SIAM J. SCI. COMPUT.
Vol. 26, No. 5, pp. 1573–1597

NUMERICAL CONSERVATION PROPERTIES OF H(div)-CONFORMING LEAST-SQUARES FINITE ELEMENT METHODS FOR THE BURGERS EQUATION*

H. DE STERCK†, THOMAS A. MANTEUFFEL‡, STEPHEN F. MCCORMICK‡, AND LUKE OLSON§

Abstract. Least-squares finite element methods (LSFEMs) for the inviscid Burgers equation are studied. The scalar nonlinear hyperbolic conservation law is reformulated by introducing the flux vector, or the associated flux potential, explicitly as additional dependent variables. This reformulation highlights the smoothness of the flux vector for weak solutions, namely, $\mathbf{f}(u) \in H(\mathrm{div}, \Omega)$. The standard least-squares (LS) finite element (FE) procedure is applied to the reformulated equations using H(div)-conforming FE spaces and a Gauss–Newton nonlinear solution technique. Numerical results are presented for the one-dimensional Burgers equation on adaptively refined space-time domains, indicating that the H(div)-conforming FE methods converge to the entropy weak solution of the conservation law. The H(div)-conforming LSFEMs do not satisfy a discrete exact conservation property in the sense of Lax and Wendroff. However, weak conservation theorems that are analogous to the Lax–Wendroff theorem for conservative finite difference methods are proved for the H(div)-conforming LSFEMs. These results illustrate that discrete exact conservation in the sense of Lax and Wendroff is not a necessary condition for numerical conservation but can be replaced by minimization in a suitable continuous norm.

Key words. least-squares variational formulation, finite element discretization, Burgers equation, nonlinear hyperbolic conservation laws, weak solutions

AMS subject classifications. 65N15, 65N30, 65N55

DOI. 10.1137/S1064827503430758

1. Introduction.

We consider finite element (FE) methods of least-squares (LS) type [7] for the scalar nonlinear hyperbolic conservation law

$$H(u) := \nabla \cdot \mathbf{f}(u) = 0 \quad \Omega, \qquad u = g \quad \Gamma_I, \tag{1.1}$$

on a domain $\Omega \subset \mathbb{R}^d$ with boundary Γ. The inflow boundary, Γ_I, is the part of the domain boundary where the characteristic curves of (1.1) enter the domain. The outflow boundary, Γ_O, is defined by $\Gamma_O = \Gamma \backslash \Gamma_I$. Variable u is the conserved quantity, $\mathbf{f}(u)$ is the flux vector, and g is the specified boundary data on Γ_I. Flux vector $\mathbf{f}(u)$ is a nonlinear function of u, and we require the components, $f_i(u)$, of $\mathbf{f}(u)$ to be Lipschitz continuous:

$$\exists\, K \quad \text{s.t.} \quad |f_i(u_1) - f_i(u_2)| \;\le\; K\; |u_1 - u_2| \quad \forall\, u_1, u_2, \quad i = 1, \ldots, d. \tag{1.2}$$

*Received by the editors June 30, 2003; accepted for publication (in revised form) June 4, 2004; published electronically April 19, 2005.
http://www.siam.org/journals/sisc/26-5/43075.html

†Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80302. Current address: Department of Applied Mathematics, University of Waterloo, Waterloo N2L 3G1, ON, Canada (hdesterck@uwaterloo.ca).
‡Department of Applied Mathematics, Campus Box 26, University of Colorado at Boulder, Boulder, CO 80302 (tmanteuf@colorado.edu, stevem@colorado.edu).
§Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80302. Current address: Division of Applied Mathematics, Brown University (lolson@dam.brown.edu).

Nonlinear hyperbolic conservation laws allow for weak solutions u that contain discontinuities. It is well known that finite difference schemes that do not satisfy an exact discrete conservation property may converge to a function that is not a weak solution of the conservation law [21, 22, 20, 26]. In [21], Lax and Wendroff show that if finite difference schemes do satisfy an exact discrete conservation property and if they converge to a function boundedly almost everywhere (a.e.), then the function is guaranteed to be a weak solution of the conservation law. The exact discrete conservation property of Lax and Wendroff for the numerical approximation u^h reads as follows:

$$\oint_{\partial\Omega_i} \mathbf{n} \cdot \overline{\mathbf{f}}(u^h)\, dl = 0 \quad \forall\, \Omega_i, \tag{1.3}$$

where $\partial\Omega_i$ is the boundary of computational cell Ω_i, $\mathbf{n}$ is the normal to the boundary, and $\overline{\mathbf{f}}$ is a consistent numerical flux function. It follows immediately that the exact discrete conservation property (1.3) holds for any discrete subdomain Ω_s that is an aggregate of computational cells Ω_i, due to exact numerical flux cancellation at discrete cell interfaces. Since Lax and Wendroff established their important weak conservation result, the exact discrete conservation property (1.3) has de facto been considered a strong requirement for numerical schemes for hyperbolic conservation laws [22, 20, 26], and numerical schemes that satisfy this property have since been called *conservative schemes*.

Hou and LeFloch [17] analyzed the behavior of nonconservative finite difference schemes, and concluded that they can converge to the solution of an inhomogeneous conservation law that contains a Borel source term, thus explaining the deviation from the correct weak solution. On the other hand, the Lax–Wendroff theorem states only that, for a numerical approximation, exact conservation at the discrete level is a *sufficient* condition for convergence to a weak solution and not a *necessary condition*. In this paper, we introduce two new FE methods for the inviscid Burgers equation that are based on LS minimization of the L^2 norm of the conservation law. The first method does not impose an exact discrete conservation property, but we show that if the method converges to a function in the L^2 sense, then it must be a weak solution. This illustrates that exact discrete conservation is not a *necessary condition* for convergence to a weak solution. The second method imposes a pointwise exact conservation property for the flux vector that is, in a sense, *stronger* than the exact discrete conservation property of Lax and Wendroff, and we show convergence to a weak solution for this method as well.

The least-squares finite element methods (LSFEMs) proposed incorporate a novel approach that reformulates the conservation law in terms of the flux vector variable or its associated flux potential. These reformulations enable a choice of FE spaces for the unknown associated with the flux vector that conforms closely to the $H(\mathrm{div})$-smoothness of the flux vector. Hence, we call the resulting FE methods $H(\mathrm{div})$-conforming. Here, $H(\mathrm{div}, \Omega)$ is the Sobolev space of vector functions with square integrable divergence. The LS approach to be presented can in principle be applied to general domains in multiple spatial dimensions and can be combined with a time-marching strategy, but here we choose to restrict our methods to space-time domains with one spatial dimension. In the one-dimensional space-time setting, $\Omega \subset \mathbb{R}^2$ in (1.1), $\nabla = (\partial_x, \partial_t)$, and g in (1.1) comprises both initial and boundary conditions on the inflow boundary of the space-time domain. We present numerical results for the one-dimensional inviscid Burgers equation, for which the generalized flux vector is given by $\mathbf{f}(u) = (u^2/2, u)$.

The numerical results demonstrate convergence, for the Burgers equation, of our $H(\mathrm{div})$-conforming LSFEMs to a function, $\hat{u}$, in the L^2 sense. The numerical results also indicate that $\hat{u}$ is an entropy weak solution of the conservation law (1.1). We prove weak conservation properties of our methods, namely, that if u^h converges to $\hat{u}$, then $\hat{u}$ is a weak solution of (1.1). This theoretical result for LSFEMs is equivalent to the Lax–Wendroff conservation result for numerical schemes that satisfy the exact discrete conservation property (1.3) [21]. The FE convergence of our LSFEMs, namely, the L^2 convergence of u^h to $\hat{u}$, and the equivalence of our LS potential method to an H^{-1} formulation, will theoretically be analyzed elsewhere (see [23] for a discussion in the context of linear hyperbolic PDEs).

Our $H(\mathrm{div})$-conforming LSFEM approach is clearly different from the numerical methods that are typically considered for nonlinear hyperbolic conservation laws. Finite volume methods (see, e.g., [22, 20, 26] and references therein), and more recently also streamline-upwind Petrov–Galerkin (SUPG) and discontinuous Galerkin (DG) FE methods (see, e.g., [20, 11] and references therein), have turned out to be highly successful for simulating hyperbolic conservation laws. These techniques generally use approximate Riemann solvers that are based on upwind discretization ideas combined with nonlinear flux limiters, and have reached high levels of sophistication. Both finite volume methods and DG FE methods satisfy the discrete conservation property of Lax and Wendroff [11]. The LSFEM approach [7] finds the optimal solution within the FE space, measured in the L^2 operator norm, and has several attractive properties. It produces symmetric positive definite (SPD) linear systems, which are well suited for iterative methods, and it provides a natural, sharp a posteriori error estimator which can be used for efficient adaptive refinement. LSFEMs are widely used for elliptic PDEs [7] but have only recently been introduced for hyperbolic PDEs [5, 6, 18, 14]. They have not been applied to the $H(\mathrm{div})$-conforming reformulation we introduce in this paper, and the weak conservation properties of LSFEMs have not been analyzed theoretically.

This paper is structured as follows. Relevant Sobolev space notation is introduced first. In the next section, the $H(\mathrm{div})$-smoothness of the flux vector is discussed in the context of weak solutions, and an $H(\mathrm{div})$-conforming reformulation of nonlinear hyperbolic conservation laws is presented. In section 3, the $H(\mathrm{div})$-conforming reformulation of the conservation law is posed as an LS minimization problem. Numerical results for the Burgers equation presented in section 4 indicate convergence of the $H(\mathrm{div})$-conforming LSFEMs to an entropy weak solution of the conservation law. In section 5, weak conservation theorems are proved for the $H(\mathrm{div})$-conforming LSFEMs that are equivalent to the Lax–Wendroff theorem for numerical schemes that satisfy an exact discrete conservation property [21]. In the final section of the paper we formulate our conclusions and point to future work.

1.1. Nomenclature. Given domain $\Omega \subset \mathbb{R}^2$, denote by $C^1(\overline{\Omega})$ the space of functions that are continuously differentiable on the closure of Ω. The space of bounded functions on Ω is denoted by $L^\infty(\Omega)$. The space of square integrable functions on Ω is denoted by $L^2(\Omega)$, with the associated L^2 inner product of functions u and v given by $\langle u, v\rangle_{0,\Omega}$ and the L^2 norm of function u given by $\|u\|_{0,\Omega}$. Sobolev space $H^1(\Omega)$ consists of L^2 functions that have square integrable partial derivatives of first order, with associated inner product $\langle u, v\rangle_{1,\Omega}$, norm $\|u\|_{1,\Omega}$, and seminorm $|u|_{1,\Omega}$. We also consider Sobolev spaces $H^s(\Omega)$ of fractional and negative orders (i.e., $s \in \mathbb{R}$), with associated inner product $\langle u, v\rangle_{s,\Omega}$ and norm $\|u\|_{s,\Omega}$ (see [2]). All of these function spaces are analogously defined on domain boundary Γ, or any part thereof.

The Sobolov space of square integrable vector functions with square integrable divergence is defined by

$$H(\mathrm{div},\Omega) := \{\mathbf{w} \in L^2(\Omega)^2 \mid \nabla \cdot \mathbf{w} \in L^2(\Omega)\}. \tag{1.4}$$

The $H(\mathrm{div})$ norm of vector function $\mathbf{w}$ is defined by

$$\|\mathbf{w}\|^2_{div,\Omega} := \|\mathbf{w}\|^2_{0,\Omega} + \|\nabla \cdot \mathbf{w}\|^2_{0,\Omega}. \tag{1.5}$$

Additional function spaces with specific boundary conditions are defined below as needed. Throughout the paper, c is a generic constant that may differ at every occurrence.

2. Flux vector and flux potential reformulations of nonlinear hyperbolic conservation laws.

2.1. Smoothness of the generalized flux vector. We first discuss $H(\mathrm{div})$-smoothness of the flux vector in the context of weak solutions of (1.1). Classical solutions of (1.1) are functions $u \in C^1(\overline{\Omega})$ that satisfy (1.1) pointwise. A weak solution of (1.1) is defined as follows.

DEFINITION 2.1 (weak solution). *Assume that $u \in L^\infty(\Omega)$. Then u is a weak solution of* (1.1) *if and only if*

$$-\langle \mathbf{f}(u), \nabla\phi\rangle_{0,\Omega} + \langle \mathbf{n}\cdot\mathbf{f}(g), \phi\rangle_{0,\Gamma_I} = 0 \quad \forall\ \phi\ \in\ C^1_{\Gamma_O}(\overline{\Omega}),$$

where $C^1_{\Gamma_O}(\overline{\Omega}) = \{\phi \in C^1(\overline{\Omega}) : \phi = 0 \text{ on } \Gamma_O\}$.

Remark 2.2. Definition 2.1 is a slight generalization of the weak solution definition for the Cauchy problem as given in [16], generalized to a domain with inflow boundary Γ_I and outflow boundary Γ_O.

Following Godlewski and Raviart [16], we restrict our consideration of weak solutions to the class of so-called piecewise C^1 functions. These functions are C^1, except at a finite number of smooth curves across which the functions, or their derivatives, may have jump discontinuities. This case covers many problems of practical interest. Note that these weak solutions $u \in H^{1/2-\epsilon}(\Omega)$ for all $\epsilon > 0$. As a consequence of Lipschitz continuity (1.2), the components of the flux vector, $f_i(u), i = 1, 2$, are in $H^{1/2-\epsilon}(\Omega)$ as well. We also require that boundary data g for (1.1) be in $H^{1/2-\epsilon}(\Gamma_I)$.

Weak solutions of (1.1) that are piecewise C^1 may be characterized in terms of the smoothness of the generalized flux vector and Sobolev space $H(\mathrm{div})$. This can easily be established formally using the following theorem from [16], and a classical lemma that characterizes piecewise C^1 vector functions of $H(\mathrm{div})$ in terms of continuity of the vector component normal to any smooth curve.

THEOREM 2.3 (see [16, Theorem 2.1, p. 16]). *Assume that $u \in L^\infty(\Omega)$ is a piecewise C^1 function. Then u is a weak solution of* (1.1) *if and only if*

1. *u satisfies* (1.1) *pointwise away from the curves of discontinuity and*
2. *the Rankine–Hugoniot relation $[\mathbf{f}(u)]_\Gamma \cdot \mathbf{n} = 0$ is satisfied along the curves of discontinuity Γ.*

Here, $\mathbf{n}$ is a unit normal on curve Γ, and $[\mathbf{w}]_\Gamma \cdot \mathbf{n}$ is a difference in the normal vector components across Γ.

LEMMA 2.4. *Assume that the vector components of $\mathbf{w}$ are piecewise C^1 functions. Then $\mathbf{w} \in H(\mathrm{div},\Omega)$ if and only if $[\mathbf{w}]_\Gamma \cdot \mathbf{n} = 0$ a.e. on any smooth curve, Γ, in Ω.*

Proof. The proof follows directly from the definition

$$\nabla \cdot \mathbf{w} = \lim_{V \to 0} \frac{\oint \mathbf{w} \cdot \mathbf{n} \, dl}{V},$$

applied to a small volume along any smooth curve, Γ, in Ω. ∎

The following theorem then allows us to characterize the smoothness of the generalized flux vector, $\mathbf{f}(u)$, for a piecewise C^1 weak solution, u, of (1.1), namely, that $\mathbf{f}(u) \in H(\mathrm{div}, \Omega)$.

THEOREM 2.5 ($H(\mathrm{div})$-smoothness of the generalized flux vector). *Assume that $u \in L^\infty(\Omega)$ is a piecewise C^1 function. Then u is a weak solution of* (1.1) *if and only if $\|\nabla \cdot \mathbf{f}(u)\|^2_{0,\Omega} = 0$ (which implies $\mathbf{f}(u) \in H(\mathrm{div}, \Omega)$) and $\|u - g\|^2_{0,\Gamma_I} = 0$.*

Proof. $\boxed{\Rightarrow}$ Assume that u is a piecewise C^1 weak solution. It follows from Theorem 2.3 and Lemma 2.4 that $\mathbf{f}(u) \in H(\mathrm{div})$, and then also that $\|\nabla \cdot \mathbf{f}(u)\|^2_{0,\Omega} = 0$ and $\|u - g\|^2_{0,\Gamma_I} = 0$.

$\boxed{\Leftarrow}$ Conversely, $\|\nabla \cdot \mathbf{f}(u)\|^2_{0,\Omega} = 0$ implies that $\mathbf{f}(u) \in H(\mathrm{div})$ and, according to Lemma 2.4, this means that $[\mathbf{f}(u)]_\Gamma \cdot \mathbf{n} = 0$ a.e. on any smooth curve, including along surfaces of discontinuity. Equalities $\|\nabla \cdot \mathbf{f}(u)\|^2_{0,\Omega} = 0$ and $\|u - g\|^2_{0,\Gamma_I} = 0$ also imply that u satisfies (1.1) pointwise away from the surfaces of discontinuity. Then, according to Theorem 2.3, u is a weak solution of (1.1). ∎

Remark 2.6. Condition $[\mathbf{w}]_\Gamma \cdot \mathbf{n} = 0$ of Lemma 2.4, with $\mathbf{w} = \mathbf{f}(u)$ and Γ the shock curve, corresponds to the classical Rankine–Hugoniot relation. As an example, consider a straight shock with speed s, shock normal direction in the x,t-plane $\mathbf{n} = (1, -s)$, and generalized flux vector $\mathbf{f}(u) = (f(u), u)$, with $f(u)$ the usual spatial flux function. Then condition $[\mathbf{f}(u)]_{shock} \cdot \mathbf{n} = 0$ gives $[(f(u), u)]_{shock} \cdot (1, -s) = 0$, or $[f(u)]_{shock} = s\,[u]_{shock}$, which is the well-known Rankine–Hugoniot relation [22].

Remark 2.7. The special case of intersecting shocks is also covered by Theorem 2.5. At the shock intersection point, $\nabla \cdot \mathbf{f}(u)$ is not defined pointwise, and a Rankine–Hugoniot condition is not satisfied because the shock normal is not defined. However, $\mathbf{f}(u) \in H(\mathrm{div}, \Omega)$ and $\|\nabla \cdot \mathbf{f}(u)\|^2_{0,\Omega} = 0$ still hold.

2.2. Flux vector reformulation. We reformulate (1.1) in terms of the flux vector variable $\mathbf{w}$ as

$$F(\mathbf{w}, u) := \begin{bmatrix} \nabla \cdot \mathbf{w} \\ \mathbf{w} - \mathbf{f}(u) \end{bmatrix} = 0 \qquad \Omega, \tag{2.1}$$
$$\mathbf{n} \cdot \mathbf{w} = \mathbf{n} \cdot \mathbf{f}(g) \quad \Gamma_I,$$
$$u = g \qquad \Gamma_I.$$

2.3. Flux potential reformulation. Letting $\nabla^\perp = (-\partial_t, \partial_x)$, note that $\|\nabla \cdot \mathbf{f}(u)\|_{0,\Omega} = 0$ implies $\mathbf{f}(u) = \nabla^\perp \psi$ for some $\psi \in H^1(\Omega)$. We can thus rewrite (1.1) as

$$G(\psi, u) := \nabla^\perp \psi - \mathbf{f}(u) = 0 \qquad \Omega, \tag{2.2}$$
$$\mathbf{n} \cdot \nabla^\perp \psi = \mathbf{n} \cdot \mathbf{f}(g) \quad \Gamma_I,$$
$$u = g \qquad \Gamma_I.$$

Function ψ is the flux potential associated with the divergence-free generalized flux vector. This approach bears similarity to potential formulations that are used in fluid dynamics and plasma physics.

Remark 2.8. The flux potential formulation can be generalized to dimensions higher than $d = 2$. Scalar potential ψ would need to be replaced by a vector potential, which is the preimage of the divergence-free generalized flux vector, $\mathbf{f}(u)$, in the de

Rham-diagram [1, 4, 13]. The de Rham-diagram is also instructive for the derivation of conforming vector FE spaces in higher dimensions [4].

Remark 2.9. Both reformulations (2.1) and (2.2) explicitly bring to the forefront the $H(\mathrm{div})$-smoothness of the generalized flux function and allow us to choose FE spaces that closely match this smoothness, as discussed in the next section. Also, the differential part of each reformulated equation is linear, with the nonlinearity shifted to the algebraic part of the equation where it may be treated more easily.

2.4. Boundedness of the Fréchet derivative operators for the reformulations. We solve the nonlinear systems (2.1) and (2.2) using a Newton approach applied to the L^2 norm of the nonlinear systems. In the Newton procedure, a general nonlinear system, $T(v) = 0$, is solved by solving a sequence of linearized equations:

$$T(v_i) + T'(v_i)(v_{i+1} - v_i) = 0, \tag{2.3}$$

where $T'(v_i)(r)$ is the Fréchet derivative at iterate v_i in direction r. It is well known that if $\|T'(v^*)\| = \infty$ at the solution, v^*, of $T(v) = 0$, then the basin of attraction of Newton's method may be only $\{v^*\}$. Convergence proofs for Newton's method usually assume that $T'(v)$ is Lipschitz continuous in a neighborhood of v^* [12]. An illustrative example is the scalar algebraic function $s(x) = |x|^{1/3}$, for which Newton's iteration is $x_{i+1} = -2\,x_i$ and the basin of attraction is just $\{0\}$. An unbounded Fréchet derivative at solution v^* implies that the function cannot be represented by a linear approximation around v^*. It is said that the function is not linearizable around v^* in this case. In what follows, we show that reformulated equations (2.1) and (2.2) are linearizable around discontinuous solutions, whereas the original formulation, (1.1), is not.

For system (2.1), the Fréchet derivative of nonlinear operator $F(\mathbf{w}, u)$ at initial guess $(\mathbf{w}_0, u_0)$ in direction $(\mathbf{w}_1 - \mathbf{w}_0, u_1 - u_0)$ is given by

$$F'(\mathbf{w}_0, u_0)(\mathbf{w}_1 - \mathbf{w}_0, u_1 - u_0) = \begin{bmatrix} \nabla\cdot & 0 \\ I & -\mathbf{f}'(u_0) \end{bmatrix} \cdot \begin{bmatrix} \mathbf{w}_1 - \mathbf{w}_0 \\ u_1 - u_0 \end{bmatrix}. \tag{2.4}$$

LEMMA 2.10. *Fréchet derivative operator* $F'(\mathbf{w}_0, u_0) : H(\mathrm{div}, \Omega) \times L^2(\Omega) \to L^2(\Omega)$ *is bounded:* $\| F'(\mathbf{w}_0, u_0) \|_{0,\Omega} \le \sqrt{1 + K^2}$, *where* K *is the Lipschitz constant of* (1.2), *and the norm on* $H(\mathrm{div}, \Omega) \times L^2(\Omega)$ *is given by* $\|(\mathbf{w}, u)\| = \max(\|\mathbf{w}\|_{div,\Omega}, \|u\|_{0,\Omega})$.

Proof.

$$\begin{aligned}
\| F'(\mathbf{w}_0, u_0) \|^2_{0,\Omega} &= \sup_{\substack{\mathbf{b} \in H(div,\Omega), v \in L^2(\Omega) \\ \|(\mathbf{b},v)\|=1}} \| F'(\mathbf{w}_0, u_0)(\mathbf{b}, v) \|^2_{0,\Omega} \\
&= \sup_{\substack{\mathbf{b} \in H(div,\Omega), v \in L^2(\Omega) \\ \|(\mathbf{b},v)\|=1}} \| \nabla \cdot \mathbf{b} \|^2_{0,\Omega} + \| \mathbf{b} - \mathbf{f}'(u_0)v \|^2_{0,\Omega} \\
&\le \sup_{\substack{\mathbf{b} \in H(div,\Omega), v \in L^2(\Omega) \\ \|(\mathbf{b},v)\|=1}} \| \nabla \cdot \mathbf{b} \|^2_{0,\Omega} + \| \mathbf{b} \|^2_{0,\Omega} + K^2 \| v \|^2_{0,\Omega} \\
&= 1 + K^2.
\end{aligned}$$

Here we used

$$|\mathbf{f}'(u_0)v| \le K|v| \quad \text{a.e.},$$

which follows directly from Lipschitz continuity (1.2). ∎

For system (2.2), the Fréchet derivative of nonlinear operator $G(\psi, u)$ at initial guess (ψ_0, u_0) in direction $(\psi_1 - \psi_0, u_1 - u_0)$ is given by

$$G'(\psi_0, u_0)(\psi_1 - \psi_0, u_1 - u_0) = \begin{bmatrix} \nabla^\perp & -\mathbf{f}'(u_0) \end{bmatrix} \cdot \begin{bmatrix} \psi_1 - \psi_0 \\ u_1 - u_0 \end{bmatrix}. \tag{2.5}$$

LEMMA 2.11. *Fréchet derivative operator $G'(\psi_0, u_0) : H^1(\Omega) \times L^2(\Omega) \to L^2(\Omega)$ is bounded: $\| G'(\psi_0, u_0) \|_{0,\Omega} \le \sqrt{1 + K^2}$, where K is the Lipschitz constant of* (1.2)*, and the norm on $H^1(\Omega) \times L^2(\Omega)$ is given by $\|(\psi, u)\| = \max(\|\psi\|_{1,\Omega}, \|u\|_{0,\Omega})$.*

Proof. The proof is analogous to the proof of Lemma 2.10. ∎

Remark 2.12. The Fréchet derivative of conservation law operator $H(u)$ in (1.1) at initial guess u_0 is given by

$$H'(u_0)(v) = \nabla \cdot (\mathbf{f}'(u_0)\, v). \tag{2.6}$$

Fréchet derivative operator $H'(u_0) : H^{1/2-\epsilon}(\Omega) \to L^2(\Omega)$, with $u_0 \in H^{1/2-\epsilon}(\Omega)$, is unbounded for most cases of flux vector $\mathbf{f}(u)$ and initial guess u_0. Indeed, in most cases

$$\begin{aligned} \| H'(u_0) \|_{0,\Omega} &= \sup_{\substack{v \in H^{1/2-\epsilon}(\Omega) \\ \|v\|_{1/2-\epsilon,\Omega} = 1}} \| H'(u_0)(v) \|_{0,\Omega} \\ &= \sup_{\substack{v \in H^{1/2-\epsilon}(\Omega) \\ \|v\|_{1/2-\epsilon,\Omega} = 1}} \| \nabla \cdot (\mathbf{f}'(u_0)\, v) \|_{0,\Omega} \\ &= \infty, \end{aligned}$$

because, in general, $(\mathbf{f}'(u_0)\, v) \notin H(\mathrm{div})$ for $u_0, v \in H^{1/2-\epsilon}(\Omega)$. For example, it is easy to see that for the Burgers equation, with $\mathbf{f}(u) = (u^2/2, u)\ \forall\ u_0 \in H^{1/2-\epsilon}(\Omega)$, for almost every $v \in H^{1/2-\epsilon}(\Omega)$ we have $(\mathbf{f}'(u_0)\, v) \notin H(\mathrm{div})$. In the appendix, it is shown that a Newton LSFEM based directly on (1.1) does not converge.

3. $H(\mathrm{div})$-conforming LSFEMs.

In this section, we describe how the general least-squares finite element methodology [7] can be applied to the reformulations of (1.1) presented above.

3.1. $H(\mathrm{div})$-conforming LSFEM.

The LSFEM for solving (2.1) consists of minimizing LS functional

$$\mathcal{F}(\mathbf{w}, u; g) = \|\nabla \cdot \mathbf{w}\|^2_{0,\Omega} + \|\mathbf{w} - \mathbf{f}(u)\|^2_{0,\Omega} + \|\mathbf{n} \cdot (\mathbf{w} - \mathbf{f}(g))\|^2_{0,\Gamma_I} + \|u - g\|^2_{0,\Gamma_I} \tag{3.1}$$

over finite-dimensional subspaces $\mathcal{W}^h \times \mathcal{U}^h \subset H(\mathrm{div}, \Omega) \times L^2(\Omega)$. Let

$$(\mathbf{w}^h_*, u^h_*) = \underset{\mathbf{w}^h \in \mathcal{W}^h, u^h \in \mathcal{U}^h}{\arg\min} \mathcal{F}(\mathbf{w}^h, u^h; g). \tag{3.2}$$

We treat (3.2) by linearizing equations (2.1) around $(\mathbf{w}^h_0, u^h_0)$, minimizing an LS functional based on the linearized equations, and iteratively repeating this procedure with $(\mathbf{w}^h_0, u^h_0)$ replaced by the newly obtained approximations until convergence. This approach is, in general, called the Gauss–Newton technique for nonlinear LS minimization [12]. The resulting weak equations are given in the following problem statement.

Problem 3.1 (Gauss–Newton $H(\mathrm{div})$-conforming LSFEM). Given $(\mathbf{w}_0^h, u_0^h) \in \mathcal{W}^h \times \mathcal{U}^h$, find $\mathbf{w}^h \in \mathcal{W}^h$ and $u^h \in \mathcal{U}^h$ s.t.

$$\begin{aligned}\langle \nabla\cdot\mathbf{w}^h, \nabla\cdot\mathbf{v}^h\rangle_{0,\Omega} + \langle \mathbf{w}^h - \mathbf{f}(u_0^h) - \mathbf{f}'(u_0^h)(u^h - u_0^h), \mathbf{v}^h\rangle_{0,\Omega} \\ + \langle \mathbf{n}\cdot(\mathbf{w}^h - \mathbf{f}(g)), \mathbf{n}\cdot\mathbf{v}^h\rangle_{0,\Gamma_I} = 0 \qquad \forall\ \mathbf{v}^h \in \mathcal{W}^h, \\ \langle \mathbf{w}^h - \mathbf{f}(u_0^h) - \mathbf{f}'(u_0^h)(u^h - u_0^h), -\mathbf{f}'(u_0^h)(s^h)\rangle_{0,\Omega} \\ + \langle u^h - g, s^h\rangle_{0,\Gamma_I} = 0 \qquad \forall\ s^h \in \mathcal{U}^h.\end{aligned}$$

A full Newton approach could alternatively be considered, in which nonlinear weak equations are derived from minimization of the nonlinear LS functional, followed by a linearization of the nonlinear weak equations [12]. However, we consider only the Gauss–Newton approach here for simplicity.

For the FE spaces, we choose Raviart–Thomas vector FEs of lowest order (RT_0) [9, 4] on quadrilaterals for the flux vector variable $\mathbf{w}^h$ and standard continuous bilinear FEs for u^h [9]. The RT_0 vector FEs, which have continuous normal vector components at element edges, are $H(\mathrm{div})$-conforming: $RT_0 \subset H(\mathrm{div}, \Omega)$.

3.2. Potential $H(\mathrm{div})$-conforming LSFEM. The LSFEM for solving (2.2) consists of minimizing LS functional

$$\mathcal{G}(\psi, u; g) := \|\nabla^\perp\psi - \mathbf{f}(u)\|_{0,\Omega}^2 + \|\mathbf{n}\cdot(\nabla^\perp\psi - \mathbf{f}(g))\|_{0,\Gamma_I}^2 + \|u - g\|_{0,\Gamma_I}^2 \tag{3.3}$$

over finite-dimensional subspaces $\Psi^h \times \mathcal{U}^h \subset H^1(\Omega) \times L^2(\Omega)$. Let

$$(u_*^h, \psi_*^h) = \underset{\psi^h\in\Psi^h, u^h\in\mathcal{U}^h}{\arg\min}\ \mathcal{G}(\psi^h, u^h; g). \tag{3.4}$$

Using the same Gauss–Newton approach as above, the resulting weak equations are given in the following problem statement.

Problem 3.2 (potential Gauss–Newton $H(\mathrm{div})$-conforming LSFEM). Given $(\psi_0^h, u_0^h) \in \Psi^h \times \mathcal{U}^h$, find $\psi^h \in \Psi^h$ and $u^h \in \mathcal{U}^h$ s.t.

$$\begin{aligned}\langle \nabla^\perp\psi^h - \mathbf{f}(u_0^h) - \mathbf{f}'(u_0^h)(u^h - u_0^h), \nabla^\perp\phi^h\rangle_{0,\Omega} \\ + \langle \mathbf{n}\cdot(\nabla^\perp\psi^h - \mathbf{f}(g)), \mathbf{n}\cdot\nabla^\perp\phi^h\rangle_{0,\Gamma_I} = 0 \qquad \forall\ \phi^h \in \Psi^h, \\ \langle \nabla^\perp\psi^h - \mathbf{f}(u_0^h) - \mathbf{f}'(u_0^h)(u^h - u_0^h), -\mathbf{f}'(u_0^h)(s^h)\rangle_{0,\Omega} \\ + \langle u^h - g, s^h\rangle_{0,\Gamma_I} = 0 \qquad \forall\ s^h \in \mathcal{U}^h.\end{aligned}$$

For the FE spaces, we choose standard continuous bilinear FEs on quadrilaterals for both ψ^h and u^h [9].

Remark 3.3. Note that the divergence-free subspace of RT_0 is spanned by $\nabla^\perp\psi^h$, where ψ^h is a bilinear function. The two $H(\mathrm{div})$-conforming methods are thus closely related.

Remark 3.4. The flux vector approximation, $\nabla^\perp\psi^h$, is pointwise a.e. divergence-free on every grid. In this sense, the potential $H(\mathrm{div})$-conforming method imposes an exact discrete conservation constraint that is stronger than Lax–Wendroff conservation condition (1.3). However, flux vector approximation $\mathbf{f}(u^h)$ does not satisfy this pointwise exact discrete conservation property because $\mathbf{f}(u^h) = \nabla^\perp\psi^h$ is only weakly enforced.

Remark 3.5. For the Burgers equation in space-time domains, $\nabla^\perp\psi = (-\partial_t\psi, \partial_x\psi) = (f(u), u)$, which means that $-\partial_t\psi^h$ is a numerical approximation

of $f(u)$, and $\partial_x \psi^h$ is a numerical approximation of u. For these approximations, $\nabla \cdot \nabla^\perp \psi^h = 0$ pointwise a.e. So if one insists on having an approximation for u that is strictly conservative in a discrete sense, then $\partial_x \psi^h$ provides such an approximation, together with the approximation $-\partial_t \psi^h$ for $f(u)$. For the approximation u^h, $\nabla \cdot (f(u^h), u^h) = 0$ does not hold in any discrete sense.

Remark 3.6. It is important to note that care should be taken in choosing the FE spaces for Problems 3.1 and 3.2. This is related to the fact that the LS functionals, $\mathcal{F}(\mathbf{w}, u; 0)$ in (3.1) and $\mathcal{G}(\psi, u; 0)$ in (3.3), do not bound $\|u\|_{0,\Omega}$ and are thus not uniformly coercive w.r.t. $\|u\|_{0,\Omega}$ (see [23]). As shown in section 5, the functionals are coercive w.r.t. the H^{-1} norm of $\nabla \cdot \mathbf{f}(u)$, but this fact does not imply L^2 convergence of u^h, as $\|\nabla \cdot \mathbf{f}(u)\|_{-1,\Gamma_O,\Omega}$ does not bound $\|u\|_{0,\Omega}$. For example, the choice of piecewise constant elements for u^h results in solutions with high-frequency oscillations. When a continuous piecewise linear space is chosen for u^h, however, the high-frequency oscillations are eliminated, and L^2 convergence of u^h is obtained for the Burgers equation. For this choice of FE spaces, the functional bounds the L^2 error. Numerical results indicating convergence of the methods for the Burgers equation are presented in section 4, but a theoretical convergence proof remains open. In future work, we plan to investigate whether the LSFEMs proposed in this paper for the Burgers equation can be extended to more general hyperbolic conservation laws. FE convergence of our LSFEMs applied to general conservation law systems, namely, L^2 convergence of u^h to a function $\hat{u}$, remains a topic for further study. In the present paper, we limit theoretical developments to the numerical conservation properties of methods (3.1) and (3.2) for nonlinear problems, as discussed in section 5.

3.3. Error estimator and adaptive refinement. LSFEMs provide a natural a posteriori local error estimator given by the element contribution to the functional. For example, for a linear PDE, $Lu = f$, the value of the functional can be rewritten in terms of error $e^h = u^h - u_*$, with u^h the current approximation and u_* the exact solution, as follows:

$$\begin{aligned} \mathcal{F}(u^h; f) &= \|Lu^h - f\|_{0,\Omega}^2 \\ &= \|Lu^h - Lu_*\|_{0,\Omega}^2 \\ &= \|L(u^h - u_*)\|_{0,\Omega}^2 \\ &= \|Le^h\|_{0,\Omega}^2. \end{aligned} \tag{3.5}$$

The LS functional value thus gives a local a posteriori error estimator that can be used for adaptive refinement. The numerical convergence results presented in section 4 for the Burgers equation indicate that functional (3.5) bounds the L^2 error, s.t. adaptive refinement based on the functional also controls L^2 error. See [3] for a detailed discussion on the sharpness of the error estimator and on adaptive refinement strategies. In the numerical simulation results below, we apply adaptive refinement to Burgers flow solutions in a global space-time domain. We start out with the whole space-time domain covered by a single FE cell and successively refine the grid. The cells with functional density above a certain user-defined threshold are marked for refinement in the next level. The functional density is the functional value over an element divided by the area. Quadrilateral cells that need to be refined are divided into four smaller quadrilateral cells. We base the error estimation on the full nonlinear functional, which can easily be evaluated. We refine after Newton convergence is achieved on a given level, so that the value of the nonlinear functional is close to

the value of the functional of the linearized equations. Adaptive refinement based on a functional density strategy is inexpensive, straightforward to implement, and amenable to parallelism. Although we achieve good performance using this approach, a detailed study of the adaptive algorithm is beyond the scope of this paper. A perhaps more optimal refinement strategy is outlined in [3]. Also, the results reported below are preliminary in that we do not attempt to separate the locally generated error from the error propagating along the characteristic curve. The performance we see is, nevertheless, promising.

4. Numerical results. Here, we present numerical results for our $H(\mathrm{div})$-conforming Gauss–Newton LSFEMs on test problems for the Burgers equation involving shocks and rarefaction waves. We investigate the solution quality in terms of smearing, oscillations, and overshoots and undershoots at discontinuities. We numerically study L^2 convergence of u^h to a function $\hat{u}$, and convergence of nonlinear functionals $\mathcal{F}(\mathbf{w}^h, u^h; g)$ and $\mathcal{G}(\psi^h, u^h; g)$. That is, we confirm that $\|u^h - u\|_{0,\Omega}^2 \to 0$, $\mathcal{F}(\mathbf{w}^h, u^h; g) \to 0$, and $\mathcal{G}(\psi^h, u^h; g) \to 0$ as $h \to 0$, and we estimate the rate of convergence. We denote the rates of convergence of the square of the L^2 error and the nonlinear functionals by α. For example, for the squared L^2 error of approximation u^h, we assume

$$\|u^h - u\|_{0,\Omega}^2 \approx \mathcal{O}(h^\alpha). \tag{4.1}$$

This implies that α can be approximated between successive levels of refinement by

$$\frac{\|u^h - u\|^2}{\|u^{2h} - u\|^2} \approx \left(\frac{1}{2}\right)^\alpha. \tag{4.2}$$

The theoretical optimal convergence rate of the squared L^2 error for solutions with discontinuities is $\alpha = 1.0$; i.e., $\|u^h - u\|_{0,\Omega}^2 \approx \mathcal{O}(h)$, or $\|u^h - u\|_{0,\Omega} \approx \mathcal{O}(h^{1/2})$. This can be seen easily by considering any interpolant (e.g., piecewise constant, continuous piecewise linear, or piecewise quadratic) with shock width proportional to h; see also [25]. Convergence properties of the Gauss–Newton procedure are also discussed. In the case of solutions with discontinuities, it is important to establish that u^h converges to a weak solution of (1.1) with the correct shock speed. In the case of problems with nonunique weak solutions, e.g., rarefaction wave problems, we study whether u^h converges to the so-called entropy weak solution, which is the unique weak solution that is stable against arbitrarily small perturbations. This is also the weak solution that satisfies an entropy inequality and can be obtained as the vanishing viscosity limit of a parabolic regularization of (1.1) with a viscosity term [22]. Finally, we investigate whether adaptive refinement, based on the LS error estimator, is an effective mechanism to counter smearing at shocks and we discuss adaptivity on full space-time domains. We combine adaptivity with grid continuation (also called nested iteration or full multigrid) for the Gauss–Newton procedure and investigate numerically the number of Newton iterations that are required on each grid level to obtain convergence up to discretization error.

4.1. Results for $H(\mathrm{div})$-conforming LSFEM. We present numerical results for the $H(\mathrm{div})$-conforming LSFEM described in Problem 3.1, applied to the inviscid Burgers equation, for which $\mathbf{f}(u) = (u^2/2, u)$ in (1.1). We consider the following model flow.

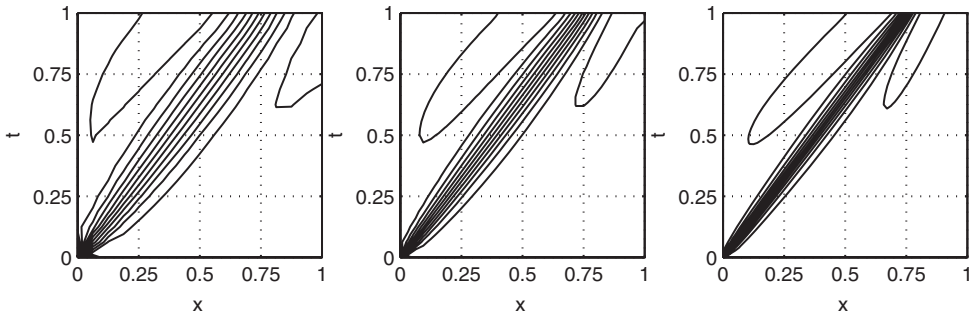


FIG. 4.1. *H(div)-conforming LSFEM, Example* 4.1 *(single shock): u^h solution contours on grids with 16^2, 32^2, and 64^2 quadrilateral elements.*

Example 4.1 (single shock, Figure 4.2). The space-time flow domain is given by $\Omega = [0,1] \times [0,1]$, with initial and boundary conditions

$$u(x,t) = \begin{cases} 0.5 & \text{if } t = 0, \\ 1.0 & \text{if } x = 0. \end{cases} \tag{4.3}$$

The unique weak solution of this problem consists of a shock propagating with shock speed $s = 3/4$ from the origin, $(x,t) = (0,0)$. Conserved quantity $u(x,t) = 1.0$ on the left of the shock and $u(x,t) = 0.5$ on the right of the shock.

Figure 4.1 shows contours of the numerical solution, u^h, for Example 4.1 on grids with decreasing h. The correct shock speed is obtained and the solution does not show excessive spurious oscillations. There are, however, small overshoots and undershoots that appear to be generated where the shock interacts with the outflow boundary. The amplitude of these overshoots and undershoots does not grow when the grid is refined. In our globally coupled space-time solution these slight oscillations seem to propagate in the direction of the characteristic curves, in accordance with the signal propagation properties of hyperbolic PDEs. These effects are reduced as the grid is refined. Figure 4.2 shows the u^h solution profile, which illustrates the well-known fact that LS methods introduce substantial smearing at shocks [5, 6, 18, 14]. However, Table 4.1 shows that both the overshoots and undershoots, as well as the smearing, disappear in the L^2 sense as the grid is refined. The numerical approximation converges to exact solution u as $h \to 0$, and the rate of convergence of $\|u^h - u\|^2_{0,\Omega}$ approaches the optimal value $\alpha = 1.0$ as grids are refined. Table 4.1 also shows that nonlinear functional $\mathcal{F}(\mathbf{w}^h, u^h; g)$ converges as $h \to 0$, with α approaching 1.00.

The left panel of Figure 4.3 shows that the Gauss–Newton method converges linearly on each grid, in accordance with the theory [12]. The right panel of Figure 4.3 shows that the discretization error on each grid level is reached after only two or three Newton iterations (indicated by the flatness in the error graphs at higher iterations), suggesting that grid continuation strategies may require only a few Newton iterations per grid level, as is confirmed below.

It is clear that numerical approximation u^h does not satisfy an exact discrete conservation property of type (1.3). Neither does the flux vector approximation $\mathbf{w}^h$. From $\mathbf{w} = (w_1, w_2) = (\mathbf{f}(u), u)$, it is apparent that w_1^h is an approximation of $\mathbf{f}(u)$, and $\mathbf{w}_2^h$ is an approximation of u. Figure 4.4 shows that $\nabla \cdot \mathbf{w}^h$ does not vanish

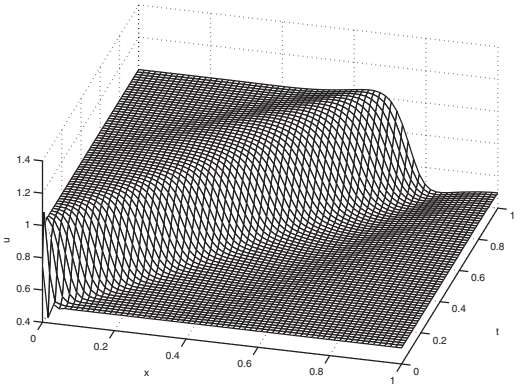
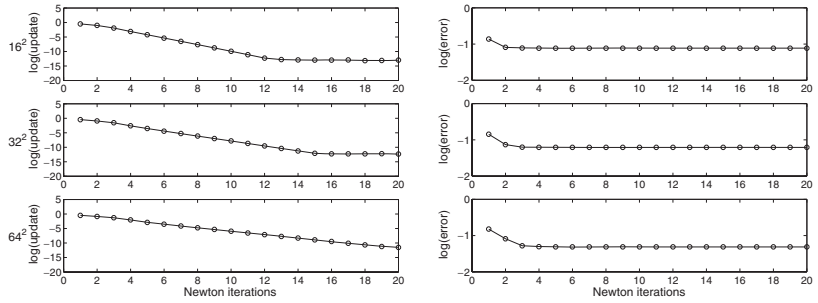


FIG. 4.2. *H(div)-conforming LSFEM, Example* 4.1 *(single shock): u^h solution profile on a grid with 32^2 quadrilateral elements.*

TABLE 4.1
H(div)-conforming LSFEM, Example 4.1 *(single shock): Convergence rates on successive grids with N^2 elements;* 20 *Newton iterations on each grid level.*

N	$\|\cdot\|^2_{0,\Omega}$	α	$\mathcal{F}$	α
16	5.96e-3		1.89e-2	
		0.58		1.03
32	3.81e-3		9.25e-3	
		0.69		1.02
64	2.36e-3		4.56e-3	
		0.77		1.01
128	1.38e-3		2.26e-3	
		0.85		1.01
256	7.66e-4		1.12e-3	



(a) Update convergence. (b) Error convergence.

FIG. 4.3. *H(div)-conforming LSFEM, Example* 4.1 *(single shock): Newton convergence. Left: $\|u^h_{i+1} - u^h_i\|^2_{0,\Omega}$ Newton update convergence. Linear convergence can be observed. Right: $\|u^h - u\|^2_{0,\Omega}$ error convergence. Discretization error is reached after just two or three Newton iterations.*

exactly in a discrete sense for our H(div)-conforming LSFEM, which illustrates that our method does not impose exact discrete conservation (1.3) of $\mathbf{w}^h$ in the sense of Lax and Wendroff [21]. Note, however, that $\nabla \cdot \mathbf{w}^h$ is very small (the scale of Figure 4.4 is 10^{-3}) and convergence of nonlinear functional $\mathcal{F}(\mathbf{w}^h, u^h; g)$ implies that

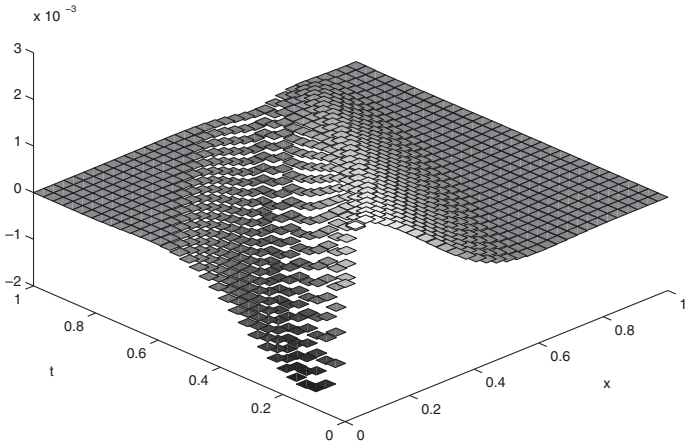


FIG. 4.4. *$H(\mathrm{div})$-conforming LSFEM, Example* 4.1 *(single shock): $\nabla \cdot \mathbf{w}^h$ on a grid with 32^2 quadrilateral elements.*

$\|\nabla \cdot \mathbf{w}^h\|_{0,\Omega} \to 0$ as $h \to 0$. Note also that $\nabla \cdot \mathbf{w}^h$ is constant in each cell since the RT_0 vector FEs are used for flux variable $\mathbf{w}$.

4.2. Results for potential $H(\mathrm{div})$-conforming LSFEM. For the potential $H(\mathrm{div})$-conforming LSFEM described in Problem 3.2, we study a more extensive set of model flows, with both shocks and rarefactions. First, consider Example 4.1, the single shock problem, for the potential $H(\mathrm{div})$-conforming LSFEM combined with adaptive refinement. A grid continuation procedure for the nonlinear solver is also used, in which the initial solution for the Newton procedure on a given grid level is obtained by interpolation from the next coarser level.

Figure 4.5 shows the u^h solution profile for the shock problem. Observe that the correct weak solution with the right shock speed is obtained, and that adaptive refinement based on the LS error estimator effectively captures and refines the shock, thus counteracting the LS smearing. Figure 4.6 shows the solution profile for potential variable ψ^h, which is continuous.

Table 4.2 shows that the convergence rates and error magnitudes for the adaptively refined grids are comparable to those of the uniformly refined grids. Squared L^2 error $\|u^h - u\|^2_{0,\Omega}$ and nonlinear functional $\mathcal{G}(\psi^h, u^h; g)$ converge to zero as $h \to 0$. We denote by $\mathcal{G}_{\mathrm{int}}$ and $\mathcal{G}_{\mathrm{bdy}}$ the interior and boundary terms of functional $\mathcal{G}(\psi^h, u^h; g)$. The convergence rates, α, appear to approach 1.0 in all cases, with the squared L^2 error approaching this rate from below and the interior functional from above.

Figure 4.7 shows a detailed view of the number of nodes used in the adaptive algorithm versus the uniform refinement algorithm. Notice that the ratio of adaptive nodes over refined nodes decreases as the grid is refined. Finally, Figure 4.8 shows that the nonlinear grid continuation strategy is very efficient: nonlinear Newton convergence can be reached with only one or two Newton iterations per grid level.

To corroborate our claims of weak solution convergence, we now consider a problem with two shocks merging into one.

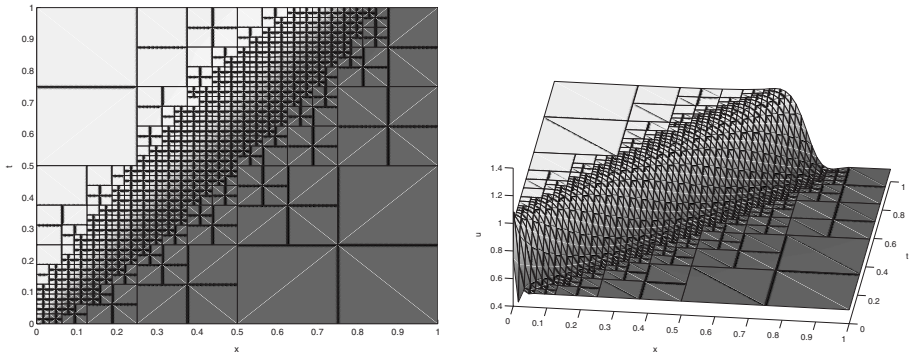


FIG. 4.5. *Potential $H(\mathrm{div})$-conforming LSFEM, Example* 4.1 *(single shock): Solution u^h on an adaptively refined grid with a resolution of $h = 1/64$ in the smallest cells.*

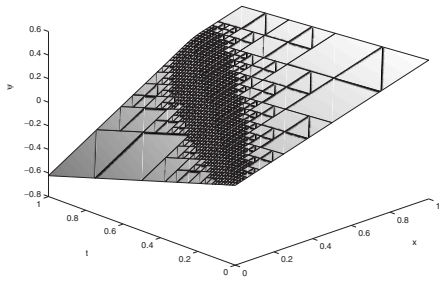
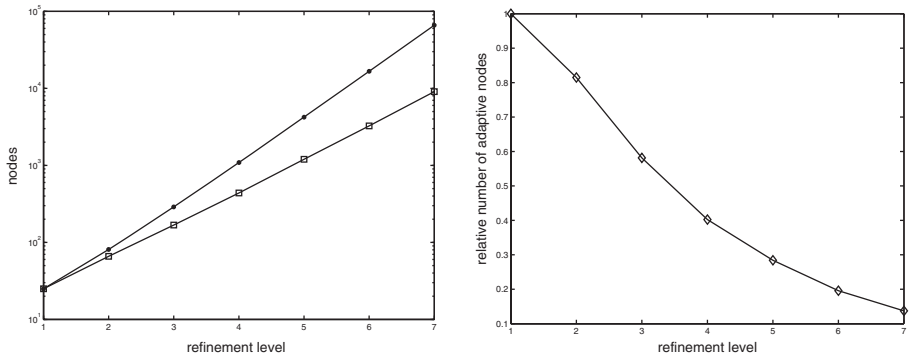


FIG. 4.6. *Potential $H(\mathrm{div})$-conforming LSFEM, Example* 4.1 *(single shock): Potential variable ψ^h.*



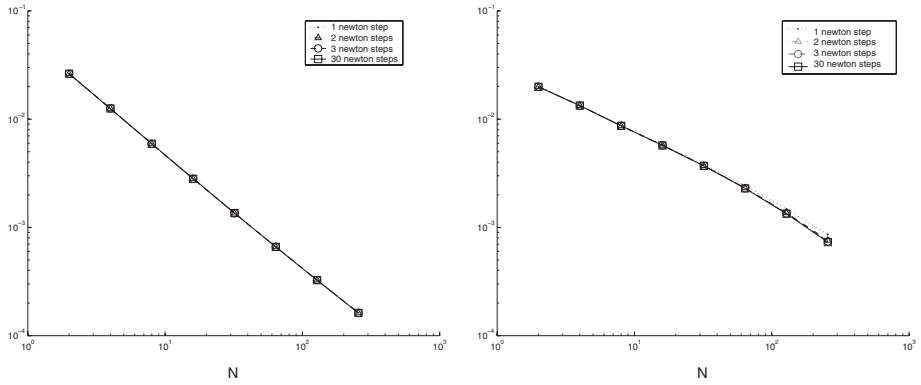
(a) Nodes on each level. (b) Ratios on each level.

FIG. 4.7. *Potential $H(\mathrm{div})$-conforming LSFEM, Example* 4.1 *(single shock): Node usage on the adaptive grids compared with the uniform grid. Left: Direct comparison of the number of nodes used at each level. • corresponds to uniform refinement and □ to adaptive refinement. Right: ◊, ratio of adaptive to uniform nodes used on level k.*

TABLE 4.2

Potential $H(\mathrm{div})$-conforming LSFEM, Example 4.1 *(single shock): Convergence rates on successive grids with $h = 1/N$ in the smallest cells;* 30 *Newton iterations on each grid level. Top section: Uniform grid refinement. Bottom section: Adaptive grid refinement.*

N	Nodes	$\|\cdot\|^2_{0,\Omega}$	α	$\mathcal{G}$	α	$\mathcal{G}_{\mathrm{int}}$	α	$\mathcal{G}_{\mathrm{bdy}}$	α
4	25	1.33e-2		1.26e-2		3.41e-3		9.15e-3	
			0.62		1.08		1.30		1.01
8	81	8.68e-3		5.92e-3		1.38e-3		4.54e-3	
			0.60		1.07		1.33		1.01
16	289	5.72e-3		2.82e-3		5.50e-4		2.26e-3	
			0.63		1.05		1.29		1.00
32	1089	3.70e-3		1.35e-3		2.25e-4		1.13e-3	
			0.69		1.03		1.22		1.00
64	4225	2.30e-3		6.61e-4		9.68e-5		5.64e-4	
			0.78		1.02		1.13		1.00
128	16641	1.34e-3		3.26e-4		4.43e-5		2.82e-4	
			0.87		1.01		1.06		1.00
256	66049	7.32e-4		1.62e-4		2.13e-5		1.41e-4	
4	25	1.33e-2		1.26e-2		3.41e-3		9.15e-3	
			0.67		1.08		1.27		1.01
8	66	8.36e-3		5.95e-3		1.41e-3		4.54e-3	
			0.62		1.08		1.33		1.01
16	168	5.46e-3		2.82e-3		5.61e-4		2.26e-3	
			0.61		1.06		1.30		1.00
32	438	3.57e-3		1.36e-3		2.28e-4		1.13e-3	
			0.71		1.04		1.21		1.00
64	1200	2.19e-3		6.63e-4		9.84e-5		5.64e-4	
			0.80		1.02		1.13		1.00
128	3258	1.25e-3		3.27e-4		4.50e-5		2.82e-4	
			0.90		1.01		1.06		1.00
256	9058	6.72e-4		1.63e-4		2.16e-5		1.41e-4	



(a) Squared L^2 error convergence. (b) Functional convergence.

FIG. 4.8. *Potential $H(\mathrm{div})$-conforming LSFEM, Example* 4.1 *(single shock): $\|u^h - u\|^2_{0,\Omega}$ error convergence (left) and functional convergence (right) on $N \times N$ grids with $N = 2^k$, where $k = 2, \ldots, 8$. A grid continuation strategy is used, and Newton convergence can be obtained with just a few Newton iterations on each level.*

Example 4.2 (double shock, Figure 4.9). The space-time flow domain is given by $\Omega = [0, 2] \times [0, 1]$ with the following initial and boundary conditions:

$$u(x,t) = \begin{cases} 1.5 & \text{if } t = 0,\ x < 0.5, \\ 0.5 & \text{if } t = 0,\ x > 0.5, \\ 2.5 & \text{if } x = 0. \end{cases} \tag{4.4}$$

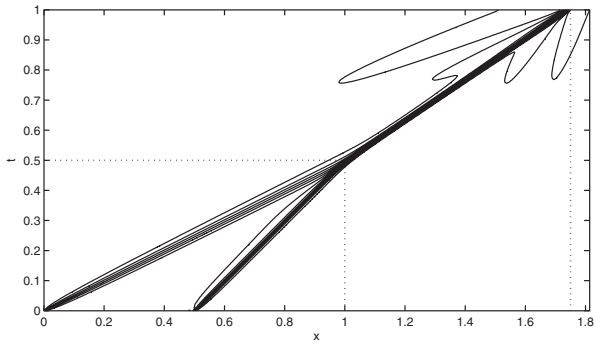


FIG. 4.9. *Potential H(div)-conforming LSFEM, Example* 4.2 *(double shock): u^h solution contours using 256^2 quadrilateral elements.*

TABLE 4.3
Potential H(div)-conforming LSFEM, Example 4.2 *(double shock): Convergence rates on successive grids with N^2 elements;* 30 *Newton iterations on each grid level.*

N	$\|\cdot\|^2_{0,\Omega}$	α	$\mathcal{G}$	α	$\mathcal{G}_{\text{int}}$	α	$\mathcal{G}_{\text{bdy}}$	α
16	2.50e-1		6.26e-2		3.06e-2		3.09e-2	
		0.82		1.10		1.19		1.02
32	1.42e-1		2.87e-2		1.34e-2		1.52e-2	
		0.86		1.05		1.10		1.01
64	7.82e-2		1.38e-2		6.27e-3		7.57e-3	
		0.91		1.02		1.03		1.00
128	4.17e-2		6.85e-3		3.07e-3		3.77e-3	
		0.93		1.00		1.00		1.00
256	2.19e-2		3.42e-3		1.54e-4		1.88e-3	

This results in two shocks that merge into one. The two original shocks emanate from $(x,t) = (0,0)$ and $(x,t) = (0,0.5)$ and travel at speeds $s = 2$ and $s = 1$, respectively. The shocks merge at $(x,t) = (1,0.5)$, and the resulting shock exits the domain at $(x,t) = (1.75,1)$ with shock speed $s = \frac{3}{2}$.

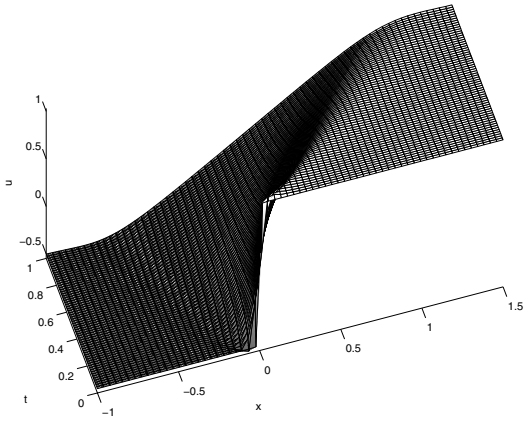
Table 4.3 shows convergence of the squared error in the L^2 sense and convergence of the functional; i.e., $\|u^h - u\|^2_{0,\Omega} \to 0$ and $\mathcal{G}(\psi^h, u^h; 0) \to 0$ as $h \to 0$. Again, we find that the convergence rates, α, approach 1.0 as grids are refined. This agrees with our findings for the single shock case. Figure 4.9 confirms convergence to the correct weak solution. The shocks merge at the correct location (indicated by the dotted lines) and the resulting single shock exits at the correct location.

We now consider a rarefaction problem for the potential H(div)-conforming LSFEM.

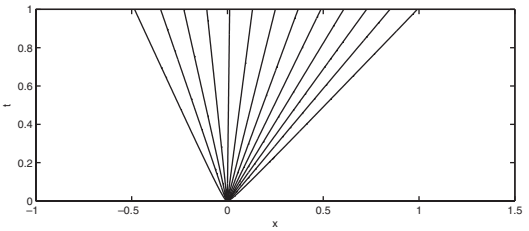
Example 4.3 (transonic rarefaction, Figure 4.10). The space-time flow domain is given by $\Omega = [-1.0, 1.5] \times [0.0, 1.0]$ with the following initial and boundary conditions:

$$u(x,t) = \begin{cases} -0.5 & \text{if } t = 0, x < 0, \\ 1.0 & \text{if } t = 0, x \geq 0, \\ -0.5 & \text{if } x = -1. \end{cases} \tag{4.5}$$

In this example, the initial discontinuity at the origin rarefies in time. There are infinitely many weak solutions, but the unique entropy solution is a rarefaction wave with $u(x,t_c)$ increasing linearly in x, at a given time t_c, from $u = -0.5$ to $u = 1.0$, between straight characteristic lines $t = -x/2$ and $t = x$.



(a) Solution profile.



(b) Solution contours.

FIG. 4.10. *Potential $H(\mathrm{div})$-conforming LSFEM, Example* 4.3 *(transonic rarefaction): Solution u^h using 64^2 quadrilateral elements.*

This rarefaction is called transonic [22] because the characteristic velocity, $f'(u) = u$, transits through zero within the rarefaction. Many higher-order numerical schemes that are based on approximate Riemann solvers or, equivalently, upwind or numerical dissipation ideas, fail to obtain the entropy weak solution for transonic rarefactions. Higher-order approximate Riemann solvers often do not capture the transonic rarefaction wave at cell interfaces appropriately, and so-called entropy fixes are necessary, as is, e.g., the case for the Roe scheme [22]. It is thus important to verify whether our $H(\mathrm{div})$-conforming LSFEMs obtain the entropy solution in the case of transonic rarefactions.

Figure 4.10 shows the u^h solution profile for the transonic rarefaction flow, confirming that the entropy solution is indeed obtained. We have to emphasize that convergence to the entropy solution has only been verified for a limited number of test problems. It remains to be seen whether the entropy solution will be obtained for all cases, or whether an entropy fix is necessary. Table 4.4 shows that squared L^2 error $\|u^h - u\|^2_{0,\Omega}$ and nonlinear functional $\mathcal{G}(\psi^h, u^h; g)$ converge to zero as $h \to 0$. The convergence rates for the squared L^2 error and for the interior functional appear to be approaching $\alpha = 1.5$ and $\alpha = 2.00$, respectively.

To conclude this section, we briefly discuss the advantages and disadvantages of the globally coupled adaptive space-time approach we have explored in our numerical

TABLE 4.4
Potential $H(\mathrm{div})$-conforming LSFEM, Example 4.3 (transonic rarefaction): Convergence rates on successive grids with N^2 elements; 30 Newton iterations on each grid level.

N	$\|\cdot\|^2_{0,\Omega}$	α	$\mathcal{G}$	α	$\mathcal{G}_{\mathrm{int}}$	α	$\mathcal{G}_{\mathrm{bdy}}$	α
16	1.27e-2		4.56e-2		4.99e-3		4.07e-2	
		1.41		1.06		1.69		1.00
32	4.79e-3		2.19e-2		1.55e-3		2.03e-2	
		1.46		1.04		1.73		1.00
64	1.74e-3		1.06e-2		4.66e-4		1.02e-2	
		1.49		1.03		1.77		1.00
128	6.21e-4		5.20e-3		1.37e-4		5.07e-3	
		1.50		1.02		1.80		1.00
256	2.20e-4		2.58e-3		3.94e-5		2.44e-3	

tests. A clear disadvantage is that the global coupling results in a large nonlinear system to be solved, with high solution complexity and large memory requirements as a result. However, efficient adaptive refinement produces nonlinear algebraic systems that are substantially smaller than those for a uniform grid system with the same effective resolution (see Table 4.2 and Figure 4.8). Consider the adaptive grid of Figure 4.5. The grid cells are concentrated in regions where the error would be large. Considering the refinement of the spatial grids as time progresses, our global space-time approach automatically takes care of refinement and coarsening. Large timesteps are taken in regions of low error, and small timesteps in the other regions. This means that the computational effort is distributed naturally in a way that is near-optimal over the whole space-time domain. This approach becomes more feasible in practice when the resulting nonlinear algebraic systems can be solved efficiently. In Figure 4.8, we show that, by using grid continuation, the nonlinear systems can be solved with nearly optimal iteration count, with just one or two Newton iterations per grid level. Combined with optimally scalable iterative solvers for the linear systems, this would result in a numerical scheme for which the work per discrete degree of freedom is bounded. Our present tests use the AMG package of Ruge [24]. Discussion of AMG convergence is beyond the scope of the present paper, but we mention that we already obtain remarkable performance using this black box linear solver. Future work includes multigrid methods designed specifically for hyperbolic problems. This could potentially make the globally coupled space-time approach competitive with explicit time-marching methods in terms of cost/accuracy ratios. It is clear that this goal has yet to be achieved, but our preliminary results suggest that it is worth further exploration. We emphasize, however, that the space-time context is not essential for the $H(\mathrm{div})$-conforming LSFEMs or their solution or conservation properties, which are the main topic of this paper. Indeed, it would be relatively straightforward to formulate our method on discontinuous timeslabs that are one FE wide in the temporal direction, e.g., as is done for the SUPG method in [19].

5. Weak conservation theorems. In this section, for our $H(\mathrm{div})$-conforming LSFEMs, we prove that if u^h converges to a function, $\hat{u}$, in the L^2 sense as $h \to 0$, then $\hat{u}$ is a weak solution of the conservation law. This statement for our LSFEMs is essentially equivalent to the Lax–Wendroff theorem for conservative finite difference methods [21]. A similar assumption on the convergence of u^h to $\hat{u}$ is made in the Lax–Wendroff theorem. The assumption that u^h converges to $\hat{u}$ for our LSFEM is made plausible, for the Burgers equation, by the numerical results of the previous section but remains an assumption that needs to be proved theoretically in future work; see also [23].

First, we relate the notion of weak solution, Definition 2.1, to the H^{-1} norm. We define the H^{-1} norm of $\nabla \cdot \mathbf{f}(u)$ as follows.

DEFINITION 5.1 (H^{-1} norm of $\nabla \cdot \mathbf{f}(u)$).

$$\|\nabla \cdot \mathbf{f}(u)\|_{-1,\Gamma_O,\Omega} = \sup_{\phi \in H^1_{\Gamma_O}(\Omega)} \left| \frac{-\langle \mathbf{f}(u), \nabla\phi\rangle_{0,\Omega} + \langle \mathbf{n}\cdot\mathbf{f}(u), \phi\rangle_{0,\Gamma_I}}{|\phi|_{1,\Omega}} \right|,$$

with $H^1_{\Gamma_O}(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ on } \Gamma_O\}$.

The following theorem identifies a relationship between weak solutions of (1.1) and the H^{-1} norm.

THEOREM 5.2. *If* $\|\nabla \cdot \mathbf{f}(u)\|_{-1,\Gamma_O,\Omega} = 0$ *and* $\|u - g\|_{0,\Gamma_I} = 0$, *then* u *is a weak solution of* (1.1).

Proof. If $\|\nabla \cdot \mathbf{f}(u)\|_{-1,\Gamma_O,\Omega} = 0$ and $\|u - g\|_{0,\Gamma_I} = 0$, then, by (1.2) and Definition 5.1, we have, $\forall\ \phi \in H^1_{\Gamma_O}(\Omega)$,

$$\begin{aligned}
|-\langle \mathbf{f}(u), \nabla\phi\rangle_{0,\Omega} + \langle \mathbf{n}\cdot\mathbf{f}(g), \phi\rangle_{0,\Gamma_I}| &= |-\langle \mathbf{f}(u), \nabla\phi\rangle_{0,\Omega} \\
&\quad + \langle \mathbf{n}\cdot(\mathbf{f}(g) - \mathbf{f}(u) + \mathbf{f}(u)), \phi\rangle_{0,\Gamma_I}| \\
&\le |-\langle \mathbf{f}(u), \nabla\phi\rangle_{0,\Omega} + \langle \mathbf{n}\cdot\mathbf{f}(u), \phi\rangle_{0,\Gamma_I}| \\
&\quad + \|\mathbf{n}\cdot(\mathbf{f}(g) - \mathbf{f}(u))\|_{0,\Gamma_I}\ \|\phi\|_{0,\Gamma_I} \\
&\le |-\langle \mathbf{f}(u), \nabla\phi\rangle_{0,\Omega} + \langle \mathbf{n}\cdot\mathbf{f}(u), \phi\rangle_{0,\Gamma_I}| \\
&\quad + K\ \|g - u\|_{0,\Gamma_I}\ \|\phi\|_{0,\Gamma_I} \\
&= 0.
\end{aligned}$$

This means, according to Definition 2.1, that u is a weak solution of (1.1). ∎

Remark 5.3. In Theorem 5.2, it is actually sufficient that $\|u - g\|_{-1/2,\Gamma_I} = 0$, because $\langle \mathbf{n}\cdot(\mathbf{f}(g) - \mathbf{f}(u)), \phi\rangle_{0,\Gamma_I} \le \|\mathbf{n}\cdot(\mathbf{f}(g) - \mathbf{f}(u))\|_{-1/2,\Gamma_I}\ \|\phi\|_{1/2,\Gamma_I}$. Condition $\|u - g\|_{0,\Gamma_I} = 0$ implies that $\|u - g\|_{-1/2,\Gamma_I} = 0$, because $\|u - g\|_{-1/2,\Gamma_I} \le \|u - g\|_{0,\Gamma_I}$.

5.1. $H(\mathrm{div})$-conforming LSFEM. We assume that the following approximation properties hold for the FE spaces $\mathcal{W}^h$, $\mathcal{U}^h$ in Problem 3.1 [25, 10]: there exist interpolants $\Pi^h\mathbf{w} \in \mathcal{W}^h, \Pi^h u \in \mathcal{U}^h$, s.t.

$$\begin{aligned}
\|\mathbf{w} - \Pi^h\mathbf{w}\|_{0,\Omega} &\le c\ h^\nu\ \|\mathbf{w}\|_{\nu,\Omega}, \\
\|u - \Pi^h u\|_{0,\Omega} &\le c\ h^\nu\ \|u\|_{\nu,\Omega}, \\
\|\mathbf{w} - \Pi^h\mathbf{w}\|_{0,\Gamma_I} &\le c\ h^\nu\ \|\mathbf{w}\|_{\nu,\Gamma_I}, \\
\|u - \Pi^h u\|_{0,\Gamma_I} &\le c\ h^\nu\ \|u\|_{\nu,\Gamma_I},
\end{aligned} \tag{5.1}$$

with $0 < \nu$. For instance, for our choice of RT_0 elements for $\mathcal{W}^h$ and bilinear elements for $\mathcal{U}^h$, (5.1) holds with $\nu \le 1$. For weak solutions with $u \in H^{1/2-\epsilon}(\Omega)$ and $\mathbf{w} \in H(\mathrm{div},\Omega) \cap (H^{1/2-\epsilon}(\Omega))^2$, $\nu = 1/2 - \epsilon$, $\forall\ \epsilon > 0$. Also, for Raviart–Thomas vector FE spaces, we can choose the interpolant in the divergence-free subspace, s.t. $\|\nabla \cdot \Pi^h\mathbf{w}\|_{0,\Omega} = 0$. We can now prove that, for solutions of minimization problem (3.2) on successively refined grids, the functional goes to zero as the grid is refined.

LEMMA 5.4. *Let* $(\mathbf{w}^h, u^h)$ *be the solution of LS minimization problem* (3.2). *Then nonlinear LS functional* $\mathcal{F}(\mathbf{w}^h, u^h; g) \to 0$ *as* $h \to 0$.

Proof. Assume that $\mathbf{w} \in H(\mathrm{div}, \Omega) \cap (H^{1/2-\epsilon}(\Omega))^2$ and $u \in H^{1/2-\epsilon}(\Omega)$ form a weak solution of (2.1). It follows that

$$\begin{aligned}
\mathcal{F}(\mathbf{w}^h, u^h; g) &\leq \mathcal{F}(\Pi^h \mathbf{w}, \Pi^h u; g) \\
&= \|\nabla \cdot \Pi^h \mathbf{w}\|^2_{0,\Omega} + \|\Pi^h \mathbf{w} - \mathbf{f}(\Pi^h u)\|^2_{0,\Omega} \\
&\quad + \|\mathbf{n} \cdot (\Pi^h \mathbf{w} - \mathbf{f}(g))\|^2_{0,\Gamma_I} + \|\Pi^h u - g\|^2_{0,\Gamma_I} \\
&\leq \|\nabla \cdot \Pi^h \mathbf{w}\|^2_{0,\Omega} + \|\mathbf{w} - \Pi^h \mathbf{w} - \mathbf{f}(u) + \mathbf{f}(\Pi^h u)\|^2_{0,\Omega} \\
&\quad + \|\mathbf{n} \cdot (\Pi^h \mathbf{w} - \mathbf{w})\|^2_{0,\Gamma_I} + \|\Pi^h u - u\|^2_{0,\Gamma_I} \\
&\leq \|\nabla \cdot \Pi^h \mathbf{w}\|^2_{0,\Omega} + \|\mathbf{w} - \Pi^h \mathbf{w}\|^2_{0,\Omega} + K \, \|u - \Pi^h u\|^2_{0,\Omega} \\
&\quad + \|\Pi^h \mathbf{w} - \mathbf{w}\|^2_{0,\Gamma_I} + \|\Pi^h u - u\|^2_{0,\Gamma_I} \\
&\leq c \, h^\nu \|u\|^2_{\nu,\Omega},
\end{aligned}$$

with $\nu = 1/2 - \epsilon \; \forall \; \epsilon > 0$, which shows that $\mathcal{F}(\mathbf{w}^h, u^h; g) \to 0$ as $h \to 0$. ∎

We can now state and prove the following theorem.

THEOREM 5.5 (weak conservation theorem for $H(\mathrm{div})$-conforming LSFEM). *Let $(\mathbf{w}^h, u^h)$ be the solution of LS minimization problem* (3.2). *If FE approximation u^h converges in the L^2 sense to $\hat{u}$ as $h \to 0$, then $\hat{u}$ is a weak solution of* (1.1). *That is, if*

$$\|u^h - \hat{u}\|_{0,\Omega} \to 0, \tag{5.2}$$

$$\|u^h - \hat{u}\|_{0,\Gamma_I} \to 0 \tag{5.3}$$

for some $\hat{u}$, then $\hat{u}$ is a weak solution.

Proof. According to Theorem 5.2, it suffices to prove that $\|\nabla \cdot \mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega} = 0$ and $\|\hat{u} - g\|_{0,\Gamma_I} = 0$. This can be obtained as follows. From (5.1), we have

$$\|\nabla \cdot \mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega} = \sup_{\phi \in H^1_{\Gamma_O}} \left| \frac{-\langle \mathbf{f}(\hat{u}), \nabla \phi \rangle_{0,\Omega} + \langle \mathbf{n} \cdot \mathbf{f}(\hat{u}), \phi \rangle_{0,\Gamma_I}}{|\phi|_{1,\Omega}} \right|.$$

Adding and subtracting $\mathbf{w}^h$ and $\mathbf{f}(u^h)$ in the interior term and using Green's formula results in

$$\begin{aligned}
\|\nabla \cdot \mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega} &= \sup_{\phi \in H^1_{\Gamma_O}} \left| \frac{-\langle \mathbf{w}^h + \mathbf{f}(u^h) - \mathbf{w}^h + \mathbf{f}(\hat{u}) - \mathbf{f}(u^h), \nabla \phi \rangle_{0,\Omega}}{|\phi|_{1,\Omega}} \right. \\
&\qquad \left. + \frac{\langle \mathbf{n} \cdot \mathbf{f}(\hat{u}), \phi \rangle_{0,\Gamma_I}}{|\phi|_{1,\Omega}} \right|, \\
&= \sup_{\phi \in H^1_{\Gamma_O}} \left| \frac{\langle \nabla \cdot \mathbf{w}^h, \phi \rangle_{0,\Omega} - \langle \mathbf{f}(u^h) - \mathbf{w}^h + \mathbf{f}(\hat{u}) - \mathbf{f}(u^h), \nabla \phi \rangle_{0,\Omega}}{|\phi|_{1,\Omega}} \right. \\
&\qquad \left. + \frac{\langle \mathbf{n} \cdot (\mathbf{f}(\hat{u}) - \mathbf{w}^h), \phi \rangle_{0,\Gamma_I}}{|\phi|_{1,\Omega}} \right|.
\end{aligned}$$

By adding and subtracting $\mathbf{f}(u^h)$ and $\mathbf{f}(g)$ in the boundary term, we have

$$\|\nabla\cdot\mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega} = \sup_{\phi\in H^1_{\Gamma_O}} \left| \frac{\langle \nabla\cdot\mathbf{w}^h,\phi\rangle_{0,\Omega} - \langle \mathbf{f}(u^h)-\mathbf{w}^h+\mathbf{f}(\hat{u})-\mathbf{f}(u^h),\nabla\phi\rangle_{0,\Omega}}{|\phi|_{1,\Omega}} + \frac{\langle \mathbf{n}\cdot(\mathbf{f}(\hat{u})-\mathbf{f}(u^h)+\mathbf{f}(u^h)-\mathbf{f}(g)-\mathbf{w}^h+\mathbf{f}(g)),\phi\rangle_{0,\Gamma_I}}{|\phi|_{1,\Omega}} \right|.$$

We now use the generalized Cauchy–Schwarz inequality,

$$\langle\psi,\phi\rangle_{0,\Gamma} \;\le\; \|\psi\|_{-1/2,\Gamma}\;\|\phi\|_{1/2,\Gamma},$$

to arrive at

$$\begin{aligned}\|\nabla\cdot\mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega} \le \sup_{\phi\in H^1_{\Gamma_O}} & \frac{\|\nabla\cdot\mathbf{w}^h\|_{0,\Omega}\;\|\phi\|_{0,\Omega}+\|\mathbf{f}(u^h)-\mathbf{w}^h\|_{0,\Omega}\;|\phi|_{1,\Omega}}{|\phi|_{1,\Omega}} \\ & + \frac{\|\mathbf{f}(\hat{u})-\mathbf{f}(u^h)\|_{0,\Omega}\;|\phi|_{1,\Omega}}{|\phi|_{1,\Omega}} \\ & + \frac{\|\mathbf{n}\cdot(\mathbf{f}(\hat{u})-\mathbf{f}(u^h))\|_{-1/2,\Gamma_I}\;\|\phi\|_{1/2,\Gamma_I}}{|\phi|_{1,\Omega}} \\ & + \frac{\|\mathbf{n}\cdot(\mathbf{f}(u^h)-\mathbf{f}(g))\|_{-1/2,\Gamma_I}\;\|\phi\|_{1/2,\Gamma_I}}{|\phi|_{1,\Omega}} \\ & + \frac{\|\mathbf{n}\cdot(\mathbf{w}^h-\mathbf{f}(g))\|_{-1/2,\Gamma_I}\;\|\phi\|_{1/2,\Gamma_I}}{|\phi|_{1,\Omega}}.\end{aligned}$$

Using the Lipschitz continuity of $\mathbf{f}(u)$, the Poincaré–Friedrichs inequality [8],

$$\exists\; c \text{ s.t.} \quad \|\phi\|_{0,\Omega}\le\; c\;|\phi|_{1,\Omega} \quad \forall\;\phi\;\in H^1_{\Gamma_O},$$

and the trace inequality [15],

$$\exists\; c \text{ s.t.} \quad \|\phi\|_{1/2,\Gamma_I}\le\; c\;|\phi|_{1,\Omega} \quad \forall\;\phi\;\in H^1_{\Gamma_O},$$

we find

$$\begin{aligned}\|\nabla\cdot\mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega} \le c\,\big(&\|\nabla\cdot\mathbf{w}^h\|_{0,\Omega}+\|\mathbf{f}(u^h)-\mathbf{w}^h\|_{0,\Omega}+\|\hat{u}-u^h\|_{0,\Omega} \\ &+\|\hat{u}-u^h\|_{-1/2,\Gamma_I}+\|u^h-g\|_{-1/2,\Gamma_I}+\|\mathbf{n}\cdot(\mathbf{w}^h-\mathbf{f}(g))\|_{-1/2,\Gamma_I}\big).\end{aligned}$$

Using our convergence assumption and the convergence of the nonlinear LSFEM functional from Lemma 5.4, it then follows by taking the limit $h\to 0$ of the right-hand side that $\|\nabla\cdot\mathbf{f}(\hat{u})\|_{-1,\Gamma_O,\Omega}=0$.

Similarly, $\|\hat{u}-g\|_{0,\Gamma_I}=0$ follows from

$$\begin{aligned}\|\hat{u}-g\|_{0,\Gamma_I} &= \|\hat{u}-u^h+u^h-g\|_{0,\Gamma_I} \\ &\le \|\hat{u}-u^h\|_{0,\Gamma_I}+\|u^h-g\|_{0,\Gamma_I},\end{aligned}$$

which proves the theorem. ∎

Remark 5.6. In Theorem 5.5, it is actually sufficient that $\|u^h-\hat{u}\|_{-1/2,\Gamma_I}\to 0$.

5.2. Potential H(div)-conforming LSFEM. We can proceed analogously for the potential H(div)-conforming LSFEM. If Ψ^h and $\mathcal{U}^h$ satisfy an approximation property, we can prove that $\mathcal{G}(u^h, \psi^h; g) \to 0$ as $h \to 0$, when (u^h, ψ^h) solves (3.4).

LEMMA 5.7. *Let (u^h, ψ^h) be the solution of LS minimization problem* (3.4)*. Then nonlinear LS functional $\mathcal{G}(u^h, \psi^h; g) \to 0$ as $h \to 0$.*

Proof. The proof is analogous to the proof of Lemma 5.4. ∎

Similar to Theorem 5.5, we have the following.

THEOREM 5.8 (weak conservation theorem for potential H(div)-conforming LSFEM). *Let (ψ^h, u^h) be the solution of LS minimization* (3.4)*. If finite element approximation u^h converges in the L^2 sense to $\hat{u}$ as $h \to 0$, then $\hat{u}$ is a weak solution of* (1.1)*. That is, if*

$$\|u^h - \hat{u}\|_{0,\Omega} \to 0, \tag{5.4}$$

$$\|u^h - \hat{u}\|_{0,\Gamma_I} \to 0 \tag{5.5}$$

for some $\hat{u}$, then $\hat{u}$ is a weak solution.

Proof. The proof is analogous to the proof of Theorem 5.5. ∎

6. Conclusions and future work. We presented two classes of H(div)-conforming LSFEMs for the Burgers equation. The methods are based on a reformulation of the conservation law in terms of the generalized flux vector, or its associated flux potential. We presented extensive numerical results, which indicate that the H(div)-conforming LSFEMs converge to an entropy weak solution of the Burgers equation. This is an interesting result, because the schemes do not satisfy an exact discrete conservation property in the sense of Lax and Wendroff [21], and because the entropy solution seems to be obtained without need for an entropy fix. We presented weak conservation theorems for our H(div)-conforming LSFEMs that are equivalent to the Lax–Wendroff conservation theorem for conservative finite difference methods. We proved that if approximation u^h converges to a function, $\hat{u}$, in the L^2 sense, then $\hat{u}$ is a weak solution of the conservation law. Our H(div)-conforming methods do not satisfy an exact discrete conservation property, but converge to a weak solution. This illustrates that the discrete conservation property of Lax and Wendroff is not a necessary condition for convergence to a weak solution. The exact discrete conservation requirement can be replaced by a minimization principle in a suitable continuous norm.

The main disadvantage of LSFEM for hyperbolic conservation laws is that there is substantial smearing at shocks. However, our results show that this smearing can be counteracted efficiently by adaptive refinement. Also, a clear disadvantage of our reformulation is that the dimensionality of the problem is increased because extra variables are introduced ($\mathbf{w}$ or ψ). Therefore, even though our approach has several attractive features, including SPD linear systems and a natural error estimator, it is not yet clear that our H(div)-conforming LSFEMs can compete in terms of cost/accuracy ratios with successful existing approaches based on the Lax–Wendroff conservation paradigm, which have reached high levels of sophistication. Nevertheless, our results indicate that our approach is promising and that research on numerical schemes for hyperbolic conservation laws does not have to be limited to the class of schemes that satisfy the discrete conservation property of Lax and Wendroff.

In future work we plan to explore the following avenues.

i. *Theoretical convergence analysis of the H(div)-conforming LSFEMs for the Burgers equation.* This entails proving convergence of u^h to $\hat{u}$ in the L^2 sense (see [23] for preliminary results) and proving convergence to an entropy weak solution.

ii. *Extension to higher-order elements.* In [14], we observed, for a different LSFEM, that higher-order elements lead to sharper shock profiles, while overshoots and undershoots do not increase. This suggests that the shock-resolving properties of our $H(\mathrm{div})$-conforming LSFEMs may also be improved by employing higher-order elements, which should be a relatively straightforward extension of the schemes presented in this paper. It is important to note that the resulting higher-order LSFEM schemes are fully linear (i.e., nonlinear limiter functions are not needed [22]), which makes the higher-order LSFEM schemes attractive for iterative solution methods. We also intend to explore the use of h, p-refinement methods.

iii. *Extension to multiple spatial dimensions and to systems of conservation laws.* The LSFEMs proposed in this paper for the Burgers equation may be more generally applicable to systems of hyperbolic conservation laws in multiple spatial dimensions. However, it remains to be investigated whether, and for which FE spaces, convergent FE methods would be obtained using the minimization framework introduced in this paper. This is a topic for future research.

iv. *Optimal AMG solvers for the linear systems.* This is work in progress; see [23] for preliminary results.

Appendix. Divergence of Gauss–Newton LSFEM for the standard conservation law formulation.

From Remark 2.12, problematic Newton convergence may be expected when applying the Newton procedure directly to the standard conservation law formulation. Define the LS functional as

$$\mathcal{H}(u;g) := \|\nabla \cdot \mathbf{f}(u)\|_{0,\Omega}^2 + \|u-g\|_{0,\Gamma_I}^2, \tag{A.1}$$

and formulate the minimization over a finite-dimensional subspace,

$$u_*^h = \arg\min_{u^h \in \mathcal{U}^h} \mathcal{H}(u^h;g), \tag{A.2}$$

with bilinear elements on quadrilaterals. The resulting Gauss–Newton LSFEM (see also [14]) fails to converge for a discontinuous solution when $\mathbf{f}(u)$ is nonlinear. Figure A.1 shows u^h contours for a Burgers shock simulation result with $u = 1$ on the left of the shock and $u = 0$ on the right. The LSFEM clearly fails to converge to the solution of the problem. On each grid level, the Newton procedure converges, but to an incorrect solution. The L^2 error and the nonlinear functional fail to converge as the grid is refined. It is plausible that this behavior of the non-$H(\mathrm{div})$-conforming Gauss–Newton LSFEM is caused by the unboundedness of the Fréchet derivative. The method seems to converge to a spurious solution, which is probably a spurious stationary point of the functional in (A.1). In fact, functional (A.1) is not continuous, whereas the reformulated functionals proposed in the body of this paper are continuous.

It is interesting to note that, by adding numerical dissipation to the standard conservation law formulation, the unboundedness of the Fréchet derivative in the Newton procedure can be remedied. This may be the reason convergence problems do not immediately arise when Newton methods are used for implicit schemes that rely on discretizations of (1.1) in which numerical dissipation is added. In the LSFEMs discussed in this paper, numerical dissipation is not added explicitly, and the Newton convergence problem appears. As is shown in the body of this paper, the $H(\mathrm{div})$-conforming reformulations result in operators with bounded Fréchet derivative, and

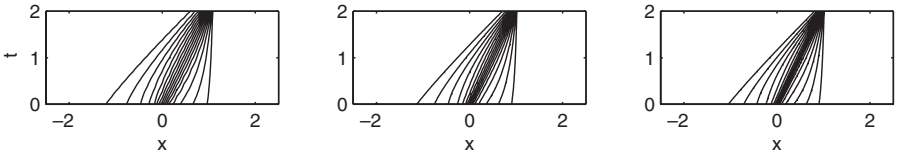


FIG. A.1. *Non-H(div)-conforming LSFEM: Shock problem on grids with 32^2, 64^2, and 128^2 elements. The LSFEM seems to converge to a spurious stationary point of the functional that is not a weak solution of the conservation law.*

the Gauss–Newton method can be applied successfully to solve the resulting nonlinear problems.

REFERENCES

[1] R. ABRAHAM, J. E. MARSDEN, AND T. RATIU, *Manifolds, Tensor Analysis, and Applications*, Appl. Math. Sci. 75, 2nd ed., Springer-Verlag, New York, 1988.

[2] R. A. ADAMS, *Sobolev Spaces*, Pure Appl. Math. 65, Academic Press, New York, London, 1975.

[3] M. BERNDT, T. A. MANTEUFFEL, AND S. F. MCCORMICK, *Local error estimate and adaptive refinement for first-order system least squares (fosls)*, Electron. Trans. Numer. Anal., 6 (1997), pp. 35–43.

[4] P. BOCHEV, J. HU, A. C. ROBINSON, AND R. TUMINARO, *Towards robust 3D Z-pinch simulations: Discretization and fast solvers for magnetic diffusion in heterogeneous conductors*, Electron. Trans. Numer. Anal., 15 (2003), pp. 186–210.

[5] P. B. BOCHEV AND J. CHOI, *A comparative study of least-squares*, SUPG *and Galerkin methods for convection problems*, Int. J. Comput. Fluid Dyn., 15 (2001), pp. 127–146.

[6] P. B. BOCHEV AND J. CHOI, *Improved least-squares error estimates for scalar hyperbolic problems*, Comput. Methods Appl. Math., 1 (2001), pp. 115–124.

[7] P. B. BOCHEV AND M. D. GUNZBURGER, *Finite element methods of least-squares type*, SIAM Rev., 40 (1998), pp. 789–837.

[8] D. BRAESS, *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*, 2nd ed., Cambridge University Press, Cambridge, UK, 2001.

[9] S. BRENNER AND L. SCOTT, *The Mathematical Theory of Finite Element Methods*, Texts Appl. Math. 15, Springer-Verlag, New York, 1994.

[10] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer-Verlag, New York, 1991.

[11] B. COCKBURN, G. E. KARNIADAKIS, AND C.-W. SHU, *The development of discontinuous Galerkin methods*, in Discontinuous Galerkin Methods (Newport, RI, 1999), Lect. Notes Comput. Sci. Eng. 11, Springer-Verlag, Berlin, 2000, pp. 3–50.

[12] J. E. DENNIS, JR. AND R. B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Classics Appl. Math. 16, SIAM, Philadelphia, 1996.

[13] H. DE STERCK, *Multi-dimensional Upwind Constrained Transport on Unstructured Grids for "Shallow Water" Magnetohydrodynamics*, AIAA Ed. Ser. 2001-2623, AIAA, Washington, DC, 2001.

[14] H. DE STERCK, T. A. MANTEUFFEL, S. F. MCCORMICK, AND L. OLSON, *Least-squares finite element methods and algebraic multigrid solvers for linear hyperbolic PDEs*, SIAM J. Sci. Comput., 26 (2004), pp. 31–54.

[15] V. GIRAULT AND P.-A. RAVIART, *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer-Verlag, Berlin, 1986.

[16] E. GODLEWSKI AND P.-A. RAVIART, *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Appl. Math. Sci. 118, Springer-Verlag, New York, 1996.

[17] T. Y. HOU AND P. G. LEFLOCH, *Why nonconservative schemes converge to wrong solutions: Error analysis*, Math. Comp., 62 (1994), pp. 497–530.

[18] B. N. JIANG, *The Least-Squares Finite Element Method*, Sci. Comput., Springer-Verlag, Berlin, 1998.

[19] C. JOHNSON AND A. SZEPESSY, *On the convergence of a finite element method for a nonlinear hyperbolic conservation law*, Math. Comp., 49 (1987), pp. 427–444.

[20] D. Kröner, *Numerical Schemes for Conservation Laws*, Wiley-Teubner Ser. Adv. Numer. Math., John Wiley and Sons Ltd., Chichester, UK, 1997.

[21] P. Lax and B. Wendroff, *Systems of conservation laws*, Comm. Pure Appl. Math., 13 (1960), pp. 217–237.

[22] R. J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge Texts Appl. Math., Cambridge University Press, Cambridge, UK, 2002.

[23] L. Olson, *Multilevel Least-Squares Finite Element Methods for Hyperbolic Partial Differential Equations*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2003.

[24] J. Ruge and K. Stüben, *Efficient solution of finite difference and finite element equations*, in Multigrid Methods for Integral and Differential Equations (Bristol, 1983), Inst. Math. Appl. Conf. Ser. New Ser. 3, Oxford University Press, New York, 1985, pp. 169–212.

[25] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[26] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer-Verlag, Berlin, 1999.