frac{w_j}{\tau_j}$. *Then*

$$\mathbf{x}^T \widetilde{\mathcal{M}} \mathbf{x} \ge \alpha |\mathbf{x}|^2 = \alpha (\sum_{i=1}^{q} x_i^2), \quad \forall \mathbf{x} \in \mathbb{R}^q. \tag{2.8}$$

Employing then similar arguments as in [KM2], we get the existence and uniqueness of the numerical approximations.

Theorem 2.1. *Let* U^n *be given in* $S_h^{n-1} \times S_h^{n-1}$. *Then for* k_n *sufficiently small there exists a unique solution* $U \in V_q^n \times V_q^n$ *of equation* (1.3).

3. Error estimates

We split the error $U - u = (U - W) + (W - u)$, where $W \in V_q \times V_q$ will be defined below, and we estimate $E = U - W$ and $u - W$. To define W, we consider the elliptic projection operator $P_E^n : H_0^1(\Omega) \to S_h^n$ defined as usual by

$$(\nabla P_E^n v, \nabla \chi) = (\nabla v, \nabla \chi), \qquad \forall \chi \in S_h^n. \tag{3.1}$$

We assume that the family of spaces S_h^n satisfies

$$\|\nabla (v - P_E^n v)\| \le c h_n^{s-1} \|v\|_s, \qquad v \in H^s \cap H_0^1, \; 2 \le s \le r, \tag{3.2}$$

and

$$\|v - P_E^n v\| \le c h_n^s \|v\|_s, \qquad v \in H^s \cap H_0^1,\ 2 \le s \le r, \tag{3.3}$$

where c is independent of n.

See [BO], [EJ], [BS, Chapter 0], and [KM1] for a discussion on these assumptions.

We will also need a temporal interpolation operator: As in [KM2] we consider interpolation at Gauss-Lobatto points. For this let $0 = \xi_0 < \cdots < \xi_q = 1$ be the $q+1$ roots of the polynomial $L(x) = \dfrac{d^{q-1}}{dx^{q-1}}[x(1-x)]^q$. The corresponding Gauss-Lobatto quadrature rule, [BS],

$$\int_0^1 g(\tau)\,d\tau \approx \sum_{j=0}^{q} b_j g(\xi_j), \tag{3.4}$$

is exact on $\mathbb{P}_{2q-1}$. As done with the Gauss-Legendre points, we can define the points $\xi^{n,j}$ and the weights $b_{n,j}$ corresponding to the interval I_n. In addition let $\mathcal{I}_{Lo}^{n,q}$ be the Lagrange interpolation operator at the $q+1$ Lobatto points $t^n = \xi^{n,0} < \cdots < \xi^{n,q} = t^{n+1}$.

We define ω and η by

$$\omega(x,t) = P_E^n u(x,t), \quad \eta = u - \omega, \quad (x,t) \in \Omega \times I_n,\ n = 0, \dots, N-1. \tag{3.5}$$

We next define $W = (W_1, W_2)$. First let

$$W_2(x,t)\big|_{I_n} = \mathcal{I}_{Lo}^{n,q} \omega_t(x,t), \qquad W_2(t^0) = P_E^0 u^1, \tag{3.6}$$

and then,

$$W_1\big|_{I_n} = \mathcal{I}_{Lo}^{n,q}\Big(\int_{t^n}^t W_2\,dt + \omega^{n+}\Big). \tag{3.7}$$

We have

Lemma 3.1. *It holds that*

$$\int_{I_n} (W_{1,t}, \varphi) dt = \int_{I_n} (W_2, \varphi) dt, \quad \forall \varphi \in V_{q-1}. \tag{3.8}$$

Proof. Let $Z\big|_{I_n} = \int_{t^n}^t W_2\,dt + \omega^{n+}$. Then $W_1\big|_{I_n} = \mathcal{I}_{Lo}^{n,q} Z$. Using the definition of $\mathcal{I}_{Lo}^{n,q}$ and the exactness of the Gauss-Lobatto integration rule, we get

$$\begin{aligned}
\int_{I_n} (W_{1,t}, \varphi) dt &= -\int_{I_n} (W_1, \varphi_t) dt + (Z,\varphi)(t^{n+1}) - (Z,\varphi)(t^{n+}) \\
&= -\sum_{j=0}^{q} b_{n,j} (Z, \varphi_t)(\xi^{n,j}) + (Z,\varphi)(t^{n+1}) - (Z,\varphi)(t^{n+}) \\
&= -\int_{I_n} (Z, \varphi_t) dt \ + (Z,\varphi)(t^{n+1}) - (Z,\varphi)(t^{n+}) \\
&= \int_{I_n} (Z_t, \varphi) dt = \int_{I_n} (W_2, \varphi) dt,
\end{aligned}$$

since $(Z, \varphi_t) \in \mathbb{P}_{2q-1}$. ∎

The basic error equation.

Lemma 3.2. *Let* $E = E|_{I_n} = U - W$ *and*

$$\begin{aligned} A_I^n &:= \mathcal{I}_{Lo}^{n,q}\big(\int_{t^n}^t (I - \mathcal{I}_{Lo}^{n,q})u_t\, ds\big), \\ A_{II}^n &:= u_{tt} - W_{2,t}, \\ A_{III}^n &:= f(W_1) - f(u) + (\mathcal{I}_{Lo}^{n,q} - I)\Delta u. \end{aligned}$$

Then for all $\Phi = (\phi_1, \phi_2)^T \in V_{q-1}^n \times V_{q-1}^n$ *and* $n = 0, 1, \dots, N-1$,

(3.9)

$$\int_{I_n} ((E_t + \mathcal{L}_h^n E, \Phi))\, dt = \int_{I_n} \big\{(f(U_1) - f(W_1), \phi_2) - (\Delta A_I^n, \phi_2) + (A_{II}^n + A_{III}^n, \phi_2)\big\}\, dt.$$

Proof. To begin, E satisfies

$$\int_{I_n} ((E_t + \mathcal{L}_h^n E, \Phi))\, dt = \int_{I_n} \big\{((F_h(U), \Phi)) - ((W_t + \mathcal{L}_h^n W, \Phi))\big\}\, dt.$$

Also note that $((F_h(U), \Phi)) = (f(U_1), \phi_2)$. On the other hand, by Lemma 3.1 we have

$$\begin{aligned} \int_{I_n} ((W_t + \mathcal{L}_h^n W, \Phi))\, dt &= \int_{I_n} \big\{(W_{1,t} - W_2, \phi_1) + (W_{2,t}, \phi_2) + (\nabla W_1, \nabla \phi_2)\big\} dt \\ &= \int_{I_n} (W_{2,t}, \phi_2) dt + \int_{I_n} (\nabla W_1, \nabla \phi_2)\, dt\,. \end{aligned}$$

By the definition of W_1 and (1.1) we obtain

$$\begin{aligned} \int_{I_n} (\nabla W_1, \nabla\phi_2) dt &= \sum_{j=0}^{q} b_{n,j} (\nabla Z, \nabla \phi_2)(\xi^{n,j}) \\ &= \sum_{j=0}^{q} b_{n,j} (\nabla (\int_{t^n}^{\xi^{n,j}} \mathcal{I}_{Lo}^{n,q} u_t ds + u^n), \nabla \phi_2(\xi^{n,j})) \\ &= \int_{I_n} (\Delta A_I^n, \phi_2) + \sum_{j=0}^{q} b_{n,j} (\nabla (\int_{t^n}^{\xi^{n,j}} u_t ds + u^n), \nabla \phi_2(\xi^{n,j})) \\ &= \int_{I_n} (\Delta A_I^n, \phi_2) + \sum_{j=0}^{q} b_{n,j} (\nabla u(\xi^{n,j})\,, \nabla \phi_2(\xi^{n,j})) \\ &= \int_{I_n} (\Delta A_I^n, \phi_2) + \int_{I_n} (\nabla u\,, \nabla\phi_2) + \int_{I_n} (\nabla(\mathcal{I}_{Lo}^{n,q} u - u)\,, \nabla \phi_2) \\ &= \int_{I_n} (\Delta A_I^n, \phi_2) + \int_{I_n} (f(u), \phi_2) - \int_{I_n} (u_{tt}, \phi_2) - \int_{I_n} ((\mathcal{I}_{Lo}^{n,q} - I)\Delta u\,, \phi_2)\,, \end{aligned}$$

and the proof is complete. □

In the next lemma we show that indeed W_1 and W_2 have the right approximation properties and we estimate $A_I^n, A_{II}^n, A_{III}^n$.

Lemma 3.3. (i) *Let* W_1 *and* W_2 *be as in* (3.7) *and* (3.6)*, respectively. Then for* $p = 2$ *and* $p = \infty$*,*

$$\|W_1 - u\|_{L^p(I_n;L^2)} \le ck_n^{q+1}\| \,|u^{(q+1)}| + |u^{(q+2)}|\, \|_{L^p(I_n;L^2)} + ch_n^r\| \,|u| + k_n|u_t| + k_n^2|u_{tt}|\, \|_{L^p(I_n;H^r)}, \tag{3.10}$$

$$\|W_2 - u_t\|_{L^p(I_n;L^2)} \le ck_n^{q+1}\|u^{(q+2)}\|_{L^p(I_n;L^2)} + ch_n^r\| \,|u_t| + k_n|u_{tt}|\, \|_{L^p(I_n;H^r)}, \tag{3.11}$$

where $u^{(m)} := \partial_t^m u$.

(ii) *Let* A_I^n*,* $A_{I\!I}^n$ *and* $A_{I\!I\!I}^n$ *be as in Lemma 3.2. Then*

$$\|\Delta A_I^n\|_{L^2(I_n;L^2)} \le ck_n^{q+1}\|\Delta u^{(q+1)}\|_{L^2(I_n;L^2)} \tag{3.12}$$

$$\Big|\int_{I_n}(A_{I\!I}^n,\varphi)dt\Big| \le c\Big(k_n^{q+1}\|u^{(q+3)}\|_{L^2(I_n;L^2)} + h_n^r\|u_{tt}\|_{L^2(I_n;H^r)}\Big)\|\varphi\|_{L^2(I_n;L^2)}, \quad \varphi \in V_{q-1}, \tag{3.13}$$

$$\|A_{I\!I\!I}^n\|_{L^2(I_n;L^2)} \le c\|W_1 - u\|_{L^2(I_n;L^2)} + Ck_n^{q+1}\|\Delta u^{(q+1)}\|_{L^2(I_n;L^2)}. \tag{3.14}$$

Proof. (i) We will consider only the case $p = 2$, the case $p = \infty$ being in fact easier. We begin by proving a stability property of the Lobatto interpolation operator $\mathcal{I}_{Lo}^{n,q}$: It is clear that $\mathcal{I}_{Lo}^{n,q}$ is stable with respect to the L^∞ norm, i.e., $\|\mathcal{I}_{Lo}^{n,q}\phi\|_{L^\infty(I_n)} \le c\|\phi\|_{L^\infty(I_n)}$, $\forall \phi \in L^\infty(I_n)$. Thus, using the Sobolev imbedding lemma and Cauchy-Schwarz, we can easily deduce that $\|\mathcal{I}_{Lo}^{n,q}\phi\|_{L^2(I_n)} \le c\|\phi\|_{L^2(I_n)} + ck_n\|\phi'\|_{L^2(I_n)}$. This in turn gives the required stability result in the $\|\cdot\|_{L^2(I_n;L^2)}$ norm:

$$\|\mathcal{I}_{Lo}^{n,q}\phi\|_{L^2(I_n;L^2)} \le c\|\phi\|_{L^2(I_n;L^2)} + ck_n\|\phi_t\|_{L^2(I_n;L^2)}. \tag{3.15}$$

We shall also use the inequality

$$\|\phi\|_{L^2(I_n;L^2)} \le ck_n\|\psi\|_{L^2(I_n;L^2)}, \quad \text{where } \phi = \int_{t^n}^t \psi\, ds. \tag{3.16}$$

Write $W_1 - u = \mathcal{I}_{Lo}^{n,q}\int_{t^n}^t (W_2 - \omega_t)\, ds + \mathcal{I}_{Lo}^{n,q}\omega - u$. From (3.15) and (3.16) it follows that

$$\|\mathcal{I}_{Lo}^{n,q}\int_{t^n}^t (W_2 - \omega_t)\, ds\|_{L^2(I_n;L^2)} \le ck_n\|W_2 - \omega_t\|_{L^2(I_n;L^2)}. \tag{3.17}$$

Now, $W_2 - \omega_t = \mathcal{I}_{Lo}^{n,q}\omega_t - \omega_t = -\mathcal{I}_{Lo}^{n,q}\eta_t - (u_t - \mathcal{I}_{Lo}^{n,q}u_t) - \eta_t$. From (3.15),

$$\begin{aligned}\|\mathcal{I}_{Lo}^{n,q}\eta_t\|_{L^2(I_n;L^2)} + \|\eta_t\|_{L^2(I_n;L^2)} &\le c\|\eta_t\|_{L^2(I_n;L^2)} + ck_n\|\eta_{tt}\|_{L^2(I_n;L^2)} \\ &\le ch_n^r\|u_t\|_{L^2(I_n;H^r)} + ck_nh_n^r\|u_{tt}\|_{L^2(I_n;H^r)}.\end{aligned} \tag{3.18}$$

The approximation properties of the operator $\mathcal{I}_{Lo}^{n,q}$ give

$$\|u_t - \mathcal{I}_{Lo}^{n,q}u_t\|_{L^2(I_n;L^2)} \le ck_n^{q+1}\|u^{(q+2)}\|_{L^2(I_n;L^2)}. \tag{3.19}$$

Writing $\mathcal{I}_{Lo}^{n,q}\omega - u = -\mathcal{I}_{Lo}^{n,q}\eta + \mathcal{I}_{Lo}^{n,q}u - u$, as above we obtain

$$\|\mathcal{I}_{Lo}^{n,q}\omega - u\|_{L^2(I_n;L^2)} \le ch_n^r\|u\|_{L^2(I_n;H^r)} + ck_n^{q+1}\|u^{(q+1)}\|_{L^2(I_n;L^2)}. \tag{3.20}$$

Inequality (3.10) now follows from (3.17)–(3.20). Similarly, writing $u_t - W_2 = u_t - \mathcal{I}_{Lo}^{n,q}u_t + \mathcal{I}_{Lo}^{n,q}\eta_t$, we see that (3.11) follows from (3.18) and (3.19).

(ii) We next estimate $\|\Delta A_I^n\|_{L^2(I_n;L^2)}$. From (3.15) and (3.16)

$$\begin{aligned}\|\Delta A_I^n\|_{L^2(I_n;L^2)} &= \|\mathcal{I}_{Lo}^{n,q}\int_{t^n}^t (I-\mathcal{I}_{Lo}^{n,q})\Delta u_t\, ds\|_{L^2(I_n;L^2)}\\ &\le ck_n\|(I-\mathcal{I}_{Lo}^{n,q})\Delta u_t\|_{L^2(I_n;L^2)}\\ &\le ck_n^{q+1}\|\Delta u^{(q+1)}\|_{L^2(I_n;L^2)}.\end{aligned}$$

Also,

$$\begin{aligned}\|A_{I\!I\!I}^n\|_{L^2(I_n;L^2)} &\le \|f(W_1)-f(u)\|_{L^2(I_n;L^2)} + \|(I-\mathcal{I}_{Lo}^{n,q})\Delta u\|_{L^2(I_n;L^2)}\\ &\le c\|W_1-u\|_{L^2(I_n;L^2)} + Ck_n^{q+1}\|\Delta u^{(q+1)}\|_{L^2(I_n;L^2)}.\end{aligned}$$

It remains to estimate $A_{I\!I}^n$. Let $\varphi\in V_{q-1}$. Then ($\eta_t = u_t-\omega_t$)

$$\int_{I_n}(A_{I\!I}^n,\varphi)dt = \int_{I_n}(u_{tt}-(\mathcal{I}_{Lo}^{n,q}u_t)_t,\varphi)dt + \int_{I_n}((\mathcal{I}_{Lo}^{n,q}\eta_t)_t,\varphi)dt =: \Gamma_1+\Gamma_2\,.$$

To estimate Γ_1, we have since the endpoints of I_n are included in the Lobatto points

$$\Gamma_1 = \int_{I_n}(u_{tt}-(\mathcal{I}_{Lo}^{n,q}u_t)_t,\varphi)dt = -\int_{I_n}(u_t-\mathcal{I}_{Lo}^{n,q}u_t,\varphi_t)dt.$$

We next let $\mathcal{I}^{n,q+1}$ denote the Lagrange interpolation operator at the $q+2$ points of $[t^n,t^{n+1}]$ consisting of the $q+1$ Lobatto points $\xi^{n,0},\dots,\xi^{n,q}$ and any number in $[t^n,t^{n+1}]$ distinct from the above, e.g., the average of any two adjacent Lobatto points. Then, $(\mathcal{I}^{n,q+1}u_t)\varphi_t$ is a polynomial of degree $2q-1$ in t and we thus obtain

$$\int_{I_n}(u_t-\mathcal{I}_{Lo}^{n,q}u_t,\varphi_t)dt = \int_{I_n}(u_t-\mathcal{I}^{n,q+1}u_t,\varphi_t)dt\,.$$

Integrating by parts, we finally get

$$\begin{aligned}|\Gamma_1| &= |\int_{I_n}(\big[u_t-\mathcal{I}^{n,q+1}u_t\big]_t,\varphi)dt|\\ &\le Ck_n^{q+1}\|u^{(q+3)}\|_{L^2(I_n;L^2)}\,\|\varphi\|_{L^2(I_n;L^2)}\,.\end{aligned}$$

Now viewing $\eta_t(t^{n+})$ as a function constant in time,

$$\Gamma_2 = \int_{I_n}\Big(\big(\mathcal{I}_{Lo}^{n,q}\big[\eta_t-\eta_t(t^{n+})\big]\big)_t,\varphi\Big)dt = \int_{I_n}\bigg(\Big(\mathcal{I}_{Lo}^{n,q}\Big[\int_{t^n}^t\eta_{tt}\Big]\Big)_t,\varphi\bigg).$$

Using an H^1-L^2 inverse property (similar to (2.5)), we obtain

$$|\Gamma_2| \le ck_n^{-1}\Big\|\mathcal{I}_{Lo}^{n,q}\Big[\int_{t^n}^t\eta_{tt}\Big]\Big\|_{L^2(I_n;L^2)}\|\varphi\|_{L^2(I_n;L^2)}.$$

Inequality (3.13) now follows from (3.15) and (3.16). The proof of the lemma is complete. □

Stability. Our intention is to derive estimates for $|\!|\!|E|\!|\!| = \big(\|\nabla E_1\|^2+\|E_2\|^2\big)^{1/2}$. For this, we take $\Phi = \mathcal{P}_t^{n,q-1}\mathcal{A}_h^nE\in V_{q-1}\times V_{q-1}$ in (3.9), where

$$\mathcal{P}_t^{n,q-1} = \begin{pmatrix}P_t^{n,q-1} & 0\\ 0 & P_t^{n,q-1}\end{pmatrix}\text{ and }\mathcal{A}_h^nE = \begin{pmatrix}A_h^nE_1\\ E_2\end{pmatrix}.$$

Then clearly,

$$\int_{I_n}((E_t,\Phi))\,dt = \frac12|\!|\!|E^{n+1}|\!|\!|^2 - \frac12|\!|\!|E^{n+}|\!|\!|^2\,. \tag{3.21}$$

On the other hand, the definition of $\mathcal{A}_h E$ and (2.6) imply

$$\int_{I_n} ((\mathcal{L}_h^n E, \mathcal{P}_t^{n,q-1} \mathcal{A}_h^n E))\, dt = \sum_{j=1}^{q} w_{n,j} ((\mathcal{L}_h^n E, \mathcal{A}_h^n E))\,(t^{n,j}) = 0\,. \tag{3.22}$$

As for the right-hand side of (3.9), note that only $\phi_2 = P_t^{n,q-1} E_2$ appears. Since f is Lipschitz, the first term is bounded by $c\|E_1\|_{L^2(I_n;L^2)}\|E_2\|_{L^2(I_n;L^2)}$ and hence by $c\|\nabla E_1\|_{L^2(I_n;L^2)}\|E_2\|_{L^2(I_n;L^2)}$ via Poincaré's inequality. The remaining terms having been estimated in Lemma 3.3 (estimates (3.12)–(3.14)), from (3.21) and (3.22) it follows that for $n = 0, \dots, N-1$,

$$\|E^{n+1}\|^2 \le \|E^{n+}\|^2 + c\|E\|^2_{L^2(I_n;L^2)} + c\Big(k_n^{q+1}\mathcal{E}_t^n + ch_n^r \mathcal{E}_x^n\Big)^2, \tag{3.23}$$

where

$$\mathcal{E}_t^n = \mathcal{E}_t^n(u,q) = \|\ |u^{(q+1)}| + |u^{(q+2)}| + |u^{(q+3)}| + |\nabla u^{(q+2)}| + |\Delta u^{(q+1)}|\ \|_{L^2(I_n;L^2)},$$
$$\mathcal{E}_x^n = \mathcal{E}_x^n(u,r) = \|\ |u| + |u_t| + |u_{tt}|\ \|_{L^2(I_n;H^r)}.$$

We have used this particular presentation of $\mathcal{E}_t^n(u,q)$ and $\mathcal{E}_x^n(u,r)$ for brevity. We have also included the term ∇u^{q+2} in anticipation of the term B_{n-1} that will appear in Lemma 3.5.

We next estimate the term $\|E\|_{L^2(I_n;L^2)}$ in (3.23)

Lemma 3.4. *For any n, $0 \le n \le N-1$, and k_n sufficiently small, it holds that*

$$\|E\|^2_{L^2(I_n;L^2)} \le ck_n\|E^{n+}\|^2 + ck_n\Big(k_n^{q+1}\mathcal{E}_t^n + ch_n^r\mathcal{E}_x^n\Big)^2. \tag{3.24}$$

Proof. Let $\widetilde{E}^{n,j} = \tau_j^{-1/2}E^{n,j}$, $j = 1, \dots, q$. Recalling that $E^{n,0} = E^{n+}$, we have

$$E(x,t) = \sum_{j=0}^{q} \hat{\ell}_{n,j}(t)E^{n,j}(x) = \sum_{j=1}^{q} \hat{\ell}_{n,j}(t)\tau_j^{1/2}\widetilde{E}^{n,j}(x) + \hat{\ell}_{n,0}(t)E^{n+}(x).$$

We then choose in (3.9) $\Phi = \Phi_E := \sum_{i=1}^{q} \ell_{n,i}(t)\tau_i^{-1/2}\mathcal{A}_h^n\widetilde{E}^{n,i}$. As in (3.22)

$$\int_{I_n} ((\mathcal{L}_h^n E, \Phi_E)) = \sum_{j=1}^{q} w_{n,j}\tau_j^{-1}((\mathcal{L}_h^n E^{n,j}, \mathcal{A}_h^n E^{n,j})) = 0\,.$$

Also, using (2.7) and Lemma 2.1

$$\begin{aligned}\int_{I_n} ((E_t, \Phi_E))dt = &\sum_{i,j=1}^{q} \widetilde{m}_{ij}\big[(\nabla\widetilde{E}_1^{n,j}, \nabla\widetilde{E}_1^{n,i}) + (\widetilde{E}_2^{n,j}, \widetilde{E}_2^{n,i})\big] \\ &+ \sum_{i=1}^{q} m_{i0}\tau_i^{-1/2}\big[(\nabla E_1^{n+}, \nabla\widetilde{E}_1^{n,i}) + (E_2^{n+}, \widetilde{E}_2^{n,i})\big] \\ &\ge c\sum_{j=1}^{q} \|\widetilde{E}^{n,j}\|^2 - c\big(\sum_{j=1}^{q} \|\widetilde{E}^{n,j}\|^2\big)^{1/2}\|E^{n+}\|.\end{aligned} \tag{3.25}$$

The quantities (norms) $\sum_{j=1}^{q} \|\widetilde{E}^{n,j}\|^2$ and $\sum_{j=1}^{q} \|E^{n,j}\|^2$ are equivalent modulo constants that depend only on the τ_i's. Similarly, it is easily seen (cf. [KM2]) that the quantities $\sum_{j=0}^{q} |\!|\!| E^{n,j} |\!|\!|^2$ and $|\!|\!| E |\!|\!|^2_{L^2(I_n;L^2)}$ are equivalent in the sense that

$$ck_n \sum_{j=0}^{q} |\!|\!| E^{n,j} |\!|\!|^2 \le |\!|\!| E |\!|\!|^2_{L^2(I_n;L^2)} \le c_2 k_n \sum_{j=0}^{q} |\!|\!| E^{n,j} |\!|\!|^2. \tag{3.26}$$

The terms on the right side of (3.9) with $\Phi = \Phi_E$ can be estimated just as before. This and (3.25), (3.26) imply (3.24). □

We proceed to estimate $|\!|\!| E^{n+} |\!|\!|^2$ which appears both in (3.23) and (3.24). Let $M_n \ge 2$ be a number depending on n which will be specified in the sequel and

$$\beta_n = \frac{\gamma_n}{M_n - 1}, \quad \gamma_n = \begin{cases} 0 & \text{if } S_h^n = S_h^{n-1}, \\ 1 & \text{otherwise,} \end{cases} \qquad n = 1, \dots, N-1.$$

Lemma 3.5. *For any n, $1 \le n \le N-1$, it holds that*

$$|\!|\!| E^{n+} |\!|\!|^2 \le (1 + \beta_n + k_{n-1}) |\!|\!| E^n |\!|\!|^2 + \gamma_n (M_n + k_{n-1}) |\!|\!| J^n |\!|\!|^2 + cB^2_{n-1} \tag{3.27}$$

where

$$|\!|\!| J^n |\!|\!|^2 = \|\nabla(\omega^{n+} - \omega^n)\|^2 + \|\omega_t^{n+} - \omega_t^n\|^2$$

and

$$B_{n-1} \le ck_{n-1}^{q+1} \|\nabla u^{(q+2)}\|_{L^2(I_{n-1};L^2)}.$$

Proof. We first consider the term $\|E_2^{n+}\|$. Then,

$$\begin{aligned} \|E_2^{n+}\| &= \|U_2^{n+} - W_2^{n+}\| = \|P^n E_2^n - P^n[\omega_t^{n+} - \omega_t^n]\| \\ &\le \|E_2^n\| + \|\omega_t^{n+} - \omega_t^n\|. \end{aligned} \tag{3.28}$$

Next for $\|\nabla E_1^{n+}\|$ we have

$$\begin{aligned} \|\nabla E_1^{n+}\| &= \|\nabla(P_E^n U_1^n - W_1^{n+})\| \\ &\le \|\nabla P_E^n (U_1^n - W_1^n)\| + \|\nabla P_E^n (W_1^{n+} - W_1^n)\| \\ &\le \|\nabla(U_1^n - W_1^n)\| + \|\nabla(W_1^{n+} - W_1^n)\|. \end{aligned} \tag{3.29}$$

Now using the definition of W_1 we have

$$\begin{aligned} &\|\nabla(W_1^{n+} - W_1^n)\| \\ &\quad \le \|\nabla(W_1^{n+} - P_E^{n-1} u^n)\| + \|\nabla P_E^{n-1}(u^n - \int_{t^{n-1}}^{t^n} \mathcal{I}_{Lo}^{n-1,q} u_t \, dt - u^{n-1})\| \\ &\quad \le \|\nabla(P_E^n - P_E^{n-1}) u^n\| + \|\nabla(u^n - \int_{t^{n-1}}^{t^n} \mathcal{I}_{Lo}^{n-1,q} u_t \, dt - u^{n-1})\| \\ &\quad = \|\nabla(\omega^{n+} - \omega^n)\| + \| \int_{t^{n-1}}^{t^n} (I - \mathcal{I}_{Lo}^{n-1,q}) \nabla u_t \| \\ &\quad \le \|\nabla(\omega^{n+} - \omega^n)\| + ck_{n-1}^{1/2} B_{n-1}. \end{aligned} \tag{3.30}$$

The result now follows from (3.28)–(3.30) and applications of the arithmetic geometric mean inequality. □

We note here that unlike the term $J[\zeta^n]$, B_{n-1} is nonzero in general even if the spaces S_h^n and S_h^{n-1} are the same.

Remark 3.1. The choice $\Pi^n = \begin{pmatrix} P_E^n & 0 \\ 0 & P^n \end{pmatrix}$ in (1.3) was used in an essential way in Lemma 3.5. Indeed in (3.28) and (3.29) we used the natural stability of P^n and P_E^n in L^2 norm and H^1 seminorm, respectively. A more convenient choice in practice would be to use a Lagrangian type interpolation operator. In that case we would have $U^{n+} = \Pi^n U^n = \begin{pmatrix} \Pi_h^n & 0 \\ 0 & \Pi_h^n \end{pmatrix} U^n$, where Π_h^n is the standard interpolation operator into S_h^n. Then the result of Lemma 3.5 would still be valid provided

$$\text{(A1)} \qquad \|\Pi_h^n E_2^n\| \le (1 + Ck_n)\|E_2^n\| \ \text{ and } \ \|\nabla \Pi_h^n E_1^n\| \le (1 + Ck_n)\|\nabla E_1^n\|.$$

Of course there are no guarantees that these bounds are valid. To retain the convergence result we could impose (A1) as an extra assumption. On the other hand (A1) might be satisfied in practice when the mesh adaptation is performed under a reasonable adaptive strategy. Indeed the influence of Π_h^n is local. The alteration of the mesh will normally consist on the part where we refine and the part where the mesh is coarsened. $\Pi_h^n E_2^n$ differs from E_2^n only in the coarsened area. But a reasonable adaptive algorithm chooses to coarsen the mesh only in areas where the error is well below a given tolerance and without strong variations. Therefore although (A1) is an extraneous condition that cannot be justified a priori, its validity might be within reason in successful adaptive computations.

We are now ready to prove the main convergence result for our scheme.

Theorem 3.1. *Let u and U be the solutions of* (1.1) *and* (1.3)*, respectively. Then*

$$\text{(3.31)} \qquad \max_{t\in[0,T]} \|E\| \le c \sum_{n=0}^{N-1} e^{c(T-t_n)} \Big\{ k_n^{q+1}\mathcal{E}_t^n + h_n^r \mathcal{E}_x^n \Big\} + ce^{cT}\sqrt{\mathcal{N}_C}\, \max_n \|J^n\|,$$

where $\mathcal{N}_C$ denotes the number of times where $S_h^j \ne S_h^{j-1}$, $j = 1, \ldots, N-1$. In addition, (1.4) *and* (1.5) *hold.*

Proof. To begin, note that $E^{0+} = (0, (P^0 - P_E^0)u^1)^T$. Hence, $\|E^{0+}\| \le ch_0^r \mathcal{E}_x^0$. Also, (3.23), (3.24) and (3.27) imply (set $B_{-1} = B_0$ and $E^0 = E^{0+}$)

$$\begin{aligned} \|E^{n+1}\|^2 \le \big(1 + ck_n\big)\Big\{ &(1 + \beta_n + k_{n-1})\|E^n\|^2 + \gamma_n(M_n + k_{n-1})\|J^n\|^2 + cB_{n-1}^2 \Big\} \\ &+ c\big(1 + ck_n\big)\big(k_n^{q+1}\mathcal{E}_t^n + h_n^r\mathcal{E}_x^n\big)^2, \quad n = 0, \ldots, N-1. \end{aligned}$$

A standard Gronwall type argument gives

$$\begin{aligned} \text{(3.32)} \quad \|E^n\|^2 \le c \sum_{m=0}^{n-1} C_{m,n-1} \Big\{ &\big(k_m^{q+1}\mathcal{E}_t^m + h_m^r \mathcal{E}_x^m\big)^2 \\ &+ \gamma_m(M_m + k_{m-1})\|J^m\|^2 \Big\}, \quad n = 1, \ldots, N, \end{aligned}$$

where $C_{m,n-1} = \prod_{j=m}^{n-1}(1 + ck_j)(1 + \beta_j + k_{j-1})$. We shall next estimate these terms.

We fix n and choose $M_m = M = \mathcal{N}_C(n-1)+1$, $m = 1,\dots,n-1$, where $\mathcal{N}_C(n-1)$ denotes the number of times where $S_h^j \neq S_h^{j-1}$, $j=1,\dots,n-1$. Then, $\beta_j = \beta = \frac{1}{M-1}$, whenever $S_h^j \neq S_h^{j-1}$, and $\beta_j = 0$ otherwise. Thus, for $m \le \ell \le n-1$,

$$\begin{aligned}\mathbb{C}_{m,n-1} &\le \prod_{j=m}^{n-1}(1+ck_j) \prod_{\substack{j=m\\ \beta<k_{j-1}}}^{n-1} (1+2k_{j-1}) \prod_{\substack{j=m\\ \beta\ge k_{j-1}}}^{n-1} (1+2\beta)\\ &\le e^{c(t_n-t_m)}\cdot e^{2(t_{n-1}-t_{m-1})}\cdot(1+2\beta)^{M-1} \le e^{c(t_n-t_{m-1})}\cdot e^2.\end{aligned}$$

Using this in (3.32), we obtain

$$\max_{1\le n\le N} |||E^n|||^2 \le c\sum_{n=0}^{N-1} e^{c(T-t_n)}\Big\{k_n^{q+1}\mathcal{E}_t^n + h_n^r\mathcal{E}_x^n\Big\}^2 + ce^{cT}\mathcal{N}_C \max_n |||J^n|||^2. \tag{3.33}$$

Inequality (3.31) now follows from the inverse inequality (2.5), (3.24), (3.27) and (3.33). Now (1.4) follows from (3.31) upon using the triangle inequality (recall that $U-\begin{pmatrix}u\\u_t\end{pmatrix} = E+W-\begin{pmatrix}u\\u_t\end{pmatrix}$), Poincaré's inequality on the first component of E and the consistency estimates of Lemma (3.3) ((3.10) and (3.11) with $p=\infty$). Finally, (1.5) can be obtained by making use of a well-known H^1-L^∞ inverse inequality that holds in 2 dimensions (cf. [Thomée, p. 67]). □

Estimates with local spatial mesh sizes. In the previous estimates we did not work with the local mesh sizes, but rather with the global h. The results are extended in the local mesh form provided we assume that the elliptic projection operator $P_E^n : H_0^1(\Omega) \to S_h^n$ defined in (3.1) satisfies the "local mesh" versions of (3.2), (3.3); i.e., we assume that

$$\|\nabla(v-P_E^n v)\| \le c\|h_n^{s-1}v\|_s, \qquad v\in H^s\cap H_0^1,\ \ 2\le s\le r, \tag{3.34}$$

and

$$\|v-P_E^n v\| \le c\|h_n^s v\|_s, \qquad v\in H^s\cap H_0^1,\ \ 2\le s\le r, \tag{3.35}$$

where c is independent of n. Here we use the "local mesh" notation

$$\|h_n^s v\|_m = \Big\{\sum_{K\in\mathcal{T}_{hn}} h_K^{2s}\|v\|_{m,K}^2\Big\}^{1/2},$$

and $\|v\|_{m,K}$ denotes the restriction of the Sobolev norm to K. Clearly the estimate (3.34) is straightforward. Inequality (3.35) requires more care. Indeed, it is known that (3.35) is valid in one dimension without any assumptions on the mesh; cf. [BO]. In higher dimensions (3.35) is valid under appropriate local quasiuniformity conditions; cf. [EJ] and the references therein and also [BS, Chapter 0] for a discussion of the one-dimensional case that hints at the difficulties of the problem. Essentially (3.35) is used only to bound the "elliptic" consistency terms in Lemma 3.3. Therefore Theorem 3.1 still holds if we replace the $h_n^r\|v\|_r$-like terms by corresponding terms involving the error of the elliptic projection in the L^2-norm. We can thus obtain by entirely similar arguments

Theorem 3.2. *Let u and U be the solutions of* (1.1) *and* (1.3), *respectively. Then*

$$\max_{t\in[0,T]} \|E\| \le c \sum_{n=0}^{N-1} e^{c(T-t_n)} \Big\{ k_n^{q+1} \mathcal{E}_t^n + \mathcal{K}_x^n(h_n^r) \Big\} + ce^{cT} \sqrt{\mathcal{N}_C} \max_n \|J^n\|, \tag{3.36}$$

where $\mathcal{N}_C$ denotes the number of times where $S_h^j \ne S_h^{j-1}$, $j = 1, \ldots, N-1$, and

$$\mathcal{K}_x^n(h_n^r) = \|h_n^r(|u| + |u_t| + |u_{tt}|)\|_{L^2(I_n;H^r)}\,.$$

In addition, the local mesh version of (1.4) holds. The L^∞ estimate requires a local mesh estimate of the elliptic projection in the L^∞ norm:

$$\|v - P_E^n v\|_\infty \le c\|h_n^r v\|_{r,\infty}\,;$$

cf., e.g., [E], [SW] and their references.

4. Estimates under conditions on the mesh

We will assume in this section that the meshes in each time step are generated by a reference (coarse) mesh $\mathcal{T}_h$. In particular starting with $\mathcal{T}_h$ we may want to refine the mesh in a part of the domain after some time steps. Then (since we may want to capture a moving singularity) we may choose to derefine in this area (and thus getting the original partition there) and to refine in a new area. This situation can be summarized by assuming that a fixed space S_h, that corresponds to the reference mesh $\mathcal{T}_h$, is a subspace of all the finite element spaces S_h^n. We denote $h = \max_{K\in\mathcal{T}_h} h_K$, and we assume that

$$\inf_{\varphi\in S_h} \|v - \varphi\|_s \le Ch^{r-s}\|v\|_r, \quad s = 0, 1. \tag{4.1}$$

The main idea in this section is that although U^n lies in a different space S_h^n for each n, we will compare it with a function (denoted again) W such that $W(t) \in S_h$ for all t. For this we split the error as $U - (u, u_t)^T = (U - W) + (W - (u, u_t)^T)$. To define W, we consider the elliptic projection operator $P_E : H_0^1(\Omega) \to S_h$ defined by

$$(\nabla P_E v, \nabla\chi) = (\nabla v, \nabla\chi), \qquad \forall\chi \in S_h. \tag{4.2}$$

Then P_E satisfies the well-known estimates

$$\|\nabla(v - P_E v)\| \le ch^{s-1}\|v\|_s, \quad \|v - P_E v\| \le ch^s\|v\|_s, \quad v \in H^s \cap H_0^1,\ 2 \le s \le r. \tag{4.3}$$

We define $\bar\omega, \bar\eta$ as

$$\bar\omega(x,t) = P_E u(x,t), \quad \bar\eta = u - \bar\omega.$$

If $W = (W_1, W_2)\big|_{I_n}$, we let

$$W_2\big|_{I_n} = \mathcal{I}_{Lo}^{n,q}\bar\omega_t, \quad W_1\big|_{I_n} = \mathcal{I}_{Lo}^{n,q}\Big(\int_{t^n}^t W_2\,dt + \bar\omega^n \Big).$$

(Note that $\bar\omega^{n+} = \bar\omega^n$.) Then the relation

$$\int_{I_n} (W_{1,t}, \varphi)dt = \int_{I_n} (W_2, \varphi)dt \quad \text{for all } \varphi \in V_{q-1}$$

still holds.

The analysis then is the same as in Section 3. The main difference is that in the place of Lemma 3.4 we now have

Lemma 4.1. *It holds that*

$$\|E^{n+}\| \le \|E^n\| + G^n$$

where

$$G^n \le c k_{n-1}^{1/2} k_{n-1}^{q+1} \|\nabla u^{(q+2)}\|_{L^2(I_{n-1};L^2)}.$$

Proof. We first consider the term $\|E_2^{n+}\|$. Then,

$$E_2^{n+} = U_2^{n+} - W_2^{n+} = P^{n+}U_2^n - \bar{\omega}_t^n = P^{n+}(U_2^n - \bar{\omega}_t^n) = P^{n+}E_2^n,$$

since by our assumption $\bar{\omega}_t^n \in S_h \subset S_h^n$. Therefore $\|E_2^{n+}\| \le \|E_2^n\|$. For $\|\nabla E_1^{n+}\|$ we have, noticing again that $W_1(t) \in S_h \subset S_h^n$,

$$\begin{aligned}\|\nabla E_1^{n+}\| &= \|\nabla(P_E^{n+}U_1^n - W_1^{n+})\| \\ &\le \|\nabla P_E^{n+}(U_1^n - W_1^n)\| + \|\nabla(W_1^{n+} - W_1^n)\| \\ &\le \|\nabla(U_1^n - W_1^n)\| + \|\nabla P_E(\int_{t^{n-1}}^{t^n} (I - \mathcal{I}_{Lo}^{n-1,q})u_t)\|.\end{aligned}$$

But then as in Lemma 3.4,

$$\begin{aligned}\|\nabla P_E(\int_{t^{n-1}}^{t^n} (I - \mathcal{I}_{Lo}^{n-1,q})u_t)\| &\le \|(\int_{t^{n-1}}^{t^n} (I - \mathcal{I}_{Lo}^{n-1,q})\nabla u_t)\| \\ &\le k_{n-1}^{1/2} c k_{n-1}^{q+1} \|\nabla u^{(q+2)}\|_{L^2(I_{n-1};L^2)},\end{aligned}$$

which completes the proof. □

Therefore since no jump terms are present in this lemma, by applying the arguments in Section 3, we get

Theorem 4.1. *Let u and U be the solutions of* (1.1) *and* (1.3)*, respectively. Suppose* $S_h \subseteq S_h^n \ \ \forall n$*, where* S_h *satisfies* (4.1)*. Then*

$$\max_{t\in[0,T]} \Big(\|u(t) - U_1(t)\| + \|u_t(t) - U_2(t)\| \Big) \le C\Big\{ \max_m k_m^{q+1} C_t(u) + h^r C_x(u) \Big\}.$$

Remark 4.1. The local mesh size version of the above result is obtained as before by replacing the assumption (4.3) by

$$\|\nabla(v - P_E v)\| \le c\|h^{s-1}v\|_s, \quad \|v - P_E v\| \le c\|h^s v\|_s, \quad v \in H^s \cap H_0^1, \ \ 2 \le s \le r.$$

Then the analog of Theorem 3.2 holds in our case without the jump terms present. Note however that the locality of the estimate in this case is "smeared out" since S_h does not include the moving refined parts of the mesh. On the other hand Theorem 4.1 establishes that in the present important case of mesh modification the optimal order of convergence is preserved.

5. Estimate for the jump of the Riesz projection

In applications, when it is needed to change the spatial mesh at some time level t^n, the two meshes will be in general incompatible but will, more often than not, differ only in a region of small area. The following simple estimates show that the corresponding difference of the Riesz projections involves a factor that depends on the measure of the region of the incompatibility of the two meshes. To define the regions of incompatibility, we will use the following characterization. The set $D_i^n \subset \Omega$ is called the *incompatibility region* for the meshes $\mathcal{T}_h^{n-1}$ and $\mathcal{T}_h^n$ if

$$\text{for all } \varphi \in S_h^{n-1} \cup S_h^n \text{ with supp } \varphi = \Omega \backslash D_i^n \text{ it holds that } \varphi \in S_h^{n-1} \cap S_h^n\,,$$

and there is no other set that is contained in D_i^n with this property. It is clear that the incompatibility region is simply the region where the meshes differ. Also, since we are using conforming elements, a transition layer is needed. This could consist of a layer, one-triangle across, surrounding D_i^n. We denote it by D_ℓ^n and we let $D^n = D_i^n \cup D_\ell^n$.

We will use the following lemma, which follows from [SW, Lemma 2.3].

Lemma 5.1. *There exists a function* $\eta \in S_h^{n-1} \cap S_h^n$, $\eta = J^n := (P_E^n - P_E^{n-1})u(x,t)$ *on* $\Omega \backslash D^n$, *and* $\operatorname{supp} \eta \subset \Omega \backslash D_i^n$, *such that*

$$\|J^n - \eta\|_{1,D_\ell^n} \le c\|J^n\|_{1,D_\ell^n}.$$

We have the following

Proposition 5.1. *If Lemma 5.1 holds, then*

$$\|\nabla J^n\| \le cm(D^n)^{1/2}\|J^n\|_{1,\infty}.$$

If in addition the space $S_h^{n-1} \cap S_h^n$ *has the approximation property*

$$\inf_{\varphi \in S_h^{n-1} \cap S_h^n} \|v - \varphi\|_1 \le ch|v|_2,$$

then

$$\|J^n\| \le cm(D^n)^{1/2}\, h\, \|J^n\|_{1,\infty}.$$

Proof. Let η be as in Lemma 5.1. Since $\eta \in S_h^{n-1} \cap S_h^n$,

$$(\nabla J^n, \nabla \eta) = 0;$$

thus

$$\begin{aligned}
(\nabla J^n, \nabla J^n) &= (\nabla J^n, \nabla J^n)_{\Omega \backslash D_i^n} + (\nabla J^n, \nabla J^n)_{D_i^n} \\
&= (\nabla J^n, \nabla (J^n - \eta))_{\Omega \backslash D_i^n} + (\nabla J^n, \nabla \eta)_{\Omega \backslash D_i^n} + (\nabla J^n, \nabla J^n)_{D_i^n} \\
&= (\nabla J^n, \nabla (J^n - \eta))_{D_\ell^n} + (\nabla J^n, \nabla \eta) + (\nabla J^n, \nabla J^n)_{D_i^n} \\
&\le \|\nabla J^n\|_{D_\ell^n} \|\nabla (J^n - \eta)\|_{D_\ell^n} + \|\nabla J^n\|^2_{D_i^n} \\
&\le c\|J^n\|^2_{1,D^n} \le cm(D^n)\|J^n\|^2_{1,\infty}.
\end{aligned}$$

For the L^2 estimate consider the function $\Psi \in H^2 \cap H_0^1$ that satisfies

$$(\nabla \Psi, \nabla v) = (J_n, v) \qquad \forall v \in H_0^1.$$

Then for any $\Psi_h \in S_h^{n-1} \cap S_h^n$, the approximation property of $S_h^{n-1} \cap S_h^n$ implies

$$\begin{aligned}
\|J^n\|^2 &= (\nabla \Psi, \nabla J^n) \\
&= (\nabla (\Psi - \Psi_h), \nabla J^n) \\
&\le ch\|\nabla J^n\|\, \|\Psi\|_2 \le ch\|\nabla J^n\|\, \|J^n\|,
\end{aligned}$$

and the proof follows. □

Remark 5.1. One may get a similar result by working directly with, e.g., Clement's interpolant, but we do not present this case here.

This result implies that, e.g., in $\mathbb{R}^d$ and if the diameter of D^n is $O(h)$, then

$$\|\nabla (P_E^n - P_E^{n-1})u(x,t)\| = O(h^{r-1+d/2}), \quad \|(P_E^n - P_E^{n-1})u_t(x,t)\| = O(h^{r+d/2}).$$

Note that the approximation assumption on $S_h^{n-1} \cap S_h^n$ is realistic. This is the case for example if S_h^n is obtained from S_h^{n-1} by refining the mesh in a part of the domain and derefining it in another part. This is what we do in most situations.

Acknowledgment

The work of O. Karakashian was partly supported by NSF grant INT-9504845.

References

[AM] A.K. Aziz and P. Monk, *Continuous finite elements in space and time for the heat equation*, Math. Comp. **52** (1989), 255–274. MR **90d:**65189

[BO] I. Babuška and J. E. Osborn, *Analysis of finite element methods for second order boundary value problems using mesh dependent norms*, Numer. Math. **34** (1980), 41–62. MR **81g:**65143

[BB] G. Baker and J. H. Bramble, *Semidiscrete and fully discrete approximations for second order hyperbolic equations*, RAIRO Anal. Numer. **13** (1979), 75–100. MR **80f:**65115

[BL] L. Bales and I. Lasiecka, *Continuous finite elements in space and time for the nonhomogeneous wave equation*, Computers Math. Applic. **27** (1994), 91–102. MR **94k:**65138

[BS] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994. MR **95f:**65001

[Ci] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978. MR **58:**25001

[DR] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, New York, 1975. MR **56:**7119

[D] T. Dupont, *Mesh modification for evolution equations*, Math. Comp. **39** (1982), 85–107. MR **84g:**65131

[E] K. Eriksson, *An adaptive finite element method with efficient maximum norm error control for elliptic problems*, Math. Models Meth. Appl. Sc. **4** (1994), 313–329. MR **95c:**65180

[EJ] 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. MR **96c:**65162

[FP] D. A. French and T. E. Peterson, *A continuous space-time finite element method for the wave equation*, Math. Comp. **65** (1996), 491–506. MR **96g:**65098

[F] D. A. French, *A space-time finite element method for the wave equation*, Comput. Methods Appl. Mech. Engrg. **107** (1993), 145–157. MR **95d:**65082

[HH] T. J. R. Hughes and G. M. Hulbert, *Space-time finite element methods for elastodynamics: Formulations and error estimates*, Comput. Methods Appl. Mech. Engrg. **66** (1988), 339–363. MR **89c:**73060

[J] C. Johnson, *Discontinuous Galerkin finite element methods for second order hyperbolic problems*, Comput. Methods Appl. Mech. Engrg. **107** (1993), 117–129. MR **95c:**65154

[KM1] O. Karakashian and Ch. Makridakis, *A space-time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method*, Math. Comp. **67** (1998), 479–499. MR **98i:**65078

[KM2] O. Karakashian and Ch. Makridakis, *A space-time finite element method for the nonlinear Schrödinger equation: the continuous Galerkin method*, SIAM J. Numer. Anal. **36** (1999), 1779–1807. MR **2000h:**65139

[M] Ch. G. Makridakis, *Finite element approximations of nonlinear elastic waves*, Math. Comp. **61** (1993), 569–594. MR **94g:**73044

[SW] A. H. Schatz and L. B. Wahlbin, *Interior maximum-norm estimates for finite element methods, Part II*, Math. Comp. **64** (1995), 907–928. MR **95j:**65143

[Thomée] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Lecture Notes in Mathematics, vol. 1054, Springer-Verlag, Berlin, 1984. MR **86k:**65006

[Y] D. Q. Yang, *Grid modification for second-order hyperbolic problems*, Math. Comp. **64** (1995), 1495–1509. MR **95m:**65173

Department of Mathematics, University of Tennessee, Knoxville, Tennessee 37966
E-mail address: ohannes@math.utk.edu

Department of Applied Mathematics, University of Crete, 714 09 Heraklion, Crete, Greece; IACM-FORTH, 711 10 Heraklion, Crete, Greece
E-mail address: makr@tem.uoc.gr