GRID AVERAGE AND NODAL VALUE IN HIGH ORDER SCHEME

Z. Liu

College of Engineering
Peking University
Beijing, China

Q. Cai*

Center for Applied Physics and Technology
Beijing, China

State Key Laboratory for Turbulence and Complex System
Beijing, China

ABSTRACT

The concepts of nodal value and grid average in cell centered finite volume method (FVM) are clarified in this work, strict distinction between the two concepts in constructing numerical schemes is made, and common fault in misidentifying the two concepts is pointed out. The expansion based on grid average, similar to Taylor's expansion, is deduced to construct correct scheme in terms of grid average and to obtain modified partial differential equation (MPDE) which determines the order of accuracy of numerical scheme theoretically. Correct high order scheme, taking QUICK (Quadratic Upstream Interpolation for Convective Kinematics) scheme as an example, is constructed in different approaches. Furthermore, the property of interpolation coefficients is analyzed. We also pointed out that for high order schemes, round-off error dominates the absolute error in fine grid and truncation error dominates the absolute error in coarse grid.

Keywords: Grid average, Nodal value, High order scheme, Error analysis.

1. INTRODUCTION

The numerical schemes can be classified in terms of accuracy. The order of accuracy of a particular scheme is defined as the highest order of polynomial which the scheme is able to capture accurately. The convergence rate is defined as the power index of mesh size to which numerical error (measured by e.g. the discrete L^2 norm) converges with mesh refinement. Theoretically, the order of accuracy is the same as the order of leading term in truncation error of the corresponding MPDE, and the convergence rate is one order lower than the order of leading term in truncation error of the corresponding MPDE. However, in CFD practice, nodal value, either element center or grid vertex is regarded as integrated value in constructing numerical schemes with conservation laws. This provides exact result only for first- and second-order schemes and fails in high order schemes such as Quadratic Upwind Interpolation Convective Kinematics (QUICK) scheme [1, 2]. It is a common fault that value on specific points are regarded as integrated value.

Numerical schemes are constructed with two types of data: nodal value and grid average. To a specific function, whether scalar or vectorial, nodal values are defined as values of the function at grid node points, and grid averages are defined as integral of the function on each grid element divided by the volume of the element (we use "integral average" to refer to this process briefly following in this work). While constructing first- and second-order schemes, piecewise constant approximation and piecewise linear approximation are assumed respectively. In those cases, value at the center of grid element is equivalent to integral average over the grid element, so it is not necessary to distinguish nodal value and grid average in construction of numerical schemes. However, in high order schemes (higher than second-order), nodal value and grid average are inconsistent, so the expected order of accuracy is achieved only if strict distinction is made. The more wildly used data is nodal value since Taylor's expansion can be applied naturally in Finite Difference Method (FDM), however grid average is more convenient in FVM since its conceptual accordance with conservation laws. In Sec. 2, similar to Taylor's expansion, a new method of expansion of grid average over different orders of derivatives of the function is

* Corresponding author (caiqd@pku.edu.cn)

DOI : 10.1017/jmech.2018.24
Copyright © 2018 The Society of Theoretical and Applied Mechanics
https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

established, which is the theoretical basis of the following analysis. When FVM schemes are constructed, grid averages are introduced naturally, but one common fault is that grid average is regarded as nodal value and Taylor's expansion is directly applied to the discrete data. This produces the same form of scheme in first- and second-order cases (the detail materials are provided in Sec. 3), but fails to produce higher order schemes in the same way. Taking QUICK scheme as example, we present how to introduce grid average into construction of numerical scheme correctly, and the counterpart with nodal value as basic variable is also discussed. Furthermore, the relationship between interpolation coefficients of different orders schemes is attained in Sec. 4.

In Sec. 5, the relationship between absolute error and grid size is studied; the theoretical analysis implies that for a specific scheme, there exists one optimum size: truncation error dominates the absolution error when grid size is larger than the optimum, and round-off error dominates the absolute error when grid size is smaller than the optimum.

2. EXPANSION BASED ON GRID AVERAGE

The Finite Difference Method (FDM) is in terms of nodal values, so it is convenient to apply Taylor's expansion to FDM. As a result, almost all widely used numerical schemes are based on Taylor's expansion. However, in FVM, the basic variables are grid averages. FVM schemes should be constructed with expansions based on grid averages instead of applying Taylor's expansion directly. In this section, a new method is presented to achieve expansion based on grid averages. We take 1-D convection equation as example:

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0,\quad x\in[0,L],\quad t\in[0,T]. \tag{1}$$

Eq. (1) can be rewritten into conservation form:

$$\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}=0 \tag{2}$$

while the flux function $f(x) = au(x)$ produces the linear equation in Eq. (1), and other PDE is attained with different flux function such as $f(x) = u^2/2$ produces a non-linear equation.

Given initial data of Eq. (1), the solution is $u = u(x, t) = u_0(x - at)$. Discretize Eq. (1) in space-time (x, t) with uniform grid, then the coordinates of grid nodes are:

$$x_j=j\Delta x,\quad j=0,1,2,\ldots,N;\quad \Delta x=\frac{L}{N}$$

$$t^n=n\Delta t,\quad n=0,1,2,\ldots,N;\quad \Delta t=\frac{T}{M}$$

There are two kinds of unknowns: nodal values $u_j^n=u(x_j,t^n)$ in FDM and grid averages in FVM which is defined as

$$\overline{u}_j^n=\frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}u(x,t^n)dx \tag{3}$$

Obviously, grid averages differ from nodal values, but the difference vanishes in first- and second-order schemes due to piecewise constant approximation and piecewise linear approximation respectively. For higher order schemes, nodal values are different from grid averages, so we have to distinguish the two concepts clearly to ensure the accuracy of schemes.

Taylor's expansion can be applied naturally to nodal values, but not for grid averages $\overline{u}_j^n$. The antiderivative of $u(x, t)$, denoted as $V(x, t)$, should be introduced to produce the proper expansion. Define the notation $V_{j+\frac{1}{2}}(t)=V\left(x_{j+\frac{1}{2}},t\right)$, then we have

$$\overline{u}_{j+k}^{n+\ell}=\frac{V_{j+\frac{1}{2}+k}^{n+\ell}-V_{j-\frac{1}{2}+k}^{n+\ell}}{\Delta x} \tag{4}$$

Since $V_{j+\frac{1}{2}+k}^{n+\ell}$ is nodal value, Taylor's expansion can be applied to $V(x, t)$ at point (j, n):

$$V_{j+\frac{1}{2}+k}^{n+\ell}=V_j^n+\sum_{m=1}^{\infty}\left[\frac{1}{m!}\left(\ell\Delta t\frac{\partial}{\partial t}+\left(k+\frac{1}{2}\right)\Delta x\frac{\partial}{\partial x}\right)^m V_j^n\right]. \tag{5}$$

Substituting Eq. (5) into Eq. (4), we have:

$$\overline{u}_{j+\frac{1}{2}+k}^{n+\ell}=\frac{1}{\Delta x}\sum_{m=1}^{\infty}\left[\frac{1}{m!}\left(\ell\Delta t\frac{\partial}{\partial t}+\left(k+\frac{1}{2}\right)\Delta x\frac{\partial}{\partial x}\right)^m V_j^n-\frac{1}{m!}\left(\ell\Delta t\frac{\partial}{\partial t}+\left(k-\frac{1}{2}\right)\Delta x\frac{\partial}{\partial x}\right)^m V_j^n\right] \tag{6}$$

The derivatives to variable t all vanish, and the derivatives of V with respect to variable x can be replaced by relation $u = V_x$:

$$\overline{u}_{j+k}^{n+\ell}=\frac{1}{\Delta x}\sum_{m=1}^{\infty}\frac{1}{m!}\left\{\sum_{p=1}^{m}\binom{m}{p}(\ell\Delta t)^{m-p}\left[\sum_{s=0}^{p}\frac{1}{2^s}\binom{p}{s}k^{p-s}(1-(-1)^s\right]\Delta x^p\frac{\partial^{m-1}u}{\partial t^{m-p}\partial x^{p-1}}\right|_j^n\Bigg\}. \tag{7}$$

Eq. (7) is series expansion of grid average, so we can attain the MPDE by substitution.

Consider several special conditions for Eq. (7). When $\ell = 0$, $k = 0$:

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

$$\overline{u}_j^n = \sum_{m=1}^{\infty} \frac{1-(-1)^m}{2^m m!} \Delta x^{m-1} \left. \frac{\partial^{m-1} u}{\partial x^{m-1}} \right|_j^n = \sum_{p=0}^{\infty} \frac{1}{4^{p(2p+1)!}} \Delta x^{2p} \left. \frac{\partial^{2p} u}{\partial x^{2p}} \right|_j^n . \tag{8}$$

The second equation in Eq. (8) is obtained because the even terms in the summation over m is definitely zero, so consider odd terms and substitute m with $m = 2p + 1$ to transfer the summation over m to summation over p. When $k = 0$ is given and ℓ is arbitrary:

$$\overline{u}_j^{n+\ell} = \frac{1}{\Delta x} \sum_{m=1}^{\infty} \frac{1}{m!} \left[\sum_{p=0}^{m} (\ell \Delta t)^{m-p} \frac{1-(-1)^p}{2^p} \Delta x^p \left. \frac{\partial^m V}{\partial t^{m-p} \partial x^p} \right|_j^n \right] \tag{9}$$

$$\overline{u}_j^{n+\ell} = \sum_{m=1}^{\infty} \frac{1}{m!} \left[\sum_{s=0}^{\left[\frac{m-1}{2}\right]} \frac{1}{4^s} \binom{m}{2s+1} (\ell \Delta t)^{m-2s-1} \Delta x^{2s} \left. \frac{\partial^{m-1} u}{\partial t^{m-2s-1} \partial x^{2s}} \right|_j^n \right] \tag{10}$$

with $\left[\frac{m-1}{2}\right]$ representing floor function of $\frac{m-1}{2}$. When given $\ell = 0$ and k is arbitrary:

$$\overline{u}_{j+k}^n = \sum_{m=1}^{\infty} \frac{\Delta x^{m-1}}{m!} \left[\sum_{p=0}^{m} \frac{1-(-1)^p}{2^p} \binom{m}{p} k^{m-p} \right] \left. \frac{\partial^{m-1} u}{\partial x^{m-1}} \right|_j^n \tag{11}$$

$$\overline{u}_{j+k}^n = \sum_{m=1}^{\infty} \frac{\Delta x^{m-1}}{m!} \left[\sum_{s=0}^{\left[\frac{m-1}{2}\right]} \frac{1}{4^s} \binom{m}{2s+1} k^{m-2s-1} \right] \left. \frac{\partial^{m-1} u}{\partial x^{m-1}} \right|_j^n \tag{12}$$

Some special terms for Eq. (6) can be deduced by computer algebra system (Mathematica code is deployed is deployed which is inspired by [5]):

$$\overline{u}_j^n = u_j^n + \frac{1}{24} \Delta x^2 u_{xx} + \frac{1}{1920} \Delta x^4 u_{xxxx} + \cdots$$

$$\overline{u}_{j-1}^n = u_j^n - \Delta x u_x + \frac{13}{24} \Delta x^2 u_{xx} - \frac{5}{24} \Delta x^3 u_{xxx} + \frac{1}{1920} \Delta x^4 u_{xxxx} + \cdots$$

$$\overline{u}_{j+1}^n = u_j^n + \Delta x u_x + \frac{13}{24} \Delta x^2 u_{xx} + \frac{5}{24} \Delta x^3 u_{xxx} + \frac{1}{1920} \Delta x^4 u_{xxxx} + \cdots$$

$$\overline{u}_j^{n+1} = u_j^n + \Delta t u_t + \frac{1}{2} \Delta t^2 u_{tt} + \frac{1}{24} \Delta x^2 u_{xx} + \frac{1}{6} \Delta t^3 u_{ttt} + \frac{1}{24} \Delta t \Delta x^2 u_{xxt} + \cdots$$

$$\overline{u}_j^{n-1} = u_j^n - \Delta t u_t + \frac{1}{2} \Delta t^2 u_{tt} + \frac{1}{24} \Delta x^2 u_{xx} - \frac{1}{6} \Delta t^3 u_{ttt} - \frac{1}{24} \Delta t \Delta x^2 u_{xxt} + \cdots$$

3. FIRST- AND SECOND-ORDER SCHEMES

Low order schemes refer to first order schemes such as first-order upwind (FOU) scheme and second-order schemes such as Lax-Wendroff scheme. FOU scheme is analyzed in this section. In nodal value case, we have:

$$u_j^{n+1} = u_j^n - c\left(u_j^n - u_{j-1}^n\right) \quad c = \frac{a\Delta t}{\Delta x}. \tag{13}$$

The MPDE of Eq. (13) is attained by Taylor's expansion:

$$u_t + a u_x = \frac{a\Delta x}{2}(1-c) u_{xx} - \frac{a\Delta x^2}{6}(1-c)(1-2c) u_{xxx} + \frac{a\Delta x^3}{24}(1-c)(1-6c+6c^2) u_{xxxx} + \cdots. \tag{14}$$

If the scheme is constructed based on grid average, we should first integrate Eq. (1) over interval $\left[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right] \times \left[t^n, t^{n+1}\right]$:

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

$$\bar{u}_j^{n+1} = \bar{u}_j^n - c\left(\hat{u}_{j+\frac{1}{2}} - \hat{u}_{j-\frac{1}{2}}\right) \tag{15}$$

with

$$\hat{u}_{j+\frac{1}{2}} = \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} u\left(x_{j+\frac{1}{2}}, t\right)dt \tag{16}$$

being the average value of $u(x, t)$ over a time interval at a fixed spacial point. Up to the above step, no approximation is introduced, and all formulae are accurate. Now we would obtain the approximate expression of $\hat{u}_{j+\frac{1}{2}}$ in an explicit scheme. Considering piecewise constant approximation on each grid $\left[x_{j+\frac{1}{2}}, x_{j-\frac{1}{2}}\right]$, there are two grid averages related to point $x_{j+\frac{1}{2}}$: $\bar{u}_{j-1}^n$ and $\bar{u}_j^n$. For $a > 0$, we take the upwind value: $\hat{u}_{j+\frac{1}{2}} = \bar{u}_j^n$ and obtain the same scheme as Eq. (13):

$$\bar{u}_j^{n+1} = \bar{u}_j^n - c\left(\bar{u}_j^n - \bar{u}_{j-1}^n\right) \quad c = \frac{a\Delta t}{\Delta x}. \tag{17}$$

The MPDE of numerical scheme (17) is obtained by Eq. (7):

$$u_t + au_x = \frac{a\Delta x}{2}(1-c)u_{xx} - \frac{a\Delta x^2}{6}(1-c)(1-2c)u_{xxx} + \frac{a\Delta x^3}{24}(1-c)(1-6c+6c^2)u_{xxxx} + \cdots. \tag{18}$$

According to Eq. (18) and Eq. (17), the schemes based on nodal values and grid averages accord with each other at internal grid points. It should be emphasized that numerical schemes based on grid averages can be constructed by Eq. (7) directly just like constructing FD schemes with Taylor's expansion. For instance, if temporal derivatives and spatial derivatives are both kept to only the first-order partial derivative, then the expansions of $\bar{u}_j^n$, $\bar{u}_j^{n+1}$, $\bar{u}_{j-1}^n$ give exactly the first-order upwind scheme.

Due to limitation of capacity of a single paper, more similar deduction on first- and second-order schemes are omitted. It is straightforward to repeat the previous step to verify the following conclusion: for low order schemes, scheme based on nodal value is the same with the corresponding scheme based on grid average, and they have the same leading term of error in MPDE. This property can be illustrated intuitively as first order schemes assume piecewise constant approximation on each grid element in which case grid average is exactly the nodal value on grid centroid. The same illustration lies in second-order schemes since they assume piecewise linear approximation on each grid element.

4. HIGH ORDER SCHEMES

In this section, we take QUICK scheme [2] as an example of high order schemes. During the construction of QUICK scheme in [2], grid average and nodal value are not well distinguished, which results in reduced order of accuracy.

The semi-discrete scheme is

$$\frac{du_j}{dt} + \frac{1}{\Delta x}\left(f_{j+\frac{1}{2}} - f_{j-\frac{1}{2}}\right) = 0, \tag{19}$$

with $f_{j+\frac{1}{2}}$ being the flux across interface $x_{j+\frac{1}{2}}$. Flux f is approximated by quadratic polynomials. According to the sign of $a_{j+\frac{1}{2}} = \frac{\partial f}{\partial u}\left(x_{j+\frac{1}{2}}\right)$, a quadratic polynomial providing $f_{j+\frac{1}{2}}$ is constructed in [2] by stencil $\{x_{j-1}, x_j, x_{j+1}\}$ or $\{x_j, x_{j+1}, x_{j+2}\}$ to satisfy upwind property. If u_j is approximated by nodal values on element centroid, flux $f_{j+\frac{1}{2}}$ at interface $x_{j+\frac{1}{2}}$ is obtained by basic Lagrange interpolation on the corresponding stencils:

$$f_{j+\frac{1}{2}} = \begin{cases} \frac{1}{2}(f_j + f_{j+1}) - \frac{1}{8}(f_{j-1} - 2f_j + f_{j+1}), & a_{j+\frac{1}{2}} > 0 \\ \frac{1}{2}(f_j + f_{j+1}) - \frac{1}{8}(f_j - 2f_{j+1} + f_{j+2}), & a_{j+\frac{1}{2}} < 0 \end{cases} \tag{20}$$

The above equation differs from the result presented following in this paper, since we use grid averages instead of nodal values on grid centroid to construct flux at cell boundary. Further analysis on wave equation with $f = au$ shows that the scheme actually has only second-order accuracy.

We rewrite the flux term into fourth-order central scheme part as well as correction term (the last term):

$$f_{1+\frac{1}{2}} = \frac{9}{16}(f_j + f_{j+1}) - \frac{1}{16}(f_{j-1} + f_{j+2}) + \frac{1}{16}|a|(u_{j+2} - 3u_{j+1} + 3u_j - u_{j-1}) \tag{21}$$

and we can deduce by Taylor's expansion that the MPDE of scheme (19) is

$$u_t + au_x = -\frac{1}{24}a\Delta x^2 u_{xxx} + \frac{1}{16}|a|\Delta x^3 u_{xxxx} + \frac{11}{380}a\Delta x^4 u_{xxxxx} + \cdots \tag{22}$$

The leading term of truncation error of Eq. (22) is second-order, which means that the scheme has only second-order accuracy instead of what is expected. The reason why third-order accuracy is not achieved is that grid average and nodal value are not strictly distinguished: the LHS of Eq. (19) should not be nodal values. The disparity on coefficients of the two parts of Eq. (21) reduces order of accuracy.

We now construct QUICK scheme based on grid average. The general form of numerical scheme for the wave equation $u_t + au_x = 0$ with conservative semi-

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

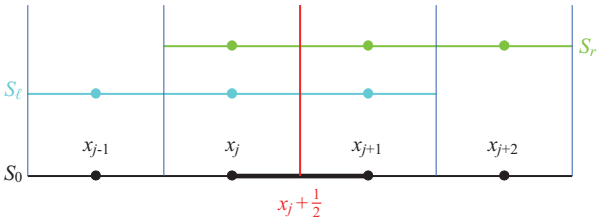


Fig. 1 The stencil for QUICK scheme.

discretization and with grid average as basic variable is:

$$\frac{d\bar{u}_j}{dt}=-\frac{1}{\Delta x}\left(f_{j+\frac{1}{2}}-f_{j-\frac{1}{2}}\right)=\sum_{k=-J}^{J-1}\alpha_k(\bar{u}_{j+k+1}-\bar{u}_{j+k}). \tag{23}$$

Eq. (23) is a scheme of $2J+1$ grid elements: $\bar{u}_j$ is grid average on $\left[x_{j-\frac{1}{2}},x_{j+\frac{1}{2}}\right]$, and $f_{j+\frac{1}{2}}$ is numerical flux in the form of $f_{j+\frac{1}{2}}=\hat{f}(\bar{u}_{j-J+1},\ldots,\bar{u}_{j+J})$ which is a function of $2J$ variables.

In order to obtain the MPDE of scheme (23) by Taylor's expansion, the antiderivative of $u(x, t)$ is introduced as mentioned before:

$$\bar{u}_{j+k}=\frac{V_{j+k+\frac{1}{2}}-V_{j+k-\frac{1}{2}}}{\Delta x} \tag{24}$$

Eq. (23) is equivalently:

$$\frac{d}{dt}\left(\frac{V_{j+\frac{1}{2}}-V_{j-\frac{1}{2}}}{\Delta x}\right)=\sum_{k=-J}^{J-1}\alpha_k\frac{V_{j+k+\frac{3}{2}}-2V_{j+k+\frac{1}{2}}+V_{j+k-\frac{1}{2}}}{\Delta x}. \tag{25}$$

The basic idea of constructing flux term in QUICK scheme is by evaluating the interpolation function at cell interface, and the interpolation function is constructed by a stencil covering several neighbor elements. The stencil is illustrated by Fig. 1. We now construct the interpolation function $p(x)$ by methods of undetermined coefficient.

Assuming $x_{j+\frac{1}{2}}=0$ and

$$p(x)=C_0+C_1x+C_2x^2,\quad x\in[x_j,x_{j+1}]=\left[-\frac{1}{2}\Delta x,\frac{1}{2}\Delta x\right]$$

we have grid average $\bar{u}_j$ on grid element $[x_a, x_b]$:

$$\bar{u}=\frac{1}{x_b-x_a}\int_{x_a}^{x_b}p(x)dx=C_0+\frac{1}{2}C_1(x_a+x_b)+\frac{1}{3}C_2(x_a^2+x_ax_b+x_b^2).$$

In case $a>0$, where stencil S_ℓ is utilized, we just need to solve C_0 in the following equations to obtain the flux $f_{j+\frac{1}{2}}$:

$$\begin{bmatrix}1 & -\frac{3}{2}\Delta x & \frac{7}{3}\Delta x^2\\ 1 & -\frac{1}{2}\Delta x & \frac{1}{3}\Delta x^2\\ 1 & \frac{1}{2}\Delta x & \frac{1}{3}\Delta x^2\end{bmatrix}\begin{bmatrix}C_0\\C_1\\C_2\end{bmatrix}=\begin{bmatrix}\bar{u}_{j-1}\\\bar{u}_j\\\bar{u}_{j+1}\end{bmatrix}.$$

Now we have

$$\begin{aligned}u_{j+\frac{1}{2}}=C_0&=\frac{1}{6}(-\bar{u}_{j-1}+5\bar{u}_j+2\bar{u}_{j+1})\\&=\frac{1}{2}(\bar{u}_j+\bar{u}_{j+1})-\frac{1}{6}(\bar{u}_{j-1}-2\bar{u}_j+\bar{u}_{j+1}),\end{aligned}$$

and finally we have

$$u_{j+\frac{1}{2}}=\begin{cases}\frac{1}{2}(\bar{u}_j+\bar{u}_{j+1})-\frac{1}{6}(\bar{u}_{j-1}-2\bar{u}_j+\bar{u}_{j+1}), & a>0\\ \frac{1}{2}(\bar{u}_j+\bar{u}_{j+1})-\frac{1}{6}(\bar{u}_j-2\bar{u}_{j+1}+\bar{u}_{j+2}), & a\le 0\end{cases} \tag{26}$$

Eq. (26) has similar form as Eq. (20), but Eq. (26) has different coefficient, the MPDE of numerical scheme corresponding to Eq. (26) should be deduced by Eq. (7).

Equation (26) can be rewritten as fourth order central scheme with an upwind correction term (the last term in Eq. (27)):

$$u_{j+\frac{1}{2}}=\begin{cases}\frac{7}{12}(\bar{u}_j+\bar{u}_{j+1})-\frac{1}{12}(\bar{u}_{j-1}+\bar{u}_{j+2})\\ \quad+\frac{1}{12}(\bar{u}_{j+2}-3\bar{u}_{j+1}+3\bar{u}_j-\bar{u}_{j-1}), & a>0\\ \frac{7}{12}(\bar{u}_j+\bar{u}_{j+1})-\frac{1}{12}(\bar{u}_{j-1}+\bar{u}_{j+2})\\ \quad-\frac{1}{12}(\bar{u}_{j+2}-3\bar{u}_{j+1}+3\bar{u}_j-\bar{u}_{j-1}), & a\le 0\end{cases} \tag{27}$$

The flux term can be expressed unified with $\kappa=\text{sgn}\,a$:

$$\begin{aligned}f_{j+\frac{1}{2}}&=\frac{7}{12}a(\bar{u}_j+\bar{u}_{j+1})-\frac{1}{12}a(\bar{u}_{j-1}+\bar{u}_{j+2})\\&\quad+\frac{1}{12}\kappa a(\bar{u}_{j+2}-3\bar{u}_{j+1}+3\bar{u}_j-\bar{u}_{j-1})\end{aligned} \tag{28}$$

By substituting Eq. (28) to Eq. (23), we have the numerical scheme of case $J=2$:

$$\begin{cases}\alpha_{-2}=\frac{1}{12\Delta x}(1+\kappa)a\\ \alpha_{-1}=-\frac{1}{12\Delta x}(7+3\kappa)a\\ \alpha_0=-\frac{1}{12\Delta x}(7-3\kappa)a\\ \alpha_1=\frac{1}{12\Delta x}(1-\kappa)a\end{cases} \tag{29}$$

The MPDE of numerical scheme in Eq. (29) is

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

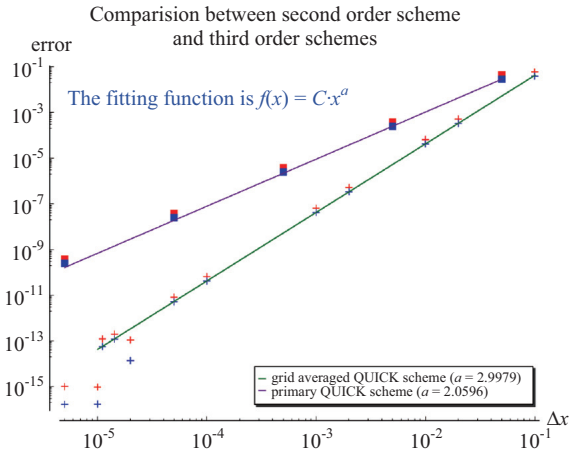


Fig. 2 Comparison between grid-averaged QUICK scheme and primary QUICK scheme. The slopes of the two fitted lines shows the convergence rate of two schemes that verify our theoretical analysis. The red marks are L^1 norm of error between numerical result and analytical result, and the blue marks are L^∞ norm of the error.

$$u_t + au_x = -\frac{1}{12}\kappa a\Delta x^3 u_{xxxx} + \frac{1}{30}a\Delta x^4 u_{xxxxx} - \frac{1}{72}\kappa a\Delta x^5 u_{xxxxxx} + \cdots, \tag{30}$$

so the scheme has third-order accuracy.

In order to verify the accuracy of scheme defined by flux (28), and compare it with scheme defined by flux (21), two numerical experiment is implemented: case 1 shows the scheme by flux (28) and case 2 shows the scheme by flux (21). Temporal domain is discretized by fourth-order Runge-Kutta method. The CFL numbers of both cases are set to 1.0. Figure 2 shows the error of the two schemes, while the error is measured by L^1 norm and L^∞ norm of deviation between the numerical solution and exact solution.

In previous discussion, Eq. (1) is discretized naturally by grid averages, however, with a little more effort, we are able to discretize it by nodal values on cell centroids. With nodal values as stencil, the flux is constructed by Eq. (21), we need to represent the first term of LHS of Eq. (19). Figure 3 shows the stencil is this case, since u_j^n no longer represent the integral value of the cell, the integral $T = I_{j^-} + I_{j^+}$ is introduced.

In Fig. 3, the minus sign in the superscript denotes that this integral value is measured on stencil S_r. Assuming $x_j = 0$ for simplicity, an interpolated polynomial can be constructed on $S_r = \{u_j, u_{j+1}, u_{j+2}\}$ (in the case of $a < 0$):

$$p(x) = C_0 + C_1 x + C_2 x^2, \quad x \in [x_j, x_{j+1}] = \left[-\frac{1}{2}\Delta x, \frac{1}{2}\Delta x\right].$$

The coefficients are determined by the nodal values:

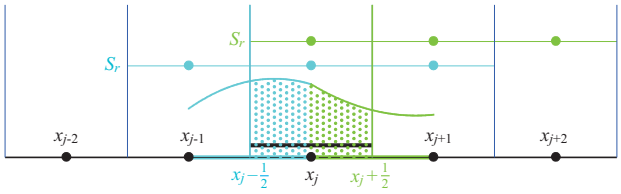


Fig. 3 The stencil for QUICK scheme based on nodal values.

$$C_0 = \frac{1}{8}(-u_j + 6u_{j+1} + 3u_{j+2}),$$

$$C_1 = \frac{u_{j+1} - u_j}{\Delta_x}, \quad C_2 = \frac{u_{j+2} - 2u_{j+1} + u_j}{2\Delta x^2}.$$

Now we have

$$I_{j^+}^- = \int_0^{\frac{1}{2}\Delta x} p(x)dx = \frac{1}{2}C_0\Delta x + \frac{1}{8}C_1\Delta x^2 + \frac{1}{24}C_2\Delta x^3 = \frac{1}{24}(2u_j + 11u_{j+1} - u_{j+2})$$

$$I_{(j+1)^-}^- = \int_{-\frac{1}{2}\Delta x}^0 p(x)dx = \frac{1}{2}C_0\Delta x - \frac{1}{8}C_1\Delta x^2 + \frac{1}{24}C_2\Delta x^3 = \frac{1}{24}(8u_j + 5u_{j+1} - u_{j+2}).$$

The integral value is

$$\overline{u}_j = T^- = I_{j^+}^- + I_{j^-}^- = \frac{\Delta x}{24}(2u_{j-1} + 19u_j + 4u_{j+1} - u_{j+2}).$$

The above calculation means that the final discretized equation while $a < 0$ is

$$\frac{1}{24}\left(2\frac{du_{j-1}}{dt} + 19\frac{du_j}{dt} + 4\frac{du_{j+1}}{dt} - \frac{du_{j+2}}{dt}\right) + \frac{1}{\Delta x}\left(f_{j+\frac{1}{2}} - f_{j-\frac{1}{2}}\right) = 0 \tag{31}$$

In case of $a > 0$, the result is symmetrically:

$$\frac{1}{24}\left(-\frac{du_{j-2}}{dt} + 4\frac{du_{j-1}}{dt} + 19\frac{du_j}{dt} + 2\frac{du_{j+1}}{dt}\right) + \frac{1}{\Delta x}\left(f_{j+\frac{1}{2}} - f_{j-\frac{1}{2}}\right) = 0 \tag{32}$$

Combining Eq. (31) and (32) together, the scheme that match all a is

$$\sum_{k=-2}^{2}\beta_k\frac{du_{j+k}}{dt} + \frac{1}{\Delta x}\left(f_{j+\frac{1}{2}} - f_{j-\frac{1}{2}}\right) = 0, \tag{33}$$

with coefficients:

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

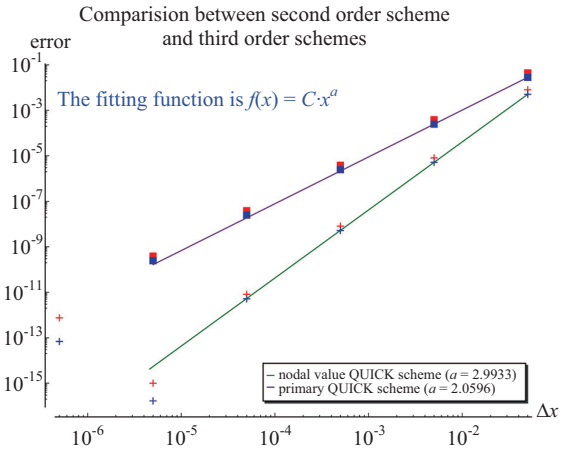


Fig. 4 Comparison between nodal value QUICK scheme and primary QUICK scheme. The slopes of the two fitted lines shows the convergence rate of two schemes that verify our theoretical analysis. The red marks are L^1 norm of error between numerical result and analytical result, and the blue marks are L^∞ norm of the error.

$$\begin{cases} \beta_{-2} = -\dfrac{1}{384}(7+8\kappa+|\kappa|) \\ \beta_{-1} = \dfrac{1}{96}(11+4\kappa+|\kappa|) \\ \beta_{0} = \dfrac{1}{192}(155-3|\kappa|) \\ \beta_{1} = \dfrac{1}{96}(11-4\kappa+|\kappa|) \\ \beta_{2} = -\dfrac{1}{384}(7-8\kappa+|\kappa|) \end{cases}$$

The MPDE of scheme (33) can be obtained by Taylor's expansion:

$$u_t + au_x = \frac{5}{48}\kappa a\Delta x^3 u_{xxxx} + \frac{(47-15|\kappa|)a\Delta x^4}{5760}u_{xxxxx} + \frac{19\kappa a\Delta x^5}{1152}u_{xxxxxx} + \cdots \quad (34)$$

which yields third-order accuracy, and this is verified by numerical experiment shown in Fig. 4.

The three numerical cases illustrated in Fig. 2 and Fig. 4 are strong and clear evidence that supports our previous theoretical analysis, and especially in Fig. 4, we can conclude that the primary QUICK scheme in literature [2] does use an improper approximation in Eq. (20) that reduces the accuracy to second-order. Our new QUIKC scheme, either gird-averaged case or nodal vale case, is capable of retaining third-order accuracy.

5. MORE ON ERROR OF HIGHER ORDER SCHEMES

The absolute error of numerical schemes consists of truncation error and round-off error. Assuming that the numerical scheme applied to a particular differential equation is compatible (which implies that the discrete scheme converges to the differential equation while Δx and Δt tend to zero), the solution to the discrete scheme converges to the solution of the differential equation [3]. The rate of convergence is measured with norm of truncation error:

$$E = \|u_h - u_e\| = C \cdot h^p \quad (35)$$

with h as the grid size, u_h as numerical solution, u_e as exact solution. In Eq. (35), p is the order of accuracy, and C is constant independent of h and p. In the ideal case, once the grid size shrinks by half, the norm of truncation error reduces to 1/2p of the previous, so the log-log plot of E-h relation is a straight line with slop p. Though a scheme with higher order will converge faster to exact solution, it is not able to converge to zero since there is round-off error in numerical computation. Round-off error dominates the absolute error instead of truncation error when the absolute error nearly reaches machine zero.

Li *et al*. [4] pointed out that, the product of round-off error Δr and truncation error Δe satisfies that

$$\Delta e + \Delta r > C$$
$$\Delta e \cdot \Delta r \geq h$$

with C and h are constants determined by the primary equations. The conclusion still holds in high order schemes, so the absolute error cannot reduce infinitely in the way mentioned above. As the grid size shrinks to a certain value which is small enough, the absolute error increases as the grid size go on shrinking since the growth of calculation leads to the growth of round-off error.

There exists an optimum grid size for all numerical schemes. The round-off error dominates the absolute error with very small grid size, and it increases while grid size decreases; the truncation error dominates the absolute error with large grid size, and it decreases while grid size decreases. Due to the above to elementary kinds of errors, there exists an optimum grid size with minimum absolute error. In Fig. 4, there is one special mesh size where $\Delta x = 5 \times 10^{-6}$. In that case, the absolute error is much higher than the truncation error implying a massive round-off error there. However, a systematical study on round-off error with double precision consumes too much computational resource, we turn to single precision float type. Figure 5 shows more examples on single precision computation. Since round-off error is very non-linear, the absolute error is distributed in a large area when the mesh size is smaller than the optimum size.

6. CONCLUSIONS

In this work, we define the grid average and nodal value rigorously, as well as the accuracy of numerical

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press

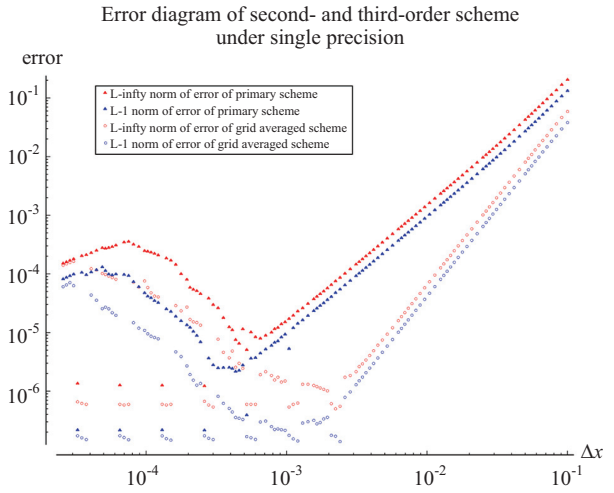


Fig. 5 The error diagram (triangles are second-order case and circles are third-order case) which shows the optimum mesh size for second-order scheme is $\Delta x = 4 \times 10^{-3}$ with CFL number is 1.0, and the optimum mesh size for third-order scheme is $\Delta x = 8 \times 10^{-2}$ with CFL number is 1.0.

schemes. By constructing the expansion based on grid average, analysis on order of accuracy of schemes based on grid average (which are ubiquitous in FVMs) becomes feasible. Theoretical analysis implies that:

1. The correct order of accuracy of a specific scheme, especially high order ones, can only be obtained by rigorous deduction of MPDE.
2. The misuse of grid average and nodal value in constructing numerical schemes will lead to reducing the order of accuracy.
3. When constructing FVM schemes, the expansion-based grid average should be utilized instead of Taylor's expansion.
4. The absolute error of high order scheme decrease very fast as grid size reduces, but due to existence of round-off error, the absolute error does not decrease infinitely. Theoretically there exists an optimum grid size which minimizes the absolute error. Numerical study on this will be done in future research.

ACKNOWLDEGEMENTS

The authors would like to acknowledge National Natural Science Foundation of China (NSFC) under Grants No. 11521091 and No. 91752202 for its support in of this work.

REFERENCES

1. Remi, A., Adam, L. and Mario, R., "Construction of Very High Order Residual Distribution Schemes for Steady Inviscid Flow Problems on Hybrid Unstructured Meshes," *Journal of Computational Physics*, **230**, pp. 4103-4136 (2011).
2. Brian, P. L., "A Stable and Accurate Convective Modelling Procedure Based on Quadratic Upstream Interpolation," *Computer Methods in Applied Mechanics and Engineering*, **19**, pp.59-98 (1979).
3. Peter, D. L. and Robert, D. R., "Survey of the Stability of Linear Finite difference equations," *Communications on Pure and Applied Mathematics*, **9**, pp. 267-293 (1956).
4. Li, J., Zeng, Q. and Chou, J., "Computational Uncertainty Principle in Nonlinear Ordinary differential Equations," *Science in China Series E: Technological Sciences*, **44**, 55-74 (2001).
5. Liu, R. and Shu, Q., *Several New Methods on Computational Fluid Dynamics* (in Chinese), Science Press (2003).

(Manuscript received November 22, 2017,
accepted for publication April 28, 2018.)

https://doi.org/10.1017/jmech.2018.24 Published online by Cambridge University Press