

## Analysis of the Temporal Discretisation of Finite Difference Schemes

M.E. Young<sup>1</sup> and A. Ooi<sup>2</sup>

<sup>1</sup>Aerospace Division, Defence Science and Technology Organisation, Victoria 3207, Australia

<sup>2</sup>Department of Mechanical Engineering, University of Melbourne, Victoria 3010, Australia

### Abstract

This paper develops a general spectral analysis that combines the effects of spatial and temporal discretisation to form a framework for the error analysis of the full-discretisation. It explores some of the implications of the full-discretisation including dissipation and group velocity errors; and the impact that varying Courant–Friedrichs–Levy (CFL) number has on these parameters for different spatial and temporal discretisations. In so doing, it reveals that the investigation of the full-discretisation schemes contains more information than what is available from analysis of the semi-discretisation alone. The application of the framework is then tested for the advection and linearised Euler equations.

### Introduction

The solution of complex systems of partial differential equations encountered in fluid mechanics generally requires the use of numerical methods that require both spatial and temporal discretisation. The dispersive (phase) and dissipative (amplitude) errors associated with the spatial semi-discretisation have been extensively studied and are well understood (e.g. [1, 4, 8]). When it comes to assessing the accuracy of a numerical method, however, the spatial semi-discretisation is only part of the story. It will be shown in this paper—using spectral analysis to assess the accuracy impacts of temporal discretisation—that it can be critical to assess the spatial discretisation *in conjunction* with the temporal discretisation method to ensure that the overall full-discretisation scheme is stable *and* accurate.

There are several methods available to spatially discretise a domain, including finite difference, finite volume, finite element, and spectral methods. The present work will concentrate on the former, where analytical tools such as Fourier analysis are readily applicable. We take as the starting point of this paper the finite difference method and its associated modified wavenumber, denoted as  $k^*\Delta$ . Refer to [4] and [8] for background on establishing the modified wavenumber.

### Finite Difference Methods

The basic structure of the finite difference equation is

$$\sum_{m=-M}^{+M} b_m \left( \frac{\partial \phi}{\partial x} \right)_{i+m} = \frac{1}{\Delta x} \sum_{l=-L}^{+L} a_l \phi_{i+l} \quad (1)$$

where  $a_l$  and  $b_m$  are the coefficients representing the weighting that determines how much nodes  $i+l$  and  $i+m$  will contribute to the calculation of the derivative at node  $i$ . For the process of determining the values of these coefficients, one is referred to [2, 3] and [6].

The result of the Fourier analysis for spatial semi-discretisations is to establish a relationship between the non-dimensionalised modified wavenumber of the finite difference scheme,  $k^*\Delta x$ , and the exact non-dimensional wavenumber,  $k\Delta x$ . The real and imaginary parts of  $k^*\Delta$  determine the dispersive and dissipative characteristics of a spatial scheme, respectively. Examples of the real part of this relationship for three different high resolution finite difference schemes are shown in figure 1. The solid line shows the ideal relationship in which all wavenumbers up to the Nyquist limit are resolved, i.e.  $k^*\Delta x = k\Delta x$ . Ideally, a spatial discretisation scheme will match this over as much of the

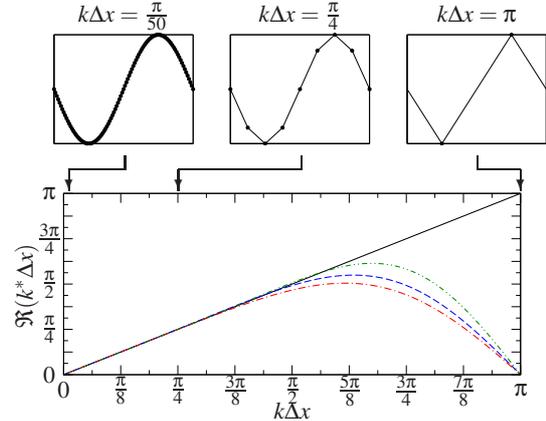


Figure 1: Modified wavenumber relationship for: --- DRP4, --- CDS6, --- WEUDS6, — exact relationship.

wavenumber range as possible. As the wavenumber increases (the effect of which is shown by the inset diagrams), the resolution of the schemes decrease.

It is the spatial discretisation, through this relationship between  $k^*\Delta$  and  $k\Delta$ , that provides the limit of resolution for a numerical scheme. That is, no matter how accurately the temporal scheme propagates the solution through time, it will always be limited by characteristics determined by the spatial scheme.

### Temporal Discretisation Methods

The range of temporal discretisation methods for the propagation of systems of ODEs (as formed from finite difference approximations of PDEs) can be divided into two categories: explicit and implicit. Explicit methods are generally more straightforward to implement and computationally less intensive; however, the advantage of implicit methods is that they can offer unconditional stability [1], and therefore allow much larger timesteps to be achieved. Nevertheless, it will be shown that stability limitations are not the only criteria one should use to determine an acceptable timestep.

The ‘standard’ fourth-order Runge–Kutta (RK4) scheme as described in [4], will be used in this work. The implicit time-stepping method chosen for analysis in this work is the Crank–Nicolson method:

$$\phi^{n+1} = \phi^n + \frac{\Delta t}{2} \left[ \frac{d\phi(\phi^{n+1}, t^{n+1})}{dt} + \frac{d\phi(\phi^n, t^n)}{dt} \right]. \quad (2)$$

### From Semi- to Full- Discretisation

Now we take the semi-discretised result, i.e.  $k^*\Delta x$ , and advance it through time with the advection equation  $\frac{\partial \phi(x,t)}{\partial t} + u \frac{\partial \phi(x,t)}{\partial x} = 0$ .

The solution at timestep  $t + \Delta t$  can be written in terms of the Fourier transform, denoted by  $\hat{\phi}$ , and values at the current timestep,  $t$ , as

$$\hat{\phi}(k, t + \Delta t) = \left[ 1 - \Delta t u i k^* + \frac{(\Delta t u i k^*)^2}{2} - \frac{(\Delta t u i k^*)^3}{6} + \frac{(\Delta t u i k^*)^4}{24} \right] \hat{\phi}(k, t), \quad (3)$$

for the RK4 scheme [1], and

$$\hat{\phi}(k, t + \Delta t) = \left[ \frac{1 + \frac{u\Delta t}{2} ik^*}{1 - \frac{u\Delta t}{2} ik^*} \right] \hat{\phi}(k, t). \quad (4)$$

for the Crank–Nicolson scheme [4]. In each instance, the terms in the square brackets represent the numerical amplification factor,  $Z$ .  $Z$  is the key to the accuracy of the full-discretisation, as it encapsulates the effects of both the spatial discretisation (since  $Z = \mathcal{F}(k^*)$ ) and the temporal discretisation (in the form of the equation for  $Z$ ).

$Z$  can be written as  $Z = |Z|e^{-i\psi}$  where  $|Z| = \sqrt{\Re(Z)^2 + \Im(Z)^2}$  and  $\psi = -\arctan\left(\frac{\Im(Z)}{\Re(Z)}\right)$  [8], giving

$$\phi(x, t_0 + n\Delta t) = \frac{1}{2\pi} \int \hat{\phi}(k, 0) |Z(k)|^n e^{i(kx - n\psi)} dk. \quad (5)$$

after  $n$  timesteps.

Now if we compare this to the exact solution:

$$\phi(x, t) = \frac{1}{2\pi} \int \hat{\phi}(k, 0) e^{i(kx - \omega t)} dk, \quad (6)$$

we can see that  $n\psi = \omega^* t = ku^* n\Delta t$ , where  $\omega^*$  denotes the numerical result. The dispersion relation of the full-discretisation is then

$$\omega_{\text{full}}^* = u^* k = \frac{\psi}{\Delta t}. \quad (7)$$

From the dispersion relation we can find the numerical group velocity [1, 7]:

$$u_g^* = \frac{\partial \omega_{\text{full}}^*}{\partial k} = \frac{1}{\Delta t} \frac{\partial \psi}{\partial k}. \quad (8)$$

In contrast, the group velocity for our model equation, defined as  $u_g = \frac{\partial \omega}{\partial k} = \frac{\partial uk}{\partial k}$ , is equal to  $u$ .

If we were to advance the semi-discretisation exactly through time (i.e. not taking into account the integration method), the resulting dispersion relation is  $\omega_{\text{semi}}^* = uk^*$ .

What, then, is the effect of the combined temporal and spatial discretisations ( $u^*k$ ) over the spatial discretisation alone ( $uk^*$ )? We'll start by evaluating what happens to the dispersion relation of the full-discretisation,  $\omega_{\text{full}}^* = \frac{\psi}{\Delta t}$  as  $\Delta t$  approaches zero, i.e. as the error associated with the temporal discretisation reduces. From [8, p.59] we can establish that  $u^*k$  approaches  $uk^*$  as  $\Delta t \rightarrow 0$ . Therefore, the spatial discretisation provides the effective limit for the accuracy of the full-discretisation.

Based, as it is, on the dispersion relation, the group velocity of the full-discretisation also approaches its respective semi-discretised value as  $\Delta t \rightarrow 0$ . However, if the timestep is not sufficiently small, then the full-discretisation will display phenomena that is not predicted under semi-discretisation. These include the possibility of (additional) dissipation or amplification of the advected waves, and numerical group velocities that differ not only from that of the underlying equation, but also that of the spatial semi-discretisation.

### Phase and Amplitude Errors

Figure 2 shows the ratio of the numerical group velocity to the exact group velocity,  $u_g^*/u$ , of the advection equation when discretised by the RK4 scheme in conjunction with the sixth-order central difference scheme (CDS6). The exact ratio is a constant value of one across the wavenumber range, as was stated above. The semi-discretisation represents the theoretical limit for group velocity as  $\Delta t \rightarrow 0$ . The full-discretisation results for several different CFL numbers<sup>1</sup> are also shown. It can be seen

<sup>1</sup>The Courant–Friedrichs–Levy (CFL) number,  $R = \frac{u\Delta t}{\Delta x}$ , is used here as a proxy for the timestep.

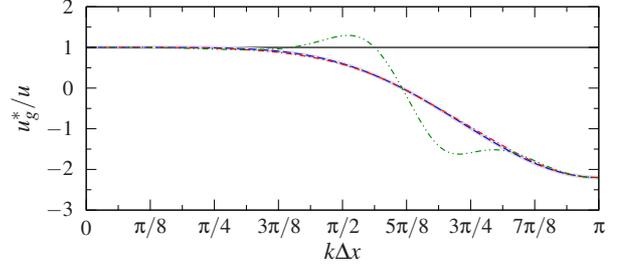


Figure 2: Group velocities of the CDS6 spatial discretisation combined with the RK4 temporal discretisation. CFL number varies: — exact; - - -  $R = 0.1$ ; ····  $R = 1.0$ ; - · - ·  $R = 1.5$ ; ■ semi-discretisation.

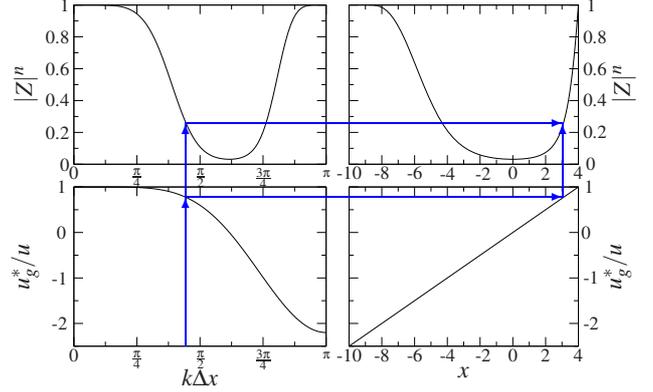


Figure 3: Illustrating the combined dissipative and dispersive effects. RK4; CDS6;  $R = 0.4R_{\text{max}}$ .

that as the CFL reduces, the full-discretisation results approach the semi-discretisation, to the point of being indistinguishable at a CFL of 0.1, confirming the earlier result.

The general notion that central schemes do not suffer from dissipation was shown earlier to not apply to full-discretisation. This phenomenon is further illustrated and combined with the effect of group velocity in the following section.

Figure 3 illustrates how one can relate both the dispersion error (in this case represented by numerical group velocity) and dissipative error (the magnitude of the numerical amplification factor) to a numerical example by combining the two effects into one figure. The plot on the lower left of figure 3 shows group velocity versus non-dimensional wavenumber. The figure on the upper left, shows the magnitude of the numerical amplification factor to the power of the number of timesteps, i.e.  $|Z|^n$ . This plot, therefore, represents the actual magnitude of the wave after  $n$  timesteps. The plot on the lower right relates the velocity to position, i.e. it is a linear relationship, the slope of which depends on the simulation time. Finally, the plot on the upper right combines both the dissipative effects (via  $|Z|^n$ ) and the dispersive effects (via  $u_g^*$  and in turn  $x$ ) on the one plot. It effectively shows where, and with what amplitude, waves of various wavenumbers will end up. The arrows overlaid on figure 3 show the relationship for one particular wavenumber of  $k\Delta x = 7\pi/16$ .

We can see the effects of the wavenumber dependent dissipation of the temporal discretisation by looking at figure 4. This figure presents the results in the same way as the upper right plot in figure 3. Here, the initial waves were all Gaussian modulated waves containing only a single wavenumber, as indicated in the legend. The solid black line represents the magnitude of the numerical amplitude over the range of resolvable wavenumbers. One can clearly see the waves at the lower and upper ends of the wavenumber spectrum remain close to an amplitude of one, whilst the middle band of wavenumbers are highly damped. And in all cases the magnitude of the single frequency

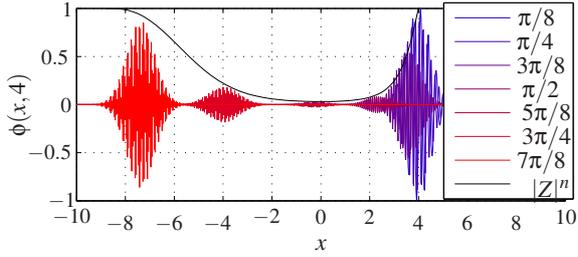


Figure 4: Illustrating the propagation and attenuation of wavenumbers from  $k\Delta x = \pi/8$  through to  $\pi$ . Overlaid on the plot is  $|Z|^n$ . The case shown is the CDS6 scheme with RK4 at a CFL of  $0.4R_{\max}$ , with advection velocity  $u = 1$ , at  $t = 4.0$ .

waves matches the theoretical predictions as represented by the  $|Z|^n$  line.

It is often suggested that it is sufficient for a numerical scheme to simply ensure stability, i.e. maintain  $R \leq R_{\max}$  or, indeed, use an implicit scheme to avoid the need for limits on CFL number at all. This next section will investigate these claims.

Figure 5 shows the numerical amplification factor and group velocity ratio for the CDS6 scheme using the RK4 and CN schemes at a CFL of 1.7 (which is just below the maximum CFL number for the CDS6/RK4 schemes combined). Firstly, let us consider the RK4 curves. It is evident that the group velocity is significantly different to the semi-discretisations. However, these discrepancies between the semi- and full-discretisations occur at wavenumbers where the semi-discretisation is already in error. Also, somewhat conveniently, the wavenumber range over which the discrepancies occur coincide with damping as shown in figure 5a, thereby probably limiting any ‘damage’ from the erroneous group velocities.

Of course, since the CN scheme is unconditionally stable, there is no such thing as  $R_{\max}$ . In practice, to take advantage of the extra stability and to make up for the additional computational time required of implicit schemes, the CFL numbers used in conjunction with such schemes would be larger than the equivalent explicit schemes. However, for illustrative purposes we have taken the same CFL numbers as used for the RK4 scheme. As expected, the numerical amplitude of the CN scheme is unaffected by the CFL number, as shown by the red dashed line in figure 5a. However, it is apparent in figure 5b, that, even at this relatively moderate CFL number, the velocity differs significantly from the semi-discretisation, even at quite small wavenumbers. More significant is the fact that they differ over the portion of the wavenumber space for which the semi-discretisations match the exact result. And, unlike the RK4 case, there is no numerical dissipation introduced by the CN scheme to damp out the erroneous group velocities.

To confirm the theoretical results for the Crank–Nicolson scheme observed in figure 5, the results of a numerical experiment (similar to that presented in figure 4) are shown in figure 6. The simulation is run for 4 time units and has a velocity of one, meaning the exact solution should be positioned at  $x = 4$ . The plot in figure 6a shows the result for a CFL number of 0.4 (consistent with the CFL number used to generate figure 4). The most obvious difference when comparing these two figures (i.e. figure 4 and figure 6) is that the amplitude is maintained for the Crank–Nicolson scheme. Upon closer inspection one can also see that the position of the longer wavelengths are not as close to the exact solution for the Crank–Nicolson scheme. This effect is exacerbated in the case with a CFL number of 1.7 shown in 6b. The effect of the increase in CFL number is to bring the group velocity of all the wavenumbers closer to zero, such that the overall spread of waves is narrowed. One can correlate the positions of the waves in figure 6b with the group velocity ratio

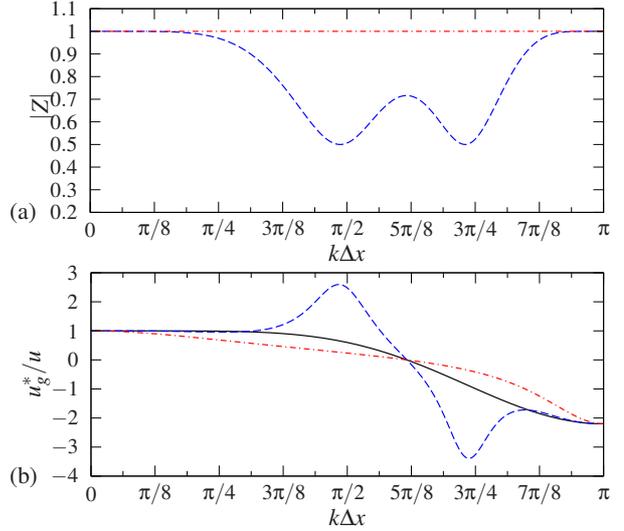


Figure 5: (a) Magnitude of the numerical amplification factor, and (b) group velocity ratio for the CDS6 scheme: --- RK4; --- CN;  $R = 1.7$ .

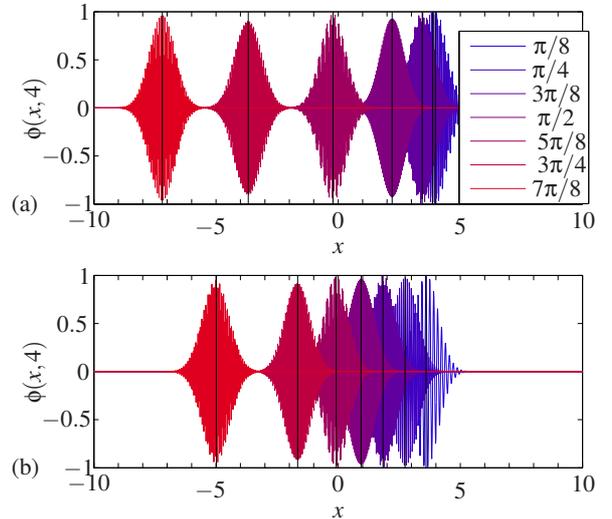


Figure 6: Propagation of waves under the advection equation using CDS6/CN schemes; (a)  $R = 0.4$ ; (b)  $R = 1.7$ .

shown in figure 5b, as shown by vertical black lines that indicate the location of the waves according to the theory.

### Examples

The linearised Euler equations in one dimension can be written as:

$$\frac{\partial U}{\partial t} + \frac{\partial E}{\partial x} = 0, \text{ where } U = \begin{bmatrix} \bar{\rho} \\ \bar{u} \\ \bar{p} \end{bmatrix}, \quad E = \begin{bmatrix} \bar{u}\bar{\rho} + \bar{\rho}\bar{u} \\ \bar{u}\bar{u} + \bar{p}/\bar{\rho} \\ \bar{u}\bar{p} + \bar{\rho}a^2\bar{u} \end{bmatrix}, \quad (9)$$

and  $\bar{\rho}$ ,  $\bar{u}$ , and  $\bar{p}$  are the perturbations of density,  $x$ -velocity, and pressure respectively,  $\bar{u}$  is the mean  $x$ -velocity,  $a$  is the mean speed of sound, and  $\bar{\rho}$  is the mean density.

The resulting dispersion relations can be written [5] as

$$\omega = k_x \bar{u} + \alpha a k_x \quad (10)$$

where  $\alpha$  can take the value of 0 or  $\pm 1$ . The semi-discretised numerical equivalent of equation (10) is

$$\omega_{\text{semi}}^* = k_x^* \bar{u} + \alpha a k_x^*. \quad (11)$$

We can now determine the effect of the full-discretisation of the linearised Euler equations by applying the principles outlined

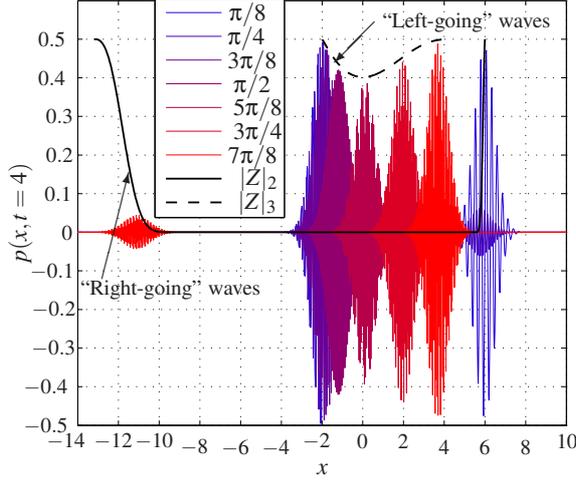


Figure 7: Illustrating wave propagation under the linearised Euler equations using CDS6/RK4 schemes;  $R = 0.75$ ;  $\bar{M}_x = 0.5$ ; and  $t = 4.0$ .

above to the dispersion relation defined by equation (11). The first step is to find the numerical amplification factor. Using the RK4 as an example,

$$\phi(x, t + \Delta t) = \left[ 1 + A + \frac{A^2}{2} + \frac{A^3}{6} + \frac{A^4}{24} \right] \phi(x, t) \quad (12)$$

where the numerical amplification factor,  $Z$  is given by the terms in the square brackets, and in which  $A = -i\Delta t \omega_{\text{semi}}^* = -iR_a (\bar{M}_x + \alpha) k_x^* \Delta x$ , where  $R_a = \Delta t a / \Delta x$  is the CFL number defined with  $a$  as the velocity term, and  $\bar{M}_x = \bar{u}/a$  is the mean Mach number.

Using the above equations, we can test the theoretical results with numerical experiments. The case presented in figure 7 (analogous to figure 4) shows the result of Gaussian modulated single frequency initial waves of the form  $p(x, 0) = e^{-2x^2} \sin(k\Delta x)$  (where  $k\Delta x$  is as indicated in the figure) after propagating for 4 seconds according to the linearised Euler equations, with  $\bar{M}_x = 0.5$ .

The exact solution of the one-dimensional linearised Euler equations have, provided the mean velocity is less than the speed of sound, acoustic (pressure) waves that propagate in both directions (corresponding to the second and third dispersion relations,  $\omega_{2,3}$ ). In the discussion that follows, right-going waves ( $\omega_2$ ) refer to the set of waves that *should* be moving right, regardless of their actual movement in the numerical solution, and vice versa.

The speed of sound is taken to be 1; therefore, the exact solution for this example has the right going wave propagating at a velocity of  $\bar{M}_x + 1$ , which results in the final position for the wave being  $x = 6$ . The left going wave propagates at  $1 - \bar{M}_x$ , giving a final position of  $x = -2$ . And the exact amplitude is 0.5, i.e. half the initial amplitude. Refer to [6] for more details of the exact solution of the linearised Euler equations.

Before commenting on the results, figure 7 requires some explanation. Blue waves indicate the well resolved end of the wavenumber spectrum, moving through purple to red for the poorly resolved waves. There is a blue wave located at  $x = 6$ , with amplitude of 0.5. This is the right-going component of the well resolved wavenumber of  $\pi/8$ . Its left-going counterpart is located at  $x = -2$ , sitting under additional higher wavenumber (purple) waves. As one would expect for the well resolved waves, these have both propagated at the correct velocity and at the correct amplitude.

At the other end of the wavenumber spectrum, the red waves

have behaved completely differently. As we saw in earlier examples, the group velocity can become negative for under resolved waves. For example, the ‘right’-going wave is located at approximately  $x = -11$ , and it also has a significantly reduced amplitude. The high wavenumber solution for the ‘left’-going wave can be seen at approximately  $x = 4$ . The remainder of the left going waves can be seen spread between  $x = -2$  and  $x = 4$ . The amplitudes of all these waves match the dashed black line representing  $|Z|_3$  (and annotated as “Left-going” waves). This line represents the theoretical magnitude of the numerical amplification factor, i.e. the magnitude of  $Z$  as defined by the terms in the square brackets in equation (12). It can be seen that, given the full-discretisation using the RK4 scheme, the theoretical prediction matches the numerical experiment, both in group velocity (i.e. the spread of waves) and their amplitudes.

## Conclusion

It was shown that the spatial semi-discretisation does not encapsulate the ‘full story’ when it comes to assessing the accuracy and behaviour of a numerical scheme. It was confirmed that the effects introduced by temporal discretisation can be minimised by ensuring the timestep, or CFL number, is sufficiently small. However, it was also shown that when this is not the case, it can adversely impact the numerical solution, often in unexpected ways. For example, central differencing schemes, for which the spatial discretisation does not incur any dissipation, can have significant dissipation when temporally under-resolved.

Implicit temporal methods are immune from stability issues and therefore have no restriction on CFL number. However, whilst implicit methods address the issue of stability, it was shown that they can both have a significant adverse impact on the accuracy of the numerical solution, even when only pursuing moderate increases in CFL number.

In order to show the applicability of the methods beyond the advection equation, the same methods were applied to the linearised Euler equations. Again it was shown that the numerical experiments confirmed the results of the theoretical framework.

## Acknowledgements

M.E. Young would like to acknowledge DSTO for supporting their candidature for a Masters of Engineering Science.

## References

- [1] Durrant, D., *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*, Springer, New York, 1999.
- [2] Lele, S., Compact finite difference schemes with spectral-like resolution, *J. Comput. Phys.*, **103**, 1992, 16–42.
- [3] Li, Y., Wavenumber-extended high order upwind biased finite difference schemes for convective scalar transport, *J. Comput. Phys.*, **133**, 1997, 235–255.
- [4] Moin, P., *Fundamentals of Engineering Numerical Analysis*, Cambridge University Press, Cambridge, 2001.
- [5] Stegeman, P., Young, M., Soria, J. and Ooi, A., Analysis of the anisotropy of group velocity error due to spatial finite difference schemes from the solution of the 2D linear Euler equations, *Int. J. Numer. Meth. Fluids*, **71**, 2013, 805–829.
- [6] Tam, C. and Webb, J., Dispersion relation preserving finite difference schemes for computational acoustics, *J. Comput. Phys.*, **107**, 1993, 262–281.
- [7] Trefethen, L., Group velocity in finite difference schemes, *SIAM Review*, **24**, 1982, 113–136.
- [8] Vichnevetsky, R. and Bowles, J., *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*, SIAM, Philadelphia, 1982.