0.0.1 Finite Element Method

Instead of a relation between G_j and υ_j as in the finite difference method above, we get $\upsilon_{j+1/2}$ 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 $\upsilon_{j+1/2}$ which will be used in the conservation equation part below as our reconstruction error coefficient for υ at the cell interface. The matrix equation of the FEM for the linearised equations (**??**) can be attained by just using $h^{+}_{j-/2} = H = h^{-}_{j+1/2}$ and so we obtain from []

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

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

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

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 $\left[x_{j-1/2}, x_{j+1/2}\right]$, 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 []

$$\sum_j \frac{\Delta x}{6} \begin{bmatrix} 2 \\ \mathcal{R}_2^- + \mathcal{R}_2^+ \end{bmatrix}^T \begin{bmatrix} G_j \\ G_j \end{bmatrix} =$$

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

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

$$\frac{\Delta x}{6} \left(\mathcal{R}_2^+ + \mathcal{R}_2^-\right) G_j =$$

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

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

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

Where the bar denotes that it is the cell average of the quantity q and $F^n_{j+1/2}$ and $F^n_{j-1/2}$ 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 $F^n_{j+1/2}$ and $F^n_{j-1/2}$ we use Kurganovs method [superscript dropped][]

$$F_{j+\frac{1}{2}} = \frac{a^+_{j+\frac{1}{2}} f\left(q^-_{j+\frac{1}{2}}\right) - a^-_{j+\frac{1}{2}} f\left(q^+_{j+\frac{1}{2}}\right)}{a^+_{j+\frac{1}{2}} - a^-_{j+\frac{1}{2}}} + \frac{a^+_{j+\frac{1}{2}}\, a^-_{j+\frac{1}{2}}}{a^+_{j+\frac{1}{2}} - a^-_{j+\frac{1}{2}}} \left[q^+_{j+\frac{1}{2}} - q^-_{j+\frac{1}{2}}\right]$$

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

$$F_{j+\frac{1}{2}} = \frac{f\left(q^-_{j+\frac{1}{2}}\right) + f\left(q^+_{j+\frac{1}{2}}\right)}{2} - \frac{\sqrt{gH}}{2}\left[q^+_{j+\frac{1}{2}} - q^-_{j+\frac{1}{2}}\right] \tag{4}$$

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

$$F^\eta_{j+\frac{1}{2}} = \frac{H\upsilon_{j+\frac{1}{2}} + H\upsilon_{j+\frac{1}{2}}}{2} - \frac{\sqrt{gH}}{2}\left[\eta^+_{j+\frac{1}{2}} - \eta^-_{j+\frac{1}{2}}\right] \tag{5}$$

The missing piece here is the error introduced by reconstruction of the edge values $\upsilon^-_{j+\frac{1}{2}}$, $\upsilon^+_{j+\frac{1}{2}}$, $\eta^-_{j+\frac{1}{2}}$ and $\eta^+_{j+\frac{1}{2}}$ from the cell averages $\bar{\upsilon}_j$ 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

$$\eta^-_{j+\frac{1}{2}} = \bar{\eta}_j + \frac{-\bar{\eta}_{j-1} + \bar{\eta}_{j+1}}{4}$$
$$\eta^+_{j+\frac{1}{2}} = \bar{\eta}_{j+1} + \frac{-\bar{\eta}_j + \bar{\eta}_{j+2}}{4}.$$

Using (??) these equations become

$$\eta^-_{j+\frac{1}{2}} = \mathcal{M}_2\eta_j + \frac{-\mathcal{M}_2\eta_j e^{-ik\Delta x} + \mathcal{M}_2\eta_j e^{ik\Delta x}}{4}$$
$$\eta^+_{j+\frac{1}{2}} = \mathcal{M}_2\eta_j e^{ik\Delta x} + \frac{-\mathcal{M}_2\eta_j + \mathcal{M}_2\eta_j e^{2ik\Delta x}}{4}.$$

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

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

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

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

$$\eta^-_{j+\frac{1}{2}} = \mathcal{R}^-_2\eta_j$$
$$\eta^+_{j+\frac{1}{2}} = \mathcal{R}^+_2\eta_j.$$

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

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

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

$$F^\eta_{j+\frac{1}{2}} = H\frac{\mathcal{G}_A}{\mathcal{G}_{FEM}}\upsilon_j - \frac{\sqrt{gH}}{2}\left[\mathcal{R}^+_2\eta_j - \mathcal{R}^-_2\eta_j\right]$$

Which becomes

$$F^\eta_{j+\frac{1}{2}} = H\frac{\mathcal{G}_A}{\mathcal{G}_{FEM}}\upsilon_j - \frac{\sqrt{gH}}{2}\left[\mathcal{R}^+_2 - \mathcal{R}^-_2\right]\eta_j$$

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

$$F^{\eta}_{j+\frac{1}{2}} = \mathcal{F}^{\eta,\upsilon}_{2}\upsilon_j + \mathcal{F}^{\eta,\eta}_{2}\eta_j. \tag{7}$$

Repeating this process for G using [] and [] we get that

$$F^{G}_{j+\frac{1}{2}} = \frac{gHh^{-}_{j+\frac{1}{2}} + gHh^{+}_{j+\frac{1}{2}}}{2} - \frac{\sqrt{gH}}{2}\left[G^{+}_{j+\frac{1}{2}} - G^{-}_{j+\frac{1}{2}}\right] \tag{8}$$

Using our reconstruction factors this becomes:

$$F^{G}_{j+\frac{1}{2}} = \frac{gH\mathcal{R}^{-}_{2}h_j + gH\mathcal{R}^{+}_{2}h_j}{2} - \frac{\sqrt{gH}}{2}\left[\mathcal{R}^{+}_{2}G_j - \mathcal{R}^{-}_{2}G_j\right]$$

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

$$F^{G}_{j+\frac{1}{2}} = gH\frac{\mathcal{R}^{-}_{2} + \mathcal{R}^{+}_{2}}{2}h_j - \frac{\sqrt{gH}}{2}\left[\mathcal{R}^{+}_{2} - \mathcal{R}^{-}_{2}\right]\mathcal{G}_a\upsilon_j$$

We then introduce the factors $\mathcal{F}^{G,\upsilon}_{2}$ and $\mathcal{F}^{G,\eta}_{2}$ so that

$$F^{G}_{j+\frac{1}{2}} = \mathcal{F}^{G,\eta}_{2}\eta_j + \mathcal{F}^{G,\upsilon}_{2}\upsilon_j \tag{9}$$

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

$$\mathcal{M}_2\eta^{n+1}_j = \mathcal{M}_2\eta^{n}_j - \frac{\Delta t}{\Delta x}\left[\left(1 - e^{ik\Delta x}\right)\left(\mathcal{F}^{\eta,\eta}_{2}h_j + \mathcal{F}^{\eta,\upsilon}_{2}\upsilon_j\right)\right]$$

$$\mathcal{M}_2G^{n+1}_j = \mathcal{M}_2G^{n}_j - \frac{\Delta t}{\Delta x}\left[\left(1 - e^{ik\Delta x}\right)\left(\mathcal{F}^{G,\eta}_{2}\eta_j + \mathcal{F}^{G,\upsilon}_{2}\upsilon_j\right)\right]$$

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

$$\eta^{n+1}_j = \eta^{n}_j - \frac{\Delta t}{\Delta x}\left[\left(1 - e^{ik\Delta x}\right)\left(\mathcal{F}^{\eta,\eta}_{2}\eta_j + \mathcal{F}^{\eta,\upsilon}_{2}\upsilon_j\right)\right]$$

$$\upsilon^{n+1}_j = \upsilon^{n}_j - \frac{1}{\mathcal{G}_{FD2}}\frac{\Delta t}{\Delta x}\left[\left(1 - e^{ik\Delta x}\right)\left(\mathcal{F}^{G,\eta}_{2}\eta_j + \mathcal{F}^{G,\upsilon}_{2}\upsilon_j\right)\right]$$

This can be written in matrix form as

$$\begin{bmatrix} \eta \\ \upsilon \end{bmatrix}^{n+1}_{j} = \begin{bmatrix} \eta \\ \upsilon \end{bmatrix}^{n}_{j} - \frac{\left(1 - e^{ik\Delta x}\right)\Delta t}{\Delta x}\begin{bmatrix} \mathcal{F}^{\eta,\eta}_{2} & \mathcal{F}^{\eta,\upsilon}_{2} \\ \frac{1}{\mathcal{G}}\mathcal{F}^{\upsilon,\eta}_{2} & \frac{1}{\mathcal{G}}\mathcal{F}^{\upsilon,\upsilon}_{2} \end{bmatrix}\begin{bmatrix} \eta \\ \upsilon \end{bmatrix}^{n}_{j}$$

Introducing

$$\boldsymbol{F}_2 = \frac{\left(1 - e^{ik\Delta x}\right)}{\Delta x} \left[\begin{array}{cc} \mathcal{F}_2^{\eta,\eta} & \mathcal{F}_2^{\eta,\upsilon} \\ \frac{1}{\mathcal{G}}\mathcal{F}_2^{\upsilon,\eta} & \frac{1}{\mathcal{G}}\mathcal{F}_2^{\upsilon,\upsilon} \end{array} \right]$$

this becomes

$$\left[\begin{array}{c} \eta \\ \upsilon \end{array} \right]_j^{n+1} = (\boldsymbol{I} - \Delta t \boldsymbol{F}_2) \left[\begin{array}{c} \eta \\ \upsilon \end{array} \right]_j^n$$