Upwinding Sources at Interfaces in conservation laws
Theodoros Katsaounis, Benoît Perthame, Chiara Simeoni

To cite this version:


HAL Id: hal-00922830
https://hal.science/hal-00922830
Submitted on 30 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Upwinding Sources at Interfaces in conservation laws

Theodoros Katsaounis (1), Benoît Perthame (2), Chiara Simeoni (1)(2)
(1) Department of Applied Mathematics, University of Crete, GR 71409 Heraklion - Crete, Greece
(2) D´ epartement de Mathématiques et Applications, École Normale Supérieure - CNRS UMR 8553
45, rue d’Ulm - 75230 Paris Cedex 05 - France

Abstract. Hyperbolic conservation laws with source terms arise in many applications, especially as a model for geophysical flows because of the gravity, and their numerical approximation leads to specific difficulties. In the context of finite volume schemes, many authors have proposed to Upwind Sources at Interfaces, i.e. the “U. S. I.” method, while a cell-centered treatment seems more natural. This note gives a general mathematical formalism for such schemes. We define consistency and give a stability condition for the “U. S. I.” method. We relate the notion of consistency to the “well-balanced” property, but its stability remains open, and we also study second order approximations as well as error estimates. The general case of a nonuniform spatial mesh is particularly interesting, motivated by two dimensional problems set on unstructured grids.

Class. No. 35L65, 65M12, 76L05

1. Introduction

We consider the Cauchy problem for a scalar conservation law, in one space dimension,

\[ \partial_t u + \partial_x A(u) + B(x, u) = 0, \quad t \in \mathbb{R}^+, \quad x \in \mathbb{R}, \]
\[ u(0, x) = u_0(x) \in L^p(\mathbb{R}), \quad 1 \leq p \leq +\infty, \]

with \( u(t, x) \in \mathbb{R} \) and we set \( a(u) = A'(u) \in C^1(\mathbb{R}) \). We focus our analysis on the source terms given by

\[ \begin{cases} 
B(x, u) = z'(x) b(u), & b \in C^1(\mathbb{R}), \quad \frac{a(u)}{b(u)} \in L^\infty(\mathbb{R}), \\
z' \in L^p(\mathbb{R}) \cap L^\infty(\mathbb{R}), & \text{for some } 1 \leq p < +\infty.
\end{cases} \]

This is suggested by the usual application of hyperbolic conservation laws as simple mathematical models for geophysical flows: in the case of the Saint-Venant equations for shallow water, for instance, \( z(x) \) describes the bottom topography (see e.g. [25] and [17]).

The equation (1) is endowed with the family of entropy inequalities, in the distributional sense,

\[ \partial_t S(u) + \partial_x \eta(u) + S'(u) B(x, u) \leq 0, \quad \eta'(u) = S'(u) a(u), \]

for any pair of a convex entropy function \( S \) and corresponding entropy flux \( \eta \) (refer to [6], [8] or [10]). Then, we are interested in computing the entropy solutions, which means that we wish to use solvers (denoted by \( A \) below) satisfying the E-condition [23], even though it does not appear as a necessary condition for our results. With the assumptions mentioned above, Kružkov’s method [19] can be applied to prove existence and uniqueness of the entropy solution to (1)-(2), in the functional space \( C(\mathbb{R}_+; L^1(\mathbb{R})) \cap L^\infty((0, T) \times \mathbb{R}) \). This relies on the \( L^1 \)-control of the space translations and the contraction property of the problem (1)-(2).
In comparison with the homogeneous case, the analytical properties of equation (1) are modified by the source term. In particular, a significant difference is the occurrence of non-constant stationary solutions, resulting from the balance between source term and internal forces. By integrating the stationary equation associated with (1), according to (3), we obtain an algebraic relation for smooth steady state solutions,

\[ D(u(x)) + z(x) = C^x, \quad D'(u) = \frac{a(u)}{b(u)}. \] (5)

It is now accepted that suitable numerical approaches to the problem (1)-(2) should preserve the steady state solutions (5), or some discrete versions at least, with enough accuracy. There are several advantages to do so; in many applications, one wishes to compute small perturbations of such steady states, schemes which do not preserve steady states exhibit a bad time decay towards these solutions, a better stability condition is generally obtained (the source term may not affect the CFL condition, as seen below), one wishes to keep the physical property that the value \( u(x) \) of the steady states is only determined by \( z(x) \) values (in the case \( D \) is monotonic).

These properties are not satisfied by centered schemes applied to the source term and thus other methods have been proposed in the literature. All of them are based on the principle of Upwinding Sources at Interfaces (the “U. S. I.” method). This idea of “U. S. I.” schemes was first used in the context of reacting flows (see [27], [1] and [2], for instance), then the source term does not depend on \( z \) and a conclusion was that it is enough to introduce a correction of the flux solver to take the source into account. In the context of gravity driven flows, many authors arrived to the conclusion that it is necessary to split, and thus upwind, the source itself at interfaces (see [7], [13], [14], [9], [21], [3], [25], [16], [22], [4] and many others). This yields interesting questions from a theoretical point of view. In this note, we formalize the “U. S. I.” method with some generality and we try to isolate the new mathematical concepts that are needed for the convergence analysis. We first present the formalism of the “U. S. I.” method, we indicate the relations with well-balanced methods have been proposed in the literature. All of them are based on the principle of Upwinding Sources at Interfaces (the “U. S. I.” method). This idea of “U. S. I.” schemes was first used in the context of reacting flows (see [27], [1] and [2], for instance), then the source term does not depend on \( z \) and a conclusion was that it is enough to introduce a correction of the flux solver to take the source into account. In the context of gravity driven flows, many authors arrived to the conclusion that it is necessary to split, and thus upwind, the source itself at interfaces (see [7], [13], [14], [9], [21], [3], [25], [16], [22], [4] and many others). This yields interesting questions from a theoretical point of view. In this note, we formalize the “U. S. I.” method with some generality and we try to isolate the new mathematical concepts that are needed for the convergence analysis. We first present the formalism of the “U. S. I.” method, we indicate the relations with well-balanced schemes and we also propose appropriate second order extensions.

2. The “U. S. I.” method

In this section, we consider the finite volume method for treating numerically hyperbolic systems of conservation laws, and the consistency analysis presented here extends to systems even though we restrict our notations to the scalar case. This method is robust and presents the advantage to be conservative (refer e.g. to [11] and [8]).

We set up a mesh on \( \mathbb{R} \) made up of cells \( C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \), with center \( x_i \), cell interfaces \( x_{i-\frac{1}{2}} \) and nonuniform length \( \Delta x_i \), for \( i \in \mathbb{Z} \). We set \( h = \text{sup}_{i \in \mathbb{Z}} \Delta x_i \). Then, we construct a piecewise constant approximation of the function \( z(x) \) on the mesh, whose coefficients are \( z_i = \frac{1}{\Delta x_i} \int_{C_i} z(x) \, dx \). We also introduce a time-step \( \Delta t \) and we set \( t_n = n \Delta t, n \in \mathbb{N} \); therefore we have to consider an additional restriction on the ratio \( \Delta t / \Delta x_i \), the usual CFL condition, to guarantee numerical stability.

In this framework, the discrete unknowns are expected to be approximations of the cell-averages of the solution, \( u^n_i \approx \frac{1}{\Delta x_i} \int_{C_i} u(t_n, x) \, dx, i \in \mathbb{Z} \) (the conservative quantities are cell-centered), while the numerical fluxes are defined at the interfaces of the mesh. On the other hand, departing from discrete data we reconstruct the piecewise constant function \( u^h(t, x) = u^n_i, t \in [t_n, t_{n+1}), x \in C_i \).

The general fully explicit, three points, “U. S. I.” scheme for equation (1) reads

\[ \frac{\Delta x_i}{\Delta t} (u^n_{i+1} - u^n_i) + (A^n_{i+\frac{1}{2}} - A^n_{i-\frac{1}{2}}) + B^n_{i-\frac{1}{2}+} + B^n_{i+\frac{1}{2}-} = 0. \] (6)

The numerical fluxes are usual consistent approximations of the analytical flux function, i.e.

\[ A^n_{i+\frac{1}{2}} = A(u^n_i, u^n_{i+1}), \quad A \in C^1, \quad A(u, u) = A(u), \] (7)
and they can be chosen in the general class of E-schemes (see [10], for instance).
Because of the choice of source terms in the form (3), the function \( z(x) \) is defined up to a constant.
Therefore, without loss of generality, we assume that the source term is discretized by

\[
B^{n, \pm}_{i+\frac{1}{2}} = B^\pm (\Delta x_i, \Delta x_{i+1}; u^n_i, u^n_{i+1}, z_{i+1} - z_i), \quad B^\pm \in C^2,
\]

\[
B^\pm (h, k; u, v, 0) = 0, \quad \frac{\partial B^\pm}{\partial u} (h, k; u, v, 0) = \frac{\partial B^\pm}{\partial v} (h, k; u, v, 0) = 0.
\]

The only dependency upon \( z_{i+1} - z_i \) is natural, because the problem is unchanged when adding a constant. Also, the last condition is natural when the problem (1)-(3) becomes homogeneous, namely \( z'(x) = 0 \), since the scheme (6)-(8) reduces to the usual finite volume approximation for scalar conservation laws. Such a discretization is also upwinded, in the sense that \( B^-_{i+\frac{1}{2}} \) represents the contribution of the waves coming from the left of the interface \( x_{i+1} \) and moving towards the cell \( C_i \) with non-positive velocity, while \( B^+_{i+\frac{1}{2}} \) represents the waves moving forward from the right of the interface \( x_{i+\frac{1}{2}} \) with nonnegative velocity.

When the spatial mesh is not regular, a general notion of consistency for the “U. S. I.” method is not obvious to formulate.

**Definition 1.** An “U. S. I.” scheme (6)-(8) is consistent with equation (1) if, locally uniformly in \((u; h, k)\), it holds that

\[
\lim_{\lambda \to 0} \frac{B^+ (h, k; u, u, \lambda) + B^- (h, k; u, u, \lambda)}{\lambda} = b(u).
\]

The following result constitutes a first stage for the convergence analysis of the “U. S. I.” method and extends the classical Lax-Wendroff theorem [20]. We note that it is valid for systems as well.

**Theorem 1.** Consider an “U. S. I.” scheme (6)-(8), which satisfies the consistency condition (9). We assume, for various constants \( C \): the CFL condition \( \Delta t \leq C h \), that \( ||u^h||_{L^\infty(\mathbb{R}_+ \times \mathbb{R})} \leq C \) and that \( u^h \rightarrow u \) in \( L^1_{loc}(\mathbb{R}_+ \times \mathbb{R}) \), as \( h \to 0 \). We also assume either that, for all bounded subsets \( \Omega \) of \( \mathbb{R}_+ \times \mathbb{R} \), it holds

\[
\sum_{(n,i) \in K_\Omega} (\Delta x_i + \Delta x_{i+1}) |u^n_{i+1} - u^n_i| \rightarrow 0, \quad as \ h \to 0,
\]

where \( K_\Omega \) denotes the set of indices such that \((t_n, x_i) \in \Omega\); or weak regularity of the mesh, that is

\[
\exists \alpha, \beta > 0 \quad so \ that \quad \alpha \Delta x_{i+1} \leq \Delta x_i \leq \beta \Delta x_{i+1}, \forall i \in \mathbb{Z}.
\]

Then \( u \) is a weak solution to the initial values problem (1)-(2).

The proof of Theorem 1 applies standard arguments for homogeneous systems of conservation laws, adapted for dealing with the source term. We refer to [26] for details. We point out that the definition of consistency (9) for the source term does not imply that the consistency error vanishes, in finite volume sense: indeed, due to the choice of a nonuniform spatial mesh, the space-step \( \Delta x_i \) could be very different from the length \( |x_{i+1} - x_i| \) of an interfacial interval. Therefore, standard techniques for uniform mesh (see e.g. [11] and [10]) do not apply in this case and the additional hypothesis (10) or (11) is required to control the spatial variations of the numerical solution in comparison with the cells of the mesh. Two examples illustrate the optimality of Theorem 1.
Example 1. We consider $A = A = 0$, $b(u) = 1$ and $z'(x) = C_{st}$ in equation (1); we set up $\Delta x_i = h$ for $i$ even, $\Delta x_i = h/2$ for $i$ odd, and the discretization for the source term is given by

$$B^-(h, k; u, v, \lambda) = B^+(h, k; u, v, \lambda) = \lambda/2,$$

which turns out to be independent from $\Delta x_i$ or $\Delta x_{i+1}$. This is a consistent scheme and its (explicit) discrete solution converges weakly but not strongly to the solution of (1)-(2). With an additional term $b(u)$, it does not converge to the correct limit.

Example 2. We consider $A = A = 0$, $b(u) = 1$ and $z'(x) = C_{st}$ in equation (1); we set up $\Delta x_i = h$ for $i$ even, $\Delta x_i = h/2$ for $i$ odd, and the discretization for the source term is given by

$$B^-_{i+1/2}(u, v, \lambda) = \Delta x_i \lambda / (\Delta x_i + \Delta x_{i+1}), \quad B^+_{i+1/2}(u, v, \lambda) = \Delta x_{i+1} \lambda / (\Delta x_i + \Delta x_{i+1}).$$

This is a consistent scheme and, again on the explicit solution, one can see that it converges strongly.

This second example enters a general compactness framework, that completes Theorem 1, using the stability conditions (we do not write the more general condition here but look for simplicity)

$$B^-_{i+1/2} = \frac{\Delta x_i}{\Delta x_i + \Delta x_{i+1}} b^+(\Delta x_{i-1}, \Delta x_i; u_{i-1}, u_i, z_i - z_{i-1}) (z_i - z_{i-1}),$$

$$B^+_{i+1/2} = \frac{\Delta x_i}{\Delta x_i + \Delta x_{i+1}} b^-(\Delta x_i, \Delta x_{i+1}; u_i, u_{i+1}, z_{i+1} - z_i) (z_{i+1} - z_i).$$

Then, the consistency condition (9) can be rewritten as follows,

$$h b^-(h, k; u, u, 0) + k b^+(h, k; u, u, 0) = (h + k) b(u).$$

It implies, by opposition to the more general condition (9), that the truncation error vanishes (also for nonuniform grids, otherwise the two conditions are equivalent).

Another statement in the same direction is the following.

Theorem 2. Assume that $z'' \in L^1(\mathbb{R})$, $b^-, b^+ \in C^1$ and $\mathcal{A}$ is an ordered scheme for the fluxes (i.e. $\mathcal{A}(u, v) \leq \min(\mathcal{A}(u), \mathcal{A}(v))$, for $u < v$, and $\mathcal{A}(u, v) \geq \max(\mathcal{A}(u), \mathcal{A}(v))$, for $u > v$; refer to [28] or [24]). Then, an “U. S. I.” scheme with the stability condition (12) satisfies, for $T = n \Delta t$,

$$\sum_{i \in \mathbb{Z}} |u^n_i - u^0_i| \leq \|u_0\|_{TV(\mathbb{R})} + C(T, \|z'\|_{BV(\mathbb{R})}).$$

A specific convergence result which does not involve $BV$ neither weak $BV$ bounds is given in [3], for uniform grids. But no general theory is available in the “weak $BV$” framework (see [8]), and especially for two-dimensional problems set on unstructured grids. We conjecture that either the form (12) is enough or, at least, the “well-balanced” property that we relate now to the above theory.

3. Well-balanced schemes

We now assume that the function $D$ in (5) is strictly monotonic, unbounded from below and above (these conditions are not satisfied when dealing with systems). For all $z'$, this ensures the existence of a unique Lipschitz continuous steady state, or a discrete steady state

$$D(u_i) + z_i = C_{st}, \quad \forall i \in \mathbb{Z}.$$

Following [14], the “well-balanced” schemes are the “U. S. I.” schemes (6)-(8) for which the above discrete stationary solutions are preserved.
Lemma 1. The “U. S. I.” method (6)-(8) for the problem (1)-(2) is well-balanced if and only if, for all \( u, v, z_-, z_+ \) such that \( D(u) + z_- = D(v) + z_+ \), we have the equalities

\[
A(u, v) - A(u) + B^-(u, v, z_+ - z_-) = 0, \quad A(v) - A(u, v) + B^+(u, v, z_+ - z_-) = 0.
\]

In particular, we can see that the mesh size \( \Delta x_i \) or \( \Delta x_{i+1} \) does not appear explicitly in most of the schemes that satisfy this property.

Several well-balanced schemes have been developed, either by taking the source term into account directly in the numerical fluxes (see e.g. [3], [12] and [14]), or by making explicitly use of discrete relations for the stationary solutions (see e.g. [16] and [21]). The simplest example, which does not extend to systems in practical situations, is to choose

\[
B_{i+1/2}^\prime = A(u_i, u_{i+1}^-) - A(u_i, u_{i+1}), \quad D(u_{i+1}^-) + z_{i+1} = D(u_i) + z_i.
\]

The following result, obtained by means of standard asymptotic expansions, guarantees the consistency with equation (1) for well-balanced schemes.

Lemma 2. A well-balanced scheme, in the sense of Lemma 1, verifies the consistency condition (9).

Of course well-balanced schemes do not satisfy the strong stability condition (12) on nonuniform grids. Therefore, a general notion of strong stability is still needed.

4. Second order “U. S. I.” schemes

In practice, higher order accuracy is needed. We propose two extensions of the finite volume scheme (6)-(8) to second order approximations. The difficulty is to take into account second order accuracy on \( z \) also, which appears to be essential for practical computations.

We consider piecewise linear reconstructions of the discrete functions on the mesh, with numerical derivatives computed by applying an appropriate slope limiter technique (refer to [10]). For the sake of simplicity, we consider only first order discretization in time, but it is easy to recover higher order accuracy by applying Runge-Kutta methods, for instance.

A first approach to second order schemes for the “U. S. I.” method reads

\[
\frac{\Delta x_i}{\Delta t} (u_{i+1}^n - u_i^n) + (A_i^{n+\frac{1}{2}} - A_i^{n-\frac{1}{2}}) + B_i^{n,+} + B_i^{n,-} + \Delta x_i B_i^n = 0,
\]

where the numerical fluxes \( A_i^{n+\frac{1}{2}} = A(u_i^{n,+}, u_{i+1}^-) \) and \( B_i^{n,+} = B^+(u_i^{n,+}, u_{i+1}, z_{i+1} - z_i) \) are defined by means of the interfacial values

\[
u_i^n = u_i^n - \frac{\Delta x_i}{2} u_i', \quad u_i^{n,+} = u_i^n + \frac{\Delta x_i}{2} u_i', \quad z_i^- = z_i - \frac{\Delta x_i}{2} z_i', \quad z_i^+ = z_i + \frac{\Delta x_i}{2} z_i'.
\]

Because of the introduction of piecewise linear approximations of \( z(x) \), the difference of interfacial values approximates the second derivative and the upwind discretization “overtakes” the source term (3); an additional centered term \( B_i^n = z_i' b(u_i^n) \), which depends on the cell-averages, is thus necessary in (15) to recover the first derivative and to achieve second order accuracy.

An alternative approach to obtain second order extensions of the “U. S. I.” method is based on improving the consistency properties of the discrete source term. We consider piecewise constant approximations of the function \( z(x) \) and piecewise linear reconstructions of the numerical solution on the mesh, whose interfacial values are used to define \( A_i^{n+\frac{1}{2}} = A(u_i^{n,+}, u_{i+1}^-) \) and \( B_i^{n,+} = B^+(u_i^{n,+}, u_{i+1}, z_{i+1} - z_i) \) in the finite volume scheme (6), as
suggested by the special form of the source term (3) given by the product of functions with different orders of derivative. To recover second order accuracy, we assume an improved consistency condition, as in Definition 1: there exists a constant $C_B$ such that
\[
\left| \frac{B^+(u, u, \lambda) + B^-(u, u, \lambda)}{\lambda} - b(u) \right| \leq C_B \lambda^2.
\] (17)

Both second order methods illustrated above are strictly related, as formally verified by means of standard asymptotic expansions on the numerical functions and simple algebraic calculations with the differences of discrete interfacial values (16).

Following a remark by F. Bouchut [5], the property of preserving steady state solutions is also satisfied by the above second order schemes, when the piecewise linear reconstruction is performed on $D(u^h_i)$ and $z_i$ also. Indeed, “symmetric” limiters will then preserve the discrete relation (14).

The second order scheme (15) is validated by the numerical results obtained for the steady state solutions of the Saint-Venant equations in [17].

4. Error estimates for “U. S. I.” schemes

Deriving error estimates, that confirm the accuracy of second order schemes, faces the specific difficulty that for transport equations only $h^{1/2}$ rate of convergence is proved. In order to avoid this difficulty we only discretize the source term, while keeping continuous the transport terms in some kind of extended semi-discrete versions of equation (1) with a linear flux,

\[
\partial_t u + \partial_x u + B(x, u) = 0, \quad t \in \mathbb{R}_+, \quad x \in \mathbb{R},
\] (18)

and

\[
\partial_t u^h + \partial_x u^h + B^h(x, u^h) = 0,
\] (19)

where $B^h(x, u^h)$ indicates an appropriate discretization of the source term (3) according to the “U. S. I.” method, by using the cell-averages $u_i^h(t) = \frac{1}{h} \int_{\mathcal{C}_i} u^h(t, x) \, dx$, $i \in \mathbb{Z}$.

For simplicity we use an uniform spatial mesh, i.e. $\Delta x_i = h$, $\forall i \in \mathbb{Z}$, and we obtain

\[
B^h(x, u^h) = \sum_{i \in \mathbb{Z}} \frac{1}{h} \left[ B^+(u_{i-1}^h, u_i^h, z_i - z_i-1) + B^-(u_i^h, u_{i+1}^h, z_i+1 - z_i) \right] \nu_{\mathcal{C}_i}(x).
\] (20)

**Theorem 3.** We assume $z \in W^{2,p}$, $1 \leq p < +\infty$, and we consider a numerical source term (20) satisfying the first order consistency condition (9). Then, for all $t \in \mathbb{R}_+$, we have the first order error estimate

\[
\|u(t) - u^h(t)\|_{L^p} \leq C(t) \left( \|u_0 - u_0^h\|_{L^p} + h\|z\|_{W^{2,p}} + h \int_0^t \|u(s)\|_{W^{2,p}} \, ds \right).
\]

The convergence properties of second order schemes are notably affected by the technique used to construct piecewise linear approximations of the numerical functions, namely the choice of the slope limiter. Without appropriate hypotheses on the coefficients of such approximations, the proof of the error estimate fails and numerical evidence shows that the discretization (15) loses second order accuracy. This leads to use discrete derivatives computed in the restricted class of slope limiters introduced in [15] and [29]. We refer to [18] for details and for the proof of these results.

**Theorem 4.** With the above restrictions, we assume $z \in W^{3,p}$, $1 \leq p < +\infty$, and we consider the numerical source term in (18) defined as in (15). Then, for all $t \in \mathbb{R}_+$, we have the second order error estimate

\[
\|u(t) - u^h(t)\|_{L^p} \leq C(t) \left( \|u_0 - u_0^h\|_{L^p} + h^2\|z\|_{W^{3,p}} + h^2 \int_0^t \|u(s)\|_{W^{2,p}} \, ds \right).
\]
Acknowledgments
This work was partially supported by the ACI (Ministère de la Recherche, France): Modélisation de processus hydrauliques à surface libre en présence de singularités (http://www-rocq.inria.fr/m3n/CatNat/) and by HYKE European programme HPRN-CT-2002-00282 (http://www.hyke.org). The authors would like to thank François Bouchut and Th. Gallouët for helpful discussions.

References


