Fluctuation-Splitting Schemes and Hyp/Ell Decompositions of the Euler Equations

Hiroaki Nishikawa

August 2004

1 Fluctuations and Forms of The Euler Equations

Consider the Euler equations in two-dimensions

$$\mathbf{u}_t + \mathbf{f}_x + \mathbf{g}_y = 0 \tag{1}$$

where

$$\mathbf{u} = [\rho, \rho u, \rho v, \rho E]^t. \tag{2}$$

In the fluctuation-splitting schemes, this is conservatively linearized using the parameter vector,

$$\mathbf{u}_t + \mathbf{A}_z \mathbf{z}_x + \mathbf{B}_z \mathbf{z}_y = 0 \tag{3}$$

where

$$\mathbf{A}_z = \begin{bmatrix} z_2 & z_1 & 0 & 0 \\ \frac{\gamma-1}{\gamma} z_4 & \frac{\gamma+1}{\gamma} z_2 & -\frac{\gamma-1}{\gamma} z_3 & \frac{\gamma-1}{\gamma} z_1 \\ 0 & z_3 & z_2 & 0 \\ 0 & z_4 & 0 & z_2 \end{bmatrix} \tag{4}$$

$$\mathbf{B}_z = \begin{bmatrix} z_3 & 0 & z_1 & 0 \\ 0 & z_3 & z_2 & 0 \\ \frac{\gamma-1}{\gamma} z_4 & -\frac{\gamma-1}{\gamma} z_2 & \frac{\gamma+1}{\gamma} z_3 & \frac{\gamma-1}{\gamma} z_1 \\ 0 & 0 & z_4 & z_3 \end{bmatrix} \tag{5}$$

Note that these are NEITHER Jacobians nor coefficient matrices for the equations for the parameter vector because we still keep the conservative variables in the time derivative. A nice thing about this form is that it can be easily linearized over a triangular element because these matrices are linear in $\mathbf{z}$. Assuming a piecewise linear variation of $\mathbf{z}$ over the element, we see that its gradients are constant within the element and therefore we can easily integrate them to obtain the fluctuation

$$\begin{aligned} \mathbf{\Phi} &= \int_T \mathbf{u}_t = -\int_T (\mathbf{A}_z \mathbf{z}_x + \mathbf{B}_z \mathbf{z}_y) &(6)\\ &= \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}}_z n_{x_i} + \tilde{\mathbf{B}}_z n_{y_i}\right) \mathbf{Z}_i &(7)\end{aligned}$$

where $\mathbf{Z}_i$ are the numarical values of $\mathbf{z}$ stored at the node i, and $\hat{\mathbf{A}}_z$ and $\hat{\mathbf{B}}_z$ are both evaluated at the arithmetic average of $\mathbf{Z}_i$. The integration is exact under the assumption. Therefore, this fluctuation represents the flux balance exactly,

$$\mathbf{\Phi} = \oint_{\partial T} \mathbf{f} dy - \mathbf{g} dx \tag{8}$$

with a certain quadrature rule which is perhaps very complicated. This is important for conservation.

Once we linearize the equations, it is a simple matter to rewrite the fluctuation in other forms. For instance, to write it in terms of the conservative Jacobians, we simply use the transformation matrix,

$$\begin{aligned} \mathbf{\Phi} &= \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}} n_{x_i} + \tilde{\mathbf{B}} n_{y_i}\right) \widetilde{\frac{\partial \mathbf{u}}{\partial \mathbf{z}}} \mathbf{Z}_i &(9)\\ &= \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}} n_{x_i} + \tilde{\mathbf{B}} n_{y_i}\right) \mathbf{U}_i &(10)\end{aligned}$$

This means that we can use the conservative Jacobians evaluated at the average state in $\mathbf{z}$ as long as we compute the nodal conservative variables from the nodal parameter vectors using the transformation matrix which is also evaluated at the same average state, i.e. we must compute $\mathbf{U}_i$ by

$$\mathbf{U}_i = \widetilde{\frac{\partial \mathbf{u}}{\partial \mathbf{z}}} \mathbf{Z}_i \tag{11}$$

Starting with this form, we obtain, for example, the fluctuation for the equations with the primitive variables $\mathbf{w}$ (although not very useful).

$$\begin{aligned} \mathbf{\Phi}^w = \widetilde{\frac{\partial \mathbf{w}}{\partial \mathbf{u}}} \mathbf{\Phi} &= \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}}_w n_{x_i} + \tilde{\mathbf{B}}_w n_{y_i}\right) \widetilde{\frac{\partial \mathbf{w}}{\partial \mathbf{u}}} \mathbf{U}_i &(12)\\ &= \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}}_w n_{x_i} + \tilde{\mathbf{B}}_w n_{y_i}\right) \mathbf{W}_i &(13)\end{aligned}$$

where

$$\mathbf{W}_i = \widetilde{\frac{\partial \mathbf{w}}{\partial \mathbf{u}}} \mathbf{U}_i = \widetilde{\frac{\partial \mathbf{w}}{\partial \mathbf{u}}} \widetilde{\frac{\partial \mathbf{u}}{\partial \mathbf{z}}} \mathbf{Z}_i = \widetilde{\frac{\partial \mathbf{w}}{\partial \mathbf{z}}} \mathbf{Z}_i \tag{14}$$

Therefore, any form can be used as long as the nodal values are computed from the parameter vector with a linearized transformation matrix.

2 Hyp/Ell Decomposition; Preliminary consideration

The two-dimensional Euler equations can be decomposed into hyperbolic and elliptic parts by using the preconditioning matrix of Van Leer-Lee-Roe,

$$\mathbf{P}_{vll} = \begin{bmatrix} \frac{\tau}{\beta^2}M^2 & -\frac{\tau}{\beta^2}M & 0 & 0 \\ -\frac{\tau}{\beta^2}M & \frac{\tau}{\beta^2}+1 & 0 & 0 \\ 0 & 0 & \tau & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{15}$$

where

$$\beta = \begin{cases} \sqrt{1-M^2} & M < 1 \\ \sqrt{M^2-1} & M \geq 1 \end{cases} \tag{16}$$

$$\tau = \begin{cases} \sqrt{1-M^2} & M < 1 \\ \sqrt{1-M^{-2}} & M \geq 1 \end{cases} \tag{17}$$

This form of the preconditioning matrix is for use in the symmetric form of the Euler equations which is obtained as follows. Conservative variables are transformed by a set of symmetrizing variables,

$$\partial\mathbf{v} = \left[\frac{\partial p}{\rho a}, \partial q, q\partial\theta, \partial p - a^2\partial\rho\right]^t \tag{18}$$

by the transformation matrices

$$\frac{\partial\mathbf{u}}{\partial\mathbf{v}} = \begin{bmatrix} \rho/a & 0 & 0 & -1/a^2 \\ u\rho/a & \rho u/q & -\rho v/q & -u/a^2 \\ v\rho/a & \rho v/q & \rho u/q & -v/a^2 \\ \rho a(\frac{1}{2}M^2 + \frac{1}{\gamma-1}) & \rho q & 0 & -\frac{1}{2}M^2 \end{bmatrix} \tag{19}$$

$$\frac{\partial\mathbf{v}}{\partial\mathbf{u}} = \begin{bmatrix} \frac{\gamma-1}{2}\frac{q^2}{\rho a} & -\frac{(\gamma-1)u}{\rho a} & -\frac{(\gamma-1)v}{\rho a} & \frac{(\gamma-1)}{\rho a} \\ -q/\rho & u/(\rho q) & v/(\rho q) & 0 \\ 0 & -v/(\rho q) & u/(\rho q) & 0 \\ \frac{\gamma-1}{2}q^2 - a^2 & -(\gamma-1)u & -(\gamma-1)v & \gamma-1 \end{bmatrix} \tag{20}$$

Then, the equations are transformed by applying $\frac{\partial\mathbf{v}}{\partial\mathbf{u}}$ from the left,

$$\frac{\partial\mathbf{v}}{\partial\mathbf{u}}\mathbf{u}_t + \frac{\partial\mathbf{v}}{\partial\mathbf{u}}\mathbf{A}\mathbf{u}_x + \frac{\partial\mathbf{v}}{\partial\mathbf{u}}\mathbf{B}\mathbf{u}_y = 0 \tag{21}$$

$$\frac{\partial\mathbf{v}}{\partial t} + \frac{\partial\mathbf{v}}{\partial\mathbf{u}}\mathbf{A}\frac{\partial\mathbf{u}}{\partial\mathbf{v}}\frac{\partial\mathbf{v}}{\partial x} + \frac{\partial\mathbf{v}}{\partial\mathbf{u}}\mathbf{B}\frac{\partial\mathbf{u}}{\partial\mathbf{v}}\frac{\partial\mathbf{v}}{\partial y} = 0 \tag{22}$$

$$\frac{\partial\mathbf{v}}{\partial t} + \mathbf{A}_{\mathbf{v}}\frac{\partial\mathbf{v}}{\partial x} + \mathbf{B}_{\mathbf{v}}\frac{\partial\mathbf{v}}{\partial y} = 0 \tag{23}$$

where the coefficient matrices are found to be

$$\mathbf{A_v} = \begin{bmatrix} u & ua/q & -va/q & 0 \\ ua/q & u & 0 & 0 \\ -va/q & 0 & u & 0 \\ 0 & 0 & 0 & u \end{bmatrix} \tag{24}$$

$$\mathbf{B_v} = \begin{bmatrix} v & va/q & ua/q & 0 \\ va/q & v & 0 & 0 \\ ua/q & 0 & v & 0 \\ 0 & 0 & 0 & v \end{bmatrix} \tag{25}$$

which become even simpler if written in the streamline coordinates (with θ denoting the flow angle),

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{A_v}\cos\theta + \mathbf{B_v}\sin\theta)\frac{\partial \mathbf{v}}{\partial s} + (-\mathbf{A_v}\sin\theta + \mathbf{B_v}\cos\theta)\frac{\partial \mathbf{v}}{\partial n} = 0 \tag{26}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{A_v^s}\frac{\partial \mathbf{v}}{\partial s} + \mathbf{B_v^n}\frac{\partial \mathbf{v}}{\partial n} = 0 \tag{27}$$

$$\mathbf{A_v^s} = \begin{bmatrix} q & a & 0 & 0 \\ a & q & 0 & 0 \\ 0 & 0 & q & 0 \\ 0 & 0 & 0 & q \end{bmatrix} \tag{28}$$

$$\mathbf{B_v^n} = \begin{bmatrix} 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{29}$$

This is a consequence of the choice of the symmetrizing variables. We would have obtained these simple matrices in $x-y$ coordinates if we had chosen

$$\partial \mathbf{v} = \left[\frac{\partial p}{\rho a}, \partial u, \partial v, \partial p - a^2\partial\rho\right]^t \tag{30}$$

Now we apply the preconditioning matrix to the symmetric system.

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{P}_{vll}\left(\mathbf{A_v^s}\frac{\partial \mathbf{v}}{\partial s} + \mathbf{B_v^n}\frac{\partial \mathbf{v}}{\partial n}\right) = 0 \tag{31}$$

Note that this is no longer the Euler equations; we are solving something different which has the optimal condition number. This system itself is not in a decomposed form. To see the decoupling, we need to express the system using other variables such as

$$\partial \mathbf{x} = \left[\partial p - a^2\partial\rho,\ \partial p + \rho q\partial q,\ \beta\partial p + \rho q^2\partial\theta,\ \beta\partial p - \rho q^2\partial\theta\right]^t \tag{32}$$

Note that the second quantity may be called enthalpy, but it is not equal to $\rho\,\partial H$: $\partial h \equiv \partial p + \rho q \partial q = \rho\,\partial H - \frac{\partial p - a^2\partial\rho}{\gamma-1} = \rho\,\partial H - \frac{\partial S}{\gamma - 1}$. ($H = E + p/\rho = e + q^2/2 + p/\rho$)

Matrices necessary to carry out the transformation are give by

$$\frac{\partial \mathbf{x}}{\partial \mathbf{v}} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ \rho a & \rho q & 0 & 0 \\ \beta\rho a & 0 & \rho q & 0 \\ \beta\rho a & 0 & -\rho q & 0 \end{bmatrix} \tag{33}$$

$$\frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 & 1/(2\rho a\beta) & 1/(2\rho a\beta) \\ 0 & 1/(\rho q) & -1/(2\rho q\beta) & -1/(2\rho q\beta) \\ 0 & 0 & 1/(2\rho q) & -1/(2\rho q) \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{34}$$

Then we obtain the decoupled system as follows.

$$\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\left(\mathbf{A}^s_{\mathbf{v}}\frac{\partial \mathbf{v}}{\partial s} + \mathbf{B}^n_{\mathbf{v}}\frac{\partial \mathbf{v}}{\partial n}\right) = 0 \tag{35}$$

$$\frac{\partial \mathbf{x}}{\partial t} + \left(\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\mathbf{A}^s_{\mathbf{v}}\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\frac{\partial \mathbf{v}}{\partial s} + \left(\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\mathbf{B}^s_{\mathbf{v}}\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\frac{\partial \mathbf{v}}{\partial n} = 0 \tag{36}$$

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}^s_{\mathbf{x}}\frac{\partial \mathbf{x}}{\partial s} + \mathbf{B}^n_{\mathbf{x}}\frac{\partial \mathbf{x}}{\partial n} = 0 \tag{37}$$

where

$$\mathbf{A}^s_{\mathbf{x}} = \begin{bmatrix} q & 0 & 0 & 0 \\ 0 & q & 0 & 0 \\ 0 & 0 & \tau q\nu^+ & -\tau q\nu^- \\ 0 & 0 & -\tau q\nu^- & \tau q\nu^+ \end{bmatrix} \tag{38}$$

$$\mathbf{B}^n_{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\tau q}{\beta} & 0 \\ 0 & 0 & 0 & -\frac{\tau q}{\beta} \end{bmatrix} \tag{39}$$

$$\nu^{\pm} = \frac{\beta^2 \pm (M^2-1)}{2\beta^2} \tag{40}$$

This is one form of hyp/ell decomposition of the Euler. Note that $\nu^- = 0$ for subsonic flows, and then the system is diagonalized completely. In the actual implementation, it may be convenient to write the system in $x-y$ coordinates.

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}_{\mathbf{x}}\frac{\partial \mathbf{x}}{\partial x} + \mathbf{B}_{\mathbf{x}}\frac{\partial \mathbf{x}}{\partial y} = 0 \tag{41}$$

where

$$\mathbf{A}_{\mathbf{x}} = \begin{bmatrix} u & 0 & 0 & 0 \\ 0 & u & 0 & 0 \\ 0 & 0 & \frac{\tau(\nu^+ u\beta - v)}{\beta} & -\tau u\nu^- \\ 0 & 0 & -\tau u\nu^- & \frac{\tau(\nu^+ u\beta + v)}{\beta} \end{bmatrix} \tag{42}$$

$$\mathbf{B_x} = \begin{bmatrix} v & 0 & 0 & 0 \\ 0 & v & 0 & 0 \\ 0 & 0 & \frac{\tau(\nu^+ v\beta + u)}{\beta} & -\tau v \nu^- \\ 0 & 0 & -\tau v \nu^- & \frac{\tau(\nu^+ v\beta - u)}{\beta} \end{bmatrix} \tag{43}$$

NOTE: This is different from what Mani used in his Euler code. He further manipulated the elliptic subsystem to make it simpler. This destroys the optimal condition number of the subsonic. If decomposition is the only interest, it is OK.

It is also important in the fluctuation-splitting schemes to know how this system is obtained from the conservative form. Tracking back the transformation, we find the decomposed system is equivalent to

$$\frac{\partial \mathbf{x}}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\frac{\partial \mathbf{v}}{\partial \mathbf{u}}\left(\mathbf{A}\frac{\partial \mathbf{u}}{\partial x} + \mathbf{B}\frac{\partial \mathbf{u}}{\partial y}\right) = 0 \tag{44}$$

where

$$\frac{\partial \mathbf{x}}{\partial \mathbf{u}} = \begin{bmatrix} \frac{\gamma-1}{2}q^2 - a^2 & -(\gamma-1)u & -(\gamma-1)v & \gamma-1 \\ \frac{\gamma-3}{2}q^2 & -(\gamma-2)u & -(\gamma-2)v & \gamma-1 \\ \frac{\gamma-1}{2}q^2\beta & (1-\gamma)u\beta - v & (1-\gamma)v\beta + u & (\gamma-1)\beta \\ \frac{\gamma-1}{2}q^2\beta & (1-\gamma)u\beta + v & (1-\gamma)v\beta - u & (\gamma-1)\beta \end{bmatrix} \tag{45}$$

$$\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\frac{\partial \mathbf{v}}{\partial \mathbf{u}} = \begin{bmatrix} \frac{\gamma-1}{2}q^2 - a^2 & -(\gamma-1)u & -(\gamma-1)v & \gamma-1 \\ -q^2 & u & v & 0 \\ \frac{\tau M^2}{\beta}(\frac{\gamma-1}{2}q^2 + a^2) & -\frac{\tau}{\beta}[(\gamma-1)uM^2 + (u+v\beta)] & -\frac{\tau}{\beta}[(\gamma-1)vM^2 + (v-u\beta)] & \frac{(\gamma-1)\tau M^2}{\beta} \\ \frac{\tau M^2}{\beta}(\frac{\gamma-1}{2}q^2 + a^2) & -\frac{\tau}{\beta}[(\gamma-1)uM^2 + (u-v\beta)] & -\frac{\tau}{\beta}[(\gamma-1)vM^2 + (v+u\beta)] & \frac{(\gamma-1)\tau M^2}{\beta} \end{bmatrix} \tag{46}$$

But perhaps more useful ones would be their inverse matrices,

$$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \begin{bmatrix} -\frac{1}{a^2} & 0 & \frac{1}{2a^2\beta} & \frac{1}{2a^2\beta} \\ -\frac{u}{a^2} & \frac{u}{q^2} & \frac{u}{2a^2\beta} - \frac{u+v\beta}{2\beta q^2} & \frac{u}{2a^2\beta} - \frac{u-v\beta}{2\beta q^2} \\ -\frac{v}{a^2} & \frac{v}{q^2} & \frac{v}{2a^2\beta} - \frac{v-u\beta}{2\beta q^2} & \frac{v}{2a^2\beta} - \frac{v+u\beta}{2\beta q^2} \\ -\frac{M^2}{2} & 1 & \frac{M^2}{4\beta} - \frac{\gamma-2}{2\beta(\gamma-1)} & \frac{M^2}{4\beta} - \frac{\gamma-2}{2\beta(\gamma-1)} \end{bmatrix} \tag{47}$$

$$\left(\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\frac{\partial \mathbf{v}}{\partial \mathbf{u}}\right)^{-1} = \begin{bmatrix} -1/a^2 & 1/q^2 & \frac{\beta}{2q^2\tau} & \frac{\beta}{2q^2\tau} \\ -u/a^2 & 2u/q^2 & -\frac{v-u\beta}{2q^2\tau} & \frac{v+u\beta}{2q^2\tau} \\ -u/a^2 & 2u/q^2 & \frac{u+v\beta}{2q^2\tau} & -\frac{u-v\beta}{2q^2\tau} \\ -\frac{M^2}{2} & \frac{3}{2} + \frac{1}{(\gamma-1)M^2} & \frac{\beta}{4\tau}[1 + \frac{2}{(\gamma-1)M^2}] & \frac{\beta}{4\tau}[1 + \frac{2}{(\gamma-1)M^2}] \end{bmatrix} \tag{48}$$

Incidentally, the determinant of $\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{P}_{vll}\frac{\partial \mathbf{v}}{\partial \mathbf{u}}$ is

$$-\frac{2\tau^2 q^2(\gamma-1)}{\beta} \tag{49}$$

which vanishes at $M = 1$ and $q = 0$.

In the actual coding, we work with the decomposed form of the Euler,

$$\mathbf{\Phi}^{\mathbf{x}} = \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}}_x n_{x_i} + \tilde{\mathbf{B}}_x n_{y_i}\right) \mathbf{X}_i \tag{50}$$

where

$$\mathbf{X}_i = \frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{z}} \mathbf{Z}_i \tag{51}$$

we distribute this to the vertices of the element, $\mathbf{\Phi}_i^{\mathbf{x}}$ $(i = 1, 2, 3)$. What variables to update with it requires consideration of conservation.

3 Hyp/Ell Decomposition; Revised

The decomposition in the previous section is not unique in the sense that we may express that in other variables. For example,

$$\partial \mathbf{x} = \left[\frac{\gamma-1}{\gamma}(\partial p - a^2\partial\rho),\, \rho\partial H,\, \beta\partial p + \rho q^2\partial\theta,\, \beta\partial p - \rho q^2\partial\theta\right]^t \tag{52}$$

This decomposes the symmetric Euler in the exactly the same way as before. The resulting decomposed form is identical to (37) - (43). This set of variables is convenient if we work with the parameter vector since the transformation matrix is simple.

$$\frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \begin{bmatrix} (1-2\gamma)z_4 + \frac{\gamma}{z_1}(z_2^2+z_3^2) & -z_2 & -z_3 & z_1 \\ -z_4 & 0 & 0 & z_1 \\ \frac{\beta(\gamma-1)}{\gamma}z_4 & -\frac{\beta(\gamma-1)}{\gamma}z_2 - z_3 & -\frac{\beta(\gamma-1)}{\gamma}z_3 + z_2 & \frac{\beta(\gamma-1)}{\gamma}z_1 \\ \frac{\beta(\gamma-1)}{\gamma}z_4 & -\frac{\beta(\gamma-1)}{\gamma}z_2 + z_3 & -\frac{\beta(\gamma-1)}{\gamma}z_3 - z_2 & \frac{\beta(\gamma-1)}{\gamma}z_1 \end{bmatrix} \tag{53}$$

This is exactly what Mani presents in his thesis (cf. p.90). But strangely, the variables he presents are different from (57). On the other hand, the inverse transformation matrix is given by

$$\frac{\partial \mathbf{z}}{\partial \mathbf{x}} = [\mathbf{r}_s,\, \mathbf{r}_H,\, \mathbf{r}_{w^+},\, \mathbf{r}_{w^-},] \tag{54}$$

where

$$\mathbf{r}_s = -\frac{1+\frac{\gamma-1}{2}M^2}{\gamma\rho q^2}\begin{bmatrix} \frac{\rho q^2}{2z_4} \\ z_2 \\ z_3 \\ \frac{\rho q^2}{2z_1} \end{bmatrix} \quad \mathbf{r}_H = \frac{1}{\rho q^2}\begin{bmatrix} 0 \\ z_2 \\ z_3 \\ \frac{\rho q^2}{z_1} \end{bmatrix} \tag{55}$$

$$\mathbf{r}_{w^+} = \frac{1}{4\beta\rho q^2}\begin{bmatrix} M^2 z_1 \\ (M^2-2)z_2 - 2\beta z_3 \\ (M^2-2)z_3 + 2\beta z_2 \\ M^2 z_4 \end{bmatrix} \quad \mathbf{r}_{w^-} = \frac{1}{4\beta\rho q^2}\begin{bmatrix} M^2 z_1 \\ (M^2-2)z_2 + 2\beta z_3 \\ (M^2-2)z_3 - 2\beta z_2 \\ M^2 z_4 \end{bmatrix} \tag{56}$$

These vectors agree with Mani's, up to a scalar factor (cf. p.91).

What if the entropy variable is not without the factor $\frac{\gamma}{\gamma-1}$?

$$\partial \mathbf{x} = \left[(\partial p - a^2 \partial \rho),\, \rho \partial H,\, \beta \partial p + \rho q^2 \partial \theta,\, \beta \partial p - \rho q^2 \partial \theta\right]^t \tag{57}$$

This again decomposes the symmetric Euler in the exactly the same way as before. The resulting decomposed form is again identical to (37) - (43). The transformation matrix is now

$$\frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \begin{bmatrix} \frac{\gamma-1}{\gamma}((1-2\gamma)z_4 + \frac{\gamma}{z_1}(z_2^2+z_3^2)) & -\frac{\gamma-1}{\gamma}z_2 & -\frac{\gamma-1}{\gamma}z_3 & \frac{\gamma-1}{\gamma}z_1 \\ -z_4 & 0 & 0 & z_1 \\ \frac{\beta(\gamma-1)}{\gamma}z_4 & -\frac{\beta(\gamma-1)}{\gamma}z_2 - z_3 & -\frac{\beta(\gamma-1)}{\gamma}z_3 + z_2 & \frac{\beta(\gamma-1)}{\gamma}z_1 \\ \frac{\beta(\gamma-1)}{\gamma}z_4 & -\frac{\beta(\gamma-1)}{\gamma}z_2 + z_3 & -\frac{\beta(\gamma-1)}{\gamma}z_3 - z_2 & \frac{\beta(\gamma-1)}{\gamma}z_1 \end{bmatrix} \tag{58}$$

so, the only the first row has been changed. The change in its inverse transformation matrix appears only in $\mathbf{r}_s$ by a factor of $\gamma/(\gamma-1)$, i.e.

$$\mathbf{r}_s = -\frac{1+\frac{\gamma-1}{2}M^2}{(\gamma-1)\rho q^2}\begin{bmatrix} \frac{\rho q^2}{2z_4} \\ z_2 \\ z_3 \\ \frac{\rho q^2}{2z_1} \end{bmatrix} \tag{59}$$

and everything else stays the same.

In the case of subsonic flows, Mani uses different set of variables for the acoustic system, which appears to be

$$\partial \mathbf{x} = \left[\frac{\gamma-1}{\gamma}(\partial p - a^2\partial\rho),\, \rho\partial H,\, \frac{\gamma-1}{\gamma}\partial p,\, \rho q^2 \partial\theta\right]^t \tag{60}$$

We then have

$$\frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \begin{bmatrix} (1-2\gamma)z_4 + \frac{\gamma}{z_1}(z_2^2+z_3^2) & -z_2 & -z_3 & z_1 \\ -z_4 & 0 & 0 & z_1 \\ z_4 & -z_2 & -z_3 & z_1 \\ 0 & -z_3 & z_2 & 0 \end{bmatrix} \tag{61}$$

For this set of variables, we can easily obtain the corresponding vectors from those in the previous results.

$$\mathbf{r}_s = -\frac{1+\frac{\gamma-1}{2}M^2}{\gamma\rho q^2}\begin{bmatrix} \frac{\rho q^2}{2z_4} \\ z_2 \\ z_3 \\ \frac{\rho q^2}{2z_1} \end{bmatrix} \quad \mathbf{r}_H = \frac{1}{\rho q^2}\begin{bmatrix} 0 \\ z_2 \\ z_3 \\ \frac{\rho q^2}{z_1} \end{bmatrix} \tag{62}$$

$$\mathbf{r}_p = \frac{(\gamma-1)(M^2-2)}{4\gamma\beta^2\rho q^2}\begin{bmatrix} \frac{M^2}{(M^2-2)}z_1 \\ z_2 \\ z_3 \\ \frac{M^2}{(M^2-2)}z_4 \end{bmatrix} \quad \mathbf{r}_\theta = \frac{1}{2\rho q^2}\begin{bmatrix} 0 \\ -z_3 \\ z_2 \\ 0 \end{bmatrix} \tag{63}$$

These vectors again agree with Mani's, up to a scalar factor (cf. p.98). With this choice, the decomposed system becomes

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}^s_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial s} + \mathbf{B}^n_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial n} = 0 \tag{64}$$

where

$$\mathbf{A}^s_{\mathbf{x}} = \begin{bmatrix} q & 0 & 0 & 0 \\ 0 & q & 0 & 0 \\ 0 & 0 & \frac{\tau(M^2-1)}{\beta^2}q & 0 \\ 0 & 0 & 0 & \tau q \end{bmatrix} \tag{65}$$

$$\mathbf{B}^n_{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \frac{\gamma}{\gamma-1}\frac{\tau}{\beta^2}q \\ 0 & 0 & \frac{\gamma-1}{\gamma}\tau q & 0 \end{bmatrix} \tag{66}$$

How about this?

$$\partial \mathbf{x} = \left[\frac{\gamma-1}{\gamma}(\partial p - a^2 \partial \rho),\, \rho \partial H,\, \partial p,\, \rho q^2 \partial \theta \right]^t \tag{67}$$

We now have

$$\frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \begin{bmatrix} (1-2\gamma)z_4 + \frac{\gamma}{z_1}(z_2^2+z_3^2) & -z_2 & -z_3 & z_1 \\ -z_4 & 0 & 0 & z_1 \\ \frac{\gamma-1}{\gamma}z_4 & -\frac{\gamma-1}{\gamma}z_2 & -\frac{\gamma-1}{\gamma}z_3 & \frac{\gamma-1}{\gamma}z_1 \\ 0 & -z_3 & z_2 & 0 \end{bmatrix} \tag{68}$$

For this set of variables, we can easily obtain the corresponding vectors from above.

$$\mathbf{r}_s = -\frac{1+\frac{\gamma-1}{2}M^2}{\gamma \rho q^2} \begin{bmatrix} \frac{\rho q^2}{2z_4} \\ z_2 \\ z_3 \\ \frac{\rho q^2}{2z_1} \end{bmatrix} \quad \mathbf{r}_H = \frac{1}{\rho q^2} \begin{bmatrix} 0 \\ z_2 \\ z_3 \\ \frac{\rho q^2}{z_1} \end{bmatrix} \tag{69}$$

$$\mathbf{r}_p = \frac{(M^2-2)}{4\beta^2 \rho q^2} \begin{bmatrix} \frac{M^2}{(M^2-2)}z_1 \\ z_2 \\ z_3 \\ \frac{M^2}{(M^2-2)}z_4 \end{bmatrix} \quad \mathbf{r}_\theta = \frac{1}{2\rho q^2} \begin{bmatrix} 0 \\ -z_3 \\ z_2 \\ 0 \end{bmatrix} \tag{70}$$

These vectors again agree with Mani's, up to a scalar factor (cf. p.98). With this choice, the decomposed system becomes

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}^s_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial s} + \mathbf{B}^n_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial n} = 0 \tag{71}$$

where

$$\mathbf{A}^s_{\mathbf{x}} = \begin{bmatrix} q & 0 & 0 & 0 \\ 0 & q & 0 & 0 \\ 0 & 0 & \frac{\tau(M^2-1)}{\beta^2}q & 0 \\ 0 & 0 & 0 & \tau q \end{bmatrix} \tag{72}$$

$$\mathbf{B}^n_{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\tau}{\beta^2}q \\ 0 & 0 & \tau q & 0 \end{bmatrix} \tag{73}$$

Finally, I've decided to use this,

$$\partial \mathbf{x} = \left[(\partial p - a^2 \partial \rho),\, \rho \partial H,\, \partial p,\, \rho q^2 \partial \theta\right]^t \tag{74}$$

We then have

$$\frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \begin{bmatrix} \frac{\gamma-1}{\gamma}((1-2\gamma)z_4 + \frac{\gamma}{z_1}(z_2^2+z_3^2)) & -\frac{\gamma-1}{\gamma}z_2 & -\frac{\gamma-1}{\gamma}z_3 & \frac{\gamma-1}{\gamma}z_1 \\ -z_4 & 0 & 0 & z_1 \\ \frac{\beta(\gamma-1)}{\gamma}z_4 & -\frac{\beta(\gamma-1)}{\gamma}z_2 & -\frac{\beta(\gamma-1)}{\gamma}z_3 & \frac{\beta(\gamma-1)}{\gamma}z_1 \\ 0 & -z_3 & z_2 & 0 \end{bmatrix} \tag{75}$$

For this set of variables, we can easily obtain the corresponding vectors from those in the previous results.

$$\mathbf{r}_s = -\frac{1+\frac{\gamma-1}{2}M^2}{(\gamma-1)\rho q^2}\begin{bmatrix} \frac{\rho q^2}{2z_4} \\ z_2 \\ z_3 \\ \frac{\rho q^2}{2z_1} \end{bmatrix} \quad \mathbf{r}_H = \frac{1}{\rho q^2}\begin{bmatrix} 0 \\ z_2 \\ z_3 \\ \frac{\rho q^2}{z_1} \end{bmatrix} \tag{76}$$

$$\mathbf{r}_p = \frac{(M^2-2)}{4\beta^2 \rho q^2}\begin{bmatrix} \frac{M^2}{(M^2-2)}z_1 \\ z_2 \\ z_3 \\ \frac{M^2}{(M^2-2)}z_4 \end{bmatrix} \quad \mathbf{r}_\theta = \frac{1}{2\rho q^2}\begin{bmatrix} 0 \\ -z_3 \\ z_2 \\ 0 \end{bmatrix} \tag{77}$$

These vectors again agree with Mani's, up to a scalar factor (cf. p.98). With this choice, the decomposed system becomes

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}^s_{\mathbf{x}}\frac{\partial \mathbf{x}}{\partial s} + \mathbf{B}^n_{\mathbf{x}}\frac{\partial \mathbf{x}}{\partial n} = 0 \tag{78}$$

where

$$\mathbf{A}^s_{\mathbf{x}} = \begin{bmatrix} q & 0 & 0 & 0 \\ 0 & q & 0 & 0 \\ 0 & 0 & \frac{\tau(M^2-1)}{\beta^2}q & 0 \\ 0 & 0 & 0 & \tau q \end{bmatrix} \tag{79}$$

$$\mathbf{B}^n_{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \frac{\tau}{\beta^2} q \\ 0 & 0 & \tau q & 0 \end{bmatrix} \tag{80}$$

In x-y coordinates,

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{A}_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial x} + \mathbf{B}_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial y} = 0 \tag{81}$$

where

$$\mathbf{A}_{\mathbf{x}} = \begin{bmatrix} u & 0 & 0 & 0 \\ 0 & u & 0 & 0 \\ 0 & 0 & \frac{\tau(M^2-1)}{\beta^2} u & -\frac{\tau}{\beta^2} v \\ 0 & 0 & -\tau v & \tau u \end{bmatrix} \tag{82}$$

$$\mathbf{B}_{\mathbf{x}} = \begin{bmatrix} v & 0 & 0 & 0 \\ 0 & v & 0 & 0 \\ 0 & 0 & \frac{\tau(M^2-1)}{\beta^2} v & \frac{\tau}{\beta^2} u \\ 0 & 0 & \tau u & \tau v \end{bmatrix} \tag{83}$$

4 Conservation Issues

Conservation property is believed to be important. It is certainly important for computing moving shocks in a time-accurate manner. But it may not as important as it seems in the steady-state calculation.

Quantities to be conserved are the conservative variables. To ensure the conservation while utilizing the Hyp/Ell decomposition, we transform the split fluctuations back and update the conservative variables.

$$\mathbf{U}^{n+1}_j = \mathbf{U}^n_j - \frac{\Delta t}{V_j} \sum_{T \in \{T_j\}} \left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}_{vll}} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right)^{-1} \mathbf{\Phi}^{\mathbf{x}}_i \tag{84}$$

so that if we sum up the changes in $\mathbf{U}_j$

$$\begin{aligned}
\sum_{allj} V_j (\mathbf{U}^{n+1}_j - \mathbf{U}^n_j) &= -\sum_{allj} \Delta t \sum_{T \in \{T_j\}} \left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}_{vll}} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right)^{-1} \mathbf{\Phi}^{\mathbf{x}}_i && (85) \\
&= -\Delta t \sum_{allT} \sum_{i \in \{i_T\}} \left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}_{vll}} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right)^{-1} \mathbf{\Phi}^{\mathbf{x}}_i && (86) \\
&= -\Delta t \sum_{allT} \left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}_{vll}} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right)^{-1} \mathbf{\Phi}^{\mathbf{x}} && (87) \\
&= -\Delta t \sum_{allT} \left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}_{vll}} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right)^{-1} \left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}_{vll}} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right) \mathbf{\Phi} && (88)
\end{aligned}$$

$$= \quad -\Delta t \sum_{allT} \mathbf{\Phi} \tag{89}$$

$$= \quad -\Delta t \sum_{allT} \oint_{\partial T} \mathbf{f} dy - \mathbf{g} dx \tag{90}$$

The contour integral is exactly expressed with the conservative linearization, and therefore the quantity on the left depends only on the boundary data, i.e. conserved. In the case of local time-stepping, the method conserves

$$\sum_{allj} \frac{V_j(\mathbf{U}_j^{n+1} - \mathbf{U}_j^n)}{(\Delta t)_j} \tag{91}$$

Note that any effect of preconditioning is cancelled at the stage of update. It was used solely for distribution purposes. Also, note that the inverse matrix is singular at stagnation points.

Another thought would be that we update the parameter vector directly from the decomposed fluctuation.

$$\mathbf{Z}_j^{n+1} = \mathbf{Z}_j^n - \frac{\Delta t}{V_j} \sum_{T \in \{T_j\}} \frac{\tilde{\partial \mathbf{z}}}{\partial \mathbf{x}} \mathbf{\Phi}_i^{\mathbf{x}} \tag{92}$$

It is easy to show that this does not conserve anything. Also, the preconditioning remains, i.e. we are solving the altered form of the Euler equations. But at a steady-state*, we have

$$\mathbf{\Phi}^{\mathbf{x}} \approx 0 \tag{93}$$

or

$$\left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}}_{vll} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right) \mathbf{\Phi} \approx 0 \tag{94}$$

which is equivalent to

$$\mathbf{\Phi} \approx 0 \tag{95}$$

provided the matrix $\left(\frac{\tilde{\partial \mathbf{x}}}{\partial \mathbf{v}} \tilde{\mathbf{P}}_{vll} \frac{\tilde{\partial \mathbf{v}}}{\partial \mathbf{u}} \right)$ is invertible. If valid, this shows that the steady-state solution is a solution of the steady Euler equations (zero flux balance for all elements). In this sense, we may perform the update by applying the transformation at vertices,

$$\mathbf{Z}_j^{n+1} = \mathbf{Z}_j^n - \frac{\Delta t}{V_j} \left(\frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right)_j \sum_{T \in \{T_j\}} \mathbf{\Phi}_j^{\mathbf{x}} \tag{96}$$

This is again not conservative in its transient phase. But the steady-state solution would be correct in the same sense as above.

*fluctuations are not exactly zero at a steady-state.

For time-accurate calculations, it is important to be conservative. The method has to be conservative at each time step. We can achieve this with non-conservative methods such as above by using the technique of dual time-stepping. We seek a steady-state solution in a pseudo-time τ of the system,

$$\mathbf{u}_\tau = -\left(\mathbf{A}\mathbf{u}_x + \mathbf{B}\mathbf{u}_y + \mathbf{u}_t\right) \tag{97}$$

treating the real-time derivative as a source term. The discrete version is

$$\mathbf{U}_j^{n+1,k+1} = \mathbf{U}_j^{n+1,k} - \frac{\Delta\tau}{V_j}\sum_{T\in\{T_j\}} \mathbf{\Phi}_j \tag{98}$$

where $\mathbf{\Phi}_j$ is a portion of the total fluctuation $\mathbf{\Phi}$,

$$\mathbf{\Phi} = \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}} n_{x_i} + \tilde{\mathbf{B}} n_{y_i}\right)\mathbf{U}_i + \int_T \mathbf{u}_t \tag{99}$$

whose last term is suitably discretized in time. Thus, at each real time-step, we solve the steady-state problem to compute $\mathbf{U}_j^{n+1}$. In solving for the steady-state, we may violate conservation. For example, we may use the decomposition method,

$$\mathbf{Z}_j^{n+1,k+1} = \mathbf{Z}_j^{n+1,k} - \frac{(\Delta\tau)_j}{V_j}\left(\frac{\partial \mathbf{z}}{\partial \mathbf{x}}\right)_j \sum_{T\in\{T_j\}} \mathbf{\Phi}_j^{\mathbf{x}} \tag{100}$$

At the steady-state, we have

$$\mathbf{\Phi}^{\mathbf{x}} = \left(\tilde{\frac{\partial \mathbf{x}}{\partial \mathbf{v}}} \tilde{\mathbf{P}_{vll}} \tilde{\frac{\partial \mathbf{v}}{\partial \mathbf{u}}}\right)\mathbf{\Phi} \approx 0 \tag{101}$$

which implies

$$\mathbf{\Phi} = \frac{1}{2}\sum_i \left(\tilde{\mathbf{A}} n_{x_i} + \tilde{\mathbf{B}} n_{y_i}\right)\mathbf{U}_i + \int_T \mathbf{u}_t \approx 0 \tag{102}$$

Therefore,

$$\sum_{allT}\int_T \mathbf{u}_t \approx \sum_{allT}\frac{1}{2}\sum_i \left(\tilde{\mathbf{A}} n_{x_i} + \tilde{\mathbf{B}} n_{y_i}\right)\mathbf{U}_i = \sum_{allT}\oint_{\partial T} \mathbf{f}dy - \mathbf{g}dx \tag{103}$$

which depends only on the boundary data. This shows that it is important to keep the conservative linearization even though the method itself is not conservative in the pseudo-transient phase.