Computers and Fluids 000 (2017) 1–7

Contents lists available at ScienceDirect

Computers and Fluids

journal homepage: www.elsevier.com/locate/compfluid

On local conservation of numerical methods for conservation laws☆

Cengke Shi, Chi-Wang Shu*

Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 29 March 2017
Revised 6 June 2017
Accepted 19 June 2017
Available online xxx

Keywords:
Local conservation property
Conservation laws
Lax–Wendroff theorem
Compact schemes
Continuous finite element Galerkin method

ABSTRACT

In this paper we introduce a definition of the local conservation property for numerical methods solving time dependent conservation laws, which generalizes the classical local conservation definition. The motivation of our definition is the Lax–Wendroff theorem, and thus we prove it for locally conservative numerical schemes per our definition in one and two space dimensions. Several numerical methods, including continuous Galerkin methods and compact schemes, which do not fit the classical local conservation definition, are given as examples of locally conservative methods under our generalized definition.

1. Introduction

Local conservation is a desired property for numerical methods solving conservation laws. The most important theoretical reason is the well-known Lax–Wendroff theorem [11]. It states that the finite volume method (in 1D) taking the conservation form would converge to a weak solution of the underlying conservation law, if the numerical solutions actually converge. It is easy to give counterexamples, e.g. in LeVeque [12], that finite volume methods not in conservation form could converge to non-weak-solutions with wrong shock speed. Kröner and Rokyta [9] and Kröner et al. [10] have extended the Lax–Wendroff theorem to finite volume methods on unstructured meshes for 2D conservation laws. Abgrall et al. [2] have further generalized the Lax–Wendroff theorem to residual schemes.

The finite volume methods and discontinuous Galerkin (DG) methods, see, e.g., [4], are by design locally conservative numerical schemes. We could easily extend the Lax–Wendroff theorem to DG methods by setting the test function to value 1 in one cell and 0 in all other cells. Continuous Galerkin (CG) methods are considered only globally conservative until Hughes et al. [8] showed that it is actually locally conservative. But their definition of flux is not consistent with ours in this paper, because in their definition at least one of the two neighboring subdomains has to be global (that is, its size is comparable to that of the whole domain and hence does not go to zero with mesh refinements) to get the uniqueness of the flux across the boundary between these two subdomains. Perot [13] showed the local conservation of the CG methods with a different flux definition that is consistent with our definition (Section 3.4.3). See also Abgrall [1] for the discussion of local conservation of CG methods.

The notion of local conservation is widely known as the rate of change of a quantity (in the classical definition this is the total mass in the cell) being equal to the sum of locally defined fluxes (exchange with neighbors), which is the idea underlying the physical conservation laws. However, the absence of a rigorous definition makes it arguable as to whether a numerical method is locally conservative or not. In this paper, we give a formal definition of the local conservation property for numerical methods in one and two space dimensions. The motivation for our definition is the requirements in the classical Lax–Wendroff theorem, and naturally we prove it for locally conservative numerical methods per our definition.

The rest of the paper is organized as follows. In Section 2, we introduce the definition of local conservation in one space dimension, and prove the corresponding Lax–Wendroff type theorem. Some examples of locally conservative numerical methods in 1D according to our definition are given at the end of the section. Section 3 is about local conservation in two space dimensions. As in Section 2, we give the formal definition, prove a Lax–Wendroff type theorem, and present some examples of locally conservative numerical methods in 2D.

☆ Research supported by ARO grant W911NF-15-1-0226 and NSF grant DMS-1418750.

* Corresponding author.
E-mail addresses: cengke_shi@brown.edu (C. Shi), shu@dam.brown.edu (C.-W. Shu).

http://dx.doi.org/10.1016/j.compfluid.2017.06.018
0045-7930/

2. One dimensional conservation laws

We first consider one dimensional scalar conservation laws:

$$\begin{aligned} &u_t + f(u)_x = 0 \text{ in } \mathbb{R}^+ \times \mathbb{R} \\ &u(\cdot, 0) = u_0 \text{ in } \mathbb{R}, \end{aligned} \tag{1}$$

where the flux function f is at least Lipschitz continuous. We would like to formally define the local conservation property for numerical schemes designed for the above conservation laws, and prove a Lax–Wendroff type theorem for locally conservative schemes.

2.1. Numerical schemes

We consider in our presentation only schemes with Euler forward time stepping since the spatial discretization is our primary concern, and thus the schemes read:

$$\frac{u_h(\cdot, t_{n+1}) - u_h(\cdot, t_n)}{\Delta t_n} = L(u_h(\cdot, t_n)), \tag{2}$$

where the time domain is discretized as $0 = t_0 < t_1 < t_2 \dots$ with $t_n \to \infty$ as $n \to \infty$. Time steps are defined as $\Delta t_n = t_{n+1} - t_n$, $\Delta t = \max_{n\geq 0} \Delta t_n$. The numerical solution u_h is a function over the time-space domain $\mathbb{R}^+ \times \mathbb{R}$, which is consistent with the initial condition in the sense of (10). Since the scheme only defines the numerical solution when $t = t_n$, $\forall n \geq 0$, we expand u_h to be a constant over each time interval, i.e. $u_h(x, t) \equiv u_h(x, t_n)$, $\forall t \in [t_n, t_{n+1})$. Therefore, we can denote functions over the space domain $u_h^n = u_h(\cdot, t)$, $\forall t \in [t_n, t_{n+1})$. For example, u_h^n is a piecewise constant function for finite difference and finite volume methods, and a piecewise polynomial for Galerkin methods.

To define the local conservation property, we need (i) a partition of the spatial domain into intervals $\mathbb{R} = \cup_{j\in\mathbb{Z}} I_j$, (ii) a locally conserved quantity on each cell, and (iii) a flux on each interval endpoint. The intervals $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, and they satisfy the regularity condition: $|I_j| > c_0 h$, where $|I_j|$ is the length of the interval, mesh size $h = \max_{j\in\mathbb{Z}} |I_j|$, and $c_0 > 0$ is independent of the mesh size. For simplicity, we now denote $\sum_{n=0}^{+\infty}$ as Σ_n, and $\sum_{j\in\mathbb{Z}}$ as Σ_j.

2.2. Definition of local conservation

We define a numerical scheme to be locally conservative if its solution satisfies the following conservation form (cf. [11]):

$$\frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t_n} + \frac{1}{|I_j|}\left(g_{j+\frac{1}{2}}(u_h^n) - g_{j-\frac{1}{2}}(u_h^n)\right) = 0. \tag{3}$$

Here $\bar{u}_j^n$ (generalized locally conserved quantity) and $g_{j+\frac{1}{2}}$ (generalized flux) both depend on the numerical solution locally. More precisely, they depend on $u_h^n(B_j)$. $B_j = \left\{x \in \mathbb{R} : |x - w_j| < c\,h\right\}$, where w_j is the midpoint of the interval I_j, and $c \geq 1$ is independent of the mesh size. Note that $\bar{u}_j^n$ here is not necessarily the cell average ($\frac{1}{|I_j|}\int_{I_j} u_h^n$) even though we use the traditional notation for cell averages. The conserved quantity and flux also need to be consistent and bounded in the following sense:

- *consistency*: if $u_h^n(x) \equiv u$, a constant, $\forall x \in B_j$, we have:

$$\bar{u}_j^n = u, \tag{4}$$

$$g_{j+\frac{1}{2}}(u_h^n) = f(u), \tag{5}$$

- *boundedness*: they are both bounded with respect to the L^∞-norm of the solution:

$$\left|\bar{u}_j^n - \bar{v}_j^n\right| \leq C \left\|u_h^n - v_h^n\right\|_{L^\infty(B_j)}, \tag{6}$$

$$\left|g_{j+\frac{1}{2}}(u_h^n) - g_{j+\frac{1}{2}}(v_h^n)\right| \leq C \left\|u_h^n - v_h^n\right\|_{L^\infty(B_j)} \tag{7}$$

for two functions (u_h and v_h) in the numerical solution space.

To get global conservation from our definition, we also require the summation of the local conservation quantities being exactly the global conservation quantity:

$$\sum_j |I_j| \bar{u}_j^n = \int_{\mathbb{R}} u_h(x, t_n)\, dx, \ \forall n \geq 0. \tag{8}$$

It is easy to see that (8) together with (3) implies the following global conservation:

$$\int_{\mathbb{R}} u_h(x, t)\, dx = \int_{\mathbb{R}} u_h(x, 0)\, dx, \quad \forall t > 0. \tag{9}$$

We now have gathered all parts, and give the formal definition of local conservation.

Definition 2.1. A numerical scheme of the form (2) is locally conservative if there are conserved quantities and fluxes, both of which locally depend on the numerical solution and satisfy (3)–(8).

2.3. Lax–Wendroff type theorem

In this section, we prove a Lax–Wendroff type theorem for locally conservative numerical schemes in the sense of Definition 2.1. We first assume that the initial condition is weakly enforced in the numerical solution as follows:

$$\int_{\mathbb{R}} \left(u_0 - u_h^0\right)\phi \to 0, \ \text{as } h \to 0, \ \forall \phi \in C_0^\infty(\mathbb{R}). \tag{10}$$

To get the convergence result, we also need an assumption similar to the bounded total variation of the numerical solution:

$$h \sum_j \max_{x\in B_j} \left|u_h^n(x) - u_h^n(w_j)\right| \to 0, \text{ as } h \to 0, \quad \forall n \geq 0 \tag{11}$$

The assumption (11) on the numerical solution may seem odd at first glance, but it is actually implied by boundedness of total variation as Proposition 2.2 shows. Recall the definition of total variation:

$$\mathrm{TV}(u_h^n) = \sup_{\mathcal{P}} \sum_{i=-\infty}^{+\infty} \left|u_h^n(x_{i+1}) - u_h^n(x_i)\right|, \tag{12}$$

where $\mathcal{P}$ denotes the set of all partitions of the real line $\mathbb{R}$.

Proposition 2.2. *If the numerical solution u_h^n has uniformly bounded total variation, i.e.*

$$\mathrm{TV}(u_h^n) < C, \quad \forall n \geq 0, \tag{13}$$

u_h^n satisfies the assumption (11).

Proof. We can deduce that there are a finite number of, say no more than p (independent of mesh size), intervals in each neighborhood B_j from the regularity of the mesh and the definition of B_j. We define the point that attains the maximal value of $|u_h^n(x) - u_h^n(w_j)|$ in B_j as m_j, which is in some interval I_k with $|j - k| < p$.

We can therefore define a series of intervals $(I_{j_1}, I_{j_2} \dots I_{j_p})$ such that $m_j \in I_{j_p}$, where $j_1 = j$, and $|j_s - j_{s+1}| \leq 1$ ($\forall 1 \leq s < p$). The summation in (11) can be estimated as follows:

$$\begin{aligned} h \sum_j \max_{x\in B_j} \left|u_h^n(x) - u_h^n(w_j)\right| &= h \sum_j \left|u_h^n(m_j) - u_h^n(w_j)\right| \\ \leq h \sum_j &\left(\left|u_h^n(m_j) - u_h^n(w_{j_p})\right| + \sum_{s=1}^{p-1} \left|u_h^n(w_{j_s}) - u_h^n(w_{j_{s+1}})\right|\right) \end{aligned}$$

Please cite this article as: C. Shi, C.-W. Shu, On local conservation of numerical methods for conservation laws, Computers and Fluids (2017), http://dx.doi.org/10.1016/j.compfluid.2017.06.018