

0.0.1 Finite Element Method

Instead of a relation between G_j and v_j as in the finite difference method above, we get $v_{j+1/2}$ by solving the finite element method. So we desire the error coefficient produced by the FEM solution of the elliptic equation given G_j and solving for $v_{j+1/2}$ which will be used in the conservation equation part below as our reconstruction error coefficient for v at the cell interface. The matrix equation of the FEM for the linearised equations (??) can be attained by just using $h_{j-1/2}^+ = H = h_{j+1/2}^-$ and so we obtain from []

$$\sum_j \frac{\Delta x}{6} \begin{bmatrix} G_{j-1/2}^+ \\ 2G_{j-1/2}^+ + 2G_{j+1/2}^- \\ G_{j+1/2}^- \end{bmatrix} = \sum_j \left(H \frac{\Delta x}{30} \begin{bmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{bmatrix} + \frac{H^3}{9\Delta x} \begin{bmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{bmatrix} \right) \begin{bmatrix} v_{j-1/2} \\ v_j \\ v_{j+1/2} \end{bmatrix}$$

Using our relations from the periodic nature of u and G , and the minmod reconstruction used on G we get that

$$\sum_j \frac{\Delta x}{6} \begin{bmatrix} e^{-ik\Delta x} \mathcal{R}_2^+ \\ 2e^{-ik\Delta x} \mathcal{R}_2^+ + 2\mathcal{R}_2^- \\ \mathcal{R}_2^- \end{bmatrix} G_j = \sum_j \left(H \frac{\Delta x}{30} \begin{bmatrix} 5i \sin(k \frac{\Delta x}{2}) + 3 \cos(k \frac{\Delta x}{2}) + 2 \\ 16 + 4 \cos(k \frac{\Delta x}{2}) \\ -5i \sin(k \frac{\Delta x}{2}) + 3 \cos(k \frac{\Delta x}{2}) + 2 \end{bmatrix} + \frac{H^3}{9\Delta x} \begin{bmatrix} 6i \sin(k \frac{\Delta x}{2}) + 8 \cos(k \frac{\Delta x}{2}) - 8 \\ -16 \cos(k \frac{\Delta x}{2}) + 16 \\ -6i \sin(k \frac{\Delta x}{2}) + 8 \cos(k \frac{\Delta x}{2}) - 8 \end{bmatrix} \right) v_j \quad (1)$$

We can now add all the terms that overlap i.e the extra contributions from the functions $\phi_{j+1/2}$ and $\phi_{j-1/2}$ from outside the cell $[x_{j-1/2}, x_{j+1/2}]$, this then gives us a relation between the sub-vectors of the total vectors of

the FEM. Doing this we can rewrite the matrix equation as []

$$\sum_j \frac{\Delta x}{6} \begin{bmatrix} 2 \\ \mathcal{R}_2^- + \mathcal{R}_2^+ \end{bmatrix}^T \begin{bmatrix} G_j \\ G_j \end{bmatrix} = \sum_j \left(H \frac{\Delta x}{30} \begin{bmatrix} 16 + 4 \cos \left(k \frac{\Delta x}{2} \right) \\ 4 \cos \left(\frac{k \Delta x}{2} \right) + 8 \cos (k \Delta x) - 2 \end{bmatrix}^T + \frac{H^3}{9 \Delta x} \begin{bmatrix} -16 \cos \left(k \frac{\Delta x}{2} \right) + 16 \\ -16 \cos \left(\frac{k \Delta x}{2} \right) + 14 \cos (k \Delta x) + 2 \end{bmatrix}^T \right) \begin{bmatrix} u_j \\ u_{j+1/2} \end{bmatrix} \quad (2)$$

So the equation for $u_{j+1/2}$ is

$$\frac{\Delta x}{6} (\mathcal{R}_2^+ + \mathcal{R}_2^-) G_j = \left(H \frac{\Delta x}{30} \left(4 \cos \left(\frac{k \Delta x}{2} \right) + 8 \cos (k \Delta x) - 2 \right) + \frac{H^3}{9 \Delta x} \left(-16 \cos \left(\frac{k \Delta x}{2} \right) + 14 \cos (k \Delta x) + 2 \right) \right) u_{j+1/2} \quad (3)$$

We have

$$G_j = \mathcal{G}_{FEM} u_{j+1/2}$$

$$\mathcal{G}_a u_j = \mathcal{G}_{FEM} u_{j+1/2}$$

$$\frac{\mathcal{G}_a}{\mathcal{G}_{FEM}} u_j = u_{j+1/2}$$

This is the error introduced by calculating $u_{j+1/2}$ in our method.

0.1 Conservation Equation

Finite volume methods have the following update scheme to approximate equations in conservation law form [] for some quantity q

$$\bar{q}_j^{n+1} = \bar{q}_j^n - \frac{\Delta t}{\Delta x} [F_{j+1/2}^n - F_{j-1/2}^n].$$

Where the bar denotes that it is the cell average of the quantity q and $F_{j+1/2}^n$ and $F_{j-1/2}^n$ are the approximations to the average fluxes across the cell boundary between the times t^n and t^{n+1} .

In our methods there is some transformation between the nodal value q_j and the cell average \bar{q}_j , which will introduce some error factor \mathcal{M} . For first and second order methods $\mathcal{M}_1 = \mathcal{M}_2 = 1$, however for higher-order methods $\mathcal{M} \neq 1$.

To calculate the fluxes $F_{j+1/2}^n$ and $F_{j-1/2}^n$ we use Kurganovs method [superscript dropped]

$$F_{j+\frac{1}{2}} = \frac{a_{j+\frac{1}{2}}^+ f(q_{j+\frac{1}{2}}^-) - a_{j+\frac{1}{2}}^- f(q_{j+\frac{1}{2}}^+)}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} + \frac{a_{j+\frac{1}{2}}^+ a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} [q_{j+\frac{1}{2}}^+ - q_{j+\frac{1}{2}}^-]$$

where $a_{j+\frac{1}{2}}^+$ and $a_{j+\frac{1}{2}}^-$ are given by the wave speed bounds [], so that

$$a_{j+1/2}^- = -\sqrt{gH}$$

$$a_{j+1/2}^+ = \sqrt{gH}.$$

Substituting these values into Kurganovs flux approximation we obtain

$$F_{j+\frac{1}{2}} = \frac{f(q_{j+\frac{1}{2}}^-) + f(q_{j+\frac{1}{2}}^+)}{2} - \frac{\sqrt{gH}}{2} [q_{j+\frac{1}{2}}^+ - q_{j+\frac{1}{2}}^-] \quad (4)$$

For η our Kurganov approximation to the flux of (??) is then

$$F_{j+\frac{1}{2}}^\eta = \frac{Hv_{j+\frac{1}{2}} + Hv_{j+\frac{1}{2}}}{2} - \frac{\sqrt{gH}}{2} [\eta_{j+\frac{1}{2}}^+ - \eta_{j+\frac{1}{2}}^-] \quad (5)$$

The missing piece here is the error introduced by reconstruction of the edge values $v_{j+\frac{1}{2}}^-$, $v_{j+\frac{1}{2}}^+$, $\eta_{j+\frac{1}{2}}^-$ and $\eta_{j+\frac{1}{2}}^+$ from the cell averages \bar{v}_j and $\bar{\eta}_j$. Because our quantities are smooth the nonlinear limiters can be neglected so we have for the second-order reconstruction of η

$$\begin{aligned} \eta_{j+\frac{1}{2}}^- &= \bar{\eta}_j + \frac{-\bar{\eta}_{j-1} + \bar{\eta}_{j+1}}{4} \\ \eta_{j+\frac{1}{2}}^+ &= \bar{\eta}_{j+1} + \frac{-\bar{\eta}_j + \bar{\eta}_{j+2}}{4}. \end{aligned}$$

Using (??) these equations become

$$\begin{aligned}\eta_{j+\frac{1}{2}}^- &= \mathcal{M}_2 \eta_j + \frac{-\mathcal{M}_2 \eta_j e^{-ik\Delta x} + \mathcal{M}_2 \eta_j e^{ik\Delta x}}{4} \\ \eta_{j+\frac{1}{2}}^+ &= \mathcal{M}_2 \eta_j e^{ik\Delta x} + \frac{-\mathcal{M}_2 \eta_j + \mathcal{M}_2 \eta_j e^{2ik\Delta x}}{4}.\end{aligned}$$

For the second order case $\mathcal{M}_2 = 1$ and these equations can be reduced to

$$\eta_{j+\frac{1}{2}}^- = \left(1 + \frac{i \sin(k\Delta x)}{2}\right) \eta_j \quad (6a)$$

$$\eta_{j+\frac{1}{2}}^+ = e^{ik\Delta x} \left(1 - \frac{i \sin(k\Delta x)}{2}\right) \eta_j. \quad (6b)$$

From these we introduce the second order reconstruction factors $\mathcal{R}_2^+ = e^{ik\Delta x} \left(1 - \frac{i \sin(k\Delta x)}{2}\right)$ and $\mathcal{R}_2^- = 1 + \frac{i \sin(k\Delta x)}{2}$ for both η and G . So that we have

$$\begin{aligned}\eta_{j+\frac{1}{2}}^- &= \mathcal{R}_2^- \eta_j \\ \eta_{j+\frac{1}{2}}^+ &= \mathcal{R}_2^+ \eta_j.\end{aligned}$$

In our numerical methods our reconstruction of v is slightly different as $v_{j+\frac{1}{2}}$ are equal as we assume v is continuous. For the second order method we have

$$v_{j+1/2} = \frac{\mathcal{G}_A}{\mathcal{G}_{FEM}} v_j$$

We now have all the pieces to substitute into (5) which for the second order method results in

$$F_{j+\frac{1}{2}}^\eta = H \frac{\mathcal{G}_A}{\mathcal{G}_{FEM}} v_j - \frac{\sqrt{gH}}{2} [\mathcal{R}_2^+ \eta_j - \mathcal{R}_2^- \eta_j]$$

Which becomes

$$F_{j+\frac{1}{2}}^\eta = H \frac{\mathcal{G}_A}{\mathcal{G}_{FEM}} v_j - \frac{\sqrt{gH}}{2} [\mathcal{R}_2^+ - \mathcal{R}_2^-] \eta_j$$

We then introduce the factors $\mathcal{F}_2^{\eta,v}$ and $\mathcal{F}_2^{\eta,\eta}$ so that

$$F_{j+\frac{1}{2}}^\eta = \mathcal{F}_2^{\eta,v} v_j + \mathcal{F}_2^{\eta,\eta} \eta_j. \quad (7)$$

Repeating this process for G using \square and \square we get that

$$F_{j+\frac{1}{2}}^G = \frac{gHh_{j+\frac{1}{2}}^- + gHh_{j+\frac{1}{2}}^+}{2} - \frac{\sqrt{gH}}{2} \left[G_{j+\frac{1}{2}}^+ - G_{j+\frac{1}{2}}^- \right] \quad (8)$$

Using our reconstruction factors this becomes:

$$F_{j+\frac{1}{2}}^G = \frac{gH\mathcal{R}_2^- h_j + gH\mathcal{R}_2^+ h_j}{2} - \frac{\sqrt{gH}}{2} [\mathcal{R}_2^+ G_j - \mathcal{R}_2^- G_j]$$

which by factoring and using the factor \mathcal{G}_{FD2} becomes

$$F_{j+\frac{1}{2}}^G = gH \frac{\mathcal{R}_2^- + \mathcal{R}_2^+}{2} h_j - \frac{\sqrt{gH}}{2} [\mathcal{R}_2^+ - \mathcal{R}_2^-] \mathcal{G}_a v_j$$

We then introduce the factors $\mathcal{F}_2^{G,v}$ and $\mathcal{F}_2^{G,\eta}$ so that

$$F_{j+\frac{1}{2}}^G = \mathcal{F}_2^{G,\eta} \eta_j + \mathcal{F}_2^{G,v} v_j \quad (9)$$

By substituting (7), (9) and \mathcal{M}_2 into \square our finite volume method can be written as

$$\begin{aligned} \mathcal{M}_2 \eta_j^{n+1} &= \mathcal{M}_2 \eta_j^n - \frac{\Delta t}{\Delta x} \left[(1 - e^{ik\Delta x}) (\mathcal{F}_2^{\eta,\eta} h_j + \mathcal{F}_2^{\eta,v} v_j) \right] \\ \mathcal{M}_2 G_j^{n+1} &= \mathcal{M}_2 G_j^n - \frac{\Delta t}{\Delta x} \left[(1 - e^{ik\Delta x}) (\mathcal{F}_2^{G,\eta} \eta_j + \mathcal{F}_2^{G,v} v_j) \right] \end{aligned}$$

Furthermore by transforming the G 's into v 's using our second order finite volume factor \mathcal{G}_{FD2} and using $\mathcal{M}_2 = 1$ we obtain

$$\begin{aligned} \eta_j^{n+1} &= \eta_j^n - \frac{\Delta t}{\Delta x} \left[(1 - e^{ik\Delta x}) (\mathcal{F}_2^{\eta,\eta} \eta_j + \mathcal{F}_2^{\eta,v} v_j) \right] \\ v_j^{n+1} &= v_j^n - \frac{1}{\mathcal{G}_{FD2}} \frac{\Delta t}{\Delta x} \left[(1 - e^{ik\Delta x}) (\mathcal{F}_2^{G,\eta} \eta_j + \mathcal{F}_2^{G,v} v_j) \right] \end{aligned}$$

This can be written in matrix form as

$$\begin{bmatrix} \eta \\ v \end{bmatrix}_j^{n+1} = \begin{bmatrix} \eta \\ v \end{bmatrix}_j^n - \frac{(1 - e^{ik\Delta x}) \Delta t}{\Delta x} \begin{bmatrix} \mathcal{F}_2^{\eta,\eta} & \mathcal{F}_2^{\eta,v} \\ \frac{1}{\mathcal{G}} \mathcal{F}_2^{v,\eta} & \frac{1}{\mathcal{G}} \mathcal{F}_2^{v,v} \end{bmatrix} \begin{bmatrix} \eta \\ v \end{bmatrix}_j^n$$

Introducing

$$\mathbf{F}_2 = \frac{(1 - e^{ik\Delta x})}{\Delta x} \begin{bmatrix} \mathcal{F}_2^{\eta,\eta} & \mathcal{F}_2^{\eta,v} \\ \frac{1}{\bar{g}}\mathcal{F}_2^{v,\eta} & \frac{1}{\bar{g}}\mathcal{F}_2^{v,v} \end{bmatrix}$$

this becomes

$$\begin{bmatrix} \eta \\ v \end{bmatrix}_j^{n+1} = (\mathbf{I} - \Delta t \mathbf{F}_2) \begin{bmatrix} \eta \\ v \end{bmatrix}_j^n$$