Pseudo-Spectral Methods for Linear Advection and Dispersive Problems

Riccardo Fazio and Salvatore Iacono*

Abstract—**The most important feature of numerical methods based on a spectral decomposition is the best convergence rate (even infinite for infinitely regular functions) with respect to all other methods used in dealing with the solution to most of the differential equations. However this is true under the mandatory condition that at each time step of the evolving numerical solution no discontinuity occurs, either advected from the initial condition or self-generated (shock wave) by the non-linearity of the problem. In the first part of this paper we will point out that also by using any second or higher order pseudo-spectral method applied to the linear advection equation, one can experience the appearance in the numerical solution of the celebrated Gibbs phenomenon, located at the discontinuity points of the initial condition. As matter of fact in order to avoid such a drawback, the only applicable methods are the first order ones. In particular for stability reasons, we chose to consider the implicit Euler method and experienced that, even the Richardson's extrapolation is not able in increasing the accuracy without falling into the same problem. As a partial remedy, we propose a novel way to improve the accuracy of the implicit Euler first order pseudo-spectral method by reducing the coefficient of its truncation error leading term via a time one-step extrapolation-like technique. The second part of the paper is instead devoted to show how for dispersive differential problems the use of pseudo-spectral methods represents a very powerful numerical approach in finding out the notorious solitons dynamics. In particular we will deal with the celebrated KdV equation in 1D and its generalization in 2D.**

Keywords: pseudo-spectral methods, advection equation, dispersive models.

*Department of Mathematics, University of Messina, Contrada Papardo, Salita Sperone 31, 98166 Messina, Italy. Email: rfazio@dipmat.unime.it iacono@dipmat.unime.it **Acknowledgement.** This work was supported by the University of Messina and partially by the Italian MUR. Date of the manuscript submission: December 3, 2008.

1 Introduction

Spectral methods are considered to be a valid or at least equivalent alternative to other numerical approaches in working out the solution to a partial differential equation. A very broad and deep treatment of spectral methods is done in various monographies and other books like the ones by Gottlieb and Orszag [9], Vichnevetsky and Bowles [11], Canuto et al. [1], Fonberg [7], Gottlieb and Hestaven [8] beyond the plenty of references quoted therein. For instance, as long as problems in acoustics or optics are concerned, in his book [10, pag.7] LeVeque states that the primary computational difficulty arises from the fact that the domain of interest is many orders of magnitude larger than the wavelength of interest and as a consequence methods with higher order of accuracy are typically used, for example, fourth order finite difference methods or spectral methods.

Dealing with an analytic function, it is proved that, by means of a spectral decomposition, it is possible to achieve a uniform convergence, exponentially increasing with the number of harmonics taken. On the contrary, if we consider functions that are piecewise smooth, even the point-wise convergence is lost and the celebrated Gibbs phenomenon arises consisting in a few overshoots located at function discontinuities. As in many practical applications the solution has to be strictly included within a predefined interval, it is obvious that the appearance of the Gibbs phenomenon might make the numerical solution lacking of physical meaning.

Our attention in this paper is devoted to numerically solving both the linear advection equation, with a constant advection velocity field, and the Korteweg-de Vries (KdV) equation. Moreover, for these models we prescribe an initial condition and consider only periodic boundary conditions, so that it is straightforward to apply the Fourier decomposition approach.

Preliminary versions of this study were presented at the SIMAI 2006 [4] and WASCOM 2007 [3] conferences.

2 Definitions and notations

Before studying the problems that will be dealt with in the subsequent sections, we need to define both the domain and the notation that will be employed.

As matter of fact, we will consider the finite space domain $\Omega = [0, L_x] \times [0, L_y] \times [0, L_z]$ which, for numerical reason is to be discretized with spacings $\Delta x = L_x/J$, $\Delta y = L_y/K$ and $\Delta z = L_z/H$, in the x, y and z direction, respectively. Here, J, K and H are integers. On the grid points $(x_j, y_k, z_h) = (j\Delta x, k\Delta y, h\Delta z)$ of the domain Ω with $j \in \{0, 1, \ldots, J-1\}$, $k \in \{0, 1, \ldots, K-1\}$ and $h \in \{0, 1, \ldots, H-1\}$, the solution $u(x_j, y_k, z_h, t)$ is approximated by $u_{j,k,h}(t)$. We denote the corresponding spectral variables by $\xi_p = 2\pi p/L_x$, $\eta_q = 2\pi q/L_y$ and $\zeta_r = 2\pi r/L_z$ with $p \in \{-J/2, \ldots, -1, 0, 1, \ldots, J/2\}$, $q \in \{-K/2, \ldots, -1, 0, 1, \ldots, K/2\}$ and $r \in \{-H/2, \ldots, -1, 0, 1, \ldots, H/2\}$. The discrete Fourier transform (DFT) is given by

$$
\begin{aligned}
v_{p,q,r} &= \mathcal{F} u_{j,k,h} = \\
&= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} \sum_{h=0}^{H-1} u_{j,k,h} e^{-i(\xi_p x_j + \eta_q y_k + \zeta_r z_h)} \ , \\
p &= -\frac{J}{2}, \ldots, -1, 0, 1, \ldots, \frac{J}{2} - 1 \ , \\
q &= -\frac{K}{2}, \ldots, -1, 0, 1, \ldots, \frac{K}{2} - 1 \ , \\
r &= -\frac{H}{2}, \ldots, -1, 0, 1, \ldots, \frac{H}{2} - 1 \ ,
\end{aligned}
$$

where i is the imaginary unit. The corresponding inverse DFT is defined by

$$
\begin{aligned}
& u_{j,k,h} = \mathcal{F}^{-1} v_{p,q,r} = \\
& = \frac{1}{JKH} \sum_{p=-J/2}^{J/2-1} \sum_{q=-K/2}^{K/2-1} \sum_{r=-H/2}^{H/2-1} u_{p,q,r} e^{-i(\xi_p x_j + \eta_q y_k + \zeta_r z_h)} \ ,
\end{aligned}
$$

$$
\begin{aligned}
j &= 0, 1, \ldots, J-1 \ , \\
k &= 0, 1, \ldots, K-1 \ , \\
h &= 0, 1, \ldots, H-1 \ .
\end{aligned}
$$

In practice the DFT and its inverse are usually carried out by means of the fast Fourier transform (FFT). As a consequence the indexes J, K and H must be always expressed as powers of 2. As the use of a pseudo-spectral method make use of the FFT (denoted by `fft`) and its inverse (indicated as `ifft`), for the subsequent sections, we will make use of the notation

$$
\begin{aligned}
v(t, \xi, \eta, \zeta) &= \texttt{fft}(u(t, x, y, z)) \ , \\
u(t, x, y, z) &= \texttt{ifft}(v(t, \xi, \eta, \zeta)) \ .
\end{aligned}
$$

Besides, in order to simplify our notation, in the following we explicitly indicate the dependence only on time for both the functions u and the v, whenever this can not cause any confusion.

3 Linear advection

For the sake of simplicity, let us show how a pseudo-spectral method is derived in the case of 1D advection equation

$$
\frac{\partial u}{\partial t} + D(au) = 0 \ , \qquad D = \frac{\partial}{\partial x} \ , \tag{1}
$$

where a is a constant different from zero. By integrating from t to $t + \Delta t$ we get

$$
u(t + \Delta t) - u(t) = -\int_t^{t+\Delta t} D(au) dt \ .
$$

At this point, according to which quadrature rule is used for the right hand-side of the above equation, we obtain the corresponding method that has the same accuracy of the chosen rule. In order to avoid any stability restriction to the time-step, Wineberg et al. [13] proposed to apply the trapezoid rule, but other possibilities are still available. For instance we can use, up to a first order of accuracy, the end-point rectangle rule (implicit Euler) to find out

$$
u(t + \Delta t) - u(t) = -\Delta t D(au(t + \Delta t)) \ ,
$$

that can be rewritten in the form

$$
(1 + a\Delta t D) u(t + \Delta t) = u(t) \ ,
$$

or symbolically, in the step form

$$
\begin{aligned}
u(t + \Delta t) &= R(\Delta t) u(t) \ , \\
R(\Delta t) &= \frac{1}{1 + a\Delta t D} \ ,
\end{aligned}
$$

where the step operator $R(\Delta t)$ has been defined. Note that, the inversion of $1 + a\Delta t D$ is straightforward since D is skew adjoint and its eigenvalues are imaginary.

In the case of the trapezoid rule (Crank-Nicolson), instead the step operator would change into

$$
R(\Delta t) = \frac{1 - 0.5 a \Delta t D}{1 + 0.5 a \Delta t D} \ .
$$

If, by the FFT, we pass to the time-spectral domain, the 1D advection equation becomes

$$v(t+\Delta t) = \hat{R}(\Delta t)v(t) \ ,$$

where $\hat{R}(\Delta t) = \texttt{fft}(R)(\Delta t)$.

In particular, recalling that $(i\xi)^n = \texttt{fft}(D^n)$, for the end point rectangle we get

$$\hat{R}(\Delta t) = \frac{1}{1+ia\xi\Delta t} \ ,$$

whereas for the trapezoid rule it results

$$\hat{R}(\Delta t) = \frac{1-0.5ia\xi\Delta t}{1+0.5ia\xi\Delta t} \ .$$

Finally it is worth pointing out that the resulting scheme is explicit despite we used an implicit approach. This is always true due to the fact that we are dealing with a linear equation.

3.1 Multi dimensional advection

By defining the further differential spacial operators

$$E = \frac{\partial}{\partial y} \ , \qquad F = \frac{\partial}{\partial z} \ ,$$

we can extend the advection equation(1) to its 3D version as follows

$$\frac{\partial u}{\partial t} + a_1 D(u) + a_2 E(u) + a_3 F(u) = 0 \ ,$$

where $\mathbf{a} = (a_1, a_2, a_3)$ is the velocity vector field assumed to be constant. Moreover by rewriting it in integral form on $[t, t+\Delta t]$, we get

$$\begin{aligned} u(t+\Delta t) = u(t) &- \int_t^{t+\Delta t} a_1 D(u) dt \\ &- \int_t^{t+\Delta t} a_2 E(u) dt - \int_t^{t+\Delta t} a_3 F(u) dt. \end{aligned}$$

By recalling the procedure carried out in the previous subsection through the application of the end-point rectangle quadrature rule, we get the implicit discrete numerical scheme

$$\begin{aligned} u(t+\Delta t) = u(t) &- \Delta t \left\{ a_1 D\left[u(t+\Delta t)\right] \right. \\ &\left. + a_2 E\left[u(t+\Delta t)\right] + a_3 F\left[u(t+\Delta t)\right] \right\} \ . \end{aligned}$$

By introducing the symbolic operator $R(\Delta t)$, this scheme can be rewritten more compactly in the form

$$\begin{aligned} u(t+\Delta t) &= R(\Delta t)u(t) \ , \\ R(\Delta t) &= \frac{1}{1+\Delta t(a_1 D + a_2 E + a_3 F)} \ . \end{aligned} \tag{2}$$

In time-spectral domain, equation (2) is rewritten like this

$$\begin{aligned} v(t+\Delta t) &= \hat{R}(\Delta t)v(t) \ , \\ \hat{R}(\Delta t) &= \frac{1}{1+i\Delta t(v_1\xi + v_2\eta + v_3\zeta)} \ . \end{aligned}$$

Just for the sake of completion, it is quite a simple matter to deduce that the step operators for the case of trapezoidal rule are

$$\begin{aligned} R(\Delta t) &= \frac{1-0.5\Delta t(a_1 D + a_2 E + a_3 F)}{1+0.5\Delta t(a_1 D + a_2 E + a_3 F)} \ , \\ \hat{R}(\Delta t) &= \frac{1-0.5i\Delta t(a_1\xi + a_2\eta + a_3\zeta)}{1+0.5i\Delta t(a_1\xi + a_2\eta + a_3\zeta)} \ . \end{aligned}$$

3.2 Advection test problems

As a test problem, we consider the following initial-boundary value problem

$$\begin{aligned} u_t - D(u) &= 0 \ , \qquad x \in [0, L_x] \\ u(0,t) &= u(L_x, t) \ , \end{aligned} \tag{3}$$

subjected to the prescribed periodic boundary conditions and alternatively to one out of the two following initial condition,

$$u(x,0) = e^{-(x-L_x/2)^2} \tag{4}$$

or

$$u(x,0) = \begin{cases} 0, & 0 < x \le L_x/2 \\ 1, & L_x/2 < x \le L_x \ . \end{cases} \tag{5}$$

These represent a Gaussian pulse or a Heaviside function. It is worth noticing that $a = -1$, so that the solution is a left traveling wave. Besides the width of the spacial domain is fixed as $L_x = 10$.

Figure 1 provides a comparison of the numerical results obtained by both the trapezoid rule and the rectangle rule for the initial condition in (4-5) and according to the numerical parameters reported in its caption. We remark that at the final time $t = 10$ the exact solution will reproduce the initial condition, so that by overlapping the initial condition to the advected numerical solution we can have a visual evaluation of the error committed in the integration interval $[0, 10]$. It can be observed that for regular functions, such as the Gaussian pulse, it is enough to use a second order method to have a very good convergence of the numerical solution to the exact one. However, it is evident that for discontinuous functions,

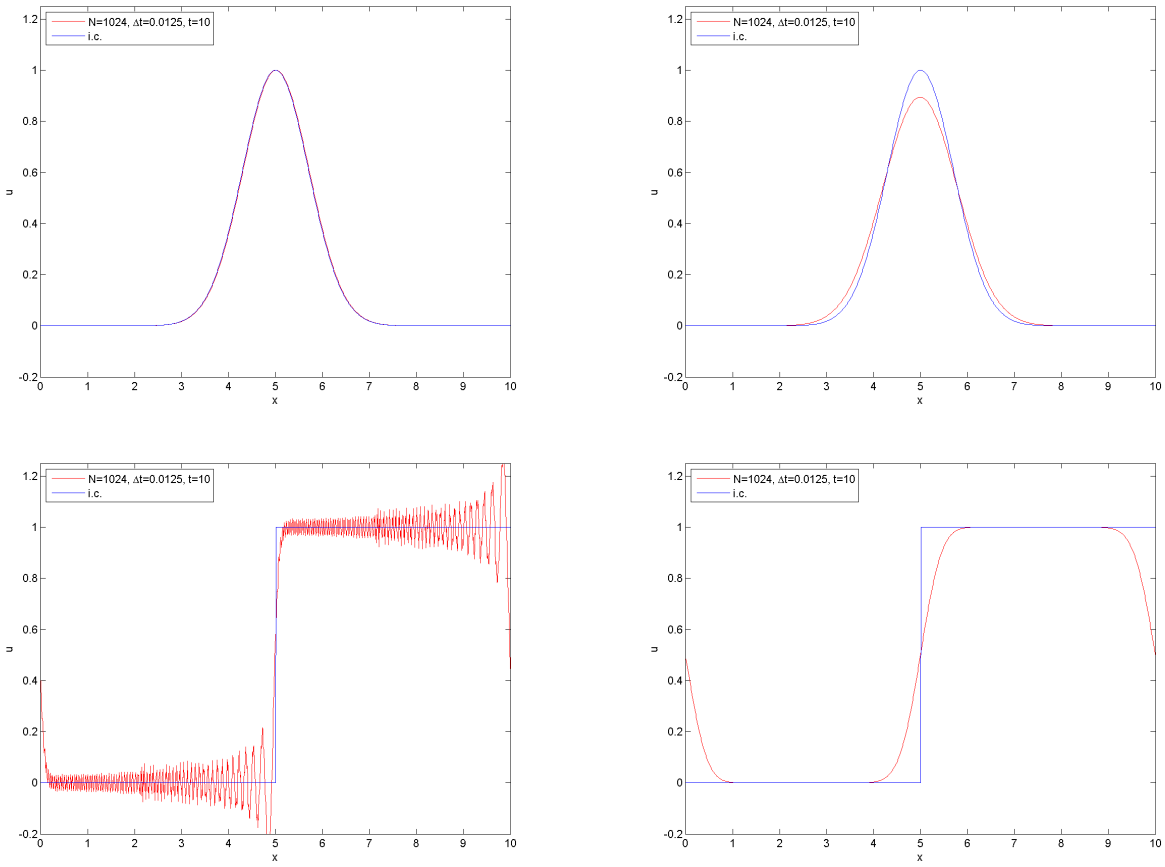

Figure 1: 1D advection equation: numerical solutions at $t = 10$ with 1024 mesh-points in the x variable and time-step $\Delta t = 0.0125$. Gaussian pulse evolution: top-left: trapezoid rule, top-right: rectangle rule. Heaviside function evolution: bottom-left: trapezoid rule, bottom-right: rectangle rule.

such as the Heaviside one, the use of methods with order greater than one results in bad degradation of the numerical solution due to the appearance of the Gibbs phenomenon. Such qualitative impression are actually confirmed also numerically by considering the discrete 2-norm of the difference between the initial condition and the advected wave as it is summarized in table 1. Just

Table 1: Values of discrete 2-norm error between initial condition and numerical solution at $t = 10$ for problems (3)-(4) and (3)-(5).

	$\|u(0) - u(10)\|_2$	
initial condition	rectangle	trapezoid
Gaussian	1.0704	0.1372
Heaviside	4.1161	2.6559

in order to improve the accuracy of the numerical solution obtained in case of dealing with discontinuous functions, in the next subsection we will propose a numerical approach based on the local error reduction through an extrapolation-like technique. The same kind of numerical test was also applied to the following 2D advection problem

$$\begin{aligned} &u_t + a_1 D(u) + a_2 E(u) = 0 \ , \\ &u(0, y, t) = u(L_x, y, t) \ , \\ &u(x, 0, t) = u(x, L_y, t) \ , \end{aligned} \tag{6}$$

with $(x, y) \in [0, L_x] \times [0, L_y]$, the prescribed periodic boundary conditions, and alternatively one out of these two initial conditions

$$u(x, y, 0) = \sin\left(\frac{2\pi}{L_x} x\right) \ \sin\left(\frac{2\pi}{L_y} y\right) \ , \tag{7}$$

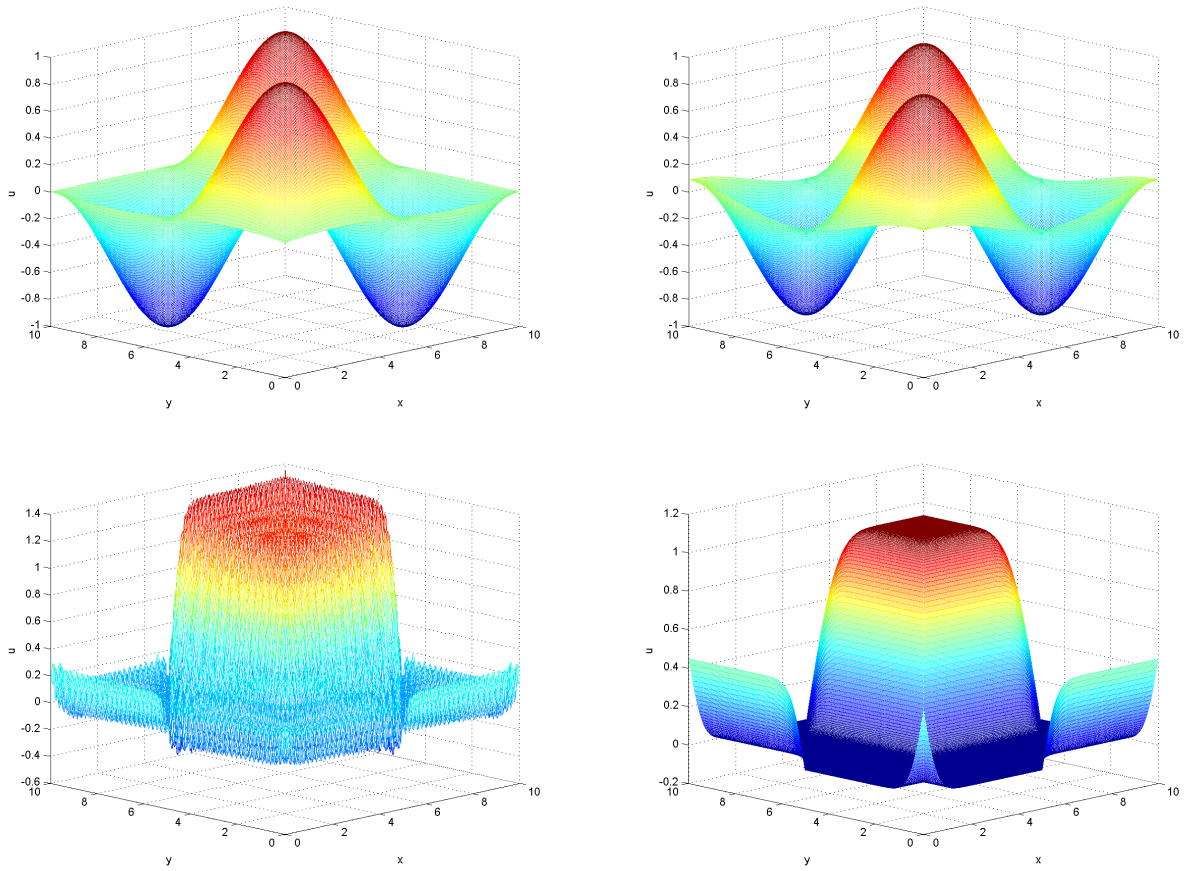

Figure 2: 2D advection equation. Sinusoidal function evolution: top-left: trapezoid rule, top-right: rectangle rule. Heaviside function evolution: bottom-left: trapezoid rule, bottom-right: rectangle rule.

or

$$u(x,y,0) = \begin{cases} 1 & \text{if } L_x/2 < x \le L_x \text{ and} \\ & \quad L_y/2 < y \le L_y \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{8}$$

For this numerical test we fixed $a_1 = 1$, $a_2 = 1$, $L_x = L_y = 10$ and an integration time $t = 10$ so that the advected numerical solution is located in correspondence of the initial condition. As long as the other integration parameters are concerned, we used 256×256 mesh-points in x and y variables and $\Delta t = 0.025$. Figure 2, display the obtained numerical results. Also in the 2D case it is possible to confirm exactly the same considerations made above for the 1D numerical test.

Finally, we remark that all of these numerical tests were carried out by means of MATLAB.

3.3 Local error reduction

The first order discrete numerical scheme at a given time provides the vector $T(\Delta t)$. Indicating with $T(0)$ the value of T obtained as the discretization parameter Δt tends to zero, we can always write

$$T(0) = T(\Delta t) + \alpha_1 \Delta t + \alpha_2 \Delta t^2 + \alpha_3 \Delta t^3 + \ldots + \alpha_m \Delta t^m + \ldots ,$$

where it is assumed that the coefficients α_i are independent from Δt. If we refer to the same scheme, but for the reduced time steps $q^k \Delta t$, with $0 < q < 1$, and the exponent $k \in \mathbb{N}$, then for each k we can also write

$$T(0) = T(q^k \Delta t) + q^k \alpha_1 \Delta t + q^{2k} \alpha_2 \Delta t^2 + q^{3k} \alpha_3 \Delta t^3 + \ldots + q^{mk} \alpha_m \Delta t^m + \ldots ,$$

By indicating with $T(k,0)$ the discrete value $T(q^k\Delta t)$, we can introduce an extrapolation scheme

$$T(k+1,n+1) = T(k,n) + \frac{T(k,n) - T(k+1,n)}{q^w - 1} \ . \quad (9)$$

As long as the parameter w is equal to the accuracy order owned by the values $T(k,0)$, such a scheme is just the one by Richardson and the value $T(k,n)$ would have actually an accuracy order increased by the integer n. Otherwise we get an extrapolation-like scheme. Let us now consider the implicit Euler method that is first order accurate, if we extrapolate properly, i.e. by using the value $w = 1$, then for any extrapolated order higher than one, the solution will go on showing the Gibbs phenomenon. On the contrary, by extrapolating once using the value $w = 2$, i.e. we are doing an extrapolation-like technique, we get discrete values that are still first order accurate, but the leading term of the truncation error becomes $q(q-1)\alpha_1\Delta t$, whereas the second term is canceled. If we want to involve the values $T(0,0)$ and a generic $T(k,0)$, then we must modify (9) by using $w = 2k$. In such a case it is easy to prove that for the resulting extrapolated value the leading term becomes $q^k(q^k-1)\alpha_1\Delta t$. As a consequence, being $0 < q < 1$, in any case we get a reduction of the local error. Finally, obtaining always a first order method, there is no point in going on extrapolating beyond the first extrapolation step.

Let us consider again the 2D advection equation with the Heaviside function as initial condition and periodic boundary conditions

$$\begin{aligned} & u_t + a_1 D(u) + a_2 E(u) = 0 \ , \quad (10) \\ & u(x,y,0) = \begin{cases} 1 & \text{if } L_x/2 < x \le L_x \text{ and} \\ & \quad L_y/2 < y \le L_y \ , \\ 0 & \text{otherwise} \ . \end{cases} \\ & u(x,0,t) = u(x,L_y,t) \ , \\ & u(0,y,t) = u(L_x,y,t) \ , \end{aligned}$$

where $(x,y) \in [0,L_x] \times [0,L_y]$. For our numerical calculation we decided to fix a square domain with $L_x = L_y = 10$ and as constant velocity vector field $\mathbf{a} = (1,1)$. As a consequence, at the final time $t = 10$, the exact solution will reproduce the initial condition. We carried out two numerical experiments, for the above 2D case and the analogous 1D case, using $q = 1/2$ and $\Delta t = 0.025$. They consisted in calculating the solution by applying one step of the extrapolation-like technique involving $T(0,0)$ and $T(k,0)$ for $1 \le k \le 5$, so that the obtained solution, $T(k,1)$, can be compared with the one obtained simply by refining the mesh by the same step reduction factor used for each $T(k,0)$. Besides, the FFT was computed involving up to 1024 harmonics for the 1D case, whereas for the 2D case it involved 128 harmonics. In both cases, there is an improvement in slope recovery at the discontinuities. This remark can be appreciated graphically by looking at figure 3, for the 1D case, and figure 4, for the 2D case. In order to prove numerically such a result, we can define err_{extra} as the difference between the initial condition and the final time solution obtained by one step of extrapolation-like technique, i.e. $err_{\text{extra}} = |T(0) - T(k,1)|$, whereas analogously err_{ref} is referred to the solution obtained by mesh refining, i.e. $err_{\text{ref}} = |T(0) - T(k,0)|$. In table 2, for 1D case, and in table 3, for the 2D case, we report the err_{extra} and err_{ref} 2-norm values for increasing values of k. From these two tables it can be seen how a real improvement is achieved. It is clear that the higher is the value taken for k the less is the numerical profit in applying the present technique and, as a consequence, the maximum profit is achieved by choosing $k = 1$.

Table 2: 1D numerical comparison for increasing k. Here $\Delta t = .025$, N=1024, and $t_{\max} = 10$.

k	$\|err_{\text{extra}}\|_2$	$\|err_{\text{ref}}\|_2$	$\|err_{\text{ref}}\|_2 - \|err_{\text{extra}}\|_2$
0	0.047818	0.047818	
1	0.036378	0.040208	0.003830
2	0.032009	0.033808	0.001799
3	0.027628	0.028425	0.000798
4	0.023558	0.023896	0.000338
5	0.019945	0.020083	0.000138

Table 3: 2D numerical comparison for increasing k. In this case $\Delta t = .025$, N=128, and $t_{\max} = 10$.

k	$\|err_{\text{extra}}\|_2$	$\|err_{\text{ref}}\|_2$	$\|err_{\text{ref}}\|_2 - \|err_{\text{extra}}\|_2$
0	0.065722	0.065722	
1	0.048061	0.053928	0.05867
2	0.041443	0.044170	0.02727
3	0.034838	0.036046	0.01208
4	0.028638	0.029163	0.00525
5	0.022861	0.023092	0.00231

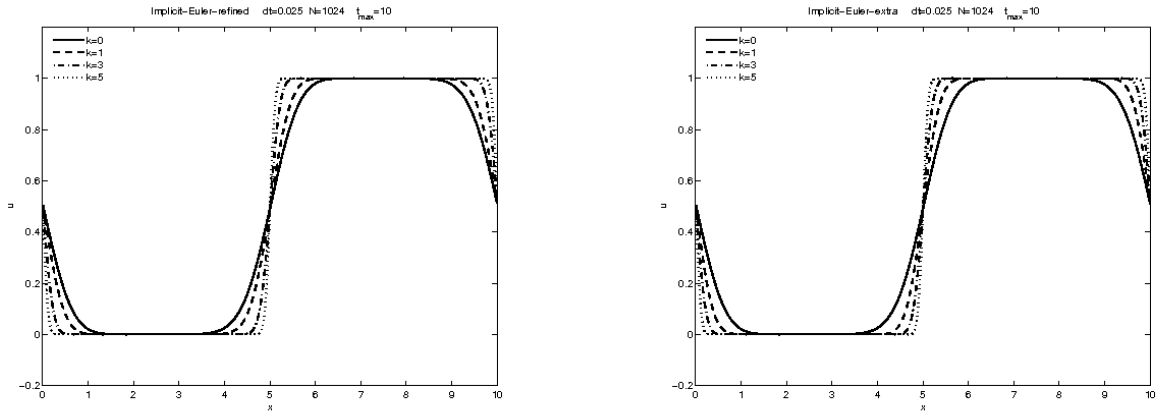

Figure 3: 1D solutions for $k = 0, 1, 3, 5$: on the left, $T(k, 0)$ obtained for mesh refining, whereas on the right, $T(k, 1)$ obtained for extrapolation-like.

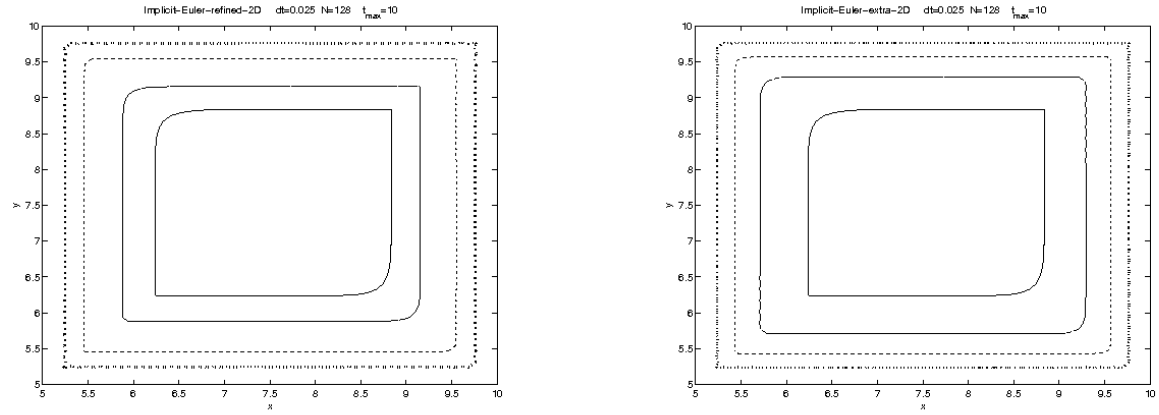

Figure 4: 2D contour plot solution at $u = 0.99$ for $k = 0, 1, 3, 5$: on the left, $T(k, 0)$ obtained for mesh refining, whereas on the right, $T(k, 1)$ obtained for extrapolation-like.

4 Dispersive problems

Pseudo-spectral methods using the trapezoid rule have been applied successfully to several problems of interest governed by nonlinear PDEs: KdV, Klein Gordon, Whitham (the equation for weak dispersion proposed in [12]), etc. As an example let us consider the KdV equation

$$\frac{\partial u}{\partial t} + D\left(\frac{u^2}{2}\right) + D^3 u = 0 \ .$$

It is a simple matter to verify that, in the trapezoid rule case, we get symbolically

$$u(t + \Delta t) = R(\Delta t)u(t) - S(\Delta t)\left(u^2(t + \Delta t) + u^2(t)\right) \ ,$$

where $R(\Delta t)$ and $S(\Delta t)$ are symbolic operators defined by

$$R(\Delta t) = \frac{1 - 0.5\Delta t D^3}{1 + 0.5\Delta t D^3} \ ,$$

and

$$S(\Delta t) = \frac{0.25\Delta t D}{1 + 0.5\Delta t D^3} \ .$$

If with the FFT, we switch to the time-spectral domain,

we have

$$v(t+\Delta t) = \hat{R}(\Delta t)v(t) - \hat{S}(\Delta t)\left(\texttt{fft}(u^2(t+\Delta t)) + \texttt{fft}(u^2(t))\right) .$$

Moreover, the introduced symbolic operators can be computed by the FFT, like it was done in the previous section. As usual, the nonlinear terms are best computed in the spatial representation, hence we transform back to the original space, make the multiplication, which is point-wise in x, and transform again. We have here an implicit method. For the solution of the nonlinear system, it is possible to apply the Newton method, but it requires the inversion of full matrices. As a consequence, Newton iterations result to be not suitable for spectral methods. On the other hand, nonlinear spectral methods are usually implemented by using, first order, but simpler, successive approximation. That is, we can apply the iterations

$$v_{n+1}(t+\Delta t) = w(t) - S(\Delta t)\left(\texttt{fft}(u_n^2(t+\Delta t))\right) ,$$

where

$$w(t) = R(\Delta t)v(t) - S(\Delta t)\left(\texttt{fft}(u^2(t))\right) ,$$

$$u_n(t+\Delta t) = \texttt{ifft}(v_n(t+\Delta t)) , \qquad u_1^2(t+\Delta t) = u^2(t) .$$

Finally, it is worth saying that the above iteration can be triggered also by a first predictor explicit step by assuming

$$u_1^2(t+\Delta t) = \texttt{ifft}(\hat{R}(\Delta t)v(t) - \hat{S}(\Delta t)\texttt{fft}(u^2(t)))$$

4.1 The 2D KdV (or KPI) equation

Let us consider now the 2D KdV equation [5] also known as KPI equation

$$D\left[\frac{\partial u}{\partial t} + D\left(3u^2\right) + D^3 u\right] - 3E^2 u = 0 . \qquad (11)$$

By passing to the time-spectral domain, through the FFT, we can rewrite (11) as

$$\frac{dv}{dt}(t) + i\xi\texttt{fft}(3u^2(t)) - \left(i\xi^3 + i3\frac{\eta^2}{\xi}\right)v(t) = 0 .$$

Then if we integrate it in time, we get

$$v(t+\Delta t) = v(t) - \int_{t+\Delta t}^{t} i\xi\texttt{fft}(3u^2(t))dt + \int_{t+\Delta t}^{t} i\left(\xi^3 + 3\frac{\eta^2}{\xi}\right)v(t)dt ,$$

that, by applying the trapezoid rule, can be rewritten more compactly as

$$v(t+\Delta t) = \hat{R}(\Delta t)v(t) - \hat{S}(\Delta t)(\texttt{fft}(u^2(t)) + \texttt{fft}(u^2(t+\Delta t))) ,$$

where

$$\hat{R}(\Delta t) = \frac{1 + 0.5\Delta t i\left(\xi^3 + \frac{\eta^2}{\xi}\right)}{1 - 0.5\Delta t i\left(\xi^3 + \frac{\eta^2}{\xi}\right)} ,$$

and

$$\hat{S}(\Delta t) = 1.5\, i\xi\Delta t .$$

4.2 Dispersive test problems

Here we report the results of numerical experiments carried out, through a pseudo-spectral method, on two test problems; the 1D and the 2D KdV equations. As a simple test problem, we consider the classical two solitons interaction discovered by Zabusky and Kruskal in the 1960's [14]. The problem to be solved is:

$$\begin{aligned} &\frac{\partial u}{\partial t} + 6pD\left(\frac{u^2}{2}\right) + D^3 u = 0 , \qquad x \in [0, L_x] \\ &u(x,0) = u_0(x) , \qquad u(0,t) = u(L_x,t) , \end{aligned} \qquad (12)$$

with initial condition

$$u_0(x) = \frac{c_1}{2p} sech^2\left(\frac{\sqrt{c_1}}{2}(x - 0.1L_x)\right) + \frac{c_2}{2p} sech^2\left(\frac{\sqrt{c_2}}{2}(x - 0.4L_x)\right) . \qquad (13)$$

In our case we fix $L_x = 100$, $p = 1$, with $c_1 = 1.5$ and $c_2 = 0.5$. Besides, the value of the parameter p can be chosen freely in order to cover all of the different versions of (12) existing in literature. A MATLAB code was used to implement the second order method defined above and to produce the numerical results reported in figure 5. By looking at frames in figure 5, although they keep maintaining their shape after merging, one can appreciate the non-linear feature of the phenomenon because when they are totally merged the amplitude does not correspond to the sum of the two of them.

As long as 2D dispersive problems are concerned, we carry out two numerical tests on the following problem

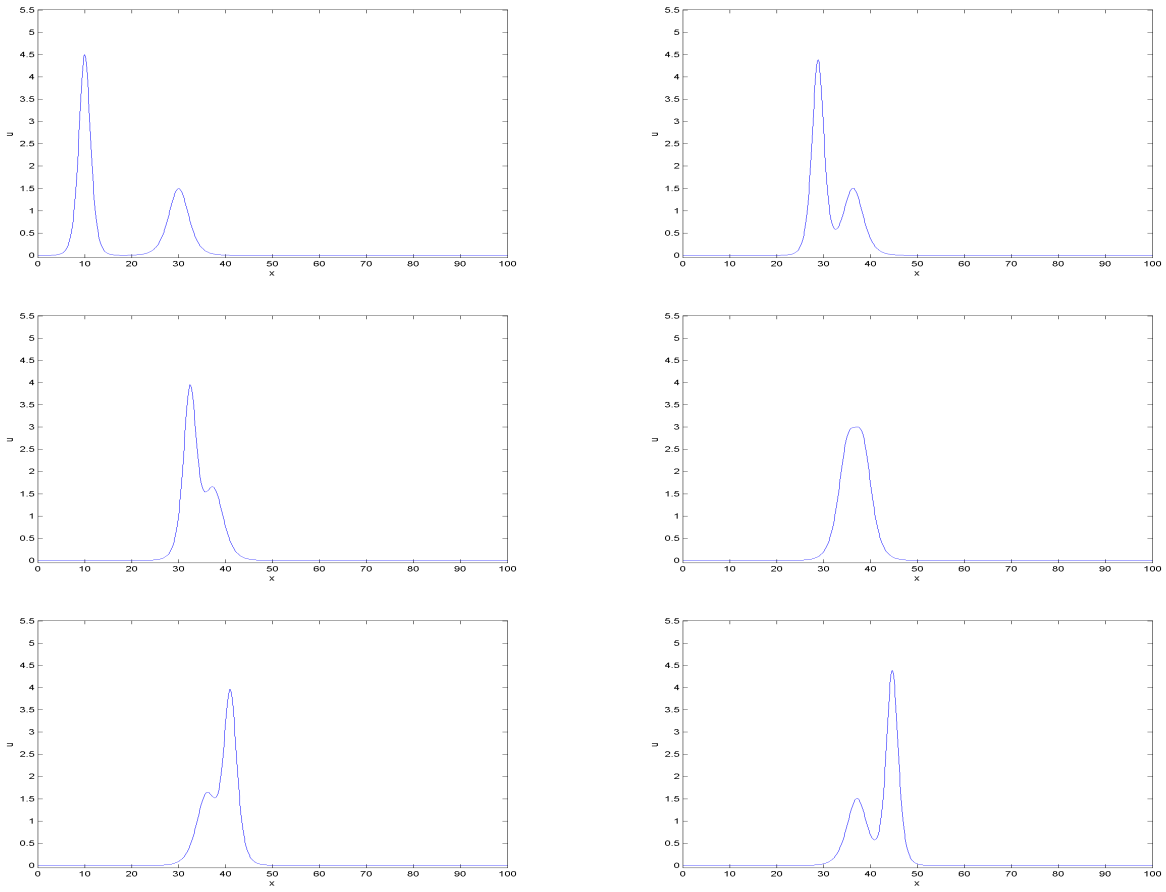

Figure 5: Interaction of two solitons for the KdV equation. Numerical solutions with 1024 mesh-points in the x variable and $\Delta t = 0.05$. Top-left: $t = 0.0$, top-right: $t = 12.55$, center-left: $t = 15.0$, center-right: $t = 17.1$, bottom-left: $t = 19.2$, and bottom-right: $t = 21.65$.

already treated by Feng et al. in [6] and in [5]

$$D\left[\frac{\partial u}{\partial t} + D\left(3u^2\right) + D^3 u\right] - 3E^2 u = 0 \ , \quad (14)$$

$$u(x,y,0) = 4\sum_{i=1}^{2} \frac{-b_i(x,y) + d_i(y)}{\left[b_i(x,y) + d_i(y)\right]^2} \ ,$$

$$u(x,0,t) = u(x,L_y,t) \ ,$$

$$u(0,y,t) = u(L_x,y,t) \ .$$

where $(x,y) \in [0,L_x] \times [0,L_y]$, $b_i(x,y) = [x - x_{0,i} + \lambda_i (y - y_{0,i})]^2$ and $d_i(y) = \mu_i^2 (y - y_{0,i})^2 + 1/\mu_i^2$. The initial condition consists of two lump-type solitons located in $x_{0,i}$ and $y_{0,i}$, whereas the parameters λ_i and μ_i determine the velocity vector field as v_i=$(3(\lambda_i^2 + \mu_i^2), -6\lambda_i)$, where $i = 1, 2$, according to [5].

As first test, we fix

$$x_{0,1} = 10 \ , \quad x_{0,2} = 18 \ , \quad y_{0,1} = y_{0,2} = 20 \ ,$$
$$L_x = L_y = 40 \ ,$$
$$\mu_1^2 = 1.5 \ , \quad \mu_2^2 = 0.75 \ , \quad \lambda_1 = \lambda_2 = 0 \ ,$$

and the numerical parameters are 256×256 mesh-points in the x and y variables and $\Delta t = 0.004$. Figure 6 shows the interactions of two lump-type solitons initially traveling along the same line at six discrete subsequent times. By looking at figure 6, we can repeat the same consideration, already said in the 1D case, concerning the amplitude of the merged lump-type solitons: at the complete interaction the amplitude is smaller than the sum of the two initial amplitudes. Moreover the non-linear feature of the phenomenon can be also appreciated from the fact that also the velocity of the two lump-type solitons is affected by their merging. Indeed it is evident that, after

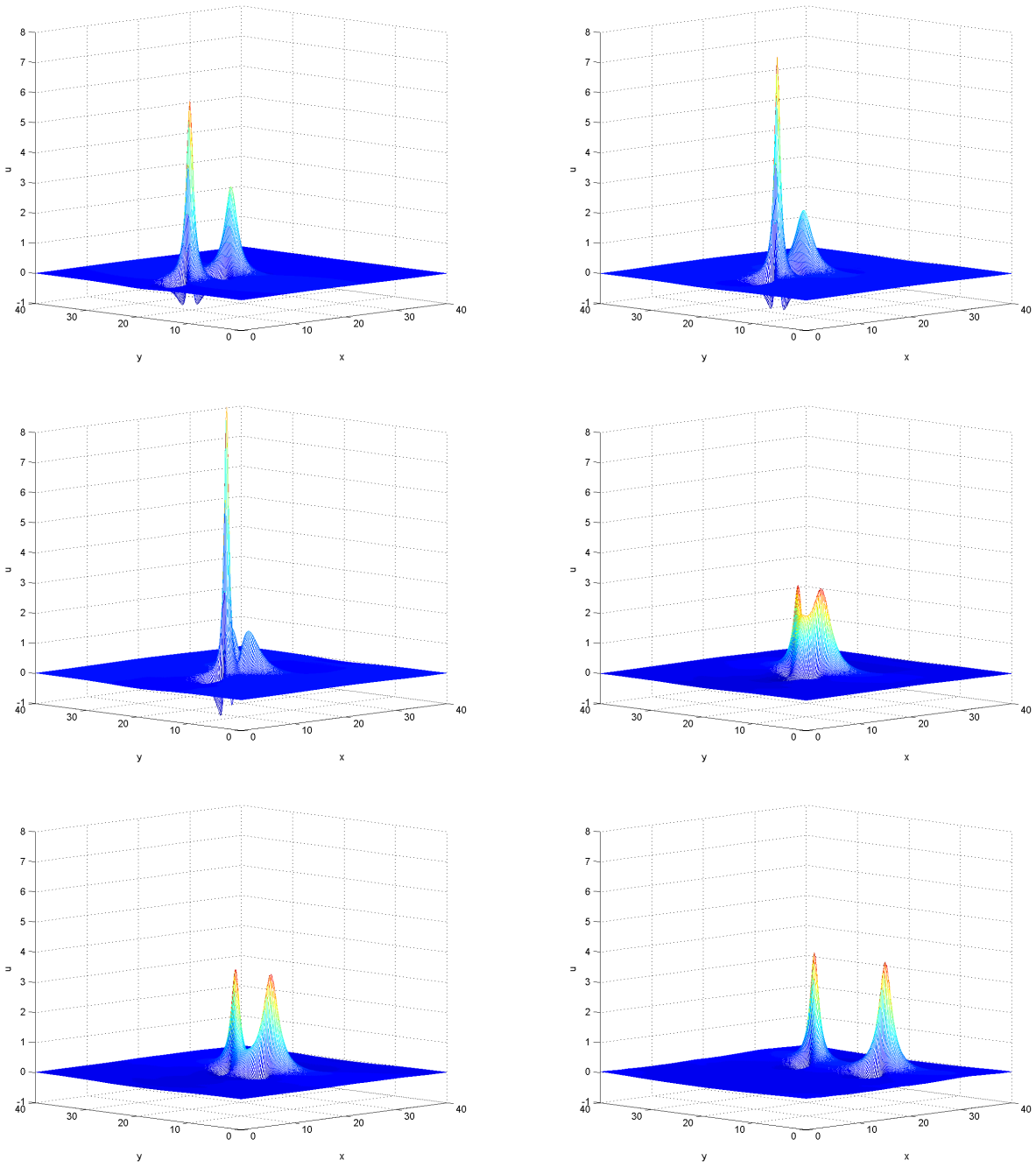

Figure 6: Interaction of two lump-type solitons marching along the same line for the KPI equation. Numerical solutions with 256×256 mesh-points in the x and y variable and $\Delta t = 0.004$. Top-left: $t = 0.0$, top-right: $t = 1.0$, center-left: $t = 1.5$, center-right: $t = 2.0$, bottom-left: $t = 2.5$, and bottom-right: $t = 4.5$.

merging, their velocity vectors divert, despite they were traveling along the same direction. In other words we have here a soliton like behavior. However, these two lump solitons undergo an inelastic collision.

As second test, we choose

$$x_{0,1} = 10\ , \quad x_{0,2} = 10\ , \quad y_{0,1} = 10\ , \quad y_{0,2} = 30\ ,$$
$$L_x = L_y = 40\ ,$$
$$\mu_1^2 = 1\ , \quad \mu_2^2 = 1\ , \quad \lambda_1 = -1\ , \quad \lambda_2 = 1\ ,$$

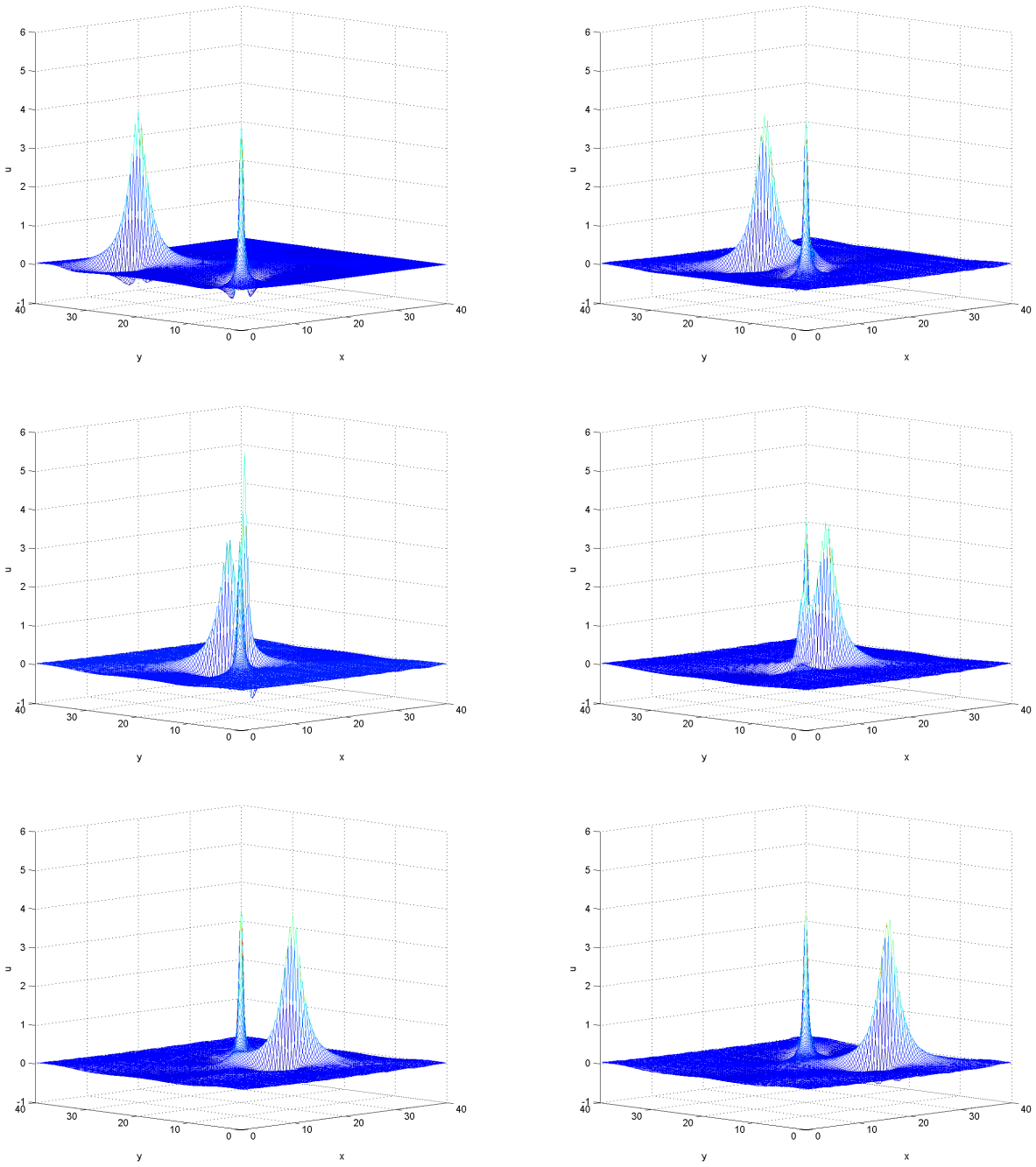

Figure 7: Interaction of two lump-type solitons traveling along orthogonal lines for the KPI equation. Numerical solutions with 128×128 mesh-points in the x and y variable and $\Delta t = 0.01$. Top-left: $t = 0.0$, top-right: $t = 1.0$, center-left: $t = 1.5$, center-right: $t = 2.0$, bottom-left: $t = 2.5$, and bottom-right: $t = 3.0$.

and in this case the numerical parameters used are 128×128 mesh-points in the x and y variables and $\Delta t = 0.01$. Figure 7 displays the interactions of two lump-type solitons traveling along orthogonal directions. In this case, the behavior of the lump-type solitons is closer to 1D case as their velocity vectors remains unchanged by their merging. In some way these two lump solitons undergo an elastic collision.

5 Concluding remarks

In this paper we have considered pseudo-spectral methods for two classes of problems of relevant interest: namely, linear advection and nonlinear dispersive problems. As far as linear advection problems are concerned, we have verified that, when dealing with discontinuous functions, the Gibbs phenomenon can be avoided by using a pseudo-spectral method of order one (in time). Moreover, even classical extrapolation techniques are not suitable to overcome the Gibbs phenomenon, that is actually a numerical artifact. However, we were able to assert the usefulness of an extrapolation like approach in order to reduce the local truncation error. It is clear that the proposed extrapolation-like error reduction technique can be easily extended to $3D$ problems.

On the other hand, we have seen that the use of our second order (in time) pseudo-spectral method for nonlinear dispersive problems is a very efficient numerical resource in order to work out and classify the different kinds of solutions belonging to such a kind of problems. In particular, we can appreciate how our numerical scheme can follow the stable propagation of the lump-type solitons without any deformation.

The part of this work dealing with the advection equation was motivated by a preliminary study, by the first author [2], concerning the implicit Euler and the second order Adams-Moulton methods in the ordinary differential context.

References

[1] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods in Fluid Dynamics*. Springer-Verlang, New York, 1988.

[2] R. Fazio. Stiffness in numerical initial-value problems: A and L-stability of numerical methods. *Int. J. Math. Educ. Sci. Technol.*, 32:752–760, 2001.

[3] R. Fazio and S. Iacono. Local error reduction for first order implicit pseudo-spectral methods applied to linear advection models. In N. Manganaro, R. Monaco, and S. Rionero, editors, *Proceedings "WASCOM 2007" 14th Conference on Waves and Stability in Continuous Media*, pages 258–263, Singapore, 2008. World Scientific.

[4] R. Fazio and A. Jannelli. Implicit pseudo-spectral methods for dispersive and wave propagation problems. Communications to SIMAI Congress, DOI: 10.1685/CSC06078, ISSN 1827-9015, Vol. 1 (2006).

[5] B.F. Feng, T. Kawahara, and Mitsui T. A consevative spectral method for several two-dimensional nonlinear wave equations. *J. Comput. Phys.*, 153:467–487, 1999.

[6] B.F. Feng and Mitsui T. A finite difference method for the Korteweg-de Vries and the Kadomtsev-Petviashvili equations. *J. Comput. Appl. Math.*, 90:95–116, 1998.

[7] B. Fornberg. *A Practical Guide to Pseudospectral Methods*. Cambridge University Press, Cambridge, 1996.

[8] D. Gottlieb and J. S. Hesthaven. Spectral methods for hyperbolic problems. *J. Comput. Appl. Math.*, 128:83–131, 2001.

[9] G. B. Gottlieb and S. A. Orszag. *Numerical Analysis of Spectral Methods: Theory and Applications*. SIAM, 1977.

[10] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Univesity Press, Cambridge, 2002.

[11] R. Vichnevetsky and J. B. Bowles. *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*. SIAM, 1982.

[12] G. B. Whitham. *Linear and Nonlinear Waves*. John Whiley & Sons, New York, 1973.

[13] S. B. Wineberg, J. F. McGrath, E. F. Gabl, L. R. Scott, and C. E. Southwell. Implicit spectral methods for wave propagation problems. *J. Comput. Phys.*, 97:311–336, 1991.

[14] Z. J. Zabusky and M. D. Kruskal. Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 15:240–243, 1965.