Directional projection with low dispersion error

SCICADE - July - 2011

J.I. Montijano

Joined work with
M. Calvo, P. Laburta and L. Rández

IUMA – Zaragoza (Spain)

Projection methods

Consider a differential system with a scalar invariant $F(y)$

$$G(y(t)) \equiv F(y) - F(y(t_0)) = 0 \quad \text{for all } t$$

- Given a numerical method that provides an approximation $\tilde{y}_{n+1}$ to the solution $y(t_{n+1})$, find the parameter λ_n such that

  $$G(\tilde{y}_{n+1} + v_n \lambda_n) = 0$$

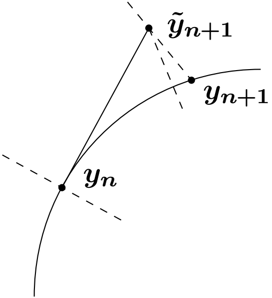
  with v_n some projection direction. The new (projected) approximation is

  $$y_{n+1} = \tilde{y}_{n+1} + v_n \lambda_n$$

- The direction vector v_n defines the projection method.

Simplified Orthogonal projection (Hairer et al.)

$$v_n = \nabla G(\tilde{y}_{n+1})^T,$$
$$G(\tilde{y}_{n+1} + \lambda_n \nabla G(\tilde{y}_{n+1})^T) = 0$$



Only a scalar nonlinear equation must be solved.

Runge–Kutta projection

Runge–Kutta schemes

$$Y_i = y_n + h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1, \dots, s$$
$$\tilde{y}_{n+1} = y_n + h \sum_{i=1}^{s} b_i f(Y_i)$$

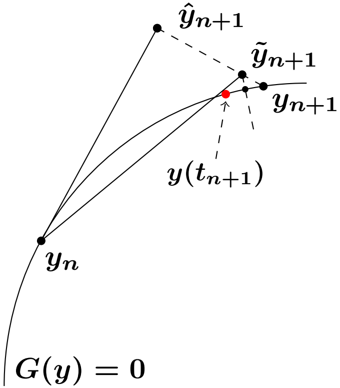
Direction of projection (based on an embedded RK method)

$$\hat{y}_{n+1} = y_n + h \sum_{i=1}^{s} \hat{b}_i f(Y_i)$$
$$v_n = \hat{y}_{n+1} - \tilde{y}_{n+1} = h \sum_{i=1}^{s} (\hat{b}_i - b_i) f(Y_i)$$
$$y_{n+1} = \tilde{y}_{n+1} + \lambda_n(\hat{y}_{n+1} - \tilde{y}_{n+1}) = (1 - \lambda_n)\tilde{y}_{n+1} + \lambda_n \hat{y}_{n+1}$$

y_{n+1} is a convex linear combination of Runge–Kutta approximations, which is also a Runge–Kutta method.

Runge–Kutta projection

$$y_{n+1} = \tilde{y}_{n+1} + \lambda_n(\hat{y}_{n+1} - \tilde{y}_{n+1}) = (1 - \lambda_n)\tilde{y}_{n+1} + \lambda_n \hat{y}_{n+1}$$



Runge–Kutta projection

Main results

Let $\tilde{y}_{n+1}$ and $\hat{y}_{n+1}$ have orders p and $q(< p)$ respectively and let $C_{q+1}(y_n)h^{q+1}$ the leading term of error of the embedded method.

- If $\nabla G(y_n) \cdot C_{q+1}(y_n) \neq 0$, for h small enough, there exist $\lambda_n = \lambda(y_n, h)$ and $y(t_{n+1}) - y_{n+1} = \mathcal{O}(h^{p+1})$.
- If $\nabla G(y_n) \cdot v_n = B(y_n)h^r + \mathcal{O}(h^{r+1})$, $r \geq 1$, $B(y_n) \neq 0$ and $2r \leq p$, then for h small enough there exist $\lambda_n = \lambda(y_n, h)$ and $y(t_{n+1}) - y_{n+1} = \mathcal{O}(h^{p-r+1})$
- The projected Runge–Kutta method preserves all linear first integrals
- The projected method is affine invariant

SIAM J. Sci. Comput. 28 (2006), no. 3, 868–885.

Runge–Kutta projection

The direction of projection

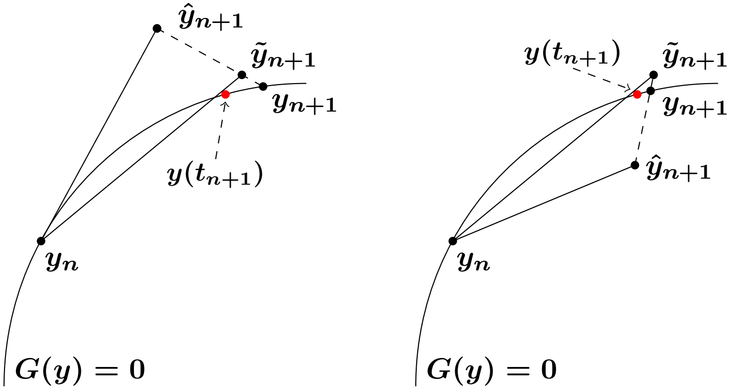
The projected approximation depends on the embedded method $\hat{y}_{n+1}$.

Can we chose it in an optimal way?

- Minimizing the error $|y(t_{n+1}) - y_{n+1}|$
- Ensuring the existence of the projection

Runge–Kutta projection

$$y_{n+1} = \tilde{y}_{n+1} + \lambda_n(\hat{y}_{n+1} - \tilde{y}_{n+1}) = (1 - \lambda_n)\tilde{y}_{n+1} + \lambda_n \hat{y}_{n+1}$$



Low dispersion error RK projection

Desired properties for $\hat{y}_{n+1}$:

- $G(\hat{y}_{n+1})$ and $G(\tilde{y}_{n+1})$ have opposite signs.
- If $\tilde{y}_{n+1}$ advances the phase with respect to $y(t_{n+1})$, then $\hat{y}_{n+1}$ must delay it, and conversely.

Low dispersion error RK projection

Linear problem

$$y' = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix} y \qquad \longleftrightarrow \qquad x' = i\omega x$$

$$G(y) = \|y\|^2 - \|y(0)\|^2 \quad \longleftrightarrow \quad g(x) = |x| - |x(0)|$$

This equivalence is valid if the method is affine invariant !

Linear problem

$$Y_i = y_n + h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1, \ldots, s$$

$$\tilde{y}_{n+1} = y_n + h \sum_{i=1}^{s} b_i f(Y_i)$$

$$\hat{y}_{n+1} = y_n + h \sum_{i=1}^{s} \hat{b}_i f(Y_i)$$

Denoting $z = i\omega h$,

$$\tilde{y}_{n+1} = \tilde{R}(z) y_n, \quad \tilde{R}(z) = 1 + z + b^T c\, z^2 + \ldots + b^T A^{s-2} c\, z^s$$

$$\hat{y}_{n+1} = \hat{R}(z) y_n, \quad \hat{R}(z) = 1 + z + \hat{b}^T c\, z^2 + \ldots + \hat{b}^T A^{s-2} c\, z^s$$

Low dispersion error RK projection

Linear problem

$$y(t_{n+1}) = y_0 e^{i\omega t_{n+1}} = e^{hi\omega} y(t_n) \quad \longleftrightarrow \quad \tilde{y}_{n+1} = \tilde{R}(ih\omega) y_n$$

Local error: $\tilde{y}_{n+1} - y(t_{n+1}; t_n, y_n) = (\tilde{R}(ih\omega) - e^{hi\omega}) y_n$

Components:

Modulus: $|\tilde{y}_{n+1}| - |y(t_{n+1}; t_n, y_n)| = |y_n|(|\tilde{R}(z)| - 1)$

Phase: $\arg \tilde{y}_{n+1} - \arg y(t_{n+1}; t_n, y_n) = \arg \tilde{R}(hi\omega) - h\omega$

- Dissipation error: $\tilde{d}(\nu) = |\tilde{R}(i\nu)| - 1$
- Dispersion error: $\tilde{\varphi}(\nu) = \arg \tilde{R}(i\nu) - \nu$

Low dispersion error RK projection

Linear problem

$$G(\tilde{y}_{n+1}) > 0 \quad \longleftrightarrow \quad \tilde{d}(h\omega) = |\tilde{R}(ih\omega)| - 1 > 0$$

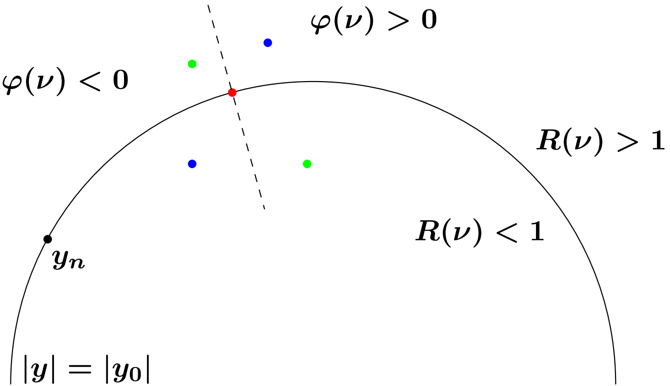
$$\tilde{y}_{n+1} \text{ advances in phase} \quad \longleftrightarrow \quad \tilde{\varphi}(h\omega) > 0$$

We search for an embedded method $\hat{y}_{n+1}$ such that

$$\tilde{d}(\nu)\hat{d}(\nu) < 0, \qquad \tilde{\varphi}(\nu)\hat{\varphi}(\nu) < 0,$$

for $\nu \in \mathbb{R}$ small enough.

Low dispersion error RK projection



Low dispersion error RK projection

3 stages, order 3

$$\tilde{R}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6}$$
$$|\tilde{R}(i\nu)|^2 - 1 = -\frac{1}{12}\nu^4 + \frac{1}{36}\nu^6$$
$$\tilde{\varphi}(\nu) = \frac{1}{30}\nu^5 - \frac{1}{252}\nu^7 + \mathcal{O}(\nu^9)$$

We want a first order method, 3 stages

$$\hat{R}(z) = 1 + z + \alpha z^2 + \beta z^3$$

In this case

$$|\hat{R}(i\nu)|^2 - 1 = (1 - 2\alpha)\nu^2 + (\alpha^2 - 2\beta)\nu^4 + \beta^2\nu^6$$
$$\hat{\varphi}(\nu) = \left(-\frac{1}{3} + \alpha - \beta\right)\nu^3 + \left(\frac{1}{5} - \alpha + \alpha^2 + \beta - \alpha\beta\right)\nu^5 + \mathcal{O}(\nu^7)$$

Low dispersion error RK projection

3 stages, order 3

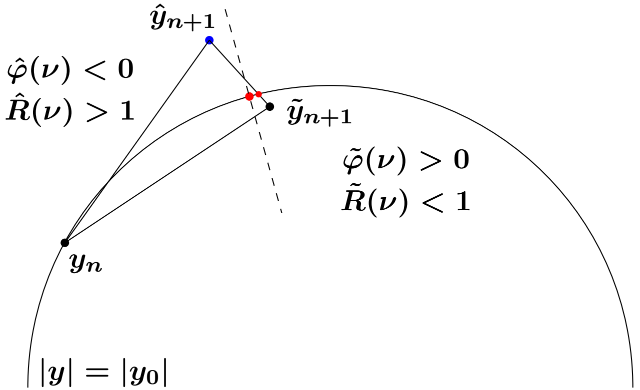
The embedded method of order 1 with 3 stages must satisfy

$$\mathbf{1 - 2\alpha > 0,} \quad \text{and} \quad \mathbf{1/3 - \alpha + \beta > 0}$$

- $\alpha = \beta = 0$ Euler method
- $\beta = 0, \quad \alpha < 1/3$ 2 stages method
- $\alpha < 1/2, \quad \beta > \alpha - 1/3$ 3 stages method

Can we chose the parameters α, β in some way ?

Low dispersion error RK projection



Low dispersion error RK projection

3 stages, order 3

The projected approximation is

$$y_{n+1} = (1-\lambda_n)\tilde{y}_{n+1} + \lambda_n \hat{y}_{n+1}$$

with λ_n the solution of

$$|R(i\nu)|^2 = |(1-\lambda_n)\tilde{R}(i\nu) + \lambda_n \hat{R}(i\nu)|^2 = 1$$

Solving for λ_n and substituting into $\varphi(\nu)$

$$\varphi(\nu) = \frac{1+3\alpha-15\beta}{180(1-2\alpha)}\nu^5 + \frac{1-24\alpha^2-4\alpha+168\alpha\beta-252\beta^2}{1512(1-2\alpha)^2}\nu^7 + \mathcal{O}(\nu^9)$$

We can chose α, β such that the projected method has dispersion error of order 6

$$\beta = (1+3\alpha)/15 \quad \Rightarrow \quad \varphi(\nu) = \mathcal{O}(\nu^7)$$

Low dispersion error RK projection

3 stages, order 3

- $\tilde{R}(z) = 1 + z + z^2/2 + z^3/6$
- $\hat{R}(z) = 1 + z + \alpha z^2 + (1 + 3\alpha)z^3/15$
- $|R(i\nu)| = 1 \Rightarrow \varphi(\nu) = -\frac{1}{12600}\nu^7 + \mathcal{O}(\nu^9)$

Conditions on the coefficients of the method

$$\begin{aligned} \hat{b}^T e &= \hat{b}_1 + \hat{b}_2 + \hat{b}_3 = 1 \\ \hat{b}^T c &= \hat{b}_2 c_2 + \hat{b}_3 c_3 = \alpha \\ \hat{b}^T A c &= \hat{b}_3 a_{32} c_2 = (1 + 3\alpha)/15 \end{aligned}$$

Bogacki and Shampine (Matlab pair 2(3))

$$Y_i = y_n + h \sum_{j=1}^{3} a_{ij} f(Y_j), \quad i = 1, \ldots, 3$$

$$\tilde{y}_{n+1} = y_n + h \sum_{i=1}^{3} b_i f(Y_i)$$

$$\hat{y}_{n+1} = y_n + h \sum_{i=1}^{3} \hat{b}_i f(Y_i)$$

$$\begin{array}{c|c} c & A \\ \hline p=3 & b^T \\ \hline q=1 & \hat{b}^T \end{array} \quad = \quad \begin{array}{c|ccc} 0 & & & \\ 1/2 & 1/2 & & \\ 3/4 & 0 & 3/4 & \\ \hline & 2/9 & 1/3 & 4/9 \\ \hline & \hat{b}_1 & \hat{b}_2 & \hat{b}_3 \end{array}$$

Low dispersion error RK projection

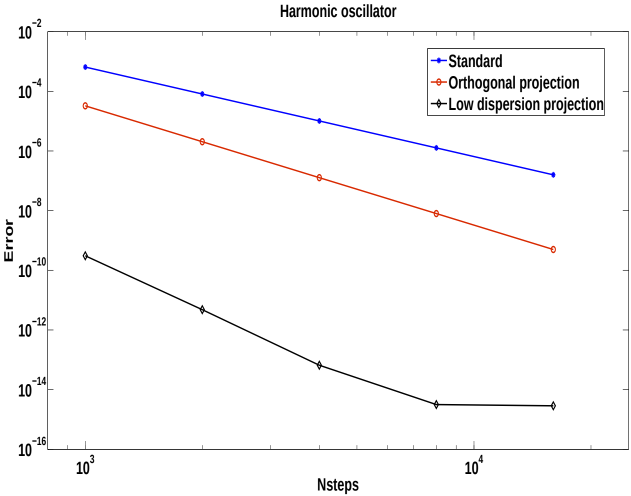
3 stages, order 3

- $b = (2/9, 1/3, 4/9)^T$
- $\hat{b} = (\hat{b}_1, \hat{b}_2, \hat{b}_3)^T$
  - $\hat{b}^T e = \hat{b}_1 + \hat{b}_2 + \hat{b}_3 = 1$
  - $\hat{b}^T c = \hat{b}_2 1/2 + \hat{b}_3 3/4 = \alpha$
  - $\hat{b}^T Ac = \hat{b}_3 (3/4)(1/2) = (1 + 3\alpha)/15$

- $\hat{b}_3 = 8(1 + 3\alpha)/45$
- $\hat{b}_2 = 2\alpha - 3\hat{b}_3/2$
- $\hat{b}_1 = 1 - \hat{b}_2 - \hat{b}_3$

For $\alpha = -1/3, \quad \hat{b} = (5/3, -2/3, 0)^T$

A Numerical experiment



$$y_1' = y_2$$

$$y_2' = -y_1$$

$$t \in [0, 20\pi]$$

Low dispersion error RK projection

General problems

$$y' = f(t, y), \quad y(t_0) = y_0, \qquad G(y(t)) = 0$$

$$\tilde{d}(\nu)\hat{d}(\nu) < 0 \quad \not\Rightarrow \quad G(\tilde{y}_{n+1})G(\hat{y}_{n+1}) < 0$$

Can we find $\hat{y}_{n+1}$ such that $G(\tilde{y}_{n+1})G(\hat{y}_{n+1}) < 0$?

- We know $G(y_n) = 0$,
- $G(\tilde{y}_{n+1})$ can be evaluated,
- $\tilde{y}_{n+1}$ has order p,
- $\hat{y}_{n+1}$ has order q,
- $G(\hat{y}_{n+1})$ depends on the free parameters (α, β) !.

Low dispersion error RK projection

$$G(\hat{y}_{n+1}) = G(\tilde{y}_{n+1} + \hat{y}_{n+1} - \tilde{y}_{n+1}) =$$
$$G(\tilde{y}_{n+1}) + \nabla G(\tilde{y}_{n+1})(\hat{y}_{n+1} - \tilde{y}_{n+1}) + \mathcal{O}(h^{2(q+1)}) =$$
$$\nabla G(\tilde{y}_{n+1})(\hat{y}_{n+1} - \tilde{y}_{n+1}) + \mathcal{O}(h^{2(q+1)}) + \mathcal{O}(h^{p+1})$$

We look for $\hat{y}_{n+1}$ such that

$$G(\tilde{y}_{n+1})\nabla G(\tilde{y}_{n+1})(\hat{y}_{n+1} - \tilde{y}_{n+1}) < 0$$

Bogacki–Shampine

$$\tilde{y}_{n+1} = (2/9)f(Y_1) + (3/9)f(Y_2) + (4/9)f(Y_3)$$
$$\hat{y}_{n+1} = \hat{b}_1 f(Y_1) + \frac{1}{2}(\alpha - 8\beta)f(Y_2) + \frac{8}{3}\beta f(Y_3)$$

Low dispersion error RK projection

Bogacki–Shampine

$$\nabla G(\tilde{y}_{n+1})(\hat{y}_{n+1} - \tilde{y}_{n+1}) = \sum_{i=1}^{3}(\hat{b}_i - b_i)K_i$$
$$K_i = \nabla G(\tilde{y}_{n+1})f(Y_i)$$

If $\beta = (1 + 3\alpha)/15$

$$\nabla G(\tilde{y}_{n+1})(\hat{y}_{n+1} - \tilde{y}_{n+1}) = a(\alpha - \frac{1}{2})$$
$$a = -2(13K_1 - 9K_2 - 4K_3)/15$$

$$G(\tilde{y}_{n+1})a > 0 \Rightarrow G(\tilde{y}_{n+1})G(\hat{y}_{n+1}) < 0$$

for all $\alpha < 1/2$ and h small enough

Low dispersion error RK projection

Bogacki–Shampine

- $G(\tilde{y}_{n+1})(-2(13K_1 - 9K_2 - 4K_3)/15) > 0$

  $\alpha < 1/2$ and $\beta = (1+3\alpha)/15 \quad \Rightarrow \quad G(\tilde{y}_{n+1})G(\hat{y}_{n+1}) < 1$

- $G(\tilde{y}_{n+1})(-2(13K_1 - 9K_2 - 4K_3)/15) < 0$

  There exist α and β such that

  - $\nabla G(\tilde{y}_{n+1})(\hat{y}_{n+1} - y_{n+1}) = \sum_{i=1}^{3}(\hat{b}_i - b_i)K_i < 0$
  - $\tilde{\varphi}(\nu)\hat{\varphi}(\nu) < 0$
  - $\tilde{d}(\nu)\hat{d}(\nu) < 0$
  - Incompatible with $\beta = (1+3\alpha)/15$

Numerical experiments: Free rigid body

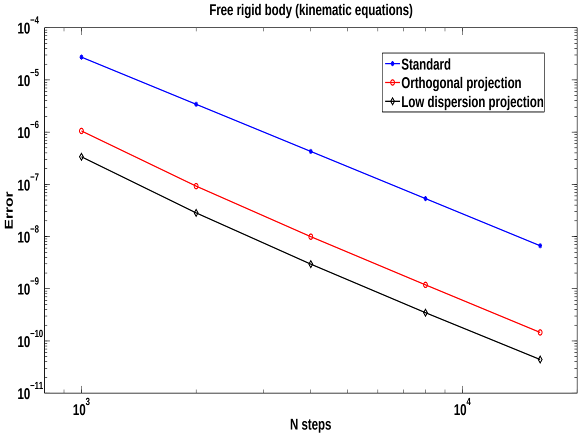
Kinematic equations in terms of quaternions

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & -w_1(t) & -w_2(t) & -w_3(t) \\ w_1(t) & 0 & w_3(t) & -w_2(t) \\ w_2(t) & -w_3(t) & 0 & w_1(t) \\ w_3(t) & w_2(t) & -w_1(t) & 0 \end{pmatrix}\begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

$w_i(t)$ solutions of Euler's equations

First integral: $q(t)^T q(t) = \|q(t)\|^2$

Numerical experiments: Free rigid body



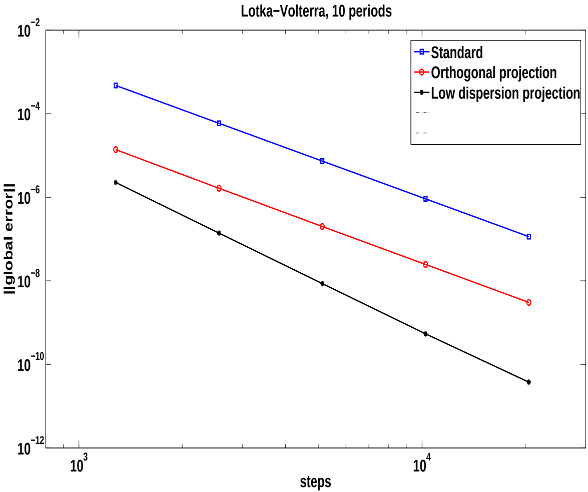
$$q(t)' = M(w(t))q(t)$$

$$G(q) = \|q\|^2 - 1$$

$$t \in [0, 46.6]$$

Integrated in the quaternions form, with exact angular velocities $w(t)$

Numerical experiments: Lotka-Volterra



$$y_1' = y_1(y_2 - 2)$$

$$y_2' = y_2(1 - y_1)$$

$$y_1(0) = 1$$

$$y_2(0) = 1$$

$$t \in [0, 46.6]$$

$$G(y) = \log y_1 - y_1 + 2\log y_2 - y_2 + 2$$

Numerical experiments: Wave equation

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R}, \quad t \geq 0,$$
$$u(x,0) = \Phi(x), \quad \lim_{|x|\to\infty} \Phi(x) = 0$$

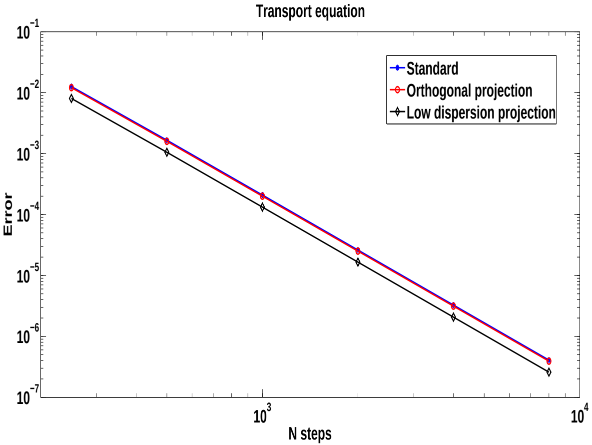
Symmetric spatial discretization ($a_l = -a_{-l}$)

- $\dfrac{\partial u(x_j,t)}{\partial x} \simeq \dfrac{1}{\Delta x}(a_{-N}u(x_{j-N},t) + \ldots + a_N u(x_{j+N},t)),$
- $\dfrac{\mathrm{d}u_j(t)}{\mathrm{d}t} + \dfrac{1}{\Delta x}\sum_{l=-N}^{N} a_l u_{j+l}(t) = 0,$

Invariants

- $\sum_j u_j(t) =$ (almost) constant
- $\sum_j u_j(t)^2 =$ constant

Numerical experiments: Wave equation



$$u' = Au$$

$$u_j(0) = \frac{e^{-x_j^2/9}}{2}$$

$$t \in [0, 80]$$

$$G(u) = \|u\|_2^2 - \|u(0)\|_2^2$$

Numerical experiments: Wave equation

Method	$G(u)$		
	$h = 0.32$	$h = 0.08$	$h = 0.02$
No projection	-7.15×10^{-3}	-1.18×10^{-4}	-1.86×10^{-6}
Orth. proj.	0	0	0
Low disp. proj.	0	0	0

Method	$\sum_j u_j - \sum_j u_j(0)$		
	$h = 0.32$	$h = 0.08$	$h = 0.02$
No projection	6.50×10^{-10}	5.99×10^{-8}	6.20×10^{-8}
Orth. proj.	1.02×10^{-2}	1.67×10^{-4}	2.69×10^{-6}
Low disp. proj.	2.80×10^{-10}	6.00×10^{-8}	6.20×10^{-8}

Thank you
for
your attention

SCICADE 2011