A Least-Squares Norm for the Euler Equations

HIROAKI NISHIKAWA*

February 2001

1 Introduction

We consider numerically solving the Euler equations in two dimensions by minimizing a norm of the form

$$\mathcal{F} = \sum_{T \in \{T\}} F_T = \frac{1}{2} \sum_{T \in \{T\}} \mathbf{\Phi}_T^t \mathbf{Q}_T \mathbf{\Phi}_T \tag{1}$$

over a set $\{T\}$ of triangles where $\mathbf{\Phi}_T$ is a vector of fluctuations for the Euler equations and Q_T is a positive definite symmetric matrix that assigns relative weight to the different equations. The most important part of the least-squares formulation is the definition of the norm to be minimized which greatly affects the numerical solutions. In the class of norms defined above, it is the choice of the matrix Q_T that determines the properties of the numerical solutions.

Assuming that we evaluate the fluctuation based on its conservative from, we derive the matrix Q_T that gives a certain number of properties to the minimization scheme. Keep in mind that we retain time dependent terms in the analysis, but our primary concern is steady state solutions.

2 A Dimensionally Consistent Set of Variables and Its Equations

We begin with the Euler equations in the conservative form.

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial x} = 0 \tag{2}$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{bmatrix}, \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho u v \\ \rho u H \end{bmatrix}, \mathbf{G} = \begin{bmatrix} \rho u \\ \rho u v \\ \rho v^2 + p \\ \rho v H \end{bmatrix} \tag{3}$$

where ρ is the density, u and v are the velocity components in the x and y direction, respectively, and p is the static pressure. The specific energy and enthalpy are given by

$$E = \frac{1}{\gamma - 1} \frac{p}{\rho} + \frac{1}{2}(u^2 + v^2) \tag{4}$$

$$H = \frac{\gamma}{\gamma - 1} \frac{p}{\rho} + \frac{1}{2}(u^2 + v^2). \tag{5}$$

One important consideration on defining a least-squares norm is the dimensional consistency. Clearly, the set of equations are not dimensionally consistent in the conservative form. To make it consistent, we introduce a set of

*Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan 48109. (hiroakin@umich.edu).

variables that is dimensionally consistent.

$$\partial \mathbf{V} = \begin{bmatrix} \partial p \\ \rho q^2 \partial\theta \\ \partial p - a^2 \partial\rho \\ \partial p + \rho q \partial q \end{bmatrix} \tag{6}$$

where θ is the flow angle, q is the flow speed, a is the speed of sound, and the third and fourth components represent the entropy and the enthalpy respectively. Note that the components now have the common physical dimension. To transform the conservative variables into the consistent variables, it is convenient to use the primitive variables.

$$\mathbf{W} = \begin{bmatrix} \rho \\ u \\ v \\ p \end{bmatrix} \tag{7}$$

which is linked with the consistent and conservative variables through the transformations

$$\partial \mathbf{V} = \mathbf{T_v} \partial \mathbf{W}, \quad \partial \mathbf{U} = \mathbf{T_u} \partial \mathbf{W} \tag{8}$$

respectively where

$$\mathbf{T_v} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -\rho\, v & \rho\, u & 0 \\ -a^2 & 0 & 0 & 1 \\ 0 & \rho\, u & \rho\, v & 1 \end{bmatrix}, \quad \mathbf{T_u} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ u & \rho & 0 & 0 \\ v & 0 & \rho & 0 \\ \frac{1}{2} q^2 & \rho\, u & \rho\, v & \frac{1}{\gamma - 1} \end{bmatrix}. \tag{9}$$

It follows that

$$\partial \mathbf{V} = \mathbf{T_v} \mathbf{T_u^{-1}} \partial \mathbf{U} \equiv \mathbf{T} \partial \mathbf{U}. \tag{10}$$

The transformation matrix T is thus given by

$$\mathbf{T} = \begin{bmatrix} \frac{1}{2}(\gamma-1)\, q^2 & -(\gamma-1)\, u & -(\gamma-1)\, v & \gamma - 1 \\ 0 & -v & u & 0 \\ \frac{1}{2}(\gamma-1)\, q^2 - a^2 & -(\gamma-1)\, u & -(\gamma-1)\, v & \gamma - 1 \\ \frac{1}{2}(\gamma-3)\, q^2 & -(\gamma-2)\, u & -(\gamma-2)\, v & \gamma - 1 \end{bmatrix}. \tag{11}$$

Now, we can transform the conservation form into a dimensionally consistent form by multiplying (2) by $\mathbf{T}$ from the left.

$$\begin{aligned} \mathbf{T}\frac{\partial \mathbf{U}}{\partial t} + \mathbf{T}\left(\frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial x}\right) &= && (12) \\ \mathbf{T}\frac{\partial \mathbf{U}}{\partial t} + \mathbf{T}\left(\mathbf{A}\frac{\partial \mathbf{U}}{\partial x} + \mathbf{B}\frac{\partial \mathbf{U}}{\partial x}\right) &= 0 && (13) \\ \mathbf{T}\frac{\partial \mathbf{U}}{\partial t} + \mathbf{T}\left(\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\frac{\partial \mathbf{U}}{\partial x} + \mathbf{B}\mathbf{T}^{-1}\mathbf{T}\frac{\partial \mathbf{U}}{\partial x}\right) &= 0 && (14) \\ \frac{\partial \mathbf{V}}{\partial t} + \mathbf{TAT}^{-1}\frac{\partial \mathbf{V}}{\partial x} + \mathbf{TBT}^{-1}\frac{\partial \mathbf{V}}{\partial x} &= 0. && (15) \end{aligned}$$

This is the form of the Euler equaitons that is dimensionally consistent in terms of the variable $\mathbf{V}$. The equations already being dimensionally consistent, in the least-squares method, we may therefore define the norm as

$$\mathcal{F} = \sum_{T \in \{T\}} F_T = \frac{1}{2} \sum_{T \in \{T\}} (\mathbf{T}\mathbf{\Phi}_T)^t \mathbf{T}\mathbf{\Phi}_T = \frac{1}{2} \sum_{T \in \{T\}} \mathbf{\Phi}_T^t \mathbf{T}^t \mathbf{T} \mathbf{\Phi}_T \tag{16}$$

which suggests

$$\mathbf{Q}_T = \mathbf{T}^t\mathbf{T}. \tag{17}$$

Note that

$$\mathbf{\Phi}_T = -\int_T \frac{\partial \mathbf{U}}{\partial t}\, dx\, dy = \int_T \left[\frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial x}\right]\, dx\, dy \tag{18}$$

which is evaluated by some quadrature rule. There are formulae that endows the scheme with a certain property such as exact shock capturing. See [2] for details.

3 A Decomposition Matrix

For simplicity, we write the dimensionally consistent Euler equations (15) in terms of the natural coordinates: the streamline and its normal.

$$\frac{\partial \mathbf{V}}{\partial t} + \mathbf{A_v}\frac{\partial \mathbf{V}}{\partial s} + \mathbf{B_v}\frac{\partial \mathbf{V}}{\partial n} = 0 \tag{19}$$

where

$$\mathbf{A_v} \;=\; \mathbf{TAT}^{-1}\cos\theta + \mathbf{TBT}^{-1}\sin\theta = \begin{bmatrix} a\frac{M^2-1}{M} & 0 & 0 & \frac{a}{M} \\ 0 & q & 0 & 0 \\ 0 & 0 & q & 0 \\ a\frac{M^2-1}{M} & 0 & 0 & a\frac{M^2+1}{M} \end{bmatrix} \tag{20}$$

$$\mathbf{B_v} \;=\; \mathbf{TBT}^{-1}\cos\theta - \mathbf{TAT}^{-1}\sin\theta = \begin{bmatrix} 0 & \frac{a}{M} & 0 & 0 \\ q & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{a}{M} & 0 & 0 \end{bmatrix}. \tag{21}$$

Note that the equation for the entropy has been decoupled, but not for the enthalpy. Now, we seek a matrix $\mathbf{P}_d$ that decouples the enthalpy as well by following the technique of preconditioning. We consider an altered system, preconditioned by $\mathbf{P}_d$.

$$\frac{\partial \mathbf{V}}{\partial t} + \mathbf{P}_d\left(\mathbf{A_v}\frac{\partial \mathbf{V}}{\partial s} + \mathbf{B_v}\frac{\partial \mathbf{V}}{\partial n}\right) = 0 \tag{22}$$

and require that

$$\mathbf{P}_d\mathbf{A_v} = \begin{bmatrix} q\frac{M^2-1}{M} & 0 & 0 & 0 \\ 0 & q & 0 & 0 \\ 0 & 0 & q & 0 \\ 0 & 0 & 0 & q \end{bmatrix}, \quad \mathbf{P}_d\mathbf{B_v} = \begin{bmatrix} 0 & -q & 0 & 0 \\ q & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{23}$$

which imply that we have two advection equations for the entropy and the enthalpy and a 2×2 acoustic subsystem in the resulting system. The matrix $\mathbf{P}_d$ can be determined from the first equation. The result is

$$\mathbf{P}_d = \begin{bmatrix} -(M^2+1) & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \tag{24}$$

and an *a posteriori* check confirms that it satisfies the other requirement for $\mathbf{B_v}$. Therefore, the system (22) becomes

$$\frac{\partial p}{\partial t} + q(1-M^2)\frac{\partial p}{\partial s} - \rho q^3\frac{\partial \theta}{\partial n} \;=\; 0 \tag{25}$$

$$\rho q^2\frac{\partial \theta}{\partial t} + \qquad \rho q^3\frac{\partial \theta}{\partial s} + \quad q\frac{\partial p}{\partial n} \;=\; 0 \tag{26}$$

$$\frac{\partial S}{\partial t} \qquad + \quad q\frac{\partial S}{\partial s} \quad = \quad 0 \tag{27}$$

$$\frac{\partial h}{\partial t} \qquad + \quad q\frac{\partial h}{\partial s} \quad = \quad 0 \tag{28}$$

where the first two compose the 2×2 acoustic subsystem and the others are completely decoupled advection equations for the entropy $\partial S = \partial p - a^2 \partial \rho$ and the enthalpy $\partial h = \partial p + \rho q \partial q$. In the least-squares method, we may therefore multiply the fluctuation by the matrix $\mathbf{P}_d$ to decompose the fluctuation, and then minimize the decomposed fluctuation in the least-squares sense.

$$\mathcal{F}\frac{1}{2}\sum_{T\in\{T\}}(\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T)^t\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T = \frac{1}{2}\sum_{T\in\{T\}}\mathbf{\Phi}_T^t\mathbf{T}^t\mathbf{P}_d^t\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T \tag{29}$$

which suggests

$$\mathbf{Q}_T = \mathbf{T}^t\mathbf{P}_d^t\mathbf{P}_d\mathbf{T}. \tag{30}$$

4 A Matrix for The Acoustic Subsystem

We consider the subsystem in its steady state form.

$$q(1-M^2)\frac{\partial p}{\partial s} - \rho q^3\frac{\partial \theta}{\partial n} \quad = \quad 0 \tag{31}$$

$$\rho q^3\frac{\partial \theta}{\partial s} + \quad q\frac{\partial p}{\partial n} \quad = \quad 0 \tag{32}$$

The system is then hyperbolic in supersonic flow ($M > 1$) and elliptic in subsonic flow ($M < 1$). Let us write the system in the matrix form.

$$\mathbf{A}_s\frac{\partial \mathbf{V}_s}{\partial s} + \mathbf{B}_s\frac{\partial \mathbf{V}_s}{\partial n} = 0 \tag{33}$$

where $\mathbf{V}_s = [p, \theta]^t$ and

$$\mathbf{A}_s = \begin{bmatrix} q(1-M^2) & 0 \\ 0 & \rho q^3 \end{bmatrix}, \quad \mathbf{B}_s = \begin{bmatrix} 0 & -\rho q^3 \\ q & 0 \end{bmatrix}. \tag{34}$$

In the hyperbolic case, we can diagonalize the system as follows. Multiplying $\mathbf{A}_s^{-1}$ from the left, we have

$$\frac{\partial \mathbf{V}_s}{\partial s} + \mathbf{A}_s^{-1}\mathbf{B}_s\frac{\partial \mathbf{V}_s}{\partial n} = 0. \tag{35}$$

The matrix $\mathbf{A}_s^{-1}\mathbf{B}_s$ has the eigenvalues $\pm 1/\beta$ where $\beta = \sqrt{M^2-1}$ and the eigenvectors which is arranged into the columns of a matrix $\mathbf{R}_s$

$$\mathbf{R}_s = \begin{bmatrix} -\frac{1}{q\beta^2} & \frac{1}{q\beta^2} \\ \frac{1}{\rho q^3\beta} & \frac{1}{\rho q^3\beta} \end{bmatrix}. \tag{36}$$

Then, the characteristic form is obtained by multiplying the system by $\mathbf{R}_s^{-1}$.

$$\mathbf{R}_s^{-1}\frac{\partial \mathbf{V}_s}{\partial s} + \mathbf{R}_s^{-1}\mathbf{A}_s^{-1}\mathbf{B}_s\mathbf{R}_s\mathbf{R}_s^{-1}\frac{\partial \mathbf{V}_s}{\partial n} \quad = \quad 0 \tag{37}$$

$$\frac{\partial \mathbf{W}_s}{\partial s} + \mathbf{\Lambda}\frac{\partial \mathbf{W}_s}{\partial n} \quad = \quad 0 \tag{38}$$

where $\partial \mathbf{W}_s = \mathbf{R}_s^{-1}\partial \mathbf{V}_s$ is the chgaracteristic variables and $\mathbf{\Lambda}$ is the diagonal matrix whose diagonal elements are the eigenvalues. Now, recall that we obtained the characteristic system by multiplying the original system (32) by

$\mathbf{A}_s^{-1}$ and then $\mathbf{R}_s^{-1}$, or the matrix $\mathbf{P}_s \equiv \mathbf{R}_s^{-1}\mathbf{A}_s^{-1}$ from the left. Then, denoting the fluctuation corresponding to the original system by $\mathbf{\Psi}$, we define the least-squares norm for the subsystem as

$$\mathcal{F}_s = \sum_{T\in\{T\}} (\mathbf{P}_s\mathbf{\Psi}_T)^t\mathbf{P}_s\mathbf{\Psi}_T = \sum_{T\in\{T\}} \mathbf{\Psi}_T^t\mathbf{P}_s^t\mathbf{P}_s\mathbf{\Psi}_T. \tag{39}$$

where

$$\mathbf{P}_s = \frac{1}{2}\begin{bmatrix} 1 & \beta \\ -1 & \beta \end{bmatrix}, \quad \mathbf{D}_s \equiv \mathbf{P}_s^t\mathbf{P}_s = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & \beta^2 \end{bmatrix}. \tag{40}$$

To deal with subsonic cases, we define, following the analysis of Roe[1],

$$\mathbf{P}_s = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{|\beta^2|} \\ -1 & \sqrt{|\beta^2|} \end{bmatrix}, \quad \mathbf{D}_s \equiv \mathbf{P}_s^t\mathbf{P}_s = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & |\beta^2| \end{bmatrix}. \tag{41}$$

With this weighting matrix, the minimization scheme recognizes the characteristic equations: the norm vanishes if the characteristic equations are satisfied for simple as well as nonsimple waves just as in the linear case. In subsonic case, with area weighting, the discretization becomes indentical to the standard finite element discretization for the equivalent second-order partial differential equations.

5 The Norm for The Euler Equations

We now put everything together. Assume that we have evaluated the fluctuation from the conservative form, $\mathbf{\Phi}_T$. First, for dimensional consistency, we multiply it by $\mathbf{T}$.

$$\mathbf{T}\mathbf{\Phi}_T \tag{42}$$

Second, to decompose the system, we multiply it by $\mathbf{T}$.

$$\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T \tag{43}$$

Finally, for the acoustic system, we multiply it by $\mathbf{P}_s$.

$$\mathbf{P}_s\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T \tag{44}$$

where $\mathbf{P}_s$ is now 4×4 matrix defined by

$$\mathbf{P}_s = \begin{bmatrix} 1/2 & \sqrt{|\beta^2|}/2 & 0 & 0 \\ -1/2 & \sqrt{|\beta^2|}/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{45}$$

We then minimize this fluctuation in the least-squares sense, i.e. minimize

$$\mathcal{F} = \sum_{T\in\{T\}} (\mathbf{P}_s\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T)^t\mathbf{P}_s\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T \tag{46}$$

or

$$\mathcal{F} = \sum_{T\in\{T\}} \mathbf{\Phi}_T^t\mathbf{Q}_T\mathbf{\Phi}_T \tag{47}$$

where

$$\mathbf{Q}_T = \mathbf{T}^t\mathbf{P}_d^t\mathbf{P}_s^t\mathbf{P}_s\mathbf{P}_d\mathbf{T}. \tag{48}$$

6 Implementation

Since the matrix $\mathbf{Q}_T$ is complicated, it would be simpler to implement the weight by computing the transformed fluctuation vector.

$$\mathbf{\Phi}_T^p = \mathbf{P}\mathbf{\Phi}_T = \mathbf{P}_s\mathbf{P}_d\mathbf{T}\mathbf{\Phi}_T \tag{49}$$

where

$$\mathbf{P} = \left[\begin{array}{cccc} -\frac{q^2}{4}\left(2+(\gamma-1)M^2\right) & \frac{u}{2}\left\{1+(\gamma-1)M^2\right\} - \frac{v}{2}\sqrt{|\beta^2|} & \frac{v}{2}\left\{1+(\gamma-1)M^2\right\} + \frac{u}{2}\sqrt{|\beta^2|} & -\frac{\gamma-1}{2}M^2 \\ \frac{q^2}{4}\left(2+(\gamma-1)M^2\right) & -\frac{u}{2}\left\{1+(\gamma-1)M^2\right\} - \frac{v}{2}\sqrt{|\beta^2|} & -\frac{v}{2}\left\{1+(\gamma-1)M^2\right\} + \frac{u}{2}\sqrt{|\beta^2|} & \frac{\gamma-1}{2}M^2 \\ \frac{\gamma-1}{2}q^2 - a^2 & -(\gamma-1)u & -(\gamma-1)v & \gamma-1 \\ -q^2 & u & v & 0 \end{array} \right]. \tag{50}$$

This matrix is evaluated within each triangle by the value in accordance with the linearization of the conservation form. Then we minimize

$$\mathcal{F} = \sum_{T\in\{T\}} F_T = \sum_{T\in\{T\}} (\mathbf{\Phi}_T^p)^t\mathbf{\Phi}_T^p \tag{51}$$

by a steepest descent method

$$\delta\mathbf{Z}_j = -\omega\frac{\partial\mathcal{F}}{\partial\mathbf{Z}_j} = -\omega\sum_{T\in\{T_j\}}\frac{\partial F_T}{\partial\mathbf{Z}_j} \tag{52}$$

where ω is a small constant and $\mathbf{Z}_j$ is a set of variables with respect to which the norm is minimized. The gradient in the cell T is given by

$$\frac{\partial F_T}{\partial\mathbf{Z}_j} = (\mathbf{\Phi}_T^p)^t\frac{\partial\mathbf{\Phi}_T^p}{\partial\mathbf{Z}_j} = (\mathbf{\Phi}_T^p)^t\mathbf{P}\frac{\partial\mathbf{\Phi}_T}{\partial\mathbf{Z}_j}. \tag{53}$$

7 Remarks

The norm derived in this paper has the following properties.

1. It is dimensionally consistent.
2. It is equivalent to a least-squares norm for the decomposed Euler equations.
3. In supersonic flows, it is equivalent to a least-squares norm for the characteristic equations, and thus the norm can vanish for simple or nonsimple wave solutions.
4. In subsonic case, with area weighting, the method becomes indentical to the standard finite element discretization for the second-order partial differential equations equivalent to the elliptic subsystem.

Shock capturing capability is endowed by a particular linearization of the conservative form via a parameter α as discussed in [2].

References

[1] Roe, P. L., Fluctuation Splitting Schemes on Optimal Grids, AIAA Paper 97-2032, June 1997.

[2] Nishikawa, H., Rad, M., Roe, P. L., Grids and Solutions from Residual Minimization, ICCFD Proceedings, Kyoto. Springer-Verlag, 2000.