AEROSP 590: Conservation and Stability

Andi Zhou

March 14, 2023

1 Introduction

The following report presents a discussion regarding conservation properties and their corresponding stability of the a typical conservation law in multiple dimensions. We first discuss conservation and stability regarding a 1-D conservation, and we then move to 2-D and eventually numerical studies and confirmations.

2 1-Dimension Conservation Law

We first look at the following basic conservation equation:

$$0 = \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}}{\partial x} \tag{1}$$

In this case, $\mathbf{u}$ represents a conservative state (such as air density ρ and temperature T) and $\mathbf{f}$ is the flux of that particular state.

2.1 1st Order Conservation

We first prove that the above equation is conserved. In finite volume method, the first step towards proving conservation starts by taking the integral of the above equation to obtain the **weak form**. The biggest difference between finite volume method and finite difference method lies in the fact that finite volume method starts with the integral form of the conservation, which generally allows for *more information*. Finite difference method involves directly approximating the above equation with finite difference stencil, which **does not account for discontinuities** in the domain. We are proving conservation in a *semi-discrete fashion*, meaning that we only investigate the law spatially.

We first define a finite domain made of N finite cells, illustrated as follow:

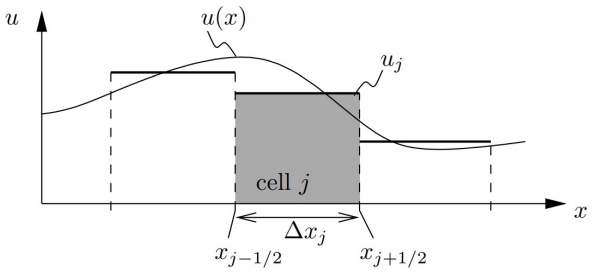

Figure 1: Caption

In finite volume methods, we store the state as a cell centered average and we use half indices $j \pm 1/2$ to represent edges.

We first begin by taking the integral of the state across one cell. Supposing we are centered on

an arbitrary cell j:

$$\begin{aligned}
\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx + \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{f}}{\partial x} dx = 0 \\
\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx + \mathbf{f}|_{x_{j-1/2}}^{x_{j+1/2}} = 0 \\
\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = -\mathbf{f}|_{x_{j-1/2}}^{x_{j+1/2}} \\
\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = -(\mathbf{f}_{x_{j+1/2}} - \mathbf{f}_{x_{j-1/2}})
\end{aligned} \tag{2}$$

We then arrive at the final semi-discrete equation for a single cell:

$$\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \mathbf{f}_{x_{j-1/2}} - \mathbf{f}_{x_{j+1/2}} \tag{3}$$

We then refer to our grid presented in Figure 1. We assume a grid of N elements, and sum over each element with respect to Equation 3 in order to examine the flux conservation. We obtain the following formula:

$$\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \sum_{j=1}^{N} (\mathbf{f}_{x_{j-1/2}} - \mathbf{f}_{x_{j+1/2}}) \tag{4}$$

We expand the summation on the right hand side, which we then obtain the following:

$$\begin{aligned}
\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \sum_{j=1}^{N} (\mathbf{f}_{x_{j-1/2}} - \mathbf{f}_{x_{j+1/2}}) \\
= (\mathbf{f}_{x_{1-1/2}} - \mathbf{f}_{x_{1+1/2}}) + (\mathbf{f}_{x_{2-1/2}} - \mathbf{f}_{x_{2+1/2}}) + (\mathbf{f}_{x_{3-1/2}} - \mathbf{f}_{x_{3+1/2}}) \dots (\mathbf{f}_{x_{N-1/2}} - \mathbf{f}_{x_{N+1/2}})
\end{aligned} \tag{5}$$

We could observe that all interior fluxes cancel out, because they all share the common edge $j \pm 1/2$. Therefore, we are left with the following:

$$\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \mathbf{f}_{1/2} - \mathbf{f}_{N+1/2} \tag{6}$$

The only flux contributions left are the boundary fluxes, since all interior fluxes cancelled out. This is called the *telescoping* property. In other words, this equation speaks that what comes through the boundaries determines the time variation of state throughout all cells. This is consistent of the general statement of finite volume conservation: *what comes in equals to what comes out.*

However, we have not taken *first order* into account here yet. First order FVM could essentially be considered as a flux approximation. In the beginning of this section, we introduced the concept that finite volume method stores state value as cell averages. However, this means that we could not obtain the flux values at the edges (locations of half indices). Therefore, we need some form of approximations. In first order methods, we approximate the flux on the cell edges as a function of the average state value $\mathbf{u}$. Then, Equation 5 becomes the following:

$$\begin{aligned}
\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = (\mathbf{f}(u_1) - \mathbf{f}(u_2)) + (\mathbf{f}(u_2) - \mathbf{f}(u_3)) + (\mathbf{f}(u_3) - \mathbf{f}(u_4)) \dots (\mathbf{f}(u_{N-1}) - \mathbf{f}(u_N)) \\
\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \mathbf{f}(u_1) - \mathbf{f}(u_N)
\end{aligned} \tag{7}$$

We see that the telescoping property still exist, and the overall solution conservation is not affected. This brings up an important statement that the *order of the approximation in FVM does not affect the general conservation statement.* In addition, as we will see in the last section, *having a conservative scheme does not guarantee solution stability.*

2.2 Diffusion

Now we look at the conservation law with the diffusion term on the right hand side. Typically, conservation with diffusion is represented as follow:

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}}{\partial x} = \frac{\partial^2 \mathbf{u}}{\partial^2 x} \tag{8}$$

To determine whether this equation is conserved, we determine whether the diffusion term could be casted as a **flux**, which in this case, it can.

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}}{\partial x} &= \frac{\partial^2 \mathbf{u}}{\partial^2 x} \\ \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}}{\partial x} &= \frac{\partial}{\partial x}(\frac{\partial \mathbf{u}}{\partial x}) \end{aligned} \tag{9}$$

In this case, $\frac{\partial \mathbf{u}}{\partial x}$ is the **diffusive flux**. We could write out its contributions again for each element $\frac{\partial \mathbf{u}}{\partial x}|_{x_{j-1/2}}^{x_{j+1/2}}$ and sum up over the whole domain. We would again have the telescoping effect and every flux contribution would cancel out except for the boundary fluxes.

It is then worthwhile to discuss what constitutes as a *conservative equation*. This strongly depends on the formulation of the *strong form*. We take a look at Burger's equation.

The conservative form of Burger's equation is:

$$\frac{d\mathbf{u}}{dx} + \frac{\partial}{\partial x}(\frac{\mathbf{u}^2}{2}) = 0 \tag{10}$$

The non-conservative form of Burger's equation is:

$$\frac{d\mathbf{u}}{dx} + \mathbf{u}\frac{\partial \mathbf{u}}{\partial x} = 0 \tag{11}$$

The flux term for the conservative Burgers is $\boxed{\frac{\partial}{\partial x}(\frac{\mathbf{u}^2}{2})}$ and $\boxed{\mathbf{u}\frac{\partial \mathbf{u}}{\partial x}}$ for the non-conservative version. Then, we use central flux and we observe the following phenomenon:

Conservative:

$$\sum \frac{\partial}{\partial x}(\frac{u^2}{2}) = \frac{u_1^2 - u_0^2}{2\Delta x} + \frac{u_2^2 - u_1^2}{2\Delta x} + \frac{u_3^2 - u_2^2}{2\Delta x} = \frac{u_3^2 - u_0^2}{2\Delta x} \tag{12}$$

Non-Conservative:

$$\sum u\frac{\partial u}{\partial x} = u_1\frac{u_1 - u_0}{\Delta x} + u_2\frac{u_2 - u_1}{\Delta x} + u_3\frac{u_3 - u_2}{\Delta x} = \frac{u_1(u_1 - u_0) + u_2(u_2 - u_1) + u_3(u_3 - u_2)}{\Delta x}) \tag{13}$$

For the conservative form, we observe the telescoping effect where all interior fluxes are cancelled out. For the non-conservative form of the Burger's equation however, there exists source terms that does not cancel out, leading to state generations/sinks within each cell. We observe that whether a conservation law is truly conservative depends on its formulation in the strong form. For Burger's equation, the conservative form have its flux within the spatial derivative $\frac{\partial}{\partial x}$. In the non-conservative form, however, there exists a state dependent term outside of the flux function, which leads to source term during flux discretizations.

2.3 Source Terms

2.4 2nd Order Conservation

We then look at the second order finite volume method and its relationship with conservation. We recall that first order FVM stores the state as cell averages per each cell, and we use the cell averages to approximate the flux on the edges. In 2nd order, we make use of neighboring cells to calculate the gradient within each cell, and use the gradient information to approximate the fluxes on the edge. In short, we have a linear representation of the state within each cell in 2nd order FVM, and the flux function now depend on cell-edge values.

Mathematically, Equation 5 becomes the following:

$$\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \sum_{j=1}^{N} (\mathbf{f}_{x_{j-1/2}} - \mathbf{f}_{x_{j+1/2}})$$

$$= (\mathbf{f}_{x_{1-1/2}} - \mathbf{f}_{x_{1+1/2}}) + (\mathbf{f}_{x_{2-1/2}} - \mathbf{f}_{x_{2+1/2}}) + (\mathbf{f}_{x_{3-1/2}} - \mathbf{f}_{x_{3+1/2}}) \dots (\mathbf{f}_{x_{N-1/2}} - \mathbf{f}_{x_{N+1/2}})$$

$$= (\mathbf{f}(\mathbf{u}_{1-1/2}) - \mathbf{f}(\mathbf{u}_{1+1/2})) + (\mathbf{f}(\mathbf{u}_{2-1/2)} - \mathbf{f}(\mathbf{u}_{2+1/2})) + (\mathbf{f}(\mathbf{u}_{3-1/2)} - \mathbf{f}(\mathbf{u}_{3+1/2})) \dots (\mathbf{f}(\mathbf{u}_{N-1/2)} - \mathbf{f}(\mathbf{u}_{N+1/2}))$$

(14)

Again with telescoping, we obtain the following equation of conservation:

$$\sum_{j=1}^{N} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{\partial \mathbf{u}}{\partial t} dx = \mathbf{f}(\mathbf{u}_{1/2}) - \mathbf{f}(\mathbf{u}_{N+1/2}) \tag{15}$$

In the above derivation, notice we replaced cell averaged value from Equation 7 $\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_N$ with cell edge values $\mathbf{u}_{1\pm1/2}, \mathbf{u}_{2\pm1/2} \dots \mathbf{u}_{N\pm1/2}$. However, the conservation of the equation is not affected as all interior fluxes are cancelled out. Therefore, solver order does not impact the conservation law.

2.5 Temporal Discretization

3 2-Dimension Conservation Law

We now turn our attention to 2-D conservation law:

$$\frac{d\mathbf{u}}{dt} + \nabla \cdot \mathbf{f} \tag{16}$$

where the divergence operator is restricted in $x - y$. We first investigate conservation for an arbitrary 2D mesh.

3.1 2D Unstructured Mesh

We prove that in an unstructured mesh, the conservation law in 2-Dimension still holds.

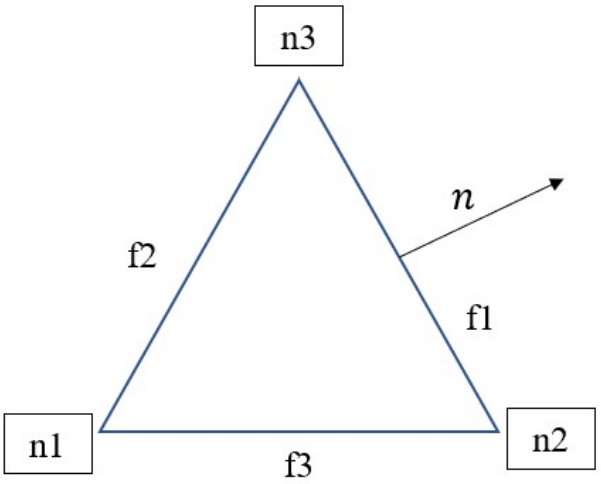

Figure 2: Illustration of a single triangle cell

In the above triangle cell, each vertex shares the same numbering as the face directly opposite from it. In addition, the normal on each face *always* points outward into the neighboring cell, unless it is a boundary element.

We again start by integrating Equation 16 according to each cell's area. This time, however, the integral becomes an area integral due to 2-D.

$$\iint_{\Omega} \frac{d\mathbf{u}}{dt} d\Omega + \iint_{\Omega} \nabla \cdot \mathbf{f} d\Omega = 0 \tag{17}$$

We use divergence theorem, which effectively turns the area integral into a line integral which we could integrate over numerically with ease:

$$\oint_{\partial\Omega} F \cdot \hat{n} ds = \iint_{\Omega} \text{div} F dA \tag{18}$$

The flux term then becomes that of the following:

$$\begin{aligned} \iint_{\Omega} \frac{d\mathbf{u}}{dt} d\Omega + \oint_{\partial\Omega} \mathbf{f} \cdot \hat{n} ds = 0 \\ \iint_{\Omega} \frac{d\mathbf{u}}{dt} d\Omega + \sum_{i=1}^{3} (F_i \cdot n_i) s_i = 0 \end{aligned} \tag{19}$$

where we replaced the line integral with a numerical summation over each edge since we have a fixed normal due to our triangular cell. Note that although triangle is used as an example in this case, technically any *closed* element could be used.

We could already observe telescoping effect from this equation. If we were to sum over every element of the mesh with Equation 19, the ideal algorithm would walk along each cell (typically in counter-clockwise fashion) and add the contribution of the flux to its state. Imagine we are walking on an interior edge, as demonstrated in Figure 3

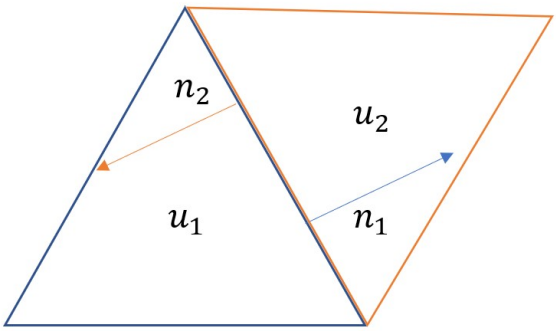

Figure 3: 2 connected elements

Assuming we are looping through on a per element basis, we would visit this edge twice. Once when we walk around element u_1, and a second time when we loop around element u_2. Since the edge is connected, we would obtain the same $\mathbf{f}$ magnitude each time. This condition is guaranteed mathematically by Equation 19 and numerically via the usage of different numerical fluxes. However, the normals are defined as pointing out of each individual element. Therefore, the two $\mathbf{f}$ quantities on each side would become an equal but opposite *vector*. This would effectively cancel out all interior fluxes within a mesh, and leaving out only the boundary contributions, analogous to Equation 3 in 1D.

3.2 2nd Order Finite Volume Scheme

In this section we introduce the Monotonic Upstream-centered Scheme for Conservation Laws (MUSCL), a second order finite volume method that is extremely popular among today's commercial solvers. We discuss the algorithm in a pseudo-code fashion and explain the gradient calculations.

Note that the above scheme pseudo-code is only for steady-state simulations. For transient simulation, all cells would take the lowest time-step calculated within the entire mesh.

Algorithm 1 MUSCL Scheme for Steady Calculations

Compute/store edge normals, cell areas, etc
Initialize the state from first-order solve (Warm-Start)
while |R| > tolerance **do**
    Initialize the gradient, $\nabla u|_i$, to zero on each element
    for f = 1:N_{face} **do**
        Let L/R be the elements adjacent to face f
        Let $\xrightarrow{n}$ be the normal pointing out of L
        Let Δl be the length of face f
        set **u** to the average of the L and R cell averages
        add $\mathbf{u} \xrightarrow{n} \Delta l$ to $\nabla u|_i$ on L and subtract it from R
    end for
    Divide each $\nabla u|_i$ by its element area, A_i
    Initialize residual and wave speed on each cell to zero
    for $f = 1 : N_{face}$ **do**
        Let L/R be the elements adjacent to face f
        Compute the L/R states on the face midpoint
        Compute the flux on face f using the *edge* states
        Increment residual on L, decrement it on R
        Add wave speed to tallies on L and R cells
    end for
    Compute the time step on each cell
    Update the state: use RK2 or higher-order method
end while