

# AEROSP 590: Entropy Stability

Andi Zhou

May 8, 2023

## 1 Entropy Condition - Origin and Formulation

In this section we explain the entropy condition, and explain why this would be necessary in solving Hyperbolic Partial Differential Equations. To explain the necessity of entropy condition, we first start by discussing a *single conservation law*.

$$u_t + f_x = 0 \quad (1.1)$$

where  $f$  is a non-linear function of  $U$ , where:

$$\frac{df}{du} = a(u) \quad (1.2)$$

Using chain rule, we could write Equation 1.1 as:

$$u_t + a(u)u_x = 0 \quad (1.3)$$

If  $u$  is constant along a characteristic trajectory, we also know that  $a$  will be constant as well.  $u$  would propagate along space in speed  $a$ . Therefore, we also define  $a$  as the signal speed. The trajectories of  $u$  would be defined as a characteristics. Equation 1.1 is known as the strong (differential) form of the PDE, where a solution to this equation would need to satisfy the differential operators at all points in space and time. A solution that matches the strong form is known as the a strong solution. However, the differential form could not admit discontinuities. We now consider the weak form:

$$\int_0^\infty \int_{-\infty}^\infty [w_t u + w_x f(u)] dx dt + \int_{-\infty}^\infty w(x, 0) \phi(x) dx = 0 \quad (1.4)$$

where  $w$  is any smooth test function with compact support (non-zero only within the specified element). The general idea of weak solution is that we could rewrite the derivatives of  $u$  as derivatives of compact test function  $w$ . From this, we basically eliminated the need for solution to be continuous. However, for hyperbolic systems (which we are interested in), the notion of weak solution **does not** guarantee uniqueness. Therefore, we need some sort of criterions to select a solution that is *physical*. We then have the **entropy** condition.

It is rather difficult to trace when the notion of entropy condition was first introduced. Earliest written work that could be traced on this subjects were done by P. Lax [citation] and O. Oleinik [citation]. Both authors have mentioned the idea of *entropy conditions* repeatedly in their text. Oleinik [citation needed] has shown that physically relevant solutions must have the following properties:

$$\frac{f(u) - f(u_L)}{u - u_L} \geq S \geq \frac{f(u) - f(u_R)}{u - u_R} \quad (1.5)$$

This is known by Oleinik as *Condition E*. If we use the Rankine-Hugoniot jump relation:

$$f(u_R) - f(u_L) = S(u_R - u_L) \quad (1.6)$$

we also arrive at the *Lax entropy condition*:

$$a(u_l) > S > a(u_r) \quad (1.7)$$

$S$  in the above equation is the speed of discontinuity. In essence, the entropy condition defines that characteristic line should ALWAYS run into the shock, not away from it.

## 2 Entropy Satisfied Schemes

Several fluxes satisfy the above entropy condition. We understood from Lax [1] that any monotone scheme satisfy the entropy condition. In this section we specifically focus on first order scheme, and list two of these fluxes below.

### 2.1 Lax-Friedrich's Scheme

The Lax-Friedrich scheme, under a specific CFL number, satisfies the entropy condition. The scheme is introduced by Lax in 1954 as a method of discretizing the first order hyperbolic equation without violating the entropy condition introduced in the above section. We introduce the general scheme below, and demonstrate that it is actually a FTCS (forward in time, centered in space) scheme with the temporal derivative replaced by a spatial average of the neighboring points.

#### 2.1.1 General Lax-Friedrich Algorithm

The general Lax-Friedrich discretization is as follow:

$$\begin{aligned} u_t &= \frac{1}{\Delta t} \left( u_j^{n+1} - \frac{u_{j-1}^n + u_{j+1}^n}{2} \right) \\ f_x &= \frac{1}{2\Delta x} (f(u_{j+1}^n) - f(u_{j-1}^n)) \end{aligned} \quad (2.1)$$

Combined together, we obtain the Lax-Friedrich scheme:

$$\begin{aligned} \frac{1}{\Delta t} \left( u_j^{n+1} - \frac{u_{j-1}^n + u_{j+1}^n}{2} \right) + \frac{1}{2\Delta x} (f(u_{j+1}^n) - f(u_{j-1}^n)) &= 0 \\ u_j^{n+1} &= \frac{1}{2}(u_{i+1}^n + u_{i-1}^n) - \frac{\Delta t}{2\Delta x} (f(u_{j+1}^n) - f(u_{j-1}^n)) \end{aligned} \quad (2.2)$$

We note that the above scheme is essentially FTCS:

$$u_j^{n+1} = \underbrace{\frac{1}{2}(u_{i+1}^n + u_{i-1}^n)}_{\text{Temporal Derivative}} - \underbrace{\frac{\Delta t}{2\Delta x} (f(u_{j+1}^n) - f(u_{j-1}^n))}_{\text{Spatial derivative (Centered in Space)}} \quad (2.3)$$

According to Lax[1], the Lax-Friedrich scheme satisfies the entropy condition under a suitable choice of the  $\lambda = \frac{\Delta t}{\Delta x}$ .

#### 2.1.2 Proof on Lax-Friedrich Satisfies the Entropy Condition

### 2.2 Engquist-Osher Scheme

The E-O scheme, or Engquist and Osher scheme could be considered as a modified version of the Cole-Murman scheme used for small disturbance equation of transonic flow. Engquist and Osher introduced this scheme in a paper published in 1980 [2] and then later generalized this scheme for general hyperbolic systems of conservation laws [3]. We discuss the formulation of this particular scheme here along with its relevant mathematical proofs.

#### 2.2.1 General E-O Algorithm

Consider a non-linear scalar conservation law in 1D:

$$w_t + f(w)_x = 0, w(x, t = 0) = \Phi(x), -\infty < x < \infty \quad (2.4)$$

The solution is approximated by a mesh function  $w_j^n$  on the mesh  $(x_j, t^n)$  with  $x_j = k\Delta x, t^n = n\Delta t, j = 0, \pm 1, \dots, n = 0, 1, \dots$ . In explicit form, the finite difference approximation is defined as:

1.

$$\begin{aligned} w_j^{n+1} &= w_j^n - \frac{\Delta t}{\Delta x} (\Delta_+ f_-(w_j^n) + \Delta_- f_+(w_j^n)) \\ w_j^0 &= \Phi(x_j), j = 0, \pm 1, \pm 2 \end{aligned} \quad (2.5)$$

2.

$$\frac{\Delta t}{\Delta x} \sup |f'| < 1 \quad (2.6)$$

We denote here as a reminder that Part 1 is the fully-discrete version of Equation 2.5 written is a per cell. We specify here as well the following notations:

$$\begin{aligned} \Delta_{\pm} w_j &= \pm(w_{j\pm 1} - w_j) \\ D_{\pm} w_j &= \frac{1}{\Delta x} \Delta_{\pm} w_j \end{aligned} \quad (2.7)$$

The auxiliary functions  $f_+$  and  $f_-$  are:

$$\begin{aligned} f_+(w) &= \int_0^w \mathcal{X}(w) f'(w) dw \\ f_-(w) &= \int_0^w (1 - \mathcal{X}(w)) f'(w) dw \end{aligned} \quad (2.8)$$

When  $f$  is convex, the definitions then reduce to:

$$\begin{aligned} f_+(w) &= f(\max(w, \bar{w})) \\ f_-(w) &= f(\min(w, \bar{w})) \end{aligned} \quad (2.9)$$

where  $\bar{u}$  is the stagnation or sonic point for which  $f'(\bar{u}) = 0$ . We make a note that when the sign is positive, Equation 2.5 becomes the *upwind* algorithm. This is also known as *flux decomposition*, where we separated the original upwinding flux into essentially a decreasing and increasing components. Decomposing this flux into different components provides us with a different perspective when looking at states near shock points.

When we consider a hyperbolic PDE with convex flux, we arrive at the following:

$$\begin{aligned} f^+(w) &= \begin{cases} f(w), & w \geq \bar{w} \\ f(\bar{w}), & w \leq \bar{w} \end{cases} \\ f^-(w) &= \begin{cases} f(w), & w \leq \bar{w} \\ f(\bar{w}), & w \geq \bar{w} \end{cases} \end{aligned} \quad (2.10)$$

If we have Burger's equation, then the flux becomes:  $f(u) = \frac{u^2}{2}$ :

$$\begin{aligned} f^+(u) &= \begin{cases} \frac{u^2}{2}, & u \geq u^* \\ 0, & u \leq u^* \end{cases} \\ f^-(u) &= \begin{cases} \frac{u^2}{2}, & u \leq u^* \\ 0, & u \geq u^* \end{cases} \end{aligned} \quad (2.11)$$

If we consider an arbitrary, following the notation from Equation 2.5, we have,

$$\Delta_- f_+ = f^+(w_j) + f^-(w_{j+1}) \quad (2.12)$$

The Engquist and Osher scheme demonstrated computational advantages over all previous schemes including Godunov's and Cole-Murman, especially for implicit calculations where a much large time steps could be taken for E-O. It also eliminates the non-physical expansion shocks (satisfies the entropy condition) that plagues the C-M and Roe's method, and also possesses *smoother* flux functions than that of Godunov's.

### 2.2.2 Proof on E-O Satisfies the Entropy Condition

TO BE ADDED

### 3 The Roe Scheme

Professor Philip Roe first introduced the Roe scheme in 1980 in his paper *Approximate Riemann Solvers, Parameter Vectors, and Difference Scheme* []. In his paper, Roe argued that an exact solution to the Riemann problem may not be necessary and we could construct an *approximate* solutions which are the *exact* solutions to an approximate problem. This is the motive behind the *Roe's Average*.

#### 3.1 Roe's Average

Consider again the one-dimensional linear advection equation:

$$u_t + au_x = 0 \quad (3.1)$$

where we construct an approximate version:

$$\mathbf{u}_t + \tilde{A}\mathbf{u}_x = 0 \quad (3.2)$$

where  $\tilde{A}$  is to be chosen such that it satisfies the *local* conditions, i.e based on  $A_L$  and  $A_R$ . The intuitive candidates are, of course:

$$\begin{aligned} \tilde{A} &= \frac{1}{2}(A_L + A_R) \\ \tilde{A} &= A\left(\frac{1}{2}(\mathbf{u}_L + \mathbf{u}_R)\right) \end{aligned} \quad (3.3)$$

However, Roe listed out the following properties that the matrix  $\tilde{A}$  has to satisfy:

1. Linear mapping from the vector space  $\mathbf{u}$  to the vector space  $\mathbf{F}$
2. As  $\mathbf{u}_L \rightarrow \mathbf{u}_R \rightarrow \mathbf{u}$ ,  $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R) \rightarrow A(\mathbf{u})$  where  $A = \frac{\partial F}{\partial \mathbf{u}}$ . In words, the approximate Jacobian  $\tilde{A}$  must match with the exact Jacobian  $A$  as  $\mathbf{u}_L$  and  $\mathbf{u}_R$  approaches  $\mathbf{u}$
3. For any  $\mathbf{u}_L, \mathbf{u}_R$ ,  $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R) \times (\mathbf{u}_L - \mathbf{u}_R) = \mathbf{F}_L - \mathbf{F}_R$ . In words, conservation must be satisfied.
4. The eigenvectors of  $\tilde{A}$  are linearly independent, which means that the PDE has to be *hyperbolic*.

The above conditions are named Property U, as they are intended to possess uniform validity across discontinuities. In addition, matrix  $\tilde{A}$  is sometimes called *Roe's matrix*. It is to be noted that none of the suggestions for  $\tilde{A}$  above satisfies the third property.

We then follow that the matrix  $\tilde{A}$  could be constructed from mean value theorems which would arrive at the following:

$$\tilde{A} = \int_0^1 A(\theta) d\theta \quad (3.4)$$

where  $\theta$  is a parameter that vary linearly between 0 and 1 along a straight path connecting each state. By choosing a suitable integration path, candidates for  $A$  that is closed and cheap to compute could be found. We could then have the following formula based on  $\tilde{A}$ :

$$F_j(\mathbf{u}_L) - F_j(\mathbf{u}_R) = \sum_i \tilde{a}_{ij} [(u_i)_L - (u_i)_R] \quad (3.5)$$

where  $a_{ij}$  is the matrix entries into  $\tilde{A}$ . The  $\tilde{a}_{ij}$  satisfies the first and third conditions of Property U.

However, the primary disadvantage of the above construction lies in that the matrix is far from unique, and the resulting formulae for eigenvalues and eigenvectors are sometimes quite cumbersome. For Euler's equation, however, the above formulae simplifies greatly due to Euler's homogeneity characteristic.

### 3.2 Roe's Scheme for Euler's Equation

For Euler's compressible flow equation (or polytropic ideal gases), Roe introduced a change of variable:

$$\mathbf{w} = \rho^{1/2}(1, u, v, w, H)^T \quad (3.6)$$

which allow us to easily integrate and satisfy the Equation 3.5. We introduce two equations:

$$\begin{aligned} (\mathbf{u}_L - \mathbf{u}_R) &= \tilde{B}(\mathbf{w}_L - \mathbf{w}_R) \\ (\mathbf{F}_L - \mathbf{F}_R) &= \tilde{C}(\mathbf{w}_L - \mathbf{w}_R) \end{aligned} \quad (3.7)$$

where the first equation provides a link for  $\mathbf{w}$  in terms of  $\mathbf{u}$ . Then, through the new change of variables conducted we could write the following  $\tilde{B}$  and  $\tilde{C}$  matrix:

$$\begin{aligned} \tilde{B} &= \begin{pmatrix} 2\bar{w}_1 & 0 & 0 & 0 & 0 \\ \bar{w}_2 & \bar{w}_1 & 0 & 0 & 0 \\ \bar{w}_3 & 0 & \bar{w}_1 & 0 & 0 \\ \bar{w}_4 & 0 & 0 & \bar{w}_1 & 0 \\ \frac{\bar{w}_5}{\gamma} & \frac{\gamma-1}{\gamma}\bar{w}_2 & \frac{\gamma-1}{\gamma}\bar{w}_3 & \frac{\gamma-1}{\gamma}\bar{w}_4 & \frac{\bar{w}_1}{\gamma} \end{pmatrix} \\ \tilde{C} &= \begin{pmatrix} \bar{w}_2 & \bar{w}_1 & 0 & 0 & 0 \\ \frac{\gamma-1}{\gamma}\bar{w}_5 & \frac{\gamma-1}{\gamma}\bar{w}_2 & -\frac{\gamma-1}{\gamma}\bar{w}_3 & -\frac{\gamma-1}{\gamma}\bar{w}_4 & \frac{\gamma-1}{\gamma}\bar{w}_1 \\ 0 & \bar{w}_3 & \bar{w}_2 & 0 & 0 \\ 0 & \bar{w}_4 & 0 & \bar{w}_2 & 0 \\ 0 & \bar{w}_5 & 0 & 0 & \bar{w}_2 \end{pmatrix} \end{aligned} \quad (3.8)$$

The  $\bar{\cdot}$  indicates the arithmetic average of the selected quantities. We then obtain  $\tilde{A}$  via:

$$\tilde{A} = \tilde{C}\tilde{B}^{-1} \quad (3.9)$$

In the process of providing an analytical equation for both the eigenvalues and eigenvectors, Roe then introduced a different form of averaging for ease of convention:

$$u = \frac{\rho_L^{1/2}u_L + \rho_R^{1/2}u_R}{\rho_L^{1/2} + \rho_R^{1/2}} = \frac{\bar{w}_2}{\bar{w}_1} \quad (3.10)$$

We could also follow similar convention for:

$$v = \frac{\bar{w}_3}{\bar{w}_1}, \quad w = \frac{\bar{w}_4}{\bar{w}_1}, \quad H = \frac{\bar{w}_5}{\bar{w}_1} \quad (3.11)$$

The above relations are known as the Roe average, in which Roe derived on obtaining a solution to the linearized Riemann problem. It is to be noted that  $\tilde{A}$  is equal to the original matrix  $A$  when evaluated with the Roe averaged states above.

## 4 Tadmor's Entropy Stability Criterion

Eitan Tadmor, in his paper written in 1987, introduced a new, general, criterion capable of evaluating entropy conservation of different numerical fluxes. This section summarizes his paper and the two main theorems that he proposed:

1. Entropy conservation by means of comparison
2. Three-point conservative schemes are entropy stable if and only if they contain *more* viscosity than their entropy conservative equivalent

We focus on the derivation and discussion for the first theorem and briefly states the second.

#### 4.1 Tadmor's Entropy Stability Criterion

We again consider a one-dimensional conservation law:

$$\frac{\partial}{\partial t} \mathbf{u} + \frac{\partial}{\partial x} [\mathbf{f}(\mathbf{u})] = 0 \quad (4.1)$$

and we introduce the entropy function  $U(u)$ , a *convex* function of  $u$  along with the entropy flux  $F(u)$ , which is also a convex function of  $u$ . We recall the entropy condition:

$$\frac{\partial V}{\partial t} + \frac{\partial G}{\partial x} \leq 0 \quad (4.2)$$

where:

$$V(\mathbf{v}) = U(\mathbf{u}(\mathbf{v})), \quad G(\mathbf{v}) = F(\mathbf{u}(\mathbf{v})) \quad (4.3)$$

We also introduce the entropy variable:

$$\mathbf{v} = \mathbf{v}(\mathbf{u}) = \frac{\partial U}{\partial \mathbf{u}}(\mathbf{u}) \quad (4.4)$$

which is one-to-one mapping to  $\mathbf{u}(\mathbf{v})$  due to the convexity of  $U$ . We could then write the conservation equation as follow:

$$\frac{\partial}{\partial t} [\mathbf{u}(\mathbf{v})] + \frac{\partial}{\partial x} [\mathbf{g}(\mathbf{v})] = 0, \quad \mathbf{g}(\mathbf{v}) = \mathbf{f}(\mathbf{u}(\mathbf{v})) \quad (4.5)$$

We discretize Equation 4.5 and 4.2 and have the following:

$$\begin{aligned} \frac{d}{dt} [\mathbf{u}(\mathbf{v}_\nu(t))] &= -\frac{1}{\Delta x} [\mathbf{g}_{\nu+1/2} - \mathbf{g}_{\nu-1/2}] \\ \frac{d}{dt} V(\mathbf{v}_\nu(t)) + \frac{1}{\Delta x} [G_{\nu+1/2} - G_{\nu-1/2}] &\leq 0 \end{aligned} \quad (4.6)$$

which are the conservation equation along with its respective entropy inequality. Our end goal is *entropy conservative* schemes, in which only the equality hold in Equation 4.2. We multiply the first equation of Equation 4.6 by the entropy variable by  $\mathbf{v}_\nu^T(t) = U_{\mathbf{u}}^T(\mathbf{u}_\nu(\mathbf{t}))$ , which we obtain the following:

$$\frac{d}{dt} V(\mathbf{v}_\nu(t)) = \frac{d}{dt} U(\mathbf{u}_\nu(t)) = -\frac{1}{\Delta x} \mathbf{v}_\nu^T [\mathbf{g}_{\nu+1/2} - \mathbf{g}_{\nu-1/2}] \quad (4.7)$$

which we could substitute into the entropy condition and have the following equality. The expression:

$$\mathbf{v}_\nu^T [\mathbf{g}_{\nu+1/2} - \mathbf{g}_{\nu-1/2}] = \mathbf{G}_{\nu+1/2} - \mathbf{G}_{\nu-1/2} \quad (4.8)$$

is conservative if and only if:

$$\Delta \mathbf{v}_{\nu+1/2}^T \mathbf{g}_{\nu+1/2} = \Psi_{\nu+1} - \Psi_\nu, \quad \Delta \mathbf{v}_{\nu+1/2} = \mathbf{v}_{\nu+1} - \mathbf{v}_\nu \quad (4.9)$$

This is Tadmor's entropy stability criterion.  $\mathbf{g}_{\nu+1/2}$  could be taken as any numerical flux, while for Euler's equation in 1D:

$$\begin{aligned} \mathbf{v} &= \begin{pmatrix} \frac{\gamma-S}{\gamma-1} - \frac{\rho u^2}{2p} \\ m \\ p \\ -\frac{\rho}{p} \end{pmatrix} \\ \Psi &= \rho u \end{aligned} \quad (4.10)$$

Tadmor's new criterion is different in that the amount of numerical viscosity present is essentially *quantified* and is then later used for Entropy stability evaluation. In addition, it is to be noted that Tadmor's criterion is extremely general, which means that it could be applied to any numerical flux. Any flux could be tested for its entropy conservation by directly substituting the relevant variables and fluxes into Equation 4.9.

## 4.2 Tadmor's Second Theorem

In [1], Tadmor only proved entropy conservation (i.e Entropy is equal across a cell interface), not stability. He later extended his idea to essentially three-point scheme and proved that: *A conservative scheme which contains more numerical viscosity than that present in the entropy conservative one is entropy stable.*

Essentially, we consider three points schemes that could be written in the following viscosity form:

$$\frac{d}{dt} [(u(\mathbf{v}_\nu(t)))] = -\frac{1}{2\Delta x} [\mathbf{g}(\mathbf{v}_{\nu+1}) - \mathbf{g}(\mathbf{v}_{\nu-1})] + \frac{1}{2\Delta x} [Q_{\nu+1/2}\Delta\mathbf{v}_{\nu+1/2} - Q_{\nu-1/2}\Delta\mathbf{v}_{\nu-1/2}] \quad (4.11)$$

Where  $Q_{\nu+1/2}$  is referred to as the numerical viscosity coefficient matrix. We note that the class of three-point schemes includes second-order accurate TVD schemes such as the High-Resolution Osher flux and a number of other schemes [1]. We then introduce a new variable:

$$D_{\nu+1/2} = Q_{\nu+1/2} - Q_{\nu+1/2}^* \quad (4.12)$$

which denote the deviation of its viscosity from that of the entropy conservative scheme, which we then have:

$$\frac{d}{dt} [(u(\mathbf{v}_\nu(t)))] = -\frac{1}{2\Delta x} [\mathbf{g}(\mathbf{v}_{\nu+1}) - \mathbf{g}(\mathbf{v}_{\nu-1})] + \frac{1}{2\Delta x} [D_{\nu+1/2}\Delta\mathbf{v}_{\nu+1/2} - D_{\nu-1/2}\Delta\mathbf{v}_{\nu-1/2}] \quad (4.13)$$

We could also rewrite the above equation in the form of entropy condition (as per last section), following essentially the same step as outlined in Equation 4.7 and Equation 4.8 and use identities outlined in [1] to obtain the following:

$$\frac{d}{dt} V(\mathbf{v}_\nu(t)) + \frac{1}{\Delta x} [G_{\nu+1/2} - G_{\nu-1/2}] = -\frac{1}{4\Delta x} [\Delta\mathbf{v}_{\nu+1/2}^T D_{\nu+1/2} \Delta\mathbf{v}_{\nu+1/2} - \Delta\mathbf{v}_{\nu-1/2}^T D_{\nu-1/2} \Delta\mathbf{v}_{\nu-1/2}] \quad (4.14)$$

We observe that if the scheme is considered to contain more viscosity, then  $D$  would be positive, and the overall right hand side of the above equation remains *negative*, which then leads to the entropy condition being satisfied (Equation 4.6). This theorem allows us to tune additional numerical viscosity such that both entropy stability and second order accuracy could be obtained.