

# Multisymplectic box schemes and the Korteweg–de Vries equation

Uri M. Ascher

*Department of Computer Science, UBC, Vancouver, Canada V6T 1Z4*  
ascher@cs.ubc.ca

Robert I. McLachlan

*IFS, Massey University, Palmerston North, New Zealand*  
R.McLachlan@massey.ac.nz

---

## Abstract

We develop and compare some geometric integrators for the Korteweg-de Vries equation, especially with regard to their robustness for large steps in space and time,  $\Delta x$  and  $\Delta t$ , and over long times. A standard, semi-explicit, symplectic finite difference scheme is found to be fast and robust. However, in some parameter regimes such schemes are susceptible to developing small wiggles. At the same instances the fully implicit and multisymplectic Preissmann scheme, written as a 12-point box scheme, stays smooth. This is accounted for by the ability of the box scheme to preserve the shape of the dispersion relation of any hyperbolic system for all  $\Delta x$  and  $\Delta t$ . We also develop a simplified 8-point version of this box scheme which maintains its advantageous features.

---

## 1 Introduction

The Korteweg–de Vries (KdV) equation needs no introduction. We consider it in the form

$$u_t = V'(u)_x + \nu u_{xxx}, \quad V(u) = \frac{\alpha}{3}u^3 + \frac{\rho}{2}u^2. \quad (1)$$

*MathSciNet* already lists more than 1000 articles on the subject. So, the first question we must address is, why another study of numerics for the KdV equation?! We started with several reasons:

1. There is growing interest in the geometric integration of PDEs. KdV is a model nonlinear hyperbolic equation which has smooth solutions for all time and is a natural test bed for such studies.
2. In elasticity and meteorology, fast-moving high frequency waves can destabilize otherwise sound schemes. Again, KdV is a model equation containing such waves. It is often so stiff that the fast waves are not resolved in time.
3. Many recent KdV schemes are fully implicit and thus cumbersome and expensive to work with. This is often necessary to preserve some of the solution structure, but would seem to be a waste on a non-parabolic, nonlinear problem. Explicit schemes may be useful only when  $\nu \ll 1$  in (1): a more generally applicable scheme must treat the stiff term  $\nu u_{xxx}$  implicitly. However, this term is linear and has a constant coefficient; so we wondered if there was a good semi-explicit (or IMEX [2,3]) scheme for KdV.
4. It has recently been proposed that hyperbolic systems which are “multi-symplectic” (symplectic in both space and time) should be discretized so as to preserve this structure [5,13]. This appears to endow the scheme with many more conservation laws than a merely symplectic structure. However, for simple equations such as the nonlinear wave equation  $u_{tt} - u_{xx} = V'(u)$ , standard discretizations (e.g. the 5-point central difference) are both symplectic and multisymplectic and so one can’t see any difference between the two properties. For KdV, symplectic and multisymplectic schemes are distinct. A multisymplectic scheme for KdV was derived in [19]. But it was only compared numerically to an ancient scheme which our semi-explicit scheme (Eq. (4) below) majorizes.

We will present results for a semi-explicit symplectic scheme, an implicit symplectic scheme, and an implicit multisymplectic (“box”) scheme. For the latter method, proposed by Preissmann [16] and developed mathematically by Zhao and Qin [19], we develop a simpler version. All methods are nondissipative as we are interested in the methods’ ability to give qualitatively correct results for very long times. While the semi-explicit scheme fulfilled our hopes, being very fast and robust, it did develop small wiggles over time which the fully implicit multisymplectic box scheme did not. Moreover, the implicit schemes did allow one to increase the time step significantly. We account for this in terms of the remarkable ability of the box scheme to preserve the shape of the dispersion relation of any linear equation for all mesh resolutions (see Proposition 1).

Before we start, we should note that KdV is *not* a typical nonlinear hyperbolic PDE; it is completely integrable. We believe this is not relevant to our results, but for us it does counteract KdV’s positive features enumerated above.

## 2 Symplectic and multisymplectic methods for the KdV equation

We first construct symplectic schemes based on a Hamiltonian semi-discretization. To construct a Hamiltonian semi-discretization (see, e.g., [14]) we start with KdV in its first Hamiltonian form,

$$u_t = \frac{\partial}{\partial x} \frac{\delta H}{\delta u}, \quad H = \int \left( V(u) - \frac{\nu}{2} (u_x)^2 \right) dx.$$

Consider a uniform spatial grid of  $N$  subintervals,  $N$  even, with spacing  $\Delta x$ . Central differences provide an antisymmetric discretization of  $\frac{\partial}{\partial x}$ ; call the finite difference operator  $D$ . Approximating the Hamiltonian by the sum

$$H = \sum_i \left( V(u_i) - \frac{\nu}{2\Delta x^2} (u_{i+1} - u_i)^2 \right) \Delta x$$

gives the standard conservative semi-discretization,  $u_t = D\nabla H(u)$ , or

$$u_t = \frac{1}{2\Delta x} [-1 \quad 0 \quad 1] V'(u) + \frac{\nu}{2(\Delta x)^3} [-1 \quad 2 \quad 0 \quad -2 \quad 1] u. \quad (2)$$

Note that the first derivatives in  $V'(u)$  and the third derivative term are approximated by wide stencils of 3 points and 5 points, respectively.

KdV has a conserved quantity  $\int (u + \tilde{\rho})^2 dx$ , where  $\tilde{\rho} = 3\rho/2\alpha$ . It is well-known (e.g. [18]) that if we consider for the nonlinear term in  $V'(u)_x$  the generalized discretization

$$\frac{\alpha}{2\Delta x} \{ \theta [-1 \quad 0 \quad 1] u^2 + 2(1 - \theta) u [-1 \quad 0 \quad 1] u \} \quad (3)$$

then the discretization  $\|u + \tilde{\rho}\|_2^2$  of the above norm is conserved only for  $\theta = 2/3$ . However, the discretization is Hamiltonian only for  $\theta = 1$ . In (2) we use  $\theta = 1$ , which is important especially for the construction of our first method (4) below. So,  $\|u + \tilde{\rho}\|_2^2$  is not conserved for any of our subsequent time discretizations. Studying the preservation and role of this conserved quantity would take us too far afield; see for example [6].

The Poisson matrix  $D$  has Casimirs (conserved quantities)  $\sum u_i$  and  $\sum u_{2i} - u_{2i+1}$ . Because  $D$  is sparse and independent of  $u$ , it is easy to construct symplectic integrators for the ODEs (2).

Our **first method** is a semi-explicit splitting method [15]. We let  $H = H_1 + H_2 + H_3$  where (omitting constants)  $H_1 = \sum_i V(u_{2i})$ ,  $H_2 = \sum_i V(u_{2i+1})$ ,

and  $H_3 = \sum_i (u_{i+1} - u_i)^2$ . The ODE  $u_t = X_3(u) = D\nabla H_3(u)$  has constant coefficients and can be solved exactly, or approximated by a symplectic, unconditionally stable scheme such as the midpoint (Crank-Nicolson) rule. The ODE  $u_t = X_2(u) = D\nabla H_2(u)$  has the form  $(u_{2i})_t = \frac{1}{2\Delta x}(V'(u_{2i+1}) - V'(u_{2i-1}))$ ,  $(u_{2i+1})_t = 0$ , which is solved exactly by Euler's method. Thus, a second order, semi-explicit symplectic integrator for (2) is provided by the composition

$$e^{\frac{\Delta t}{2}X_3}e^{\frac{\Delta t}{2}X_2}e^{\Delta t X_1}e^{\frac{\Delta t}{2}X_2}e^{\frac{\Delta t}{2}X_3}, \quad (4)$$

where  $e^{tX}$  denotes the time- $t$  flow of the vector field  $X$ . In detail, assume periodic boundary conditions (so  $u_0 = u_N$ ,  $u_{-1} = u_{N-1}$ ) and let  $M$  be the chosen approximation to the solution matrix for  $e^{\frac{\Delta t}{2}X_3}$ , evaluated once. We either store the full matrix  $M$  or (in case that  $N$  is too large) apply an  $LU$ -decomposition technique for a sparse representation. Define also  $odd = 1 : 2 : N - 1$  and  $even = 2 : 2 : N = odd + 1$ . Then one time step from  $u^n$  to  $u^{n+1}$  reads as follows:

$$\begin{aligned} U &= Mu^n; \\ U_{odd} &= U_{odd} + \frac{\Delta t}{2\Delta x}(V'(U_{odd+1}) - V'(U_{odd-1})); \\ U_{even} &= U_{even} + \frac{\Delta t}{2\Delta x}(V'(U_{even+1}) - V'(U_{even-1})); \\ u^{n+1} &= MU. \end{aligned}$$

Our **second method** is fully implicit. It uses the same semi-discretization (2) as our first method, but for the time integration uses the midpoint rule

$$\frac{u^{n+1} - u^n}{\Delta t} = D\nabla H \left( \frac{u^n + u^{n+1}}{2} \right), \quad (5)$$

which provides a symplectic (more correctly, Poisson) integrator for any system  $u_t = D\nabla H(u)$  when  $D$  is constant and antisymmetric. This method is basically Crank-Nicolson with conservative differences in space, except that it is based on midpoint rather than trapezoidal quadrature.

Our **third method** is multisymplectic [5,13,19]. KdV in the form (1) does not come from a variational principle. However, introducing the primitive  $\phi$  where  $\phi_x = u$ , it is equivalent to the stationarity of the Lagrangian

$$\mathcal{L} = \int \left( \frac{1}{2}\phi_t\phi_x + \frac{\nu}{2}\phi_{xx}^2 - V(\phi_x) \right) dx dt.$$

There are two standard ways of constructing multisymplectic schemes. One is to approximate  $\mathcal{L}$  by a sum and take variations. The second is to write the PDE as a first order system

$$Kz_x + Lz_t = \nabla S(z) \quad (6)$$

where the matrices  $K$  and  $L$  are antisymmetric, and apply one of some known integrators. Here we use the fact that the box scheme applied to (6) is multisymplectic [17]. In terms of the finite difference operators

$$\begin{aligned} D_x u_i^n &= \frac{u_{i+1}^n - u_i^n}{\Delta x}, \\ D_t u_i^n &= \frac{u_i^{n+1} - u_i^n}{\Delta t}, \\ M_x u_i^n &= \frac{u_i^n + u_{i+1}^n}{2}, \\ M_t u_i^n &= \frac{u_i^n + u_i^{n+1}}{2}, \end{aligned}$$

the box scheme is

$$KD_x M_t z + LD_t M_x z = \nabla S(M_t M_x z).$$

The form (6) of KdV, and corresponding box discretization, is

$$\begin{aligned} u_x &= -\frac{1}{\nu}v, & D_x M_t u &= -\frac{1}{\nu}M_t M_x u, \\ v_x + \frac{1}{2}\phi_t &= w + V'(u), & D_x M_t v + \frac{1}{2}D_t M_x \phi &= M_t M_x w + V'(M_t M_x u), \\ w_x + \frac{1}{2}u_t &= 0, & D_x M_t w + \frac{1}{2}D_t M_x u &= 0, \\ \phi_x &= u, & D_x M_t \phi &= M_t M_x u. \end{aligned} \quad (7)$$

This is the Preissmann box scheme. Applying  $M_t D_x^2$  to the first equation,  $M_x D_x M_t^2$  to the second,  $M_x^2 M_t$  to the third, and  $M_x D_t M_x$  to the fourth, and noting that the four operators mutually commute, gives

$$\begin{aligned} D_x^3 M_t^2 u &= -\frac{1}{\nu}M_t^2 D_x^2 M_x v, \\ M_x D_x^2 M_t^2 v + \frac{1}{2}D_x D_t M_x^2 M_t \phi &= D_x M_t^2 M_x^2 w + D_x M_t M_x V'(M_t M_x u), \\ D_x M_t^2 M_x^2 w + \frac{1}{2}D_t M_x^3 M_t u &= 0, \\ D_t D_x M_x^2 M_t \phi &= D_t M_x^3 M_t u. \end{aligned}$$

Further eliminating  $v$ ,  $w$  and  $\phi$  yields

$$M_t D_t M_x^3 u = M_t D_x M_x V'(M_t M_x u) + \nu M_t^2 D_x^3 u, \quad (8)$$

which is the multisymplectic box scheme derived in [19]. Expressed using finite difference stencils, the resulting 12-point box scheme is

$$\begin{aligned} \frac{1}{16\Delta t} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & -3 & -3 & -1 \end{bmatrix} u &= \frac{1}{4\Delta x} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} V' \left( \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} u \right) \\ &+ \frac{\nu}{4(\Delta x)^3} \begin{bmatrix} -1 & 3 & -3 & 1 \\ -2 & 6 & -6 & 2 \\ -1 & 3 & -3 & 1 \end{bmatrix} u. \end{aligned} \quad (9)$$

This scheme has several features that appear unusual at a first glance. It couples 3 time levels, when 2 would be expected for a PDE which is first order in time. Surely this allows spurious modes in time, or at least, the possibility of  $u_i^{n+1} - u_i^n$  becoming large so that the scheme no longer approximates the PDE? It has a 4-point space average in the approximation of  $u_t$ , and a 3-point time average in the approximation of  $u_{xxx}$ . At least the compact 4-point spatial difference for  $u_{xxx}$  stands out as a good thing, compared to the 5 points used in (2). As we'll see later this scheme turns out to be rather stable, a fact already hinted at by the compact differencing of (7).

Note that every term in (8) contains a factor  $M_t$ . Although  $M_t$  isn't invertible, this suggests the derivation of the following related scheme. The first equation in (7) contains no time derivatives, so it can be more accurately discretized by omitting the time average  $M_t$  to give  $D_x u = -\frac{1}{\nu} M_x u$ . Similarly, one can use  $D_x \phi = M_x u$  for the fourth equation in (7). Combining these with the second and third equations of (7) gives the 8-point box scheme

$$D_t M_x^3 u = D_x M_x V'(M_t M_x u) + \nu M_t D_x^3 u$$

or

$$\begin{aligned} \frac{1}{8\Delta t} \begin{bmatrix} 1 & 3 & 3 & 1 \\ -1 & -3 & -3 & -1 \end{bmatrix} u &= \frac{1}{2\Delta x} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} V' \left( \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} u \right) \\ &+ \frac{\nu}{2(\Delta x)^3} \begin{bmatrix} -1 & 3 & -3 & 1 \\ -1 & 3 & -3 & 1 \end{bmatrix} u. \end{aligned} \quad (10)$$

This is more obviously related to the midpoint rule in time. The 8-point scheme (10) is the equivalent one-step method of (9); if (9) is started by one application of (10), the two will yield identical results up to round-off error.

This approach to constructing compact box schemes can be applied to many multisymplectic PDEs  $Kz_x + Lz_t = \nabla S(z)$ . Let  $z = Ay$  so that  $A^T K A y_x + A^T L A y_t = A^T \nabla S(Ay) = \nabla \tilde{S}(y)$  where  $\tilde{S}(y) = S(Ay)$ . The matrix  $A$  can be chosen, for example, to put the lowest-rank of  $K$  and  $L$  in its Darboux normal form, with the maximum number of zero rows. These rows can then be treated with more compact differences.

Before we turn to numerical experiments let us mention also another scheme we tried, a variant on the box scheme (10) but with narrower differences in the nonlinear term,

$$\begin{aligned} \frac{1}{2\Delta t} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} u = \frac{1}{\Delta x} \begin{bmatrix} -1 & 1 \end{bmatrix} V' \left( \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \right) \\ + \frac{\nu}{2(\Delta x)^3} \begin{bmatrix} -1 & 3 & -3 & 1 \\ -1 & 3 & -3 & 1 \end{bmatrix} u. \end{aligned} \quad (11)$$

### 3 Numerical results

For our numerical experiments we considered the KdV equation (1) on the spatial domain  $[-1, 1]$ , with periodic boundary conditions and initial condition  $u(x, 0) = \cos(\pi x)$ . We used

$$\alpha = -\frac{3}{8}, \quad \nu = -\frac{2}{3} \times 10^{-3}, \quad \rho = -\frac{1}{10}.$$

These values were used in [10]; for us their only important feature is that  $\nu \ll 1$ . If we set  $\nu = 0$  then a shock develops well before  $t$  reaches the value  $T = 2$ . See [12,9,18]. In the sequel we use spatial resolutions where  $\nu/\Delta x^2$  is not large.

The initial condition breaks up into a train of solitons and radiation which repeatedly interact; occasionally the energy concentrates in a single peak with amplitude approximately 2.6, so we have estimated the Courant number  $2|\alpha u_{\max}| \Delta t / \Delta x$  as  $2\Delta t / \Delta x$ . See Figure 1.

Initially we found that in well-resolved simulations (sufficiently small  $\Delta x$  and  $\Delta t / \Delta x$ ), all 3 schemes performed well. The energy and norm  $\|u + \tilde{\rho}\|_2$  did not drift and clean soliton interactions were observed; the schemes could apparently be run for a very long time.

In this context we note that the semi-explicit scheme (4) is by far the fastest of the three. Per time step it is almost as fast as an explicit scheme. In contrast, for the fully implicit schemes we must at each step solve a nonlinear system of equations (we simply used linear extrapolation for the initial guess and an

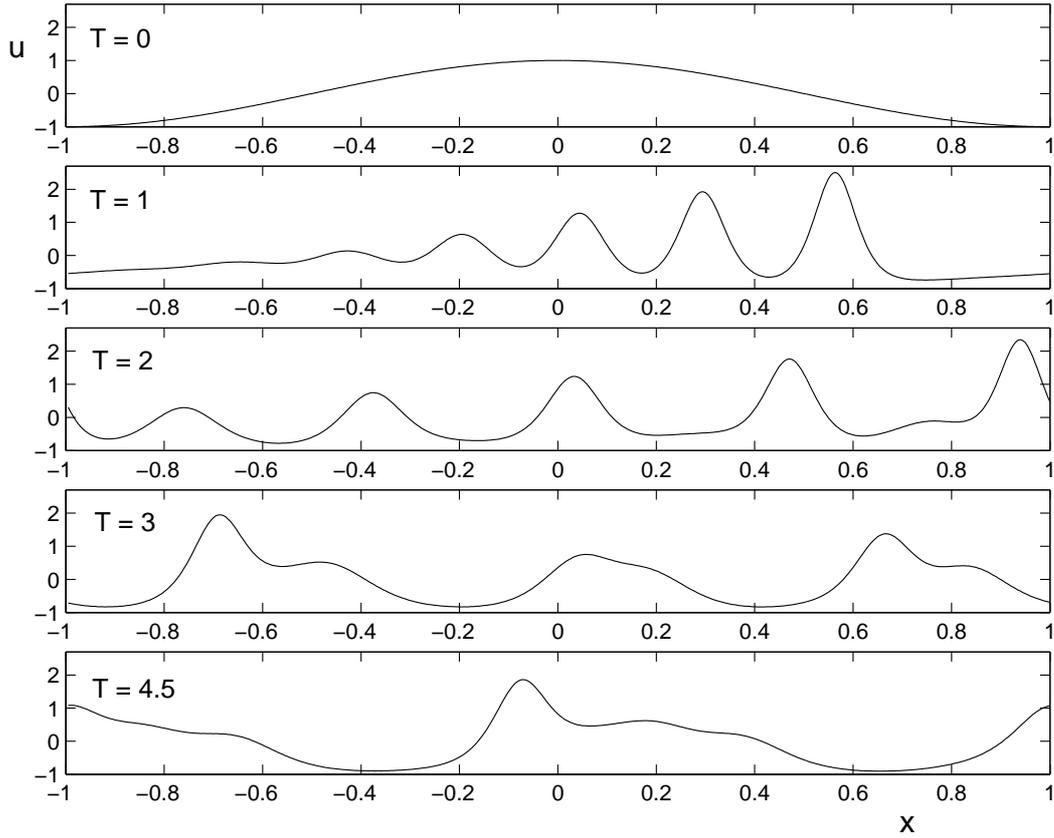


Fig. 1. Evolution of the KdV equation (1) with  $u(x, 0) = \cos(\pi x)$ ,  $u(-1, t) = u(1, t)$ . Shown are snapshots at  $t = T$  for  $T = 0, 1, 2, 3, 4.5$ .

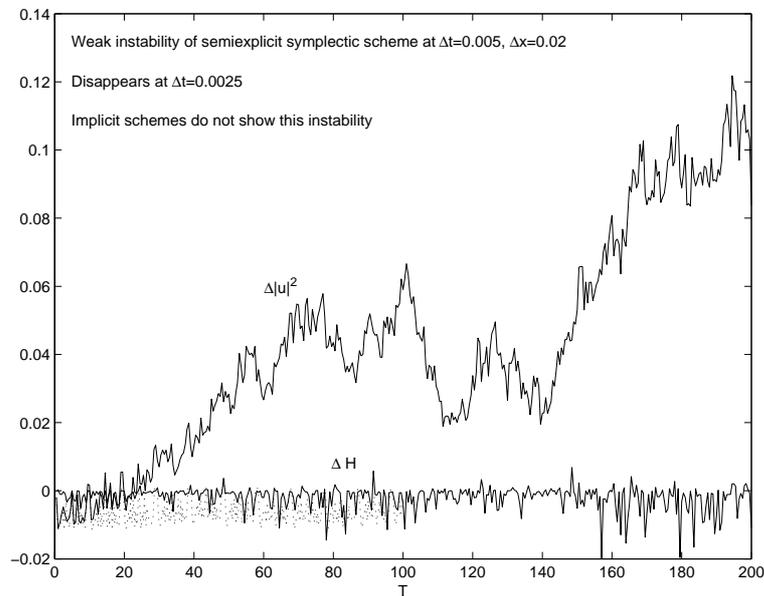


Fig. 2. Changes in energy and norm during a semi-explicit integration Eq. (4) of KdV. For Courant number approx. 0.5, the norm but not the energy drifts. This drift disappears at a Courant number of 0.25 (dots).

iteration based on Newton’s method with the nonlinear terms omitted from the Jacobian). Noting that we have hardly optimized our implementations, at a reasonable Courant number of 0.5 and using 200 spatial grid points the midpoint rule (5) is about 3.5, and the 12-point box scheme (9) about 5.5, times slower than (4). But a study of relative efficiency is not our goal, and we proceed to study our methods for coarser discretizations.

Thus, we increased  $\Delta x$  (generally to 0.02, or 100 grid points) so that the solution is poorly-resolved spatially and then we increased  $\Delta t$  until the schemes fell over.

At a time step of  $\Delta t = 0.01$  (Courant number approximately 1), the semi-explicit scheme (4) eventually blows up, although this takes a fairly long time to happen ( $T \approx 17$ ). The performance does not depend on whether the linear ( $u_{xxx}$ ) term is treated exactly or approximately (via the midpoint rule). However, even at a smaller time step of 0.005, this scheme still shows a weak instability manifested as a drift of  $\|u + \tilde{\rho}\|_2^2$ . See Figure 2. By  $T = 200$  it has drifted by about 10%; eventually  $u$  could blow up. Note that the discrete energy does not drift. This instability may be a result of an interaction between the fast waves (due to  $u_{xxx}$ ) and the nonlinear term. For smaller time steps, e.g. 0.0025, the instability disappears and  $\|u + \tilde{\rho}\|_2^2$  does not drift. However, tiny wiggles are still visible in the solution. We shall discuss these more later.

We then turned to the two implicit schemes, the symplectic midpoint rule and the multisymplectic 12-point box scheme. We were surprised that both of them are stable for Courant numbers much larger than 1. They both go unstable at a time step of around 0.08 (Courant number 8). (Initially, oscillations appear in the solution which are not present at smaller time steps. For still larger time steps, e.g. 0.1, our iterative nonlinear solver no longer converges.) We did extensive tests at  $\Delta t = 0.05$  and both schemes appeared able to run for a long time maintaining the typical pattern of soliton interactions. However, we did discover one significant difference between the two schemes. The symplectic scheme shows tiny spatial wiggles while the multisymplectic scheme does not; it remains remarkably smooth for a very long time. We tested this point further by progressively increasing  $\Delta x$ . At  $\Delta x = 2/75$ ,  $\|u + \tilde{\rho}\|_2^2$  drifts slowly for the symplectic scheme, while the multisymplectic scheme remains smooth. At  $\Delta x = 2/50$ , the symplectic scheme blows up for all  $\Delta t$ . Only at  $\Delta x = 2/25$  (so that the solitons are essentially 3-point sawtooths) does the multisymplectic scheme begin to degrade, showing a drift of  $\|u + \tilde{\rho}\|_2^2$ .

This ability of the 12-point box scheme to avoid wiggles is impressive. Remember, all the schemes are non-dissipative; a wiggle, once created, will probably persist forever. Because  $\nu/\Delta x^2$  is not large, KdV is close here to the inviscid Burgers’ equation, for which a non-dissipative scheme *certainly* can’t avoid developing wiggles as soon as a shock forms [9,18]. The situation is summarized

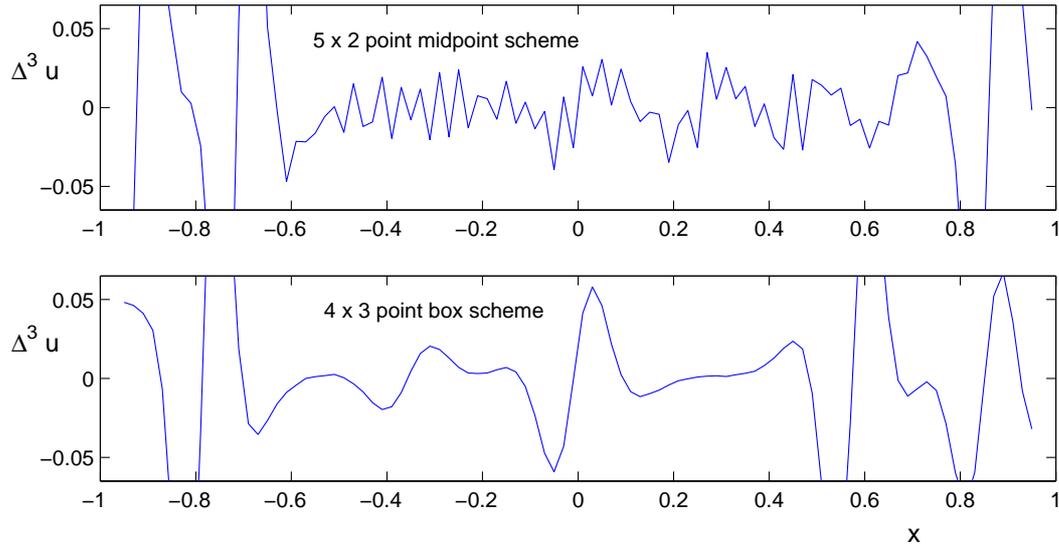


Fig. 3. The solutions of two implicit schemes are compared at  $T = 30$ , with  $\Delta x = 0.02$  and  $\Delta t = 0.05$  (Courant number 5). Top: the symplectic midpoint rule based on a  $5 \times 2$  finite difference, Eq. (5). Bottom: The 12 point box scheme, Eq. (9). To illustrate the tiny wiggles generated by the symplectic scheme we display the third difference of  $u$ . These wiggles persist for smaller  $\Delta t$  but are not present in the box scheme.

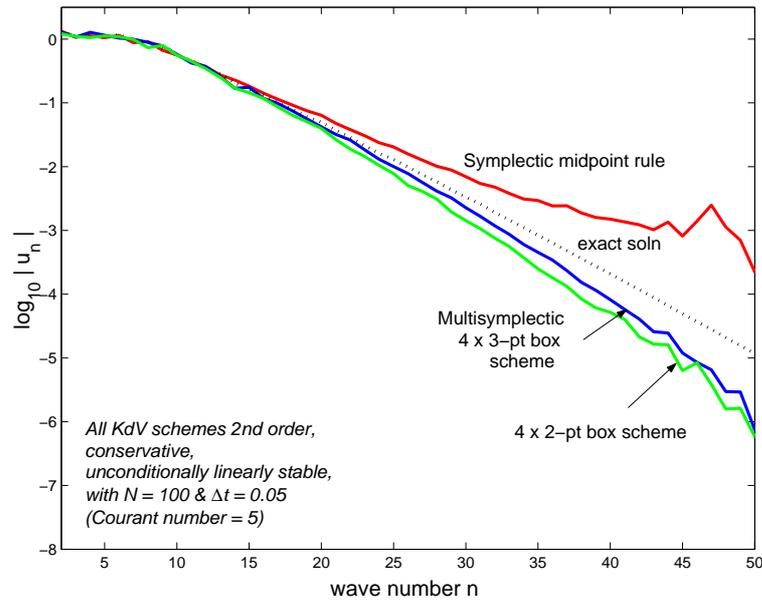


Fig. 4. Time averaged ( $T = 30$ ) discrete Fourier spectra. Despite a poorly resolved solution, the box schemes stay remarkably smooth, even smoother than the exact solution.

in Figure 4, which shows the time-averaged Fourier spectra of the solution for all schemes. Scheme (5) puts 2 to 3 orders of magnitude more energy in the high (top 1/3) frequencies than scheme (9). This appears to be true for all  $\Delta t$  and for all  $T$ , although a slow growth in the high modes is evident for all

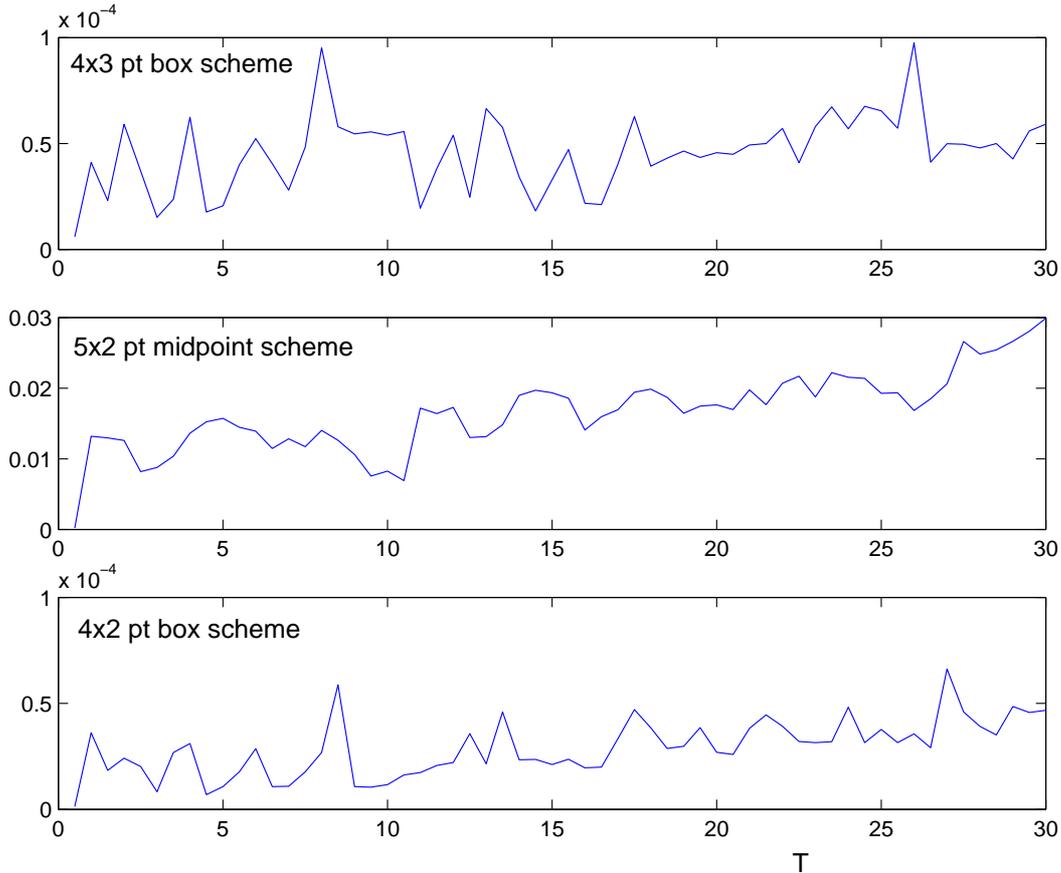


Fig. 5. Time evolution of energy in the 4 highest wave numbers; all schemes show some secular drift. Note different scale on axes.

schemes (Figure 5). In fact, the solution to the 12-point box scheme is even smoother than the exact solution! Nor does it develop sawtooth waves in time, despite being a (neutrally stable) 3 time level scheme.

To explain this we considered various alternatives. It is likely that the  $N$  independent multisymplectic conservation laws of the box scheme make it exceptionally robust. However, we know of no precise analysis using these conservation laws which can elaborate on this. We wondered if the box scheme treats the Burgers' part of KdV well: but it is basically a conservative central difference, and, if applied to inviscid Burgers', it does develop wiggles at the same rate as the other schemes. It is known [8] that for the two-point boundary value problem  $au_x + bu_{xx} = f(u)$ ,  $u(0) = c$ ,  $u(1) = d$ , the box scheme and standard central differences develop wiggles at the same value of  $\Delta x$ . We also wondered whether the spatial smoothing on the left hand side of (9) plays a role; however, note that  $[1 \ 3 \ 3 \ 1]u = 8u(x) + 3(\Delta x)^2 u_{xx} + \mathcal{O}((\Delta x)^4)$ ; the filtering of high wavenumbers here has the wrong sign and should act to *magnify* any wiggles on the right hand side.

The one essential feature of (2) which the multisymplectic box schemes improve upon is the wide stencils used by the central differencing of  $\frac{\partial}{\partial x}$ . We next try to capture this by a dispersion analysis.

#### 4 Preserving the dispersion relation

We now present a possible explanation based on a traditional linear analysis of numerical dispersion relations.

First, consider the one-way wave equation  $u_t = u_x$  and the effect of the three discretizations on a wave of the form  $u_j^n = e^{i(\xi j + \omega n)}$ . (Their dispersion relations are shown in Fig. 6, top left.) For central differences in space and (three time level) leapfrog in time, the dispersion relation is  $\sin \omega = \lambda \sin \xi$ , where  $\lambda = \Delta t / \Delta x$ . For each real  $\xi \in [-\pi, \pi]$ , the scheme supports two temporal wavenumbers  $\omega$ , the second, high-frequency one being traditionally described as a parasitic wave. For central differences in space and the midpoint rule in time, the dispersion relation is  $2 \tan(\omega/2) = \lambda \sin \xi$ . There are no temporal parasitic waves, but for small  $\omega$ , there are two real values of  $\xi$  that satisfy the dispersion relation. One of them is close to a sawtooth ( $\xi = \pi$ ) and it, too, can be called a parasitic wave. It doesn't look like one in this equation, but in a nonlinear equation, such a feature means that low temporal frequencies in the solution may generate spatial sawtooth waves—exactly what is observed in KdV. Indeed, the philosophy of the multisymplectic approach to wave equations is that space and time should be treated on an equal footing. Parasites in space or time can both be harmful!

Now consider the dispersion relation of the box scheme, which we spell out in full: the scheme is

$$\frac{1}{2\Delta t} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} u = \frac{1}{2\Delta x} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} u, \quad (12)$$

which gives the dispersion relation

$$\frac{2}{\Delta t} \cos(\xi/2) \sin(\omega/2) = \frac{2}{\Delta x} \sin(\xi/2) \cos(\omega/2).$$

The  $\cos(\xi/2)$  term on the left, which vanishes at  $\xi = \pi$ , indicates the spatial average apparently pumping up sawtooths; but it is exactly cancelled by the  $\cos(\omega/2)$  on the right, to give the dispersion relation

$$\tan(\omega/2) = \lambda \tan(\xi/2). \quad (13)$$

Now for all values of  $\lambda$ , to each  $\xi$  there is precisely one  $\omega$  and vice versa, just as in the true dispersion relation. In terms of integrators, just as the stencil of 3-level is too wide for the equation—for an ODE we would say it operates on the wrong phase space, allowing parasitic modes—the stencil of central differences in space is too wide and also allows parasitic modes. Only the box scheme avoids this problem in both space and time.

The good dispersion relation of the box scheme in this example is in fact universal.

**Proposition 1** *The box scheme qualitatively preserves the dispersion relation of any system of linear first-order PDEs  $Kz_x + Lz_t = Sz$ . Specifically, there are diffeomorphisms  $\psi_1$  and  $\psi_2$  which conjugate the exact and numerical dispersion relations such that to each pair  $(\xi, \omega)$  satisfying the numerical dispersion relation there corresponds a pair  $(\psi_1(\xi), \psi_2(\omega))$  satisfying the exact dispersion relation.*

**PROOF.** The exact dispersion relation is obtained by setting  $z = e^{i(\xi x + \omega t)}y$ , where  $y$  is a constant vector, giving the generalized eigenvalue problem

$$(i\xi K + i\omega L - S)y = 0.$$

The dispersion relation is the polynomial  $\det(i\xi K + i\omega L - S) = 0$ . For the numerical relation we put  $z_j^n = e^{i(\xi j + \omega n)}y$  to get the numerical dispersion relation (as in (13)),

$$\det\left(i\frac{2}{\Delta x}\tan(\xi/2)K + i\frac{2}{\Delta t}\tan(\omega/2)L - S\right) = 0.$$

Thus, the required diffeomorphisms are

$$\psi_1: (-\pi, \pi) \rightarrow \mathbb{R}, \quad \psi_1(\xi) = \frac{2}{\Delta x}\tan(\xi/2)$$

and

$$\psi_2: (-\pi, \pi) \rightarrow \mathbb{R}, \quad \psi_2(\omega) = \frac{2}{\Delta t}\tan(\omega/2).$$

This completes the proof.

Note that the multisymplectic property (namely that  $K$  and  $L$  are antisymmetric and  $S$  is symmetric) is not used. It is hard to imagine better functions than  $\psi_{1,2}$  for the purpose. The dispersion relation of a large PDE system can be arbitrarily complicated (with many branches, changing number of solutions etc.) and it is remarkable that the box scheme can capture all of this with

merely a shift in frequencies, unconditionally, for all  $\Delta t$  and  $\Delta x$ . Proposition 1 also holds for PDEs with any number of space dimensions.

This feature is not shared by all multisymplectic schemes. Of course the explicit ones are only conditionally stable. But even the multisymplectic implicit Runge-Kutta schemes have a problem. Consider the one-dimensional ODE  $y' = i\lambda y$  for which 4th order (symplectic) Gaussian Runge-Kutta gives  $y^{n+1} = e^{i\theta} y^n$ ,  $\tan(\theta/2) = 6Z/(12 - Z^2)$ ,  $Z = \lambda\Delta t$ . As  $Z$  increases from 0 to  $\infty$ , the angle  $\theta$  increases from 0 to  $2\pi$ . The high frequencies  $Z > \sqrt{12}$  are mapped into  $(\pi, 2\pi)$ , and hence in the dispersion relation, are aliased by  $\theta \mapsto \theta - 2\pi$  into low frequencies in  $(-\pi, 0)$ . In a discretization of a wave equation these correspond to parasitic sawtooth waves.<sup>1</sup>

We illustrate Proposition 1 for the model equation  $u_t = \rho u_x + \nu u_{xxx}$ , corresponding to a linearization of the KdV equation. That is, we consider the PDE (1) with  $V(u) = \frac{\rho}{2}u^2$ . Figure 6 (top right) shows how the box scheme preserves the shape of the dispersion relation. To each of the two temporal wave numbers  $\omega$  shown corresponds the correct number of spatial wave numbers  $\xi$ . The whole high-frequency part of the relation is squeezed into the corner near  $\xi = \omega = \pi$ . By contrast, the standard symplectic discretization (2), (5), although unconditionally stable, has the parasitic waves shown in Figure 6 (bottom left) for small  $\Delta x$ . Of course the solution is rich in low temporal frequencies and these will generate the high spatial frequencies—the wiggles observed earlier. (It's not so clear what the lack of *any* real  $\xi$  corresponding to the larger  $\omega$ 's means. All the solutions for  $\xi$  are complex, but these growing modes are presumably suppressed by the periodic boundary conditions.) When the  $u_{xxx}$  term dominates, i.e. for small  $\Delta x$ , one finds that the parasitic waves occupy the top 1/3 of the wavenumbers, as observed in the simulations (Figure 4).

For the box scheme, the dispersion relation has the same shape for all  $\Delta t$  and  $\Delta x$ . Other schemes can undergo bifurcations. For example, when  $C := \frac{\rho}{\nu}(\Delta x)^2 > 4$ , the symplectic  $5 \times 2$  point midpoint scheme (2),(5) has only one real  $\xi$  for each  $\omega$  (Figure 6, bottom right)—a kind of loss of stability in the  $x$ -direction. The box scheme (11) has a correct dispersion relation for  $C < 4$ , but also undergoes a bifurcation and has the wrong dispersion relation for  $C > 4$ . (Figure 6, bottom right.) This scheme also developed wiggles. (Consider the steady state solutions. Scheme (11) now looks like an explicit integrator in the  $x$ -direction and can be expected to be only conditionally stable. By contrast, scheme (10) is fully implicit in the  $x$ -direction.)

This analysis partly explains why the 12-point box scheme can be so robust.

<sup>1</sup> One is tempted to define yet another kind of Runge-Kutta stability, that of mapping the positive imaginary axis diffeomorphically into the top half of the unit circle.

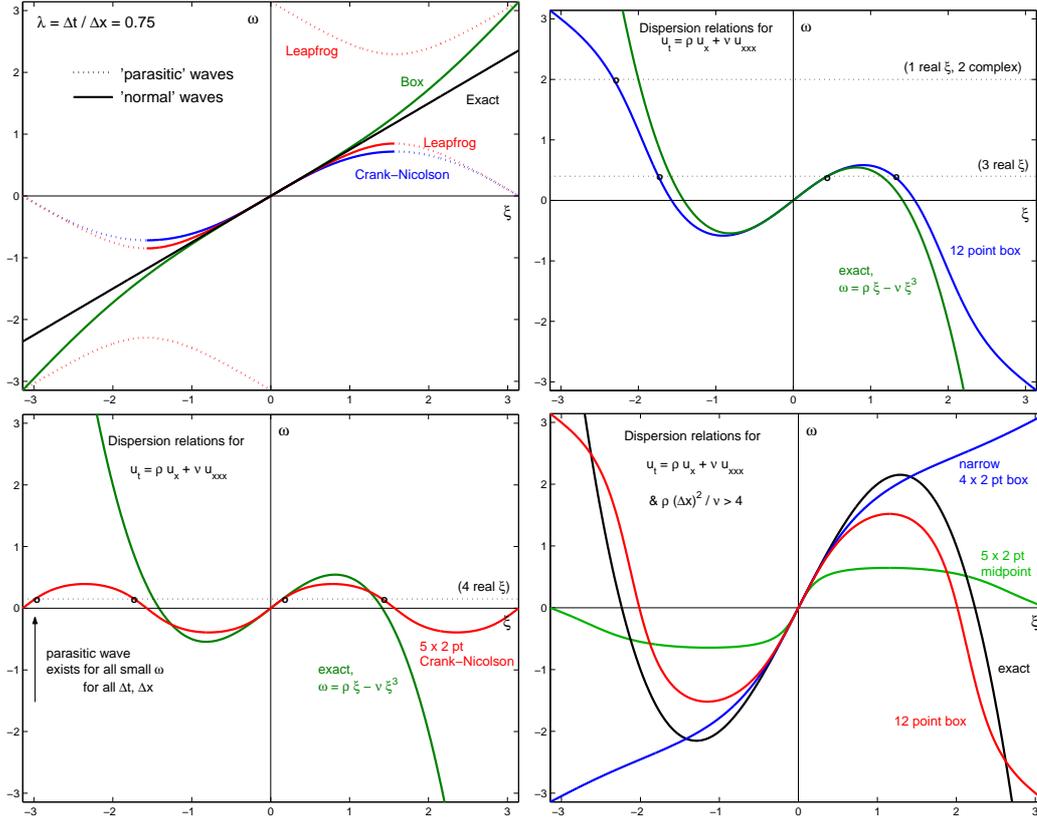


Fig. 6. Numerical dispersion relations. Top left: for the one-way wave equation  $u_t = u_x$ ,  $(3 \times 3)$  point leapfrog,  $(3 \times 2)$  point Crank-Nicolson, and  $(2 \times 2)$  point box (Eq. (12)) schemes compared. Only the box scheme avoids parasitic waves. Top right: The box scheme preserves the shape of the dispersion relation for  $u_t = \rho u_x + \nu u_{xxx}$ . Bottom left: the symplectic scheme (5) does not. Bottom right: the relation for some schemes is even more wrong for large  $\Delta x$ , but not that of the box scheme. The ‘narrow’ box is Eq. (11).

Despite coupling 3 time levels, it contains (in a linear analysis) no parasitic temporal waves. Nonlinear effects, of course, could still cause  $u_i^{n+1} - u_i^n$  to grow, causing a breakdown of the solution after a very long time. This could be overcome by (10). Note that the 8 point box scheme (10) also preserves dispersion relations unconditionally—the omitted  $\cos(\omega/2)$  terms coming from the time averaging merely cancel out anyway—and is multisymplectic. For large enough time steps (as well as for very many time steps, so many that linear roundoff accumulation becomes a factor), nonlinear effects can still potentially destabilize the scheme.

We conclude by illustrating the size of the parasitic waves in our KdV simulations for increasing spatial resolution in Figure 7. Whether parasitic waves of the amplitude shown by the semi-explicit splitting method are a problem or not, and whether they are contributing to the global error, depends on the

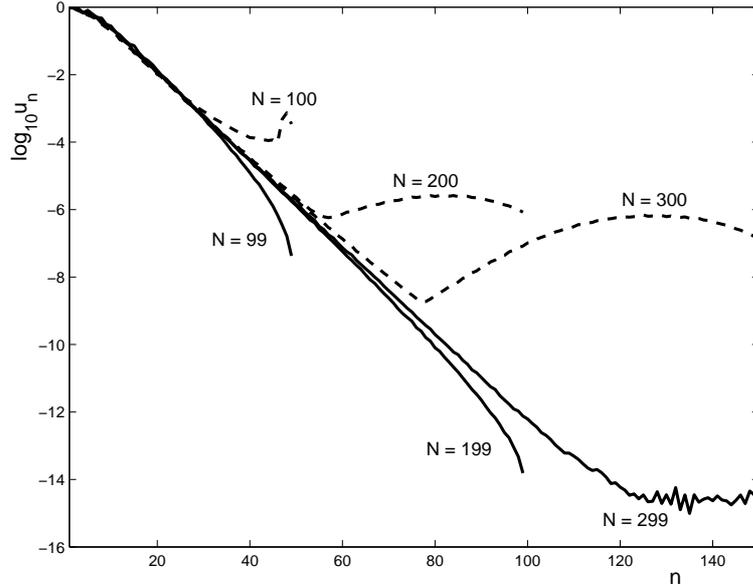


Fig. 7. Time-averaged Fourier spectra vs. spatial resolution. The solid line shows the 8-point box scheme (10), which preserves dispersion relations as in Proposition 1, for  $N = 99, 199,$  and  $299$  grid points, with Courant number fixed at 1. The dashed line shows the semi-explicit symplectic integrator (2), (4), for  $N = 100, 200,$  and  $300$  grid points, with Courant number fixed at 0.25. The simulation time is a relatively short  $T = 5$ .

application.

The box scheme has been around for more than 40 years, being variously known as the Preissmann [16] or Keller [11] box scheme, but it has only really caught on in hydrology [1]. It would be nice to know if its robustness on KdV also holds for other nonlinear problems, larger systems, diverse boundary conditions, and for PDEs with more space dimensions and/or complex geometry.

## Acknowledgements

The first author thanks Andre Nachbin for many helpful discussions; he also thanks NSERC Canada and CNPq Brazil for their support.

The second author thanks the Marsden Fund of the Royal Society of New Zealand, the Centre for Advanced Study, Oslo, and the organizers of the ‘Innovative Integrators for PDEs’ workshop for their support.

## References

- [1] M. B. Abbott and D. R. Basco, *Computational fluid dynamics. An introduction for engineers*, Wiley, New York, 1990.
- [2] U. Ascher, S. Ruuth and B. Wetton, Implicit-explicit methods for time-dependent PDE's, *SIAM J. Numer. Anal.* **32** (1995), 797–823.
- [3] U. Ascher, S. Ruuth and R. Spiteri, Implicit-Explicit Runge-Kutta Methods for Time-Dependent Partial Differential Equations, *Appl. Numer. Math.* **25** (1997), 151–167.
- [4] C.J. Budd and M. D. Piggott, Geometric integration and its applications, *Handbook of Numerical Analysis vol. XI*, to appear.
- [5] T. J. Bridges and S. Reich, Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity, *Phys. Lett. A* **284** (2001), no. 4-5, 184–193.
- [6] J. B. Chen, H. Munthe-Kaas, and M. Z. Qin, Square-conservative schemes for a class of evolution equations using Lie-group methods, *SIAM J. Numer. Anal.* **39** (2002), no. 6, 2164–2178.
- [7] J.-P. Croisille, Keller's box-scheme for the one-dimensional stationary convection-diffusion equation, *Computing* **68** (2002), no. 1, 37–63.
- [8] R. S. Fisk, On an oscillation phenomenon in the numerical solution of the diffusion-convection equation, *SIAM J. Numer. Anal.* **19** (1982), no. 4, 721–724.
- [9] T.Y. Hou and P.D. Lax, Dispersive approximation in fluid dynamics, *Comm. Pure Appl. Math.* **44** (1991), 1-40,
- [10] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge, 1996.
- [11] H. B. Keller, A new difference scheme for parabolic problems, in *Numerical Solution of Partial Differential Equations, II (SYNSPADE 1970)*, (Proc. Sympos., Univ. of Maryland, College Park, Md., 1970), pp. 327–350, Academic Press, New York, 1971.
- [12] P.D. Lax and C.D. Levermore, The small dispersion limit of the Korteweg-deVries Equation: I,II,III, *Comm. Pure Appl. Math.* **36** (1983) 253-290, 571-593, 809-829
- [13] J. E. Marsden, G. W. Patrick, and S. Shkoller, Multisymplectic geometry, variational integrators, and nonlinear PDEs, *Comm. Math. Phys.* **199** (1998), no. 2, 351–395.
- [14] R. I. McLachlan, Symplectic integration of Hamiltonian wave equations, *Numer. Math.* **66** (1994), no. 4, 465–492.

- [15] R. I. McLachlan and G. R. W. Quispel, Splitting methods, *Acta Numerica* **11** (2002), 341–434.
- [16] A. Preissmann, Propagation des intumescences dan les canaux et rivières, *First Congress French Assoc. for Computaton*, Grenoble, 1961.
- [17] S. Reich, Multi-symplectic Runge-Kutta collocation methods for Hamiltonian wave equations, *J. Comput. Phys.* **157** (2000), no. 2, 473–499.
- [18] C.V. Turner and R.R. Rosales, The small dispersion limit for a nonlinear semidiscrete system of equation, *Studies in Appl. Math.* **99** (1997), 205–254.
- [19] P. F. Zhao and M. Z. Qin, Multisymplectic geometry and multisymplectic Preissmann scheme for the KdV equation, *J. Phys. A* **33** (2000), no. 18, 3613–3626.