Conservative Discontinuous Galerkin Methods For The Generalized Korteweg-de Vries equation

J.L. Bona†, H. Chen‡, O. Karakashian* and Y.Xing**

(†) Department of Mathematics UIC
(‡) Department of Mathematics, Memphis
(*)The University of Tennessee, Knoxville
(**) The University of Tennessee, Knoxville and JICS

NSF support gratefully acknowledged

The Problem

$$\begin{cases} u_t + (u^{p+1})_x + \epsilon u_{xxx} & = & 0, & 0 < x < 1,\ t > \\ u(x,0) & = & u^0(x), & 0 < x < 1, \end{cases}$$

ϵ Positive parameter

$p \geq 1$ integer

We are interested in periodic solutions
Conserved quantities

$$\int_0^1 u\, dx, \qquad \int_0^1 u^2\, dx.$$

The meshes

$\mathcal{T}_h$ is a partition of [0,1]

$$0 = x_0 < x_1 < \cdots < x_M = 1.$$

We consider functions that are discontinuous.

Jumps and averages:

$$[u]_m = u_m^+ - u_m^- \quad \text{is the jump of } u \text{ at } x_m$$

$$\{u\}_m = \frac{1}{2}(u_m^+ + u_m^-) \quad \text{is the average of } u \text{ at } x_m$$

"periodicity":

$$x_0^- := x_M^-, \quad x_M^+ = x_0^+, \quad u_0^- := u_M^+$$

Broken Sobolev Spaces: $W^{s,p}(\mathcal{T}_h) = \Pi_{I \in \mathcal{T}_h} W^{s,p}(I)$.

Discontinuous finite element spaces: $V_h^q = \Pi_{I \in \mathcal{T}_h} \mathcal{P}_q(I)$.

Sobolev embedding/ trace inequality:

$$\|v\|_{L^\infty(I)} \le c h_I^{-1/2} \|v\|_{L^2(I)} + c h_I^{1/2} \|v_x\|_{L^2(I)}, \quad \forall v \in H^1(I),$$

Approximation results:

$$|v - \chi|_{j,I} \le c h_I^{i-j} |v|_{i,I}, \quad 0 \le j \le i \le q+1,$$

Inverse inequality:

$$|\chi|_{j,I} \le h_I^{-j} |\chi|_I, \quad \forall \chi \in \mathcal{P}_q.$$

The weak formulation

Multiplying the nonlinear term in the equation by $v \in H^1(\mathcal{T}_h)$, integrating by parts and then summing over cells we obtain

$$\begin{aligned}\sum_{I\in\mathcal{T}_h}((u^{p+1})_x, v)_I &= -\sum_{I\in\mathcal{T}_h}(u^{p+1}, v_x)_I + \sum_{m=0}^{M-1}\left[(u^-_{m+1})^{p+1}\, v^-_{m+1} - (u^+_m)^{p+1}\, v^+_m\right] \\ &= -\sum_{I\in\mathcal{T}_h}(u^{p+1}, v_x)_I - \sum_{m=0}^{M-1}[u^{p+1}v]_m.\end{aligned}$$

$$\hat{\phi}(u^+_m, u^-_m) = \frac{1}{p+2}\sum_{j=0}^{p+1}(u^+_m)^{p+1-j}(u^-_m)^j.$$

With this in mind, we define the nonlinear operator $\mathcal{N} : H^1(\mathcal{T}_h) \rightarrow V^q_h$ by

$$(\mathcal{N}(u), v) = -\sum_{I\in\mathcal{T}_h}(u^{p+1}, v_x)_I - \sum_{m=0}^{M-1}\hat{\phi}(u^+_m, u^-_m)\,[v_m]. \qquad (1)$$

The operator $\mathcal{N}$ is well defined by virtue of the Riesz Representation Theorem. In addition we have,

Lemma 0.1. *(i) The nonlinear form defined with the choice of $\hat{\phi}$ indicated above is consistent in the sense that for all 1-periodic functions u in $C^1[0,1]$ there holds*

$$(\mathcal{N}(u), v) = ((u^{p+1})_x, v), \quad \forall v \in H^1(\mathcal{T}_h).$$

(ii) The nonlinear term $\mathcal{N}$ is conservative in the sense that

$$(\mathcal{N}(v), v) = 0 \quad \forall v \in H^1(\mathcal{T}_h).$$

A bilinear form for the dispersive term

we perform integration by parts twice to obtain

$$\sum_{I\in\mathcal{T}_h}(u_{xxx},v)_I = \sum_{I\in\mathcal{T}_h}(u_x,v_{xx})_I - \sum_{m=0}^{M-1}[u_{xx}v]_m + \sum_{m=0}^{M-1}[u_x v_x]_m.$$

Jump identities

$$[\phi\psi]_m = \phi_m^+[\psi]_m + [\phi]_m\psi_m^- = \phi_m^-[\psi]_m + [\phi]_m\psi_m^+ = \{\phi\}_m[\psi]_m + [\phi]_m\{\psi\}_m.$$

We define the operator $\mathcal{D} : H^3(\mathcal{T}_h) \to V_h^q$ by

$$(\mathcal{D}(u),v) = \sum_{I\in\mathcal{T}_h}(u_x,v_{xx})_I - \sum_{m=0}^{M-1}\left(u_{xx}^+[v]_m - [u]_m v_{xx}^+\right) + \sum_{m=0}^{M-1}\{u_x\}_m[v_x]_m.$$

Lemma 0.2. *(i) The form $\mathcal{D}$ is consistent in the sense that for all 1-periodic functions u in $C^2[0,1] \cap H^3(\mathcal{T}_h)$ there holds*

$$(\mathcal{D}(u), v) = (u_{xxx}, v), \quad \forall v \in H^3(\mathcal{T}_h).$$

(ii) The form $\mathcal{D}$ is conservative in the sense that

$$(\mathcal{D}(v), v) = 0 \quad \forall v \in H^3(\mathcal{T}_h).$$

Semidiscrete approximation

Define $u_h : [0,T] \to V_h^q$ by

$$
\begin{aligned}
(u_{ht}, v) + (\mathcal{N}(u_h), v) + \epsilon(\mathcal{D}(u_h), v) &= 0, \quad \forall v \in V_h^q,\ t \in [0,T], \\
u_h(0) &= Pu^0,
\end{aligned}
$$

P is some projection operator into V_h^q with optimal $O(h^{q+1})$ approximation properties.

The semidiscrete approximation u_h satisfies

$$\|u_h(t)\| = \|u_h(0)\|, \quad t > 0.$$

A unique global in time solution exists.

Note: The priodic B.C.'s are NOT enforced explicitely on the F.E. subspace or u_h. Rather, it is incorporated in the formulation via the assumption that u is periodic.

Error estimates

Given $u \in H^3(\mathcal{T}_h)$, need some function $w \in V_h^q$ which is optimally close to u and such that

$$(\mathcal{D}(w), v) = (\mathcal{D}(u), v) \quad \forall v \in V_h^q$$

Cheng-Shu projection: $q \geq 2$

$$\begin{aligned}
(\tilde{w}, v)_I &= (u, v)_I, && \forall v \in \mathcal{P}_{q-3}(I),\ I \in \mathcal{T}_h \\
\tilde{w}(x_m^-) &= u(x_m^-), && m = 1, \ldots, M, \\
\tilde{w}_x(x_m^+) &= u_x(x_m^+) && m = 0, \ldots, M-1, \\
\tilde{w}_{xx}(x_m^+) &= u_{xx}(x_m^+), && m = 0, \ldots, M-1,
\end{aligned}$$

$$\|u - \tilde{w}\|_{W^{j,p}(I)} \leq c h_I^{q+1-j} |u|_{W^{q+1,p}(I)}, \quad I \in \mathcal{T}_h, \quad j = 0, 1, \quad p = 2, \infty.$$

The projection $\tilde{w}$ defined by Cheng-Shu does not satisfy the required identity above but satisfies it if $\mathcal{D}$ is replaced by their (dissipative) formulation.

The "conservative" projection

$$
\begin{aligned}
&(w,v)_I = (u,v)_I, && \forall v \in \mathcal{P}_{q-3}(I),\ I \in \mathcal{T}_h\\
&w(x_m^-) = u(x_m^-), && m = 1,\ldots,M,\\
&\{w_x\}_m = \{u_x\}_m = u_x(x_m) && m = 0,\ldots,M-1,\ \text{or}\ m = 1,\ldots,M,\\
&w_{xx}(x_m^+) = u_{xx}(x_m^+), && m = 0,\ldots,M-1.
\end{aligned}
$$

- The difference with the Cheng-Shu projection is only in the third condition. BUT leads to huge complications! Reason: It is not a local projection anymore, but a global one.
- The projection defined above satisfies $(\mathcal{D}(w),v) = (\mathcal{D}(u),v), \quad \forall v \in H^3(\mathcal{T}_h).$

Proposition 0.1. *Assume that u is sufficiently smooth and periodic. Assume further that $q \geq 2$ is even and that the number M of cells in $\mathcal{T}_h$ is odd. Then there exists a unique w satisfying the above conditions. Furthermore for $j = 0, 1$, $p = 2, \infty$ and for all $I \in \mathcal{T}_h$,*

$$
\|u-w\|_{W^{j,p}(I)} \leq c h_I^{1-j} \left(\sum_{I \in \mathcal{T}_h^N} h_I^q \|u\|_{W^{q+1,\infty}(I)} + \sum_{I \in \mathcal{T}_h \setminus \mathcal{T}_h^N} h_I^{q+1} \|u\|_{W^{q+2,\infty}(I)} \right),
$$

where $\mathcal{T}_h^N$ is the set of cells whose length differs from that of its two neighbours.

Proof

1. Compare w to $\tilde{w}$ so let $e = w - \tilde{w}$
2. Expand e in Legendre polynomials and .....

$$
\begin{aligned}
&(e, v)_I = 0, && \forall v \in \mathcal{P}_{q-3}(I),\ \forall I \in \mathcal{T}_h \\
&e(x_m^-) = 0, && m = 1, \ldots, M, \\
&e_x(x_m^-) + e_x(x_m^+) = u_x(x_m) - \tilde{w}_x(x_m^-) && m = 0, \ldots, M-1, \text{ or } m = 1, \ldots, M, \\
&e_{xx}(x_m^+) = 0, && m = 0, \ldots, M-1.
\end{aligned}
$$

With e_m denoting the restriction of e to I_m, we have

$$e_m(x) = \sum_{\ell=0}^{q} \alpha_{m,\ell} P_{m,\ell}(x) = \sum_{\ell=0}^{q} \alpha_{m,\ell} P_\ell(t), \quad m = 0, \ldots, M-1,$$

we see that the first equation and the orthogonality of the Legendre polynomials imply

$$\alpha_{m,\ell} = 0, \quad \ell = 0, \ldots, q-3, \quad m = 0, \ldots, M-1.$$

Using the second and fourth equations:

$$\alpha_{m,q-2} = -\frac{q(q+1)}{(q-2)(q-1)}\alpha_{m,q}, \quad \alpha_{m,q-1} = \frac{2(2q-1)}{(q-2)(q-1)}\alpha_{m,q}$$

Using the third equation:

$$\begin{aligned}\frac{\alpha_{m-1,q}}{h_{m-1}} + \frac{\alpha_{m,q}}{h_m} &= \frac{q-2}{2q(2q-1)}\Big(u_x(x_\ell) - \tilde{w}_x(x_\ell^-)\Big), \quad m = 1,\ldots,M,\\ &= \frac{q-2}{2q(2q-1)}h_{m-1}^q \sum_{j=0}^{q-2} \rho_j u^{(q+1)}(\zeta_{m-1,j}),\end{aligned}$$

where $\alpha_{M,q} := \alpha_{0,q}$ and where the constants ρ_j, $j = 0,\ldots,q-2$ depend only on q and the values $\zeta_{m-1,j}$, $j = 0,\ldots,q-2$ belong to the cell I_{m-1}.
The coefficient matrix of this system is an $M \times M$ circulant matrix with first row $[1, 1, 0, \ldots, 0]$. This matrix is invertible if and only if M is odd, whence its inverse, also circulant, has $\frac{1}{2}[1, -1, 1, -1, \ldots, -1, 1]$ as its first row. Thus, we have

$$\hat{\alpha}_{m,q} = \frac{q-2}{4q(2q-1)}\Big(\eta_m - \sum_{\ell\in\sigma_m}\big(\eta_\ell - \eta_{\ell+1}\big)\Big) \quad m = 0,\ldots,M-1,$$

If $h_\ell = h_{\ell+1}$, then $|\eta_\ell - \eta_{\ell+1}| \le c h_\ell^{q+1} \|u^{q+2}\|_{L^\infty(I_\ell \cup I_{\ell+1})}$ by the MVT
Hence, it follows that

$$|\alpha_{m,q}| \le c(1+\nu)h^{q+1}\|u^{q+1}\|_{L^\infty(0,1)} + ch^{q+1}\|u^{q+2}\|_{L^\infty(0,1)}, \quad m = 0, \ldots, M-1.$$

Theorem 0.1. *Assume that the solution of problem is sufficiently regular and that $u_h(0)$ is chosen to satisfy $\|u^0 - u_h(0)\| = O(h^q)$. Then, there exist h_0 and a constant c both depending on u, p and T such that for all $h < h_0$ there holds*

$$\|(u - u_h)(t)\| \le ch^q, \quad 0 \le t \le T.$$

Note the estimate is suboptimal due to the derivative in the nonlinear term. This is also true for the Cheng-Shu scheme.

Conservative Fully discrete approximations

Midpoint rule

$$(u^{n,1} - u^n, v) + \kappa\big(\mathcal{N}(u^{n,1}), v\big) + \kappa\epsilon\big(\mathcal{D}(u^{n,1}), v\big) = 0, \quad \forall v \in V_h^q,$$

$$u^{n+1} = u^n + 2(u^{n,1} - u^n).$$

κ is time stepsize.
Easy to see that $\|u^n\| = \|u^0\|$. Also can prove second order convergence in time

$$\|u(t_n) - u^n\| \le c(h^q + \kappa^2).$$

We have implemented two methods for solving the nonlinear systems

1. Newton's method
2. explicit-implicit iteration

$$(u^{n,1}_{\ell+1} - u^n, v) + \kappa\big(\mathcal{N}(u^{n,1}_{\ell}), v\big) + \kappa\epsilon\big(\mathcal{D}(u^{n,1}_{\ell+1}), v\big) = 0, \quad \forall v \in V_h^q,$$

All results hold under the assumprion that $\kappa h^{-3/2}$ is bounded.

Numerical Experiments

We use two test functions which are solutions of $u_t + uu_x + \epsilon u_{xxx} = 0$:

Cnoidal wave

$$u(x,t) = A\,\mathrm{cn}^2\left(4K(x - vt)\right),$$

1. $0 < m < 1$
2. $K = K(m)$ is the complete elliptic integral of the first kind
3. $A = 192 m\epsilon K^2$, $v = 64\epsilon(2m - 1)K^2$,
4. cn is a Jacobi elliptic function

In the experiments we have $m = 0.9$ and $\epsilon = 24^{-2}$.

Solitary wave

$$u(x,t) = A\,\mathrm{sech}^2\left(K(x - vt - x_0)\right)$$

A is the amplitude, $K = \frac{1}{2}\sqrt{\frac{A}{3\epsilon}}$, $x_0 = 1/2$.

In the experiments $A = 1$, $\epsilon = 24^{-2}$.

Also, we use a priodic version of u.

Convergence rates

Table 1: Convergence rates, the cnoidal wave, uniform mesh, $q = 2$.

	N	Δt	L^2 error	order	L^∞ error	order
C-C method	10	4.0E-02	1.3169E-00		1.9388E-00	
	20	1.0E-02	1.2735E-00	0.0483	2.1475E-00	-0.1475
	40	2.5E-03	1.7869E-01	2.8333	3.0294E-01	2.8256
	80	6.25E-04	1.2017E-02	3.8943	2.0728E-02	3.8694
	160	1.5625E-04	7.6271E-04	3.9778	1.3499E-03	3.9407
	320	3.90625E-05	4.8290E-05	3.9813	9.2342E-05	3.8697
NC-NC method	10	4.0E-02	7.1270E-01		1.2985E-00	
	20	1.0E-02	5.9638E-01	0.2571	1.1130E-00	0.2224
	40	2.5E-03	5.7218E-01	0.0598	1.0403E-00	0.0975
	80	6.25E-04	1.0466E-00	-0.8712	1.6738E-00	-0.6861
	160	1.5625E-04	2.0404E-01	2.3588	3.4832E-01	2.2646
	320	3.90625E-05	2.6643E-02	2.9370	4.5632E-02	2.9323

Table 2: Convergence rates, cnoidal wave, uniform mesh, $q = 3$.

	N	Δt	L^2 error	order	L^∞ error	order
C-C method	10	4.0E-02	1.2083E-00		2.1869E-00	
	20	1.0E-02	1.5809E-01	2.9342	3.5795E-01	2.6110
	40	2.5E-03	1.2153E-02	3.7014	3.3732E-02	3.4076
	80	6.25E-04	1.2048E-03	3.3344	3.3640E-03	3.3259
	160	1.5625E-04	1.3999E-04	3.1054	3.6877E-04	3.1894
NC-NC method	10	4.0E-02	9.7806E-01		1.6220E-00	
	20	1.0E-02	7.4734E-01	0.3882	1.2326E-00	0.3961
	40	2.5E-03	3.6619E-02	4.3511	6.2686E-02	4.2974
	80	6.25E-04	1.3171E-03	4.7972	2.2584E-03	4.7948
	160	1.5625E-04	4.8798E-05	4.7544	8.3729E-05	4.7534

Table 3: Convergence rates, cnoidal wave, uniform mesh and $q = 4$.

C-C method $\Delta t = C\Delta x^3$	10	4.0E-02	8.6715E-01		1.4038E-00	
	20	5.0E-03	7.3741E-03	6.8777	1.5963E-02	6.4585
	40	6.25E-04	2.2954E-04	5.0056	3.8679E-04	5.3670
	80	7.8125E-05	3.6186E-06	5.9872	6.1312E-06	5.9792
	160	9.765625E-06	5.6694E-08	5.9961	1.0019E-07	5.9354
NC-NC method	10	4.0E-02	1.1452E-00		1.7881E-00	
	20	5.0E-03	2.6086E-02	5.4562	4.4169E-02	5.3392
	40	6.25E-04	3.3303E-04	6.2914	5.6778E-04	6.2816
	80	7.8125E-05	4.4421E-06	6.2283	7.5900E-06	6.2251
	160	9.765625E-06	6.3137E-08	6.1366	1.0900E-07	6.1217

Convergence rates: Nonuniform mesh

Table 4: cnoidal wave, $q = 2$, non-uniform mesh of type $2\Delta x, \Delta x, \cdots, 2\Delta x, \Delta x$.

	N	Δt	L^2 error	order	L^∞ error	order
C-C method	10	4.0E-02	1.3340E-00		5.8547E-00	
	20	1.0E-02	9.1940E-01	0.5370	1.6786E-00	1.8023
	40	2.5E-03	6.1914E-01	0.5704	1.0938E-00	0.6179
	80	6.25E-04	2.3766E-01	1.3814	3.9930E-01	1.4538
	160	1.5625E-04	6.5006E-02	1.8703	1.1072E-01	1.8506
	320	3.90625E-05	1.6573E-02	1.9718	2.8665E-02	1.9496
NC-NC method	10	4.0E-02	6.8821E-01		1.2660E-00	
	20	1.0E-02	6.5336E-01	0.0750	1.2087E-00	0.0668
	40	2.5E-03	9.7878E-01	-0.5831	1.6384E-00	-0.4388
	80	6.25E-04	1.2109E-00	-0.3070	1.8813E-00	-0.1994
	160	1.5625E-04	3.2924E-01	1.8787	5.5988E-01	1.7485
	320	3.90625E-05	4.4494E-02	2.8875	7.6207E-02	2.8771

solution profiles

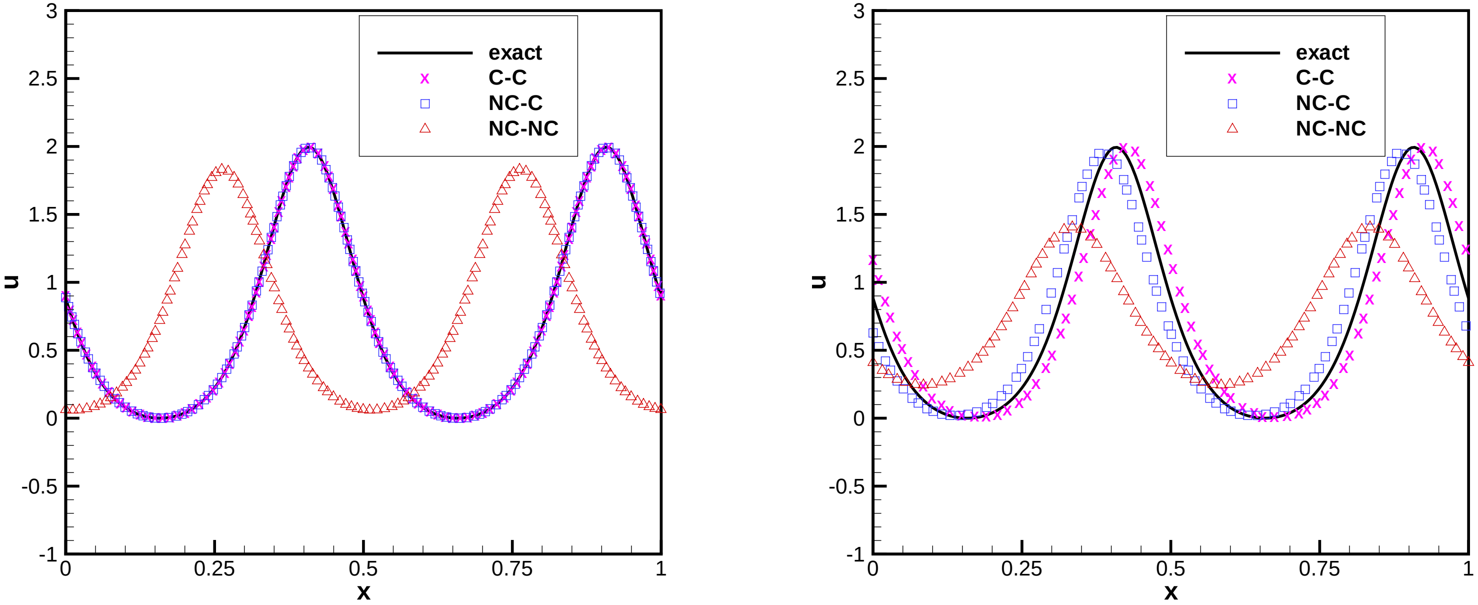


Figure 1: Comparison of C-C, NC-C and NC-NC methods, cnoidal wave problem $t = 10$ with $q = 2$. Left: 80 uniform cells; Right: 40 uniform cells.

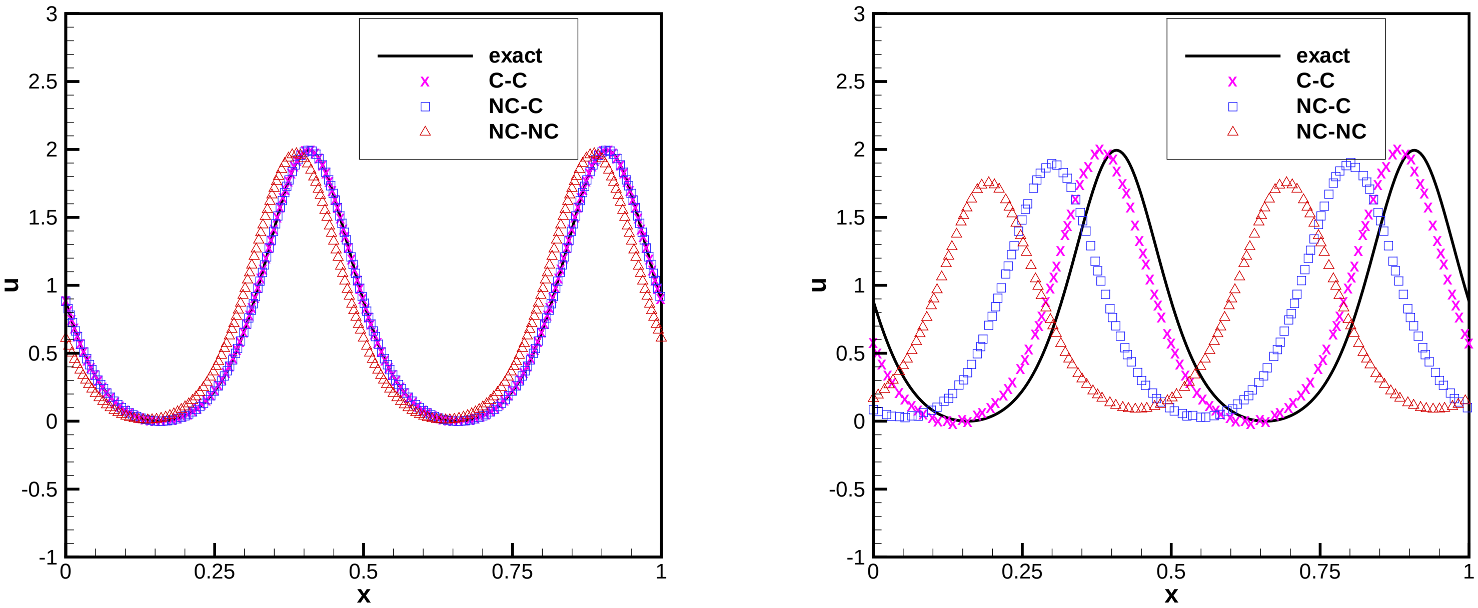


Figure 2: comparison of C-C, NC-C and NC-NC methods, cnoidal wave problem, time $t = 10$ with $q = 2$. Left: 160 uniform cells; Right: 80 nonuniform cells.

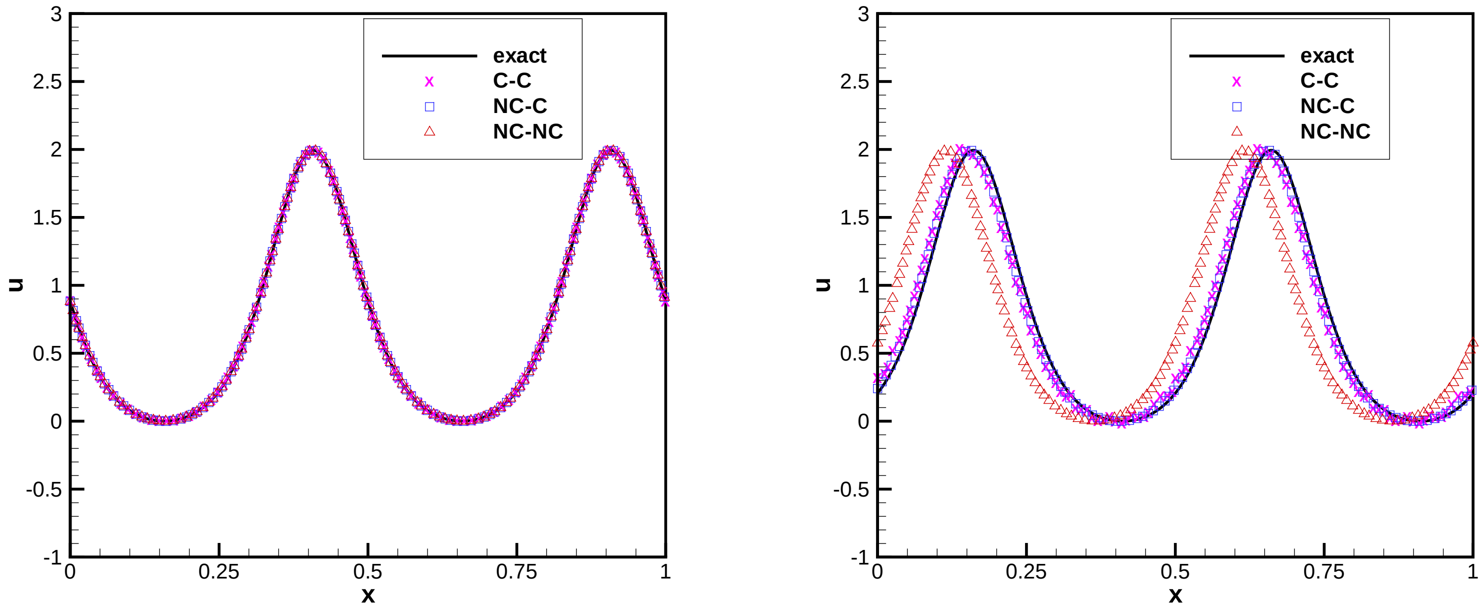


Figure 3: comparison of C-C, NC-C and NC-NC methods, cnoidal wave problem, $q = 3$ and 80 uniform cells. Left: time $t = 10$; Right: $t = 200$.

solitary wave solution

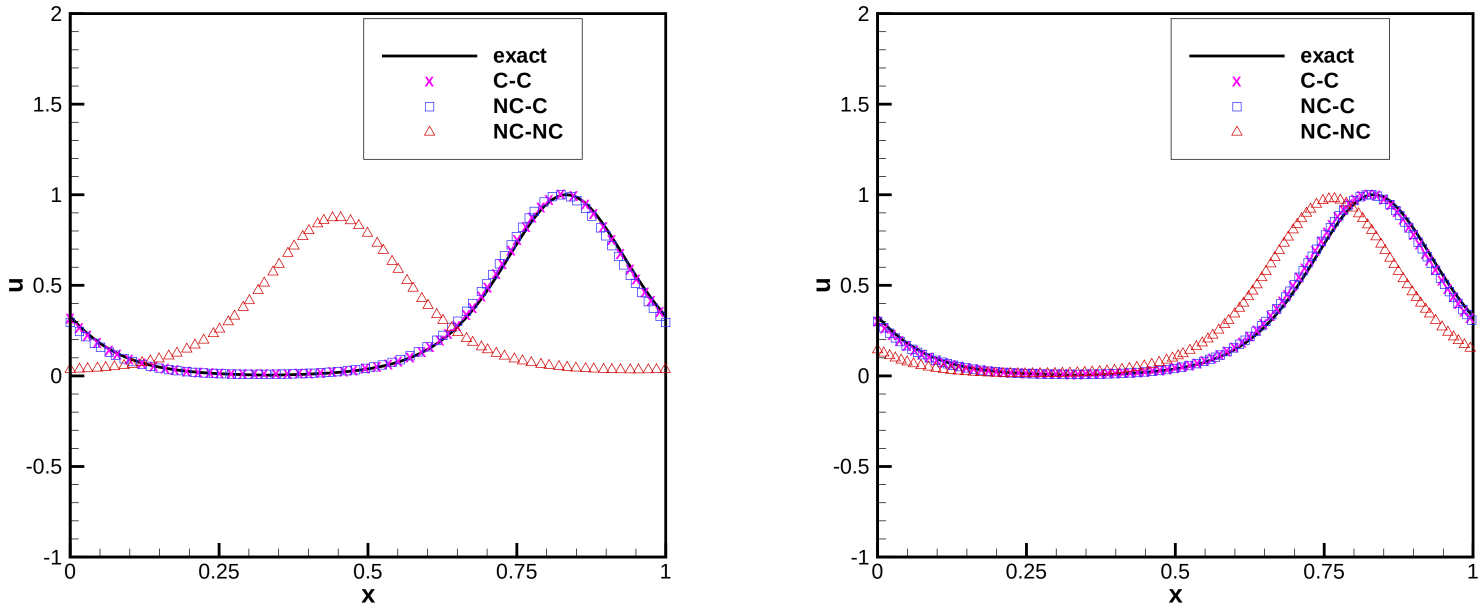


Figure 4: comparison of C-C, NC-C and NC-NC methods, solitary wave problem, $t = 25$ with $q = 2$. Left: 40 uniform cells; Right: 80 uniform cells.

Dependence of L^2 error on t

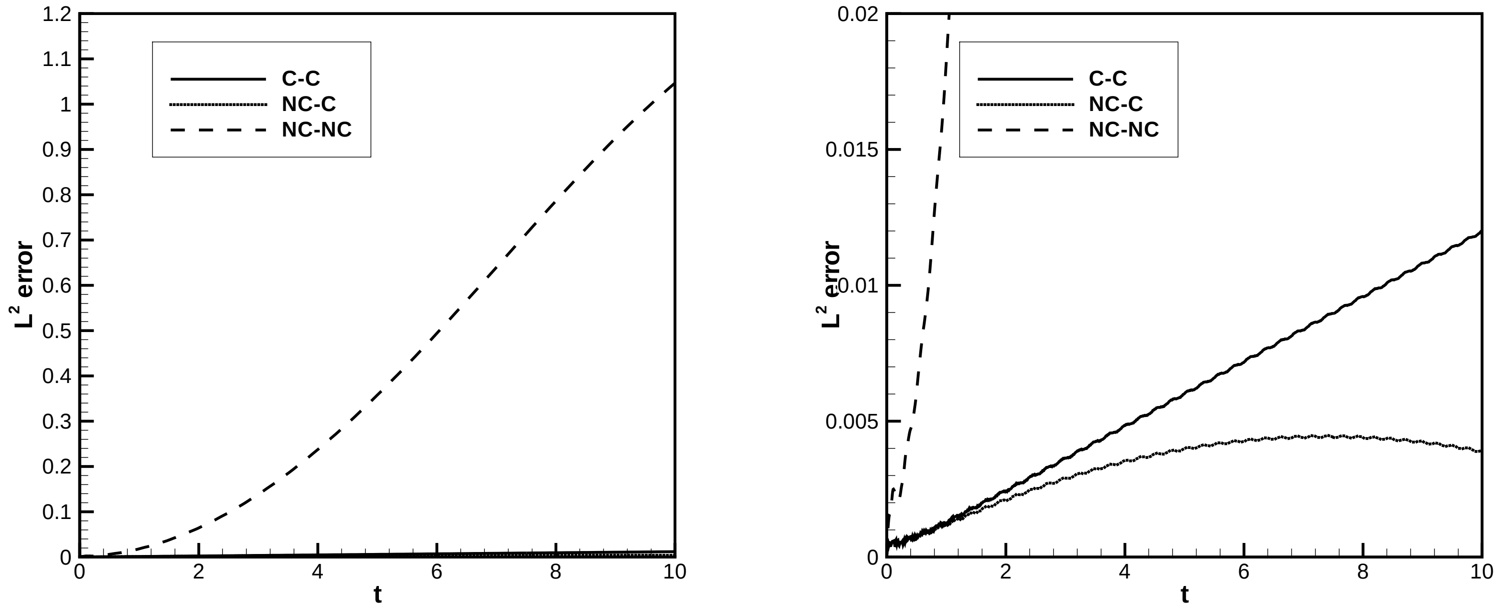


Figure 5: cnoidal wave problem, $q = 2$ and 80 uniform cells. Right graph is a zoom-in version of the left graph.

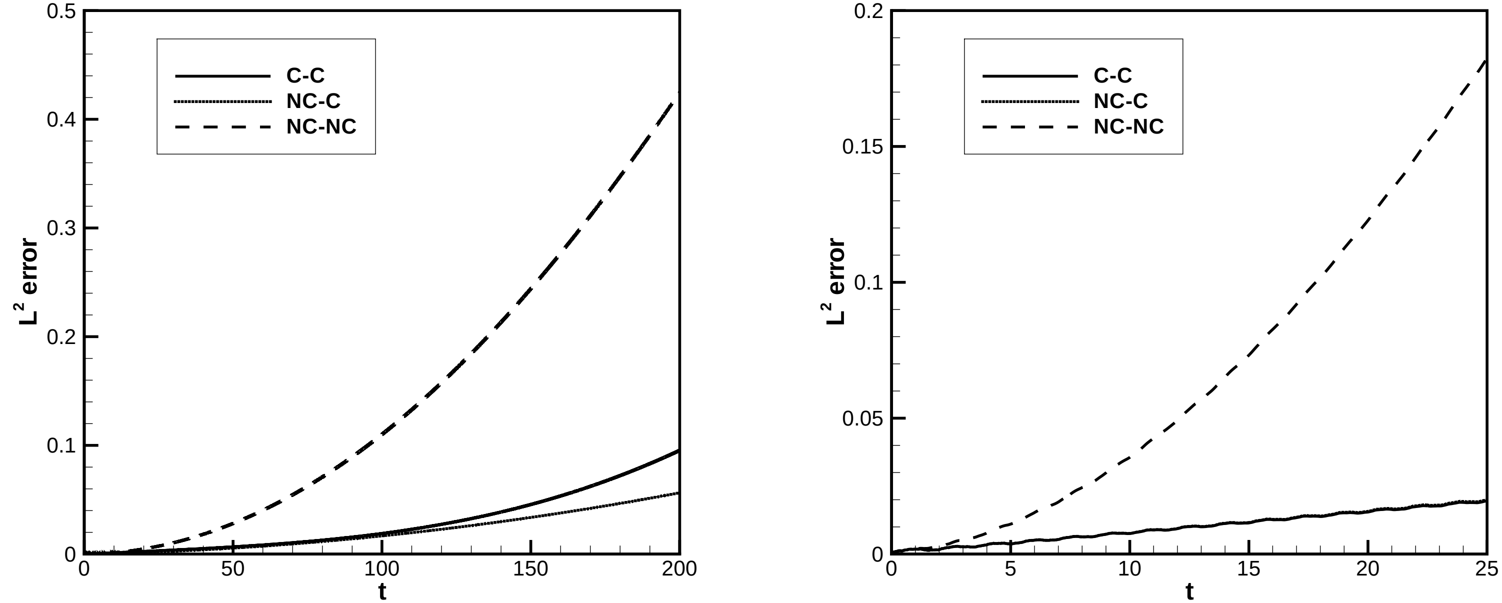


Figure 6: Left: cnoidal wave, $q = 3$ and 80 uniform cells; Right: solitary wave, $q = 2$ and 80 uniform cells.

Shape error

We define "Shape error" as $\hat{e}(x,t) = \min_{\xi \in [-0.5,0.5]} ||u_h(x,t) - u(x+\xi,t)||$

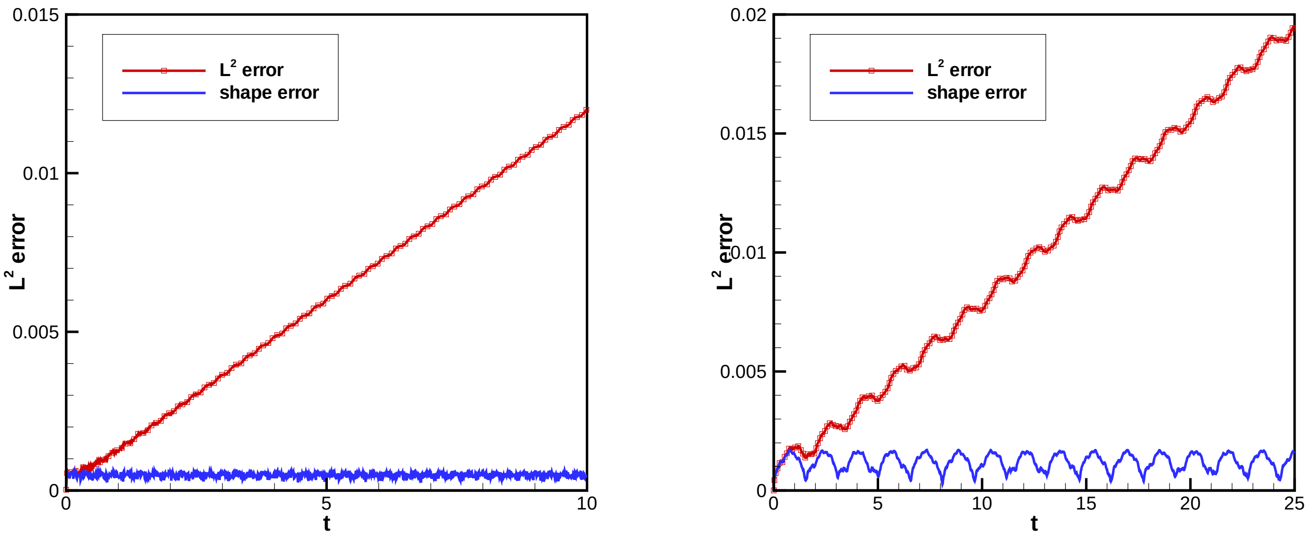


Figure 7: Time history of L^2 error and shape error of conservative method with $q = 2$ and 80 uniform cells. Left: the cnoidal wave problem, Right: the solitary wave

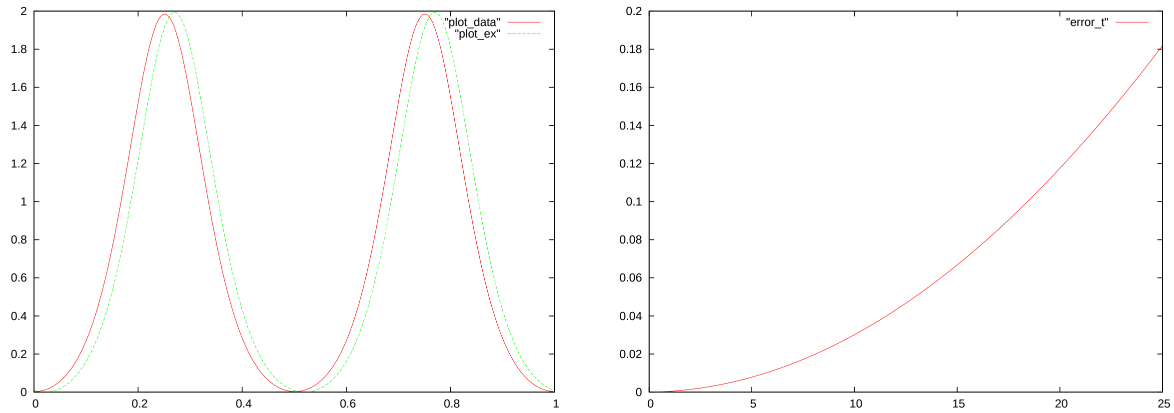


Figure 8: Dissipative method with $q = 3$ and 41 uniform cells.

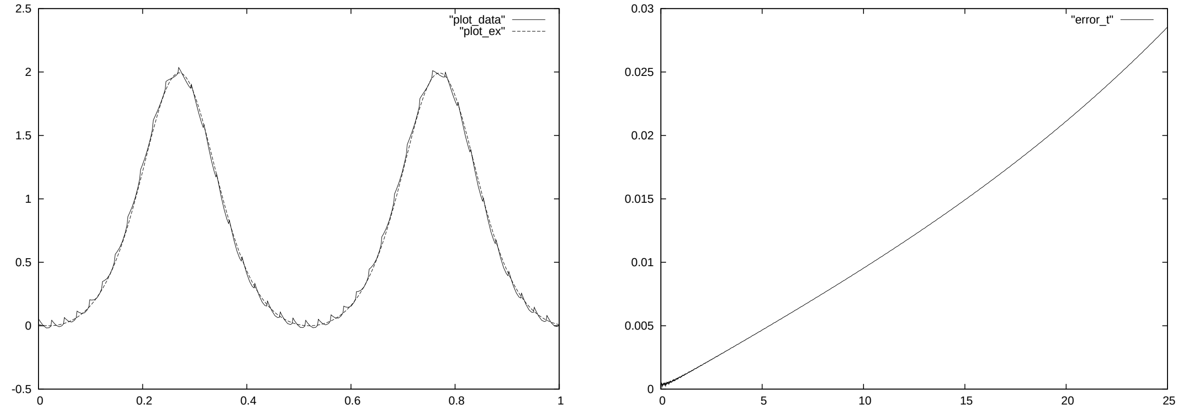


Figure 9: Conservative method with $q = 3$ and 41 uniform cells.

Finite-time blowup

$p = 5$

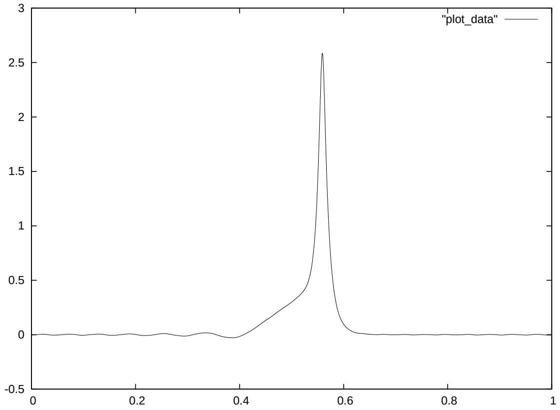
Initial profile: Perturbed solitary wave, amplitude=2.02

Adaptive code, refinement/coarsening for spatial mesh

adaptive stepsize selection

Initial stepsize $\Delta t = 10^{-4}$

M = 200

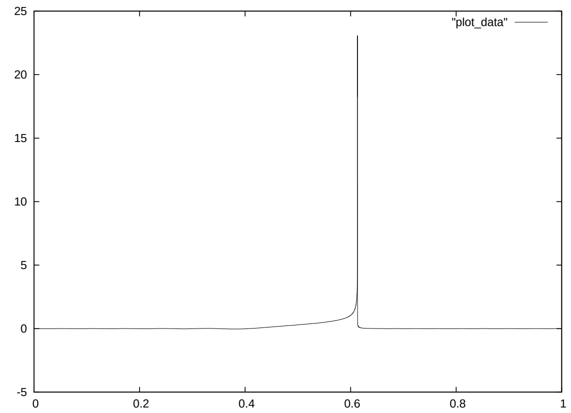


$t = .02$

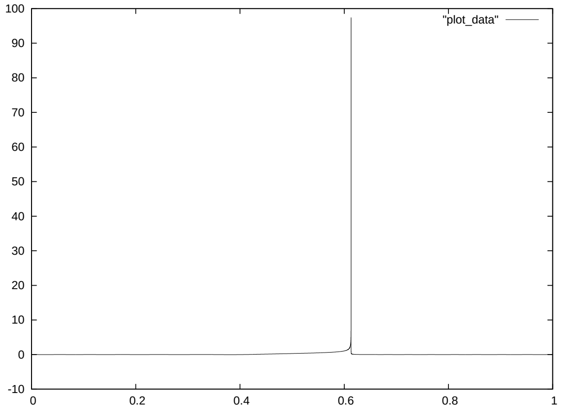
amplitude = 2.5

last $\Delta t = 10^{-4}$

M = 337



$t = .0225542$
amplitude = 23,
last $\Delta t = 1.2 \times 10^{-11}$
M = 477



$t = 0.0225542$
amplitude = 96
last $\Delta t = 1.8 \times 10^{-16}$
M = 556