Modified differential equation analysis on the RK3+central scheme applied on inviscid Burgers equation

Jinyuan Liu*

December 30, 2023

Abstract: In this note we perform an analysis of the stability and dispersion properties of a numerical scheme with second-order difference in space and third-order Runge-Kutta in time, for a one-dimensional linear advection equation. This combination is commonly used in incompressible CFD solvers. We demonstrate on a one-dimensional nonlinear convection equation (Burgers equation) that this scheme will lead to considerable numerical dispersion which can not be reduced by only refining the grid, which is a note of caution for its application.

1 Linear advection equation

Consider the linear advection equation

$$u_t + cu_x = 0 \tag{1}$$

as a substitute to the nonlinear convection equation (Burgers equation, as a prototype for Navier-Stokes/Euler in numerical analysis)

$$u_t + uu_x = 0 \tag{2}$$

for the sake of convenience in numerical analysis.

2 Runge-Kutta time-stepper

Consider the following family of RK3 scheme where u^n denotes the value of $u(x,t)$ at t_n,

$$k_1 = f(u^n, t_n) \tag{3}$$

$$k_2 = f(u^n + a_{21}\Delta t k_1, t_n + a_{21}\Delta t) \tag{4}$$

$$k_3 = f(u^n + a_{31}\Delta t k_1 + a_{32}\Delta t k_2, t_n + (a_{31} + a_{32})\Delta t) \tag{5}$$

$$u^{n+1} = u^n + \Delta t(s_1 k_1 + s_2 k_2 + s_3 k_3) \tag{6}$$

where the coefficients are generic.

The Runge-Kutta-Wray scheme has

$$a_{21} = \frac{8}{15},\ a_{31} = \frac{1}{4},\ a_{32} = \frac{5}{12},\ s_1 = \frac{1}{4},\ s_2 = 0,\ s_3 = \frac{3}{4} \tag{7}$$

while the Runge-Kutta-Ralston scheme has

$$a_{21} = \frac{1}{2},\ a_{31} = 0,\ a_{32} = \frac{3}{4},\ s_1 = \frac{2}{9},\ s_2 = \frac{1}{3},\ s_3 = \frac{4}{9}. \tag{8}$$

Note that for all RK3 schemes

$$s_1 + s_2 + s_3 = 1. \tag{9}$$

*email address: wallturb@gmail.com

3 Central difference

Semi-discretize Eq. (1) we have

$$RHS = \mathcal{F}(u_j) = -\frac{c}{2\Delta x}(u_{j+1} - u_{j-1}) \tag{10}$$

where $\mathcal{F}$ is the finite difference operator.

We also prepare the following finite difference relations (from Taylor expansion around u_j^n) before deriving the modified differential equation (MDE). Expansions till the orders shown below are necessary for later calculations.

$$u_{j+1} - u_{j-1} = 2\Delta x u_x + \frac{1}{3}\Delta x^3 u_{xxx} + \mathcal{O}(\Delta x^5) \tag{11}$$

$$u_{j+2}^n - 2u_j^n + u_{j-2}^n = 4\Delta x^2 u_{xx} + \frac{4}{3}\Delta x^4 u_{xxxx} + \mathcal{O}(\Delta x^6) \tag{12}$$

$$u_{j+3}^n - 3u_{j+1}^n + 3u_{j-1}^n - u_{j-3}^n = 8\Delta x^3 u_{xxx} + 4\Delta x^5 u_{xxxxx} + \mathcal{O}(\Delta x^7) \tag{13}$$

We also establish the following relations from Eq. (1)

$$u_t = -cu_x \tag{14}$$

$$u_{tt} = (-cu_x)_t = (-cu_t)_x = (c^2 u_x)_x = c^2 u_{xx} \tag{15}$$

$$u_{ttt} = -c^3 u_{xxx} \tag{16}$$

$$u_{tttt} = c^4 u_{xxxx} \tag{17}$$

Without further specifying, all derivatives at evaluated at u_j^n.

4 Modified differential equation

Apply RK3, we have

$$k_{1,j} = -\frac{c}{2\Delta x}(u_{j+1}^n - u_{j-1}^n) \tag{18}$$

$$k_{2,j} = -\frac{c}{2\Delta x}[u_{j+1}^n + a_{21}(-\frac{c\Delta t}{2\Delta x})(u_{j+2}^n - u_j^n) - u_{j-1}^n + a_{21}(\frac{c\Delta t}{2\Delta x})(u_j^n - u_{j-2}^n)] \tag{19}$$

$$k_{3,j} = -\frac{c}{2\Delta x}\{u_{j+1}^n + a_{31}(-\frac{c\Delta t}{2\Delta x})(u_{j+2}^n - u_j^n) \tag{20}$$

$$+ a_{32}(-\frac{c\Delta t}{2\Delta x})[u_{j+2}^n + a_{21}(-\frac{c\Delta t}{2\Delta x})(u_{j+3}^n - u_{j+1}^n) - u_j^n + a_{21}(\frac{c\Delta t}{2\Delta x})(u_{j+1}^n - u_{j-1}^n)] \tag{21}$$

$$- u_{j-1}^n + a_{31}(\frac{c\Delta t}{2\Delta x})(u_j^n - u_{j-2}^n) \tag{22}$$

$$+ a_{32}(\frac{c\Delta t}{2\Delta x})[u_j^n + a_{21}(-\frac{c\Delta t}{2\Delta x})(u_{j+1}^n - u_{j-1}^n) - u_{j-2}^n + a_{21}(\frac{c\Delta t}{2\Delta x})(u_{j-1}^n - u_{j-3}^n)]\} \tag{23}$$

where $k_{l,j}$ denotes the l-th ($l = 1, 2, 3$) step for x_j in RK3.

From Eq. (6) we have

$$LHS = \frac{u_j^{n+1} - u_j^n}{\Delta t} = u_t + \frac{\Delta t}{2}u_{tt} + \frac{\Delta t^2}{6}u_{ttt} + \frac{\Delta t^3}{24}u_{tttt} + \mathcal{O}(\Delta t^5) \tag{24}$$

$$RHS = s_1 k_1 + s_2 k_2 + s_3 k_3 \triangleq RHS_1 + RHS_2 + RHS_3 \tag{25}$$

Replace the differences with derivatives, we have

$$RHS_1 = -\, s_1 c[u_x + \frac{1}{6}\Delta x^2 u_{xxx} + \mathcal{O}(\Delta x^4)] \tag{26}$$

$$RHS_2 = -\, s_2 c[u_x + \frac{1}{6}\Delta x^2 u_{xxx} + \mathcal{O}(\Delta x^4)] + s_2 c^2 a_{21}\Delta t[u_{xx} + \frac{1}{3}\Delta x^2 u_{xxxx} + \mathcal{O}(\Delta x^4)] \tag{27}$$

$$RHS_3 = -\, s_3 c[u_x + \frac{1}{6}\Delta x^2 u_{xxx} + \mathcal{O}(\Delta x^4)] + s_3 c^2 (a_{31} + a_{32})\Delta t[u_{xx} + \frac{1}{3}\Delta x^2 u_{xxxx} + \mathcal{O}(\Delta x^4)] \tag{28}$$

$$- s_3 c^3 \Delta t^2 a_{21} a_{32}[u_{xxx} + \frac{1}{2}\Delta x^5 u_{(5)} + \mathcal{O}(\Delta x^4)] \tag{29}$$

Replace the temporal derivatives in LHS with spatial derivatives:

$$LHS = u_t + \frac{c^2\Delta t}{2}u_{xx} - \frac{c^3\Delta t^2}{6}u_{xxx} + \frac{c^4\Delta t^3}{24}u_{xxxx} + \mathcal{O}(\Delta t^5) \tag{30}$$

Rearrange LHS such that

$$LHS = u_t \tag{31}$$

and the rest are put into RHS according to the order of spatial derivative:

$$RHS_1 = -\,(s_1 + s_2 + s_3)cu_x = -cu_x \tag{32}$$

$$RHS_2 = c^2\Delta t[-\frac{1}{2} + s_2 a_{21} + s_3(a_{31} + a_{32})]u_{xx} \tag{33}$$

$$RHS_3 = c\Delta x^2[-s_3 a_{21} a_{32}\frac{c^2\Delta t^2}{\Delta x^2} - \frac{1}{6}(s_1 + s_2 + s_3)]u_{xxx} \tag{34}$$

$$RHS_4 = c^2\Delta t\Delta x^2[-\frac{c^2\Delta t^2}{24\Delta x^2} + \frac{1}{3}(s_2 a_{21} + s_3(a_{31} + a_{32}))]u_{xxxx} \tag{35}$$

The truncation error should be of the order $O(\Delta t^3, \Delta x^2)$ to reach third-order in time and second-order in space of RK3+central. The restrictions above give the constraints on RK3 parameters.

4.1 Numerical diffusion, dissipation, and stability

Consider Eq. (33). The necessary condition for this method to be stable is the numerical diffusivity (coefficient of u_{xx}) greater than or equal to zero, and to archive third order accuracy in time (implied by (34), assuming $\Delta t \propto \Delta x$ with $O(CFL) = 1$) we require:

$$-\frac{1}{2} + s_2 a_{21} + s_3(a_{31} + a_{32}) = 0, \tag{36}$$

which is the case for Runge-Kutta-Wray/Ralston shown in (7)-(8), implying that these two RK3 methods have no numerical diffusion (the coefficient proportional to u_{xx}) when applied on linear advection equation.

But we hope to have another physical constraint - the numerical dissipation (coefficient of u_{xxxx}) greater than zero, i.e.,

$$-\frac{c^2\Delta t^2}{24\Delta x^2} + \frac{1}{3}(s_2 a_{21} + s_3(a_{31} + a_{32})) < 0 \tag{37}$$

as well, for the energy of the system not to pile up. Plugging in the coefficients (from either RK-Wray or RK-Ralston), the above equation reads

$$-CFL^2 + 4 < 0 \tag{38}$$

implying

$$CFL > 2 \text{ or } CFL < -2 \tag{39}$$

and since $CFL > 0$ the restriction is

$$CFL = \frac{c\Delta t}{\Delta x} > 2 \tag{40}$$

which is not practical. Hence, the dissipation term u_{xxxx} is always injecting energy, but the larger the CFL is, the smaller the energy injection. In nonlinear problems we tend to be more conservative and use roughly $CFL = 0.5$ instead (since the CFL restrictions were typically established in linear problems/analysis).

Consider the Fourier representation

$$u(x,t) = \hat{u}_k(t)e^{ikx} \tag{41}$$

of the solution of

$$u_t = Du_{xxxx}, \tag{42}$$

where the constant D is a dissipation coefficient. We have

$$\hat{u}_k'(t) = D(ik)^4\hat{u}_k(t) \tag{43}$$

implying

$$\hat{u}_k(t) = \hat{u}_k(0)e^{Dk^3}, \tag{44}$$

with a stable solution requiring

$$D < 0. \tag{45}$$

In this context,

$$D_{\text{num}} = c^2\Delta t\Delta x^2[-\frac{c^2\Delta t^2}{24\Delta x^2} + \frac{1}{3}(s_2a_{21} + s_3(a_{31} + a_{32}))] \tag{46}$$

in (35) is positive when $CFL \leq 2$ and corresponds to numerical amplification rather than dissipation.

4.2 Numerical dispersion

Consider the numerical dispersivity (coefficient of u_{xxx}):

$$\gamma_{\text{num}} = -c\Delta x^2(s_3a_{21}a_{32}\frac{c^2\Delta t^2}{\Delta x^2} + \frac{1}{6}) < 0 \tag{47}$$

simplified with

$$s_1 + s_2 + s_3 = 1. \tag{48}$$

Unfortunately numerical dispersion is never zero.

Consider the Fourier representation (41) of the solution of

$$u_t = \gamma u_{xxx} \tag{49}$$

we have

$$\hat{u}_k'(t) = \gamma(ik)^3\hat{u}_k(t) \tag{50}$$

implying

$$\hat{u}_k(t) = \hat{u}_k(0)e^{-i\gamma k^3}. \tag{51}$$

Since the numerical dispersivity γ_{num} in Eq. (47) is negative, the phase error in computing $\hat{u}_k(t)$ will be positive (leading) and will be an increasing function of wavenumber k (finer perturbations being more dispersed). Since the exponential term in (51) is pure imaginary, there is no amplitude change as a numerical consequence.

This results above appear to be independent on the specific RK3 scheme as well.

5 Numerical experiments on Burgers equation

Consider now the non-linear case. The inviscid Burger's equation, written in conservative form, is

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial u^2}{\partial x} = 0. \tag{52}$$

Assume the following initial condition

$$u(x,0) = \begin{cases} e^{-\frac{(x-1.5)^2}{0.15}} & \text{if } x < 2.5 \\ 0 & \text{if } x \geq 2.5 \end{cases} \tag{53}$$

in the domain $0 \leq x \leq 5$ and $0 \leq t \leq 5$ with one boundary condition $u(x=0,t)=0$. Use central+RK3 to obtain the solution within $0 \leq t \leq 5$ and with $\Delta x = 0.01$.

5.1 Effect of decreasing Δt

Below are results from central+RK3 applied on Burgers. We can see that there is a considerable amount of dispersion in Figs. 1-2, as expected from the analysis above (on linear advection equation). The comparison also confirms the statements made before that the larger the CFL is the smaller the dispersive error is, assuming the associated velocity scale is $c = 1 = \max u(x,0)$.

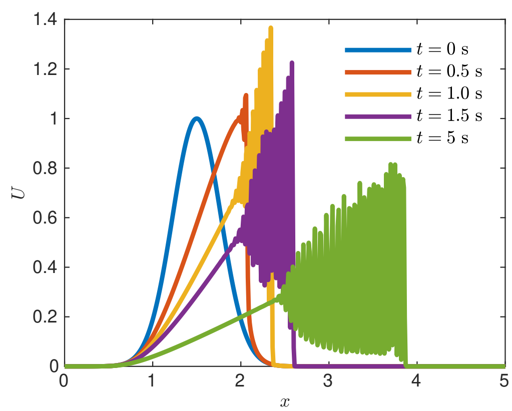

Figure 1: $\Delta t = 0.01, CFL = 1$.

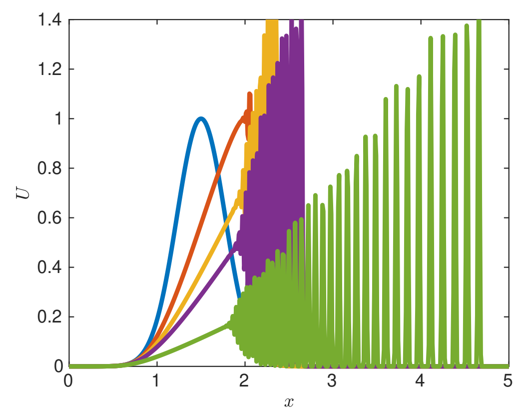

Figure 2: $\Delta t = 0.001, CFL = 0.1$.

But the effect of decreasing Δt is not obvious yet. Reducing Δt actually leads to larger dispersion error, which is **counter-intuitive**. Since numerical dispersion is always nonzero, the only controllable factor is dissipation. By selecting a larger Δt, it can be shown from Eq. (35) that the actual amplification factor is smaller and hence the amplitude of the dispersion is more controlled. In Burgers equation, the convection speed depends on the amplitude (c is not a constant), hence smaller amplification error also implies smaller dispersion error. **This is not the case in a linear advection problem!**

We've learned that controlling dissipation error is as important as controlling dispersion error in nonlinear problems where these two are closely related. (an important lesson from nonlinearity: convection speed $\propto$ amplitude; $c = u$)

5.2 Effect of decreasing Δx

According to (47), $|\gamma_{\text{num}}|$ should decrease as Δx decreases, if the equation is linear, which appear to be true in Figs. 3-4 as compared to Figs. 1-2 when CFL is large. But when CFL is small, as mentioned, the dispersion error and amplitude error go hand-in-hand, hence the effect of further reducing Δx would not be as wanted.

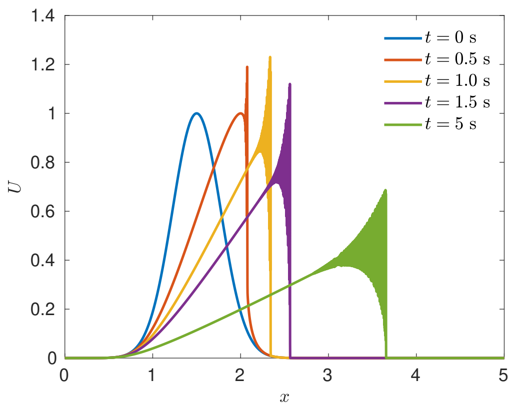

Figure 3: $\Delta x = 0.001, CFL = 1$.

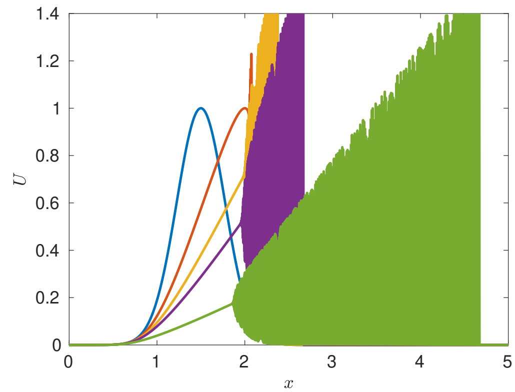

Figure 4: $\Delta x = 0.001, CFL = 0.1$.

5.3 Effect of artificial dissipation

We are probably at a point when **actual/artificial dissipation** is needed, unless TVD schemes are involved, in such a convective problem. Considering Eqns. (32)-(35), multiplying a_{32} by a constant factor seems to be a good choice (which was found due to a coding mistake...) as shown in the experiment in Figs. 5-6.

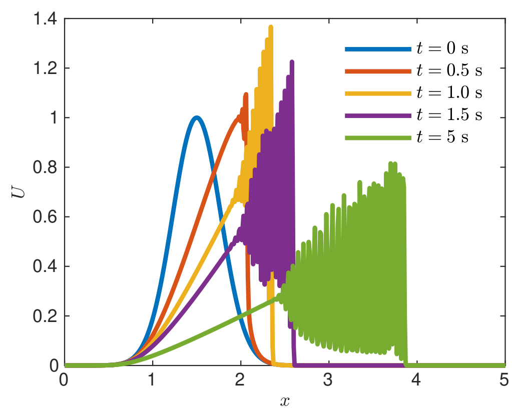

Figure 5: $\Delta t = 0.01, CFL = 1$.

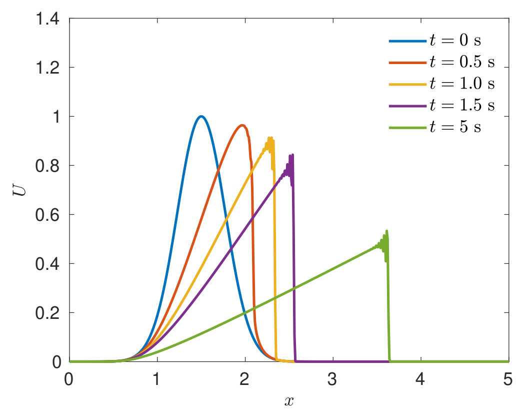

Figure 6: Same, but with $a'_{32} = 3a_{32}$.

6 MATLAB code: central/upwind+RK3 for Burgers

Could be distributed upon reasonable request (e.g. you are not currently a student in MAE 290B).