



# Contents

<b>List of Figures</b>	<b>iii</b>
<b>1 Theory</b>	<b>1</b>
1.1 Finite differences method . . . . .	1
1.1.1 Energy conservative scheme of order 2 (Hc2) . . . . .	1
1.1.2 Energy conservative scheme of order 4 (Hc4) . . . . .	2
1.1.3 Compact finite difference schemes with spectral-like resolution . . . . .	2
1.1.4 Nonlinear discretization schemes . . . . .	3
1.2 Finite element method . . . . .	5
1.2.1 Linear Lagrange finite element P1 . . . . .	5
1.2.2 Cubic Lagrange finite element P3 . . . . .	6
1.2.3 Cubic Hermite finite element H3 . . . . .	8
1.2.4 Fifth order Hermite finite element H5 . . . . .	9
1.3 Time integration . . . . .	11
1.4 Dispersive error . . . . .	11
1.4.1 Modified wavenumbers for the different spatial discretizations . . . . .	12
1.4.2 Result . . . . .	14
<b>2 Numerical results</b>	<b>19</b>
2.1 Shock formation from a sinus wave . . . . .	19
2.1.1 Physical behaviour . . . . .	19
2.1.2 Convergence study . . . . .	21
2.1.3 A note on the maximum time step . . . . .	22
2.2 Turbulent flow . . . . .	23
2.2.1 Under-resolved DNS . . . . .	23
2.2.2 Resolved DNS . . . . .	24
<b>Bibliography</b>	<b>28</b>

# List of Figures

1.1	Shape functions for the linear Lagrange element . . . . .	6
1.2	Shape functions for the cubic Lagrange element . . . . .	7
1.3	Shape functions for the cubic Hermite element . . . . .	8
1.4	Shape functions for the fifth order Hermite element . . . . .	10
1.5	(a)Modified numerical wavenumber $\eta_{num}$ as a function of $\eta$ for the Hermite elements H3. (b)-(d) Eigenvectors of Hermite H3 for $K = 0, 1$ and $2$ . . . . .	15
1.6	(a)Modified numerical wavenumber $\eta_{num}$ as a function of $\eta$ for the Hermite elements H5. (b)-(d) Eigenvectors of Hermite H5 for $K = 0, 1$ and $2$ (the parasite modes have been magnified by a factor 100 for visibility). . . . .	16
1.7	(top left) Modified numerical wavenumber $\eta_{num}$ as a function of $\eta$ , (top right) the relative error between the two wavenumbers and (bottom) the phase velocity. . .	17
2.1	Formation of a shock from a sinusoidal initial condition . . . . .	20
2.2	Formation of a shock from a sinusoidal initial condition - Convergence curves for three values of the viscosity ( $\nu = 0.01, 0.1$ and $1m^2/s$ ). . . . .	21
2.3	Under-resolved DNS - Energy spectra for (left) $\nu = 0.0035m^2/s$ and (right) $\nu = 0.0075m^2/s$ . . . . .	25
2.4	Under-resolved DNS - Energy spectra for (left) $\nu = 0.0035m^2/s$ and (right) $\nu = 0.0075m^2/s$ . . . . .	26
2.5	Resolved DNS - Matrix size=2048 - Energy spectra for (left) $\nu = 0.0035m^2/s$ and (right) $\nu = 0.0075m^2/s$ . . . . .	27

# Chapter 1

## Theory

1D Burgers equation with forcing term:

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) = \nu \frac{\partial^2 u}{\partial x^2} + f(x, t) \quad (1.1)$$

This equation is solved with a periodic condition in a domain of length  $L$ . The forcing term is also periodic with a random phase  $\Phi$  (white noise) and acts on the wave numbers  $k=2$  and  $3$ . This nodal forcing term is

$$f(x, t) = \sqrt{\Delta t} (\cos(2k_0 x + \Phi_2) + \cos(3k_0 x + \Phi_3)) \quad (1.2)$$

with  $k_0 = 2\pi/L$  and  $\Delta t$  is the increment in time used in the discretization.

### 1.1 Finite differences method

#### 1.1.1 Energy conservative scheme of order 2 (Hc2)

Centered finite difference schemes of order 2 are used for the first and second derivatives. The second derivative is computed through

$$\frac{\partial^2 u}{\partial x^2} \Big|_{x=x_i} \simeq \frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} \quad (1.3)$$

The nonlinear convective term is discretized by the skew-symmetric form  $Hc2 = (2Hd2 + Ha2)/3$ , which is a combination of two energy non-conservative schemes. These two forms are based on respectively the divergence (Hd2) and advective (Ha2) formulation of the nonlinear term. Both these schemes are based on a second order centered finite difference scheme for the first derivative. These three schemes are given by

$$\text{Divergent form Hd2 :} \quad \frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} \simeq \frac{u_{i+1}^2 - u_{i-1}^2}{4\Delta x} \quad (1.4)$$

$$\text{Advective form Ha2 :} \quad \frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} = u \frac{\partial u}{\partial x} \Big|_{x=x_i} \simeq u_i \frac{u_{i+1} - u_{i-1}}{2\Delta x} \quad (1.5)$$

$$\text{Skew-symmetric form Hc2 :} \quad \frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} \simeq \frac{(u_{i+1} + u_i + u_{i-1})(u_{i+1} - u_{i-1})}{6\Delta x} \quad (1.6)$$

It can be shown that both the Hd2 and Ha2 schemes are not energy conservative by multiplying the nonlinear term at node  $i$  by the nodal velocity value  $u_i$  at the same node and summing up

all the contributions. The energy production by the nonlinear term for the divergent form Hd2 is

$$\begin{aligned}
E_{\text{Hd2}}^{\text{nonlinear}} &= \dots + h \frac{u_{i-1}}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_{i-1}} + h \frac{u_i}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} + h \frac{u_{i+1}}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_{i+1}} + \dots \\
&= \frac{h}{4\Delta} (\dots + u_{i-1}(u_i^2 - u_{i-2}^2) + u_i(u_{i+1}^2 - u_{i-1}^2) + u_{i+1}(u_{i+2}^2 - u_i^2) + \dots) \\
&= \frac{h}{4\Delta} (\dots + u_i u_{i-1}(u_i - u_{i-1}) + u_{i+1} u_i(u_{i+1} - u_i) + \dots)
\end{aligned} \tag{1.7}$$

On the other hand, the energy production by the nonlinear term for the advective form Ha2 is

$$\begin{aligned}
E_{\text{Ha2}}^{\text{nonlinear}} &= \dots + h \frac{u_{i-1}}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_{i-1}} + h \frac{u_i}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} + h \frac{u_{i+1}}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_{i+1}} + \dots \\
&= \frac{h}{2\Delta} (\dots + u_{i-1}^2(u_i - u_{i-2}) + u_i^2(u_{i+1} - u_{i-1}) + u_{i+1}^2(u_{i+2} - u_i) + \dots) \\
&= \frac{h}{2\Delta} (\dots - u_i u_{i-1}(u_i - u_{i-1}) - u_{i+1} u_i(u_{i+1} - u_i) - \dots)
\end{aligned} \tag{1.8}$$

From that, it is clear that the energy production of the nonlinear term for the skew-symmetric form Hc2 = (2Hd2 + Ha2)/3 should be equal to zero.

### 1.1.2 Energy conservative scheme of order 4 (Hc4)

Centered finite difference schemes of order 4 are used for the first and second derivatives. The second derivative is computed through

$$\frac{\partial^2 u}{\partial x^2} \Big|_{x=x_i} \simeq \frac{-u_{i-2} + 16u_{i-1} - 30u_i + 16u_{i+1} - u_{i+2}}{12\Delta x^2} \tag{1.9}$$

The nonlinear convective term is discretized by the skew-symmetric form Hc4 = (2Hd4 + Ha4)/3, which is a combination of two energy non-conservative schemes. These two forms are based on respectively the divergence (Hd4) and advective (Ha4) formulation of the nonlinear term. Both these schemes are based on a fourth order centered finite difference scheme for the first derivative. These two schemes are given by

$$\text{Divergent form Hd4 : } \quad \frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} \simeq \frac{u_{i-2}^2 - 8u_{i-1}^2 + 8u_{i+1}^2 - u_{i+2}^2}{24\Delta x} \tag{1.10}$$

$$\text{Advective form Ha4 : } \quad \frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} = u \frac{\partial u}{\partial x} \Big|_{x=x_i} \simeq u_i \frac{u_{i-2} - 8u_{i-1} + 8u_{i+1} - u_{i+2}}{12\Delta x} \tag{1.11}$$

In the same way as it has been done in the previous section, it can be shown that both these schemes are not energy conservative by multiplying the nonlinear term at node  $i$  by the nodal velocity value  $u_i$  at the same node and summing up all the contributions. However, the errors in the kinetic energy associated to the divergence and advective schemes have similar forms. It is thus possible to combine these two schemes to generate a kinetic conserving scheme Hc4 = (2Hd4 + Ha4)/3.

### 1.1.3 Compact finite difference schemes with spectral-like resolution

These schemes allow a better representation of the shorter length scales [1]. The first derivative at node  $i$  is computed through the resolution of a matrix system:

	$\alpha$	$\beta$	$a$	$b$	$c$
Optimal	0.5771439	0.0896406	1.3025166	0.99355	0.03750245
Order 10	0.5	1/20	17/12	101/150	0.01
Order 8	4/9	1/36	40/27	25/54	0
Order 6	1/3	0	$(9 + \alpha - 20\beta)/6$	$(-9 + 32\alpha + 62\beta)/15$	$(1 - 3\alpha + 12\beta)/10$
Order 4	0	0	$2(\alpha + 2)/3$	$(4\alpha - 1)/3$	0
Order 2	0	0	1	0	0

**Table 1.1:** Values of the coefficients  $\alpha$ ,  $\beta$ ,  $a$ ,  $b$  and  $c$  for the first derivative.

	$\alpha$	$\beta$	$a$	$b$	$c$
Optimal	0.50209266	0.05569169	0.21564935	1.723322	0.17659730
Order 10	334/899	43/1798	1065/1798	1038/899	79/1798
Order 8	344/1179	$(38\alpha - 9)/214$	$(696 - 1191\alpha)/428$	$(2454\alpha - 294)/535$	0
Order 6	2/11	0	$(6 - 9\alpha - 12\beta)/4$	$(-3 + 24\alpha - 6\beta)/5$	$(2 - 11\alpha + 124\beta)/20$
Order 4	0	0	$4(1 - \alpha)/3$	$(10\alpha - 1)/3$	0
Order 2	0	0	1	0	0

**Table 1.2:** Values of the coefficients  $\alpha$ ,  $\beta$ ,  $a$ ,  $b$  and  $c$  for the second derivative.

$$\begin{aligned}
& \beta \frac{\partial f}{\partial x_{i-2}} + \alpha \frac{\partial f}{\partial x_{i-1}} + \frac{\partial f}{\partial x_i} + \alpha \frac{\partial f}{\partial x_{i+1}} + \beta \frac{\partial f}{\partial x_{i+2}} \\
&= \frac{c}{6\Delta x} (f_{i+3} - f_{i-3}) + \frac{b}{4\Delta x} (f_{i+2} - f_{i-2}) + \frac{a}{2\Delta x} (f_{i+1} - f_{i-1}) \quad (1.12)
\end{aligned}$$

The coefficients  $\alpha$ ,  $\beta$ ,  $a$ ,  $b$  and  $c$  are derived by matching the Taylor series coefficients of various orders. Note that the orders 2 and 4 match the central schemes Hc2 and Hc4 introduced in the previous sections. As for the standard central schemes, a skew-symmetric discretization is used for the non-linear term. This term is chosen equal to

$$u \frac{\partial u}{\partial x_{\text{skew-symmetric}}} = \frac{1}{3} \left( u \frac{\partial u}{\partial x} + \frac{\partial u^2}{\partial x} \right) \quad (1.13)$$

and it has been verified that the energy production of this term is nearly equal to 0 in the case of the forced Burgers equation, for all the orders given in Table 1.1.

The second derivative at node  $i$  is computed through the resolution of another matrix system:

$$\begin{aligned}
& \beta \frac{\partial^2 f}{\partial x^2_{i-2}} + \alpha \frac{\partial^2 f}{\partial x^2_{i-1}} + \frac{\partial^2 f}{\partial x^2_i} + \alpha \frac{\partial^2 f}{\partial x^2_{i+1}} + \beta \frac{\partial^2 f}{\partial x^2_{i+2}} \\
&= \frac{c}{9\Delta x^2} (f_{i+3} - 2f_i + f_{i-3}) + \frac{b}{4\Delta x^2} (f_{i+2} - 2f_i + f_{i-2}) + \frac{a}{\Delta x^2} (f_{i+1} - 2f_i + f_{i-1}) \quad (1.14)
\end{aligned}$$

Note that the orders 2 and 4 match the central schemes Hc2 and Hc4 introduced in the previous sections.

#### 1.1.4 Nonlinear discretization schemes

As discussed in [2], nonlinear discretization schemes are an interesting option to smooth out the numerical oscillations appearing around shocks or regions of high gradients. Several schemes exist and are based on the reduction of the order of the discretization near regions of high gradients.

The oldest operators are the upwind schemes which involves a discretization stencil biased in the direction determined by the sign of the convection speed. The main drawback of this method is the excessive numerical dissipation.

Another class of nonlinear operators are the Total Variation Diminishing (TVD) schemes. They are based on slope limiters which ensures the monotonicity of the solution at all times. The second order central discretization is switched to a first or second order upwind discretization in the region of high gradients. The main disadvantage is that the upwind scheme is also engaged near extrema, which leads to unnecessary dissipation.

ENO (Essential Non-Oscillatory) and WENO (Weighted ENO) schemes were introduced to obtain a higher order of accuracy. These schemes basically select the smoothest interpolating polynomial from a hierarchy of candidates with varying stencil supports. Then, a smooth approximation of the derivatives can be constructed. As observed by [3], these schemes are linearly unstable when coupled with one or two-steps Runge-Kutta time integrators but are linearly stable for three or four-steps integrators. Moreover these schemes generate numerical small-scale turbulence (see [4]) which pollutes the tail of the energy-spectrum and thus limit the resolution range. Improvements are observed when the filters are applied on the solution.

The class of Dynamic Finite Difference (DFD) schemes was introduced to achieve higher accuracy by minimizing the dispersion error and not through the smoothness of the solution as done by the other schemes.

As demonstrated by [2], nonlinear operators do not ensure the conservation of kinetic energy: TVD and WENO operators are over-dissipative whereas the DFD scheme lacks dissipation. Moreover the nonlinear mechanism in these schemes leads to an erratic behaviour of the modified wavenumber in the entire wavenumber range  $\eta \in [0, \eta_{\max}]$  (see section 1.4). This error on the modified wavenumber led to a severe reduction of accuracy in the low-wavenumber range (reduction to 1st-order). We see at this point that nonlinear operators offer mitigating results and this is why we will not put the stress on these methods in the following. The 3rd-order upwind scheme and the DFD operator were implemented and tested.

#### 1.1.4.1 3rd order upwind scheme (Up3)

$$\frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} = \frac{v_{i+1/2}^+(2u_{i+1} + 5u_i - u_{i-1}) + v_{i+1/2}^-(2u_i + 5u_{i+1} - u_{i+2})}{12\Delta x} - \frac{v_{i-1/2}^+(2u_i + 5u_{i-1} - u_{i-2}) + v_{i-1/2}^-(2u_{i-1} + 5u_i - u_{i+1})}{12\Delta x} \quad (1.15)$$

where the transport velocity  $v$  at the intermediate positions  $i \pm 1/2$  is

$$v_{i+1/2} = \frac{u_i + u_{i+1}}{2} \quad v_{i-1/2} = \frac{u_i + u_{i-1}}{2} \quad (1.16)$$

and the positive and negative contributions of these transport velocities are denoted by  $v^+ = \max(v, 0)$  and  $v^- = \min(v, 0)$  respectively. With some developments, it can be shown that the upwind schemes are equivalent to a combination of a standard central discretization and a dissipative correction term that ensures stability.

#### 1.1.4.2 Central DFD scheme

$$\frac{1}{2} \frac{\partial u^2}{\partial x} \Big|_{x=x_i} = \frac{u_{i+1}^2 + c_{i+1}(u_{i+2}^2 - 2u_{i+1}^2 + u_i^2) - u_{i-1}^2 - c_{i-1}(u_i^2 - 2u_{i-1}^2 + u_{i-2}^2)}{4\Delta x} + \mathcal{O}(\Delta x^2) \quad (1.17)$$

and the dynamic coefficient  $c$  is determined by the transported scalar  $u$ :

$$c_i = -\frac{1}{6} \left( 1 + f \max \left( \min \left( \frac{u_{i+2} - 4u_{i+1} + 6u_i - 4u_{i-1} + u_{i-2}}{u_{i+1} - 2u_i + u_{i-1}}, 0 \right), -3 \right) \right)^{-1} \quad (1.18)$$

and the parameter  $f = 0.2$  is obtained by calibrating the scheme to obtain a maximum accuracy for a field with prescribed inertial range. The DFD scheme is a combination of a central discretization and a dispersive correction term.

## 1.2 Finite element method

The Galerkin finite element formulation is defined by

$$\int_0^L N_i(x) \left( \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) - \nu \frac{\partial^2 u}{\partial x^2} - f(x, t) \right) dx = 0 \quad (1.19)$$

where the unknown is interpolated by polynomials such that

$$u(x, t) = \sum_{j=1}^N u_j(t) N_j(x) \quad (1.20)$$

The shape functions are defined such that

$$N_i(x_j) = \delta_{ij} \quad \text{and} \quad \sum_i N_i(x) = 1 \quad (1.21)$$

The discretization is inserted in the Galerkin finite element definition and the viscous terms are integrated by parts in order to reduce the second derivative to a first one. One gets

$$M_{ij} \frac{du_j}{dt} = -C_{ijk} u_j(t) u_k(t) - K_{ij} u_j(t) + M_{ij} F_j(x, t) \quad (1.22)$$

where

$$M_{ij} = \int_0^L N_i(x) N_j(x) dx \quad (1.23)$$

$$K_{ij} = \nu \int_0^L \frac{dN_i}{dx} \frac{dN_j}{dx} dx \quad (1.24)$$

$$C_{ijk} = \int_0^L \frac{N_i}{2} \frac{d(N_j N_k)}{dx} dx \quad (1.25)$$

These different matrices depend only on the shape functions and can be analytically evaluated once and for all. These local matrices are evaluated for each element and inserted in the global system of equation for all the nodes. In the next sections, three types of shape functions are considered and their matrices computed. Note that the formulation used for the non-linear term is not important for the finite element method. Indeed both the divergent and advective formulations introduced in the finite difference methods provide the same discretization for the finite element method.

A major disadvantage for the finite element methods in comparison with the finite differences and volumes is that a system of equation must be solved in order to compute the solution at the next time step. Due to the local nature of the shape functions, the global mass matrix constructed by taking into account of all the elements is sparse. It is sometimes useful to circumvent this necessity of solving a system of equations by 'lumping' the mass matrix in order to obtain a pure diagonal lumped mass matrix. The error induced in the solution is negligible as shown later in the results.



### 1.2.1 Linear Lagrange finite element P1

The interpolation inside an element bounded by the nodes  $i$  and  $i + 1$  is

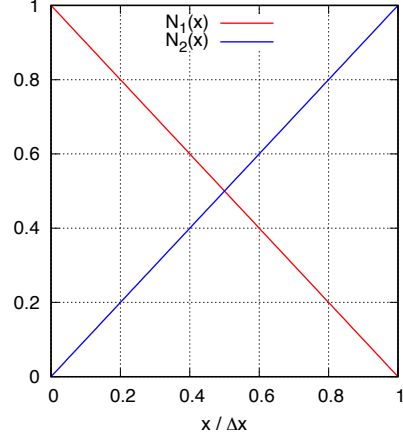
$$u(x, t) = u_i(t)N_1(x) + u_{i+1}(t)N_2(x) \quad (1.26)$$

and the shape functions are

$$N_1(x) = 1 - \xi$$

$$N_2(x) = \xi$$

with  $0 \leq \xi = x/\Delta x \leq 1$  the nondimensional spatial coordinate



**Figure 1.1:** Shape functions for the linear Lagrange element

The transposed global unknown vector is

$$u^t = ( \cdots \quad u_{i-1} \quad u_i \quad u_{i+1} \quad \cdots )^t \quad (1.27)$$

The local matrices for the element bounded by the nodes  $i$  and  $i + 1$  are given by

$$M_{ij} = \frac{\Delta x}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad (1.28)$$

$$K_{ij} = \frac{\nu}{\Delta x} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (1.29)$$

and the nonlinear convective term over the element is

$$C_{ijk}u_ju_k = \frac{1}{6} \begin{pmatrix} -2u_i^2 + u_iu_{i+1} + u_{i+1}^2 \\ -u_i^2 - u_iu_{i+1} + 2u_{i+1}^2 \end{pmatrix} \quad (1.30)$$

#### 1.2.1.1 Lumped mass matrix

The direct lumping technique is used. This technique states that the diagonal components of the lumped mass matrix are equal to the sum of the components on each line ( $M_{ii}^L = \sum_j M_{ij}$ ). The sum over all the diagonal terms of the lumped matrix should be equal to the length of the element ( $\sum_i M_{ii}^L = \Delta x$ ). The lumped mass matrix is thus given by

$$M_{ij} = \Delta x \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad (1.31)$$

Finally a major observation is that the Lagrange P1 finite element method with lumped mass matrix provides exactly the same discretization as the conservative finite difference scheme Hc2. The spatially discretized equation at node  $i$  for these two particular schemes is

$$\frac{\partial u_i}{\partial t} = -\frac{u_{i+1}^2 + u_i(u_{i+1} - u_{i-1}) - u_{i-1}^2}{6\Delta x} + \nu \frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} + F_i(t) \quad (1.32)$$

### 1.2.2 Cubic Lagrange finite element P3

The interpolation inside an element bounded by the nodes  $i$  and  $i + 1$  is

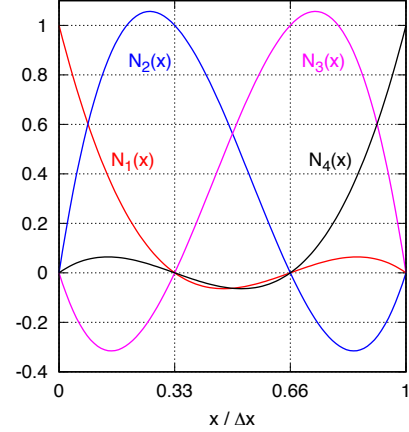
$$u(x, t) = u_i(t)N_1(x) + u_{i+\frac{1}{3}}(t)N_2(x) + u_{i+\frac{2}{3}}(t)N_3(x) + u_{i+1}(t)N_4(x) \quad (1.33)$$

Two intermediate nodes are inserted at  $x_{i+\frac{1}{3}} = x_i + \Delta x/3$  and  $x_{i+\frac{2}{3}} = x_i + 2\Delta x/3$  in order to build the third order polynomials. For the same number of elements as for the linear Lagrange element, the size of the system of equations is multiplied by three for the cubic Lagrange element.

The shape functions are

$$\begin{aligned} N_1(x) &= \frac{1}{2}(1 - \xi)(2 - 3\xi)(1 - 3\xi) \\ N_2(x) &= \frac{9}{2}\xi(1 - \xi)(2 - 3\xi) \\ N_3(x) &= \frac{9}{2}\xi(1 - \xi)(3\xi - 1) \\ N_4(x) &= \frac{1}{2}\xi(2 - 3\xi)(1 - 3\xi) \end{aligned}$$

with  $0 \leq \xi = x/\Delta x \leq 1$  the nondimensional spatial coordinate



**Figure 1.2:** Shape functions for the cubic Lagrange element

The transposed global unknown vector is

$$u^t = \left( \cdots \quad u_{i-1} \quad u_{i-\frac{2}{3}} \quad u_{i-\frac{1}{3}} \quad u_i \quad u_{i+\frac{1}{3}} \quad u_{i+\frac{2}{3}} \quad u_{i+1} \quad \cdots \right)^t \quad (1.34)$$

The local matrices for the element bounded by the nodes  $i$  and  $i + 1$  are given by

$$M_{ij} = \frac{\Delta x}{1680} \begin{pmatrix} 128 & 99 & -36 & 19 \\ 99 & 648 & -81 & -36 \\ -36 & -81 & 648 & 99 \\ 19 & -36 & 99 & 128 \end{pmatrix} \quad (1.35)$$

$$K_{ij} = \frac{\nu}{40\Delta x} \begin{pmatrix} 148 & -189 & 54 & -13 \\ -189 & 432 & -297 & 54 \\ 54 & -297 & 432 & -189 \\ -13 & 54 & -189 & 148 \end{pmatrix} \quad (1.36)$$

and the nonlinear convective term over the element for the lines corresponding respectively to the lines of the unknowns  $u_i$ ,  $u_{i+\frac{1}{3}}$ ,  $u_{i+\frac{2}{3}}$  and  $u_{i+1}$

$$\begin{aligned} C_{ijk}u_ju_k|_{x=x_i} &= \frac{1}{6720} \left( -2240u_i^2 + 1611u_iu_{i+\frac{1}{3}} + 2025u_{i+\frac{1}{3}}^2 - 630u_iu_{i+\frac{2}{3}} - 1053u_{i+\frac{1}{3}}u_{i+\frac{2}{3}} \right. \\ &\quad \left. - 162u_{i+\frac{2}{3}}^2 + 139u_iu_{i+1} + 180u_{i+\frac{1}{3}}u_{i+1} - 9u_{i+\frac{2}{3}}u_{i+1} + 139u_{i+1}^2 \right) \end{aligned} \quad (1.37)$$

$$\begin{aligned} C_{ijk}u_ju_k|_{x=x_{i+\frac{1}{3}}} &= \frac{3}{2240} \left( -179u_i^2 - 225u_iu_{i+\frac{1}{3}} + 72u_iu_{i+\frac{2}{3}} + 243u_{i+\frac{1}{3}}u_{i+\frac{2}{3}} + 243u_{i+\frac{2}{3}}^2 \right. \\ &\quad \left. - 21u_iu_{i+1} - 18u_{i+\frac{1}{3}}u_{i+1} - 45u_{i+\frac{2}{3}}u_{i+1} - 70u_{i+1}^2 \right) \end{aligned} \quad (1.38)$$

$$C_{ijk}u_ju_k|_{x=x_{i+\frac{2}{3}}} = \frac{3}{2240} \left( 70u_i^2 + 45u_iu_{i+\frac{1}{3}} + 18u_iu_{i+\frac{2}{3}} + 21u_iu_{i+1} - 243u_{i+\frac{1}{3}}^2 - 243u_{i+\frac{1}{3}}u_{i+\frac{2}{3}} - 72u_{i+\frac{1}{3}}u_{i+1} + 225u_{i+\frac{2}{3}}u_{i+1} + 179u_{i+1}^2 \right) \quad (1.39)$$

$$C_{ijk}u_ju_k|_{x=x_{i+1}} = \frac{1}{6720} \left( -139u_i^2 + 9u_iu_{i+\frac{1}{3}} - 180u_iu_{i+\frac{2}{3}} - 139u_iu_{i+1} + 162u_{i+\frac{1}{3}}^2 + 1053u_{i+\frac{1}{3}}u_{i+\frac{2}{3}} + 630u_{i+\frac{1}{3}}u_{i+1} - 2025u_{i+\frac{2}{3}}^2 - 1611u_{i+\frac{2}{3}}u_{i+1} + 2240u_{i+1}^2 \right) \quad (1.40)$$

### 1.2.2.1 Lumped mass matrix

The direct lumping technique is used. This technique states that the diagonal components of the lumped mass matrix are equal to the sum of the components on each line ( $M_{ii}^L = \sum_j M_{ij}$ ). The sum over all the diagonal terms of the lumped matrix should be equal to the length of the element ( $\sum_i M_{ii}^L = \Delta x$ ). The lumped mass matrix is thus given by

$$M_{ij} = \frac{\Delta x}{1680} \begin{pmatrix} 210 & 0 & 0 & 0 \\ 0 & 630 & 0 & 0 \\ 0 & 0 & 630 & 0 \\ 0 & 0 & 0 & 210 \end{pmatrix} \quad (1.41)$$

### 1.2.3 Cubic Hermite finite element H3

The interpolation inside an element bounded by the nodes  $i$  and  $i+1$  is

$$u(x, t) = u_i(t)H_1^0(x) + \frac{\partial u_i(t)}{\partial x}H_1^1(x) + u_{i+1}(t)H_2^0(x) + \frac{\partial u_{i+1}(t)}{\partial x}H_2^1(x) \quad (1.42)$$

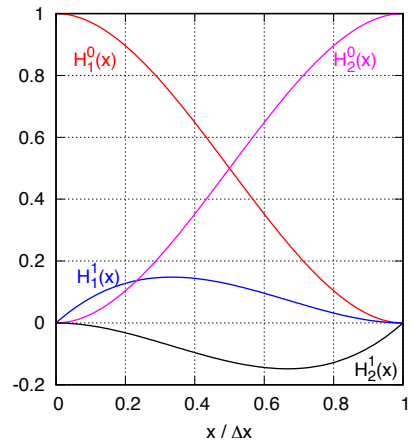
The interpolation by the Hermite third order polynomials is based on the knowledge of the solution and the first derivative of the solution at the nodes. This Hermite interpolation ensures the continuity of the solution up to the first spatial derivative through the  $H_i^1$  functions. The Lagrange interpolation do not ensure the continuity of the derivative of the solution. For the same number of elements as for the linear Lagrange element, the size of the system of equations is multiplied by two for the cubic Hermite element. The conditions that the Hermite polynomials must respect are

$$H_i^0(x_j) = \delta_{ij} \quad \frac{\partial H_i^0(x_j)}{\partial x} = 0 \quad H_i^1(x_j) = 0 \quad \frac{\partial H_i^1(x_j)}{\partial x} = \delta_{ij} \quad (1.43)$$

The shape functions are

$$\begin{aligned} H_1^0(x_j) &= 1 - 3\xi^2 + 2\xi^3 \\ H_1^1(x_j) &= \Delta x (\xi - 2\xi^2 + \xi^3) \\ H_2^0(x_j) &= 3\xi^2 - 2\xi^3 \\ H_2^1(x_j) &= \Delta x (-\xi^2 + \xi^3) \end{aligned}$$

with  $0 \leq \xi = x/\Delta x \leq 1$  the nondimensional spatial coordinate



**Figure 1.3:** Shape functions for the cubic Hermite element

The transposed global unknown vector is

$$u^t = \left( \cdots \quad u_{i-1} \quad \frac{\partial u_{i-1}}{\partial x} \quad u_i \quad \frac{\partial u_i}{\partial x} \quad u_{i+1} \quad \frac{\partial u_{i+1}}{\partial x} \quad \cdots \right)^t \quad (1.44)$$

The local matrices for the element bounded by the nodes  $i$  and  $i + 1$  are given by

$$M_{ij} = \frac{\Delta x}{420} \begin{pmatrix} 156 & 22\Delta x & 54 & -13\Delta x \\ 22\Delta x & 4\Delta x^2 & 13\Delta x & -3\Delta x^2 \\ 54 & 13\Delta x & 156 & -22\Delta x \\ -13\Delta x & -3\Delta x^2 & -22\Delta x & 4\Delta x^2 \end{pmatrix} \quad (1.45)$$

$$K_{ij} = \frac{\nu}{30\Delta x} \begin{pmatrix} 36 & 3\Delta x & -36 & 3\Delta x \\ 3\Delta x & 4\Delta x^2 & -3\Delta x & -\Delta x^2 \\ -36 & -3\Delta x & 36 & -3\Delta x \\ 3\Delta x & -\Delta x^2 & -3\Delta x & 4\Delta x^2 \end{pmatrix} \quad (1.46)$$

and the nonlinear convective term over the element for the lines corresponding respectively to the lines of the unknowns  $u_i$ ,  $\partial u_i/\partial x$ ,  $u_{i+1}$  and  $\partial u_{i+1}/\partial x$

$$C_{ijk}u_ju_k|_{u_i} = \frac{1}{840} \left[ -280u_i^2 + 140u_iu_{i+1} + 140u_{i+1}^2 + 50u_i \frac{\partial u_i}{\partial x} \Delta x - 50u_{i+1} \frac{\partial u_{i+1}}{\partial x} \Delta x - 34u_i \frac{\partial u_{i+1}}{\partial x} \Delta x \right. \\ \left. + 34u_{i+1} \frac{\partial u_i}{\partial x} \Delta x + 5 \left( \frac{\partial u_i}{\partial x} \Delta x \right)^2 - 8 \frac{\partial u_i}{\partial x} \frac{\partial u_{i+1}}{\partial x} \Delta x^2 + 5 \left( \frac{\partial u_{i+1}}{\partial x} \Delta x \right)^2 \right] \quad (1.47)$$

$$C_{ijk}u_ju_k|_{u'_i} = \frac{\Delta x}{840} \left[ -50u_i^2 + 16u_iu_{i+1} + 34u_{i+1}^2 - 5u_i \frac{\partial u_i}{\partial x} \Delta x - 11u_{i+1} \frac{\partial u_{i+1}}{\partial x} \Delta x \right. \\ \left. - 3u_i \frac{\partial u_{i+1}}{\partial x} \Delta x + 5u_{i+1} \frac{\partial u_i}{\partial x} \Delta x - \frac{\partial u_i}{\partial x} \frac{\partial u_{i+1}}{\partial x} \Delta x^2 + \left( \frac{\partial u_{i+1}}{\partial x} \Delta x \right)^2 \right] \quad (1.48)$$

$$C_{ijk}u_ju_k|_{u_{i+1}} = \frac{1}{840} \left[ -140u_i^2 - 140u_iu_{i+1} + 280u_{i+1}^2 - 50u_i \frac{\partial u_i}{\partial x} \Delta x + 50u_{i+1} \frac{\partial u_{i+1}}{\partial x} \Delta x + 34u_i \frac{\partial u_{i+1}}{\partial x} \Delta x \right. \\ \left. - 34u_{i+1} \frac{\partial u_i}{\partial x} \Delta x - 5 \left( \frac{\partial u_i}{\partial x} \Delta x \right)^2 + 8 \frac{\partial u_i}{\partial x} \frac{\partial u_{i+1}}{\partial x} \Delta x^2 - 5 \left( \frac{\partial u_{i+1}}{\partial x} \Delta x \right)^2 \right] \quad (1.49)$$

$$C_{ijk}u_ju_k|_{u'_{i+1}} = \frac{\Delta x}{840} \left[ 34u_i^2 + 16u_iu_{i+1} - 50u_{i+1}^2 - 11u_i \frac{\partial u_i}{\partial x} \Delta x + 5u_{i+1} \frac{\partial u_{i+1}}{\partial x} \Delta x \right. \\ \left. - 5u_i \frac{\partial u_{i+1}}{\partial x} \Delta x + 3u_{i+1} \frac{\partial u_i}{\partial x} \Delta x + \left( \frac{\partial u_i}{\partial x} \Delta x \right)^2 - \frac{\partial u_i}{\partial x} \frac{\partial u_{i+1}}{\partial x} \Delta x^2 \right] \quad (1.50)$$

#### 1.2.4 Fifth order Hermite finite element H5

The interpolation inside an element bounded by the nodes  $i$  and  $i + 1$  is

$$u(x, t) = u_i(t)H_1^0(x) + \frac{\partial u_i(t)}{\partial x}H_1^1(x) + \frac{\partial^2 u_i(t)}{\partial x^2}H_1^2(x)$$

$$+ u_{i+1}(t)H_2^0(x) + \frac{\partial u_{i+1}(t)}{\partial x}H_2^1(x) + \frac{\partial^2 u_{i+1}(t)}{\partial x^2}H_2^2(x) \quad (1.51)$$

The interpolation by the fifth order Hermite polynomials is based on the knowledge of the solution, its first derivative and second derivatives at the nodes. This Hermite interpolation ensures the continuity of the solution up to the second spatial derivative through the  $H_i^1$  and  $H_i^2$  shape functions. The Lagrange interpolation do not ensure the continuity of the derivative of the solution. For the same number of elements as for the linear Lagrange element, the size of the system of equations is multiplied by three for the fifth order Hermite element. The conditions that the Hermite polynomials must respect are

$$\begin{aligned} H_i^0(x_j) &= \delta_{ij} & \frac{\partial H_i^0(x_j)}{\partial x} &= 0 & \frac{\partial^2 H_i^0(x_j)}{\partial x^2} &= 0 \\ H_i^1(x_j) &= 0 & \frac{\partial H_i^1(x_j)}{\partial x} &= \delta_{ij} & \frac{\partial^2 H_i^1(x_j)}{\partial x^2} &= 0 \\ H_i^2(x_j) &= 0 & \frac{\partial H_i^2(x_j)}{\partial x} &= 0 & \frac{\partial^2 H_i^2(x_j)}{\partial x^2} &= \delta_{ij} \end{aligned} \quad (1.52)$$

The shape functions are

$$\begin{aligned} H_1^0(x_j) &= 1 - 10\xi^3 + 15\xi^4 - 6\xi^5 \\ H_1^1(x_j) &= \Delta x (\xi - 6\xi^3 + 8\xi^4 - 3\xi^5) \\ H_1^2(x_j) &= \frac{\Delta x^2}{2} (\xi^2 - 3\xi^3 + 3\xi^4 - \xi^5) \\ H_2^0(x_j) &= 10\xi^3 - 15\xi^4 + 6\xi^5 \\ H_2^1(x_j) &= \Delta x (-4\xi^3 + 7\xi^4 - 3\xi^5) \\ H_2^2(x_j) &= \frac{\Delta x^2}{2} (\xi^3 - 2\xi^4 + \xi^5) \end{aligned}$$

with  $0 \leq \xi = x/\Delta x \leq 1$  the nondimensional spatial coordinate

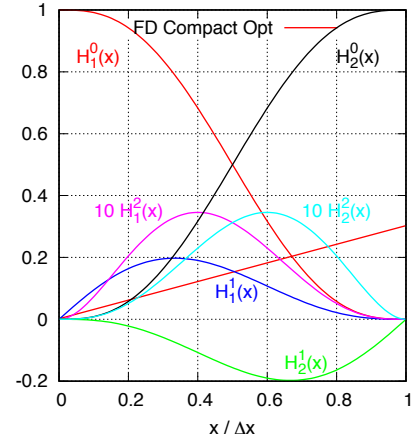
The transposed global unknown vector is

$$u^t = \left( \cdots \quad u_{i-1} \quad \frac{\partial u_{i-1}}{\partial x} \quad \frac{\partial^2 u_{i-1}}{\partial x^2} \quad u_i \quad \frac{\partial u_i}{\partial x} \quad \frac{\partial^2 u_i}{\partial x^2} \quad u_{i+1} \quad \frac{\partial u_{i+1}}{\partial x} \quad \frac{\partial^2 u_{i+1}}{\partial x^2} \quad \cdots \right)^t \quad (1.53)$$

The local matrices for the element bounded by the nodes  $i$  and  $i+1$  are given by

$$M_{ij} = \frac{\Delta x}{55440} \begin{pmatrix} 21720 & 3732\Delta x & 281\Delta x^2 & 6000 & -1812\Delta x & 181\Delta x^2 \\ 3732\Delta x & 832\Delta x^2 & 69\Delta x^3 & 1812\Delta x & -532\Delta x^2 & 52\Delta x^3 \\ 281\Delta x^2 & 69\Delta x^3 & 6\Delta x^4 & 181\Delta x^2 & -52\Delta x^3 & 5\Delta x^4 \\ 6000 & 1812\Delta x & 181\Delta x^2 & 21720 & -3732\Delta x & 281\Delta x^2 \\ -1812\Delta x & -532\Delta x^2 & -52\Delta x^3 & -3732\Delta x & 832\Delta x^2 & -69\Delta x^3 \\ 181\Delta x^2 & 52\Delta x^3 & 5\Delta x^4 & 281\Delta x^2 & -69\Delta x^3 & 6\Delta x^4 \end{pmatrix} \quad (1.54)$$

$$K_{ij} = \frac{\nu}{1260\Delta x} \begin{pmatrix} 1800 & 270\Delta x & 15\Delta x^2 & -1800 & 270\Delta x & -15\Delta x^2 \\ 270\Delta x & 288\Delta x^2 & 21\Delta x^3 & -270\Delta x & -18\Delta x^2 & 6\Delta x^3 \\ 15\Delta x^2 & 21\Delta x^3 & 2\Delta x^4 & -15\Delta x^2 & -6\Delta x^3 & \Delta x^4 \\ -1800 & -270\Delta x & -15\Delta x^2 & 1800 & -270\Delta x & 15\Delta x^2 \\ 270\Delta x & -18\Delta x^2 & -6\Delta x^3 & -270\Delta x & 288\Delta x^2 & -21\Delta x^3 \\ -15\Delta x^2 & 6\Delta x^3 & \Delta x^4 & 15\Delta x^2 & -21\Delta x^3 & 2\Delta x^4 \end{pmatrix} \quad (1.55)$$



**Figure 1.4:** Shape functions for the fifth order Hermite element

### 1.3 Time integration

The integration in time used in this project is the well-known fourth-order Runge-Kutta method with four steps RK4. This method allows to reach a high accuracy in time while having a limit on the Courant-Friedrich-Lewy number (defined in Eq. (1.56)) which is higher than other standard methods. This limit is identical for the purely advective and purely viscous forms of the Burgers equation.

$$\text{CFL} = \frac{\max(u)\Delta t}{\Delta x} \lesssim 2.8 \quad (1.56)$$

In the case of a generic equation  $\partial u / \partial t = f(u, t)$ , the RK4 time integrator provides the next time step solution through

$$\begin{aligned} k_1 &= \Delta t f(u^n, t^n) \\ k_2 &= \Delta t f(u^n + \frac{k_1}{2}, t + \frac{\Delta t}{2}) \\ k_3 &= \Delta t f(u^n + \frac{k_2}{2}, t + \frac{\Delta t}{2}) \\ k_4 &= \Delta t f(u^n + k_3, t + \Delta t) \\ u^{n+1} &= u^n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned} \quad (1.57)$$

### 1.4 Dispersive error

The study of the dispersive error is based on the linear advection equation

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0 \quad (1.58)$$

It is supposed that there exists an exact solution  $u(x, t) = \hat{u} \exp(Ikx)$ . Injecting this exact solution leads to the ordinary differential equation

$$\frac{\partial \hat{u}}{\partial t} = \lambda_{ex} \hat{u} \quad (1.59)$$

with  $\lambda_{ex} = Ika$  the exact eigenvalue of the previous equation. The spatial discretization of the linear advection equation introduces a set of equations to be solved on each nodes in order to extract the discretized solution  $u^h$ :

$$\frac{\partial u^h}{\partial t} = Au^h \quad (1.60)$$

At the position  $x_m$ , the eigenvectors of the matrix  $A$  take the form  $v_m^j = \exp(Ik_j x_m)$  for each wavenumber  $k_j = 2\pi j/L$  with  $j = 0 \rightarrow N - 1$ . For each eigenvector, an eigenvalue  $\lambda^j$  is associated. The exact eigenvalue  $\lambda_{ex}$  and the one coming from the discretization  $\lambda^j$  can then be compared on the basis of the numerical wave propagation speed  $a^j = I\lambda^j/k_j = a_r + Ia_i$ , which is generally a complex number. The comparison is then

$$\frac{\lambda^j}{\lambda_{ex}} = \frac{a^j}{a} = \frac{a_r^j}{a} + I \frac{a_i^j}{a} \quad (1.61)$$

In the following analysis, only the dispersive error  $\epsilon_\Phi = a_r^j/a$  will be considered. The reason is that all the considered discretizations are symmetrical and do not introduce any imaginary part  $a_i^j$ . Thus in the present study, no numerical dissipation is studied. A wave of wavenumber  $k_j$  propagates at the phase velocity  $\lambda^j$  which is different from the exact phase velocity  $\lambda$ . The phase velocity is approximated through the spatial discretization. In general, the wave propagates more slowly than the exact solution.

### 1.4.1 Modified wavenumbers for the different spatial discretizations

#### Finite difference scheme Hc2

The discretized equation reads

$$\frac{\partial u_i}{\partial t} + a \frac{u_{i+1} - u_{i-1}}{2\Delta x} = 0 \quad (1.62)$$

With the assumption of a periodic solution, one can deduce the solution at nodes  $i + 1$  and  $i - 1$  from the solution at node  $i$ :

$$u_{i+1} = \hat{u}(t) \exp(Ik(x + \Delta x)) = \underbrace{\hat{u}(t) \exp(Ikx)}_{u_i} \exp(I \underbrace{k\Delta x}_{\eta}) = u_i \exp(I\eta) \quad (1.63a)$$

$$u_{i-1} = \hat{u}(t) \exp(Ik(x - \Delta x)) = \underbrace{\hat{u}(t) \exp(Ikx)}_{u_i} \exp(-I \underbrace{k\Delta x}_{\eta}) = u_i \exp(-I\eta) \quad (1.63b)$$

so that the discretized equation is now written as

$$\frac{\partial u_i}{\partial t} + \frac{Ia}{\Delta x} \sin(\eta) u_i = 0 \quad (1.64)$$

Concerning the time dependency, it is assumed that the solution takes the form  $u = \hat{u} \exp(-IKt)$ , so that

$$I \left( \frac{a}{\Delta x} \sin(\eta) - K \right) u_i = 0 \quad (1.65)$$

If we define  $\eta_{num} = K\Delta x/a$  as the nondimensional modified numerical wavenumber that results from the discretization, one gets the expression

$$\eta_{num} = \sin(\eta) \quad (1.66)$$

with  $\eta \in [0, \pi]$ .

#### Finite difference scheme Hc4

Similarly to what has been done for the Hc2, one gets the modified wavenumber

$$\eta_{num} = \frac{8 \sin(\eta) - \sin(2\eta)}{6} \quad (1.67)$$

with  $\eta \in [0, \pi]$ .

#### Compact finite difference schemes with spectral-like resolution

$$\eta_{num} = \frac{a \sin(\eta) + (b/2) \sin(2\eta) + (c/3) \sin(3\eta)}{1 + 2\alpha \cos(\eta) + 2\beta \cos(2\eta)} \quad (1.68)$$

with  $\eta \in [0, \pi]$ .

#### Finite element P1

The methodology is identical to the one explained for the finite difference scheme Hc2 but one major difference appears: the time derivatives are also applied on the nodes  $i - 1$  and  $i + 1$  because of the mass matrix  $M_{ij}$ . The spatially discretized equation is thus written

$$\frac{1}{6} \left( \frac{\partial u_{i-1}}{\partial t} + 4 \frac{\partial u_i}{\partial t} + \frac{\partial u_{i+1}}{\partial t} \right) + a \frac{u_{i+1} - u_{i-1}}{2\Delta x} = 0 \quad (1.69)$$

These times derivatives can be reduced to the time derivative at node  $i$  through

$$\frac{\partial u_{i+1}}{\partial t} = \frac{\partial u_i}{\partial t} \exp(I\eta) \quad \frac{\partial u_{i-1}}{\partial t} = \frac{\partial u_i}{\partial t} \exp(-I\eta) \quad (1.70)$$

The modified wavenumber is then equal to

$$\eta_{num} = \frac{3 \sin(\eta)}{2 + \cos(\eta)} \quad (1.71)$$

with  $\eta \in [0, \pi]$ .

### Finite element P3

Similarly to what has been done for the element P1, one gets the modified wavenumber

$$\eta_{num} = 21 \frac{57 \sin(\eta/3) - 24 \sin(2\eta/3) + 7 \sin(\eta)}{128 + 99 \cos(\eta/3) - 36 \cos(2\eta/3) + 19 \cos(\eta)} \quad (1.72)$$

with  $\eta \in [0, 3\pi]$  for the original mass matrix and

$$\eta_{num} = \frac{57 \sin(\eta/3) - 24 \sin(2\eta/3) + 7 \sin(\eta)}{10} \quad (1.73)$$

for the lumped mass matrix. As it will be shown later, the denominator of Eq. (1.72) possess a zero in the range  $\eta \in [0, 3\pi]$ . This leads to an infinite value of  $\eta_{num}$  at this zero. Lumping the mass matrix allows to get rid of this zero and thus to keep limited values for  $\eta_{num}$ .

### Finite element H3

The discretized equation at node  $i$  for the unknowns  $u_i$  and  $u'_i = \partial u_i / \partial x$  is written as

$$\begin{aligned} \frac{\Delta x}{420} \begin{pmatrix} 54 & 13\Delta x & 312 & 0 & 54 & -13\Delta x \\ -13\Delta x & -3\Delta x^2 & 0 & 8\Delta x^2 & 13\Delta x & -3\Delta x^2 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} u_{i-1} \\ u'_{i-1} \\ u_i \\ u'_i \\ u_{i+1} \\ u'_{i+1} \end{pmatrix} \\ + \frac{a}{60} \begin{pmatrix} -30 & -6\Delta x & 0 & 12\Delta x & 30 & -6\Delta x \\ 6\Delta x & \Delta x^2 & -12\Delta x & 0 & 6\Delta x & -\Delta x^2 \end{pmatrix} \begin{pmatrix} u_{i-1} \\ u'_{i-1} \\ u_i \\ u'_i \\ u_{i+1} \\ u'_{i+1} \end{pmatrix} = 0 \quad (1.74) \end{aligned}$$

Then, after using  $u = \hat{u} \exp(-IKt)$  and Eqs. (1.63) and (1.70), one gets

$$\begin{pmatrix} I \left( \sin(\eta) - \frac{26 + 9 \cos(\eta)}{35} \eta_{num} \right) & \frac{\Delta x(1 - \cos(\eta))}{5} - \frac{13\Delta x \sin(\eta)}{210} \eta_{num} \\ \frac{13\Delta x \sin(\eta)}{210} \eta_{num} - \frac{\Delta x(1 - \cos(\eta))}{5} & -I \left( \frac{\Delta x^2(4 - 3 \cos(\eta))}{210} \eta_{num} + \frac{\Delta x^2 \sin(\eta)}{30} \right) \end{pmatrix} \begin{pmatrix} u_i \\ u'_i \end{pmatrix} = 0 \quad (1.75)$$



The solution is non-trivial  $(u_i, u'_i) \neq 0$  if the determinant of the previous matrix is equal to zero. This leads to an equation of second order on  $\eta_{num}$ , which is solved for  $\eta \in [0, 2\pi]$ . The equation on  $\eta_{num}$  is given by

$$\begin{aligned} & (-131 + 72 \cos(\eta) - \cos(2\eta))\eta_{num}^2 + (-384 \sin(\eta) + 12 \sin(2\eta))\eta_{num} \\ & + 966 - 1008 \cos(\eta) + 42 \cos(2\eta) = 0 \end{aligned} \quad (1.76)$$

As shown in Fig. 1.5a, one of the solution is negative and should be rejected, the other one corresponding to the physical situation. To each  $\eta$  corresponds two eigenvectors that can be reconstructed. The three first eigenvectors (for  $K = 0, 1$  and  $2$ ) are plotted in Fig. 1.5b-1.5d. One sees that one of the two eigenvectors represents the physical behaviour whereas the second mode behaves like a parasite oscillation. This second mode reaches its maximum at the locations where the first mode has the steepest variation. As the modified wavenumber  $K$  increases, the amplitude of the first eigenvector diminishes down to reaching the amplitude of the second parasite eigenvector.

### Finite element H5

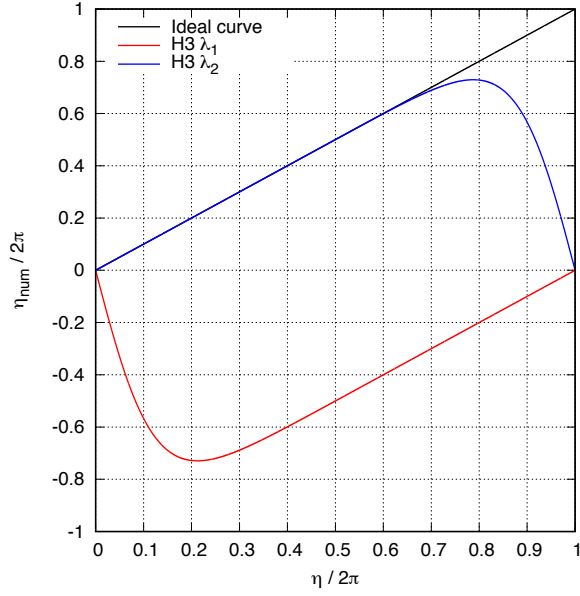
Similarly to what has been done for the Hermite H3 element, one can derive a cubic equation on  $\eta_{num}$  with  $\eta \in [0, 3\pi]$ .

$$\begin{aligned} & (144116 + 81159 \cos(\eta) + 1524 \cos(2\eta) + \cos(3\eta))\eta_{num}^3 \\ & + (-630279 \sin(\eta) - 25020 \sin(2\eta) - 27 \sin(3\eta))\eta_{num}^2 \\ & + (-4627800 - 4172130 \cos(\eta) - 181080 \cos(2\eta) - 270 \cos(3\eta))\eta_{num} \\ & + 7988310 \sin(\eta) + 495000 \sin(2\eta) + 990 \sin(3\eta) = 0 \end{aligned} \quad (1.77)$$

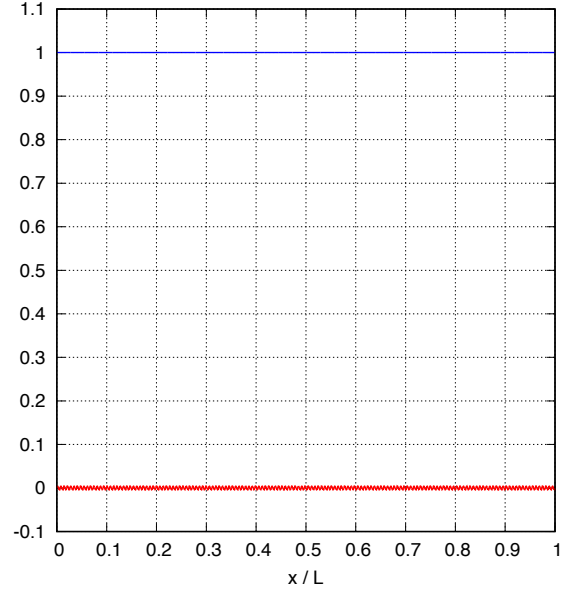
This cubic equation provides three solutions for  $\eta_{num}$  among which only one represents the physics and is thus acceptable (see Fig. 1.6a). To each  $\eta$  corresponds three eigenvectors that can be reconstructed. The three first eigenvectors (for  $K = 0, 1$  and  $2$ ) are plotted in Fig. 1.6b-1.6d. One sees that one of the three eigenvectors represents the physical behaviour whereas the two other modes behave like a parasite oscillation. These two parasite modes reach their maximum at the locations where the first mode has the steepest variation. As the modified wavenumber  $K$  increases, the amplitude of the first eigenvector diminishes down to reaching the amplitude of the two other parasite eigenvectors.

### 1.4.2 Result

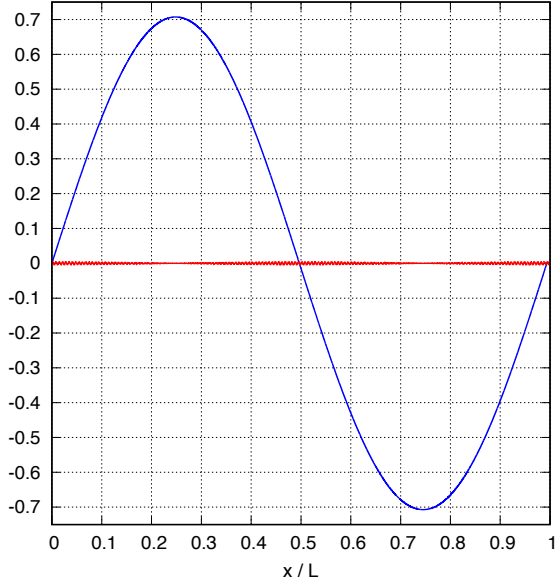
The different modified wavenumbers  $\eta_{num}$  are plotted in Fig. 1.7a as a function of the theoretical wavenumber  $\eta$ . The ranges are nondimensionalized by the maximum values  $\eta_{max}$  taken by the spatial discretizations. Ideally the spatial discretization should capture the theoretical wavenumber  $\eta$ . In practice, the discretization introduces dispersive errors. A direct consequence is that the numerical scheme will capture a correct wavenumber until a given  $\eta_{div}$  where the numerical wavenumber diverges from the theoretical one. The ratio of phase velocity  $\eta_{num}/\eta$  is plotted in Fig. 1.7b. One clearly see that most of the schemes tend to slow down the propagation of waves as the ratio becomes smaller than 1. However the element P3 present an infinite value at  $\eta_{div}/\eta_{max} \simeq 0.8838$  which forces the numerical wavenumber  $\eta_{num}$  to break off rapidly from the ideal curve. The waves propagate faster than the physical wave for  $\eta_{num}/\eta_{max} < 0.8838$  whereas they propagate in the opposite direction to the physical waves for  $\eta_{num}/\eta_{max} > 0.8838$ . The lumped version of the element P3 shows a region where the waves propagate more slowly and then a region where they propagate faster than the physical waves. From  $\eta_{num}/\eta_{max} \simeq 0.7458$ ,



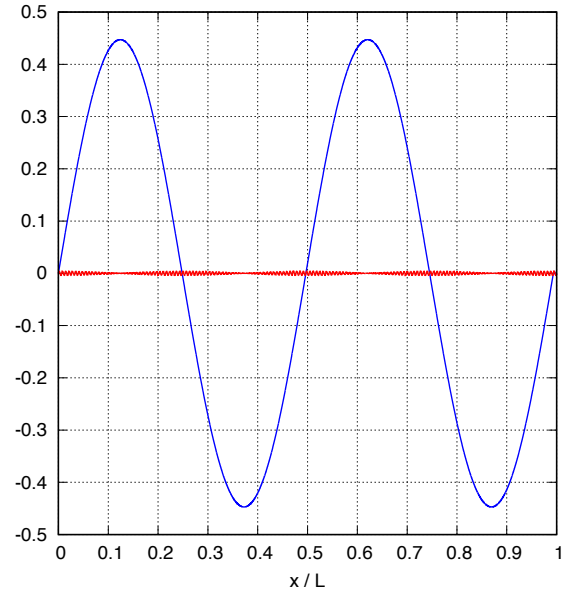
(a) Modified numerical wavenumber



(b) K=0

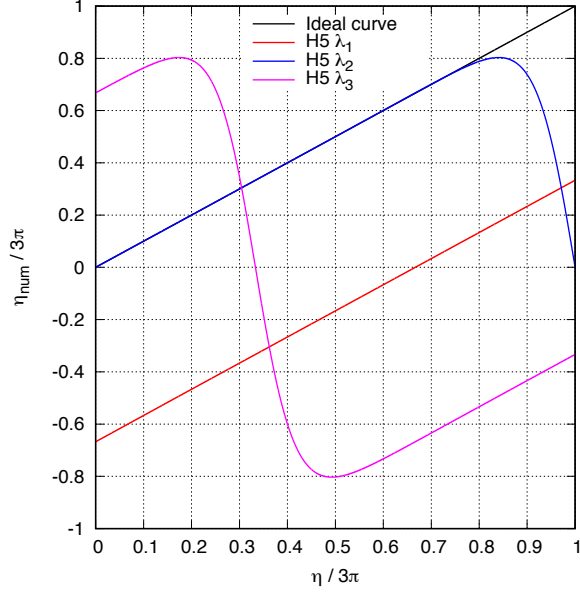


(c) K=1

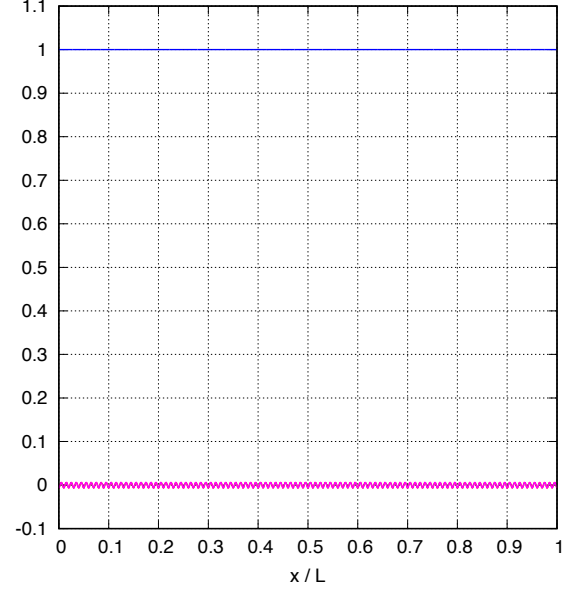


(d) K=2

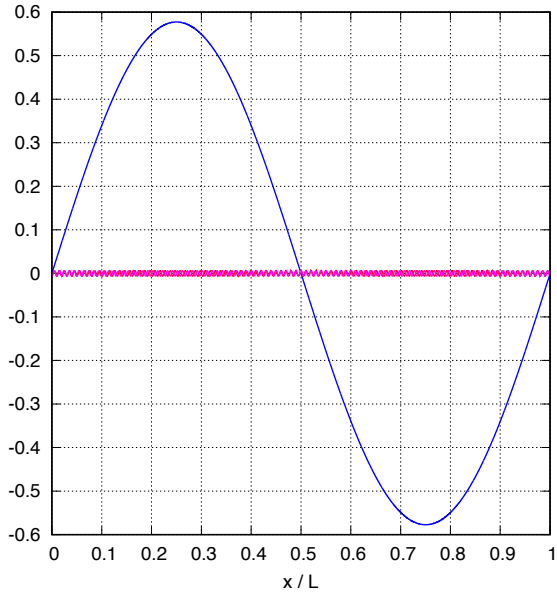
**Figure 1.5:** (a) Modified numerical wavenumber  $\eta_{num}$  as a function of  $\eta$  for the Hermite elements H3. (b)-(d) Eigenvectors of Hermite H3 for  $K = 0, 1$  and  $2$ .



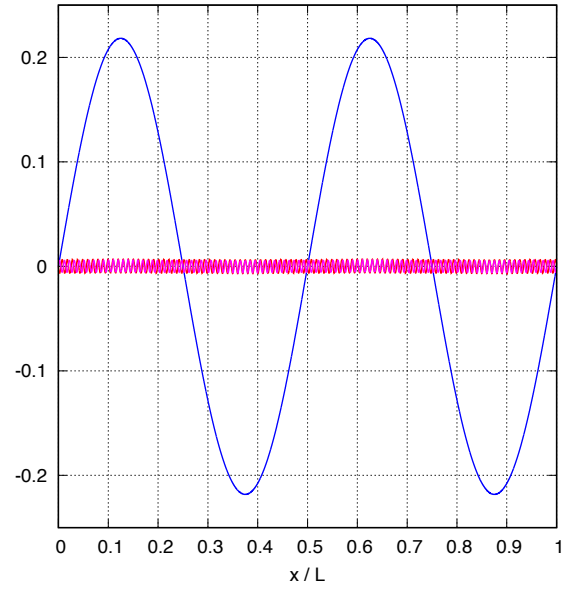
(a) Modified numerical wavenumber



(b) K=0



(c) K=1

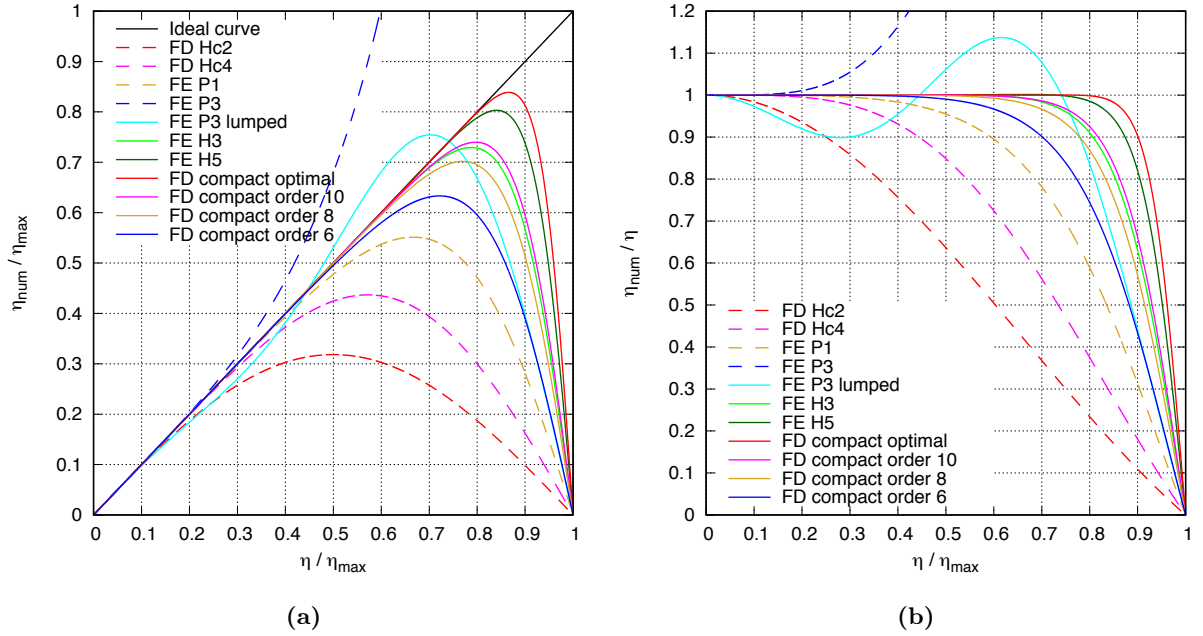


(d) K=2

**Figure 1.6:** (a) Modified numerical wavenumber  $\eta_{num}$  as a function of  $\eta$  for the Hermite elements H5. (b)-(d) Eigenvectors of Hermite H5 for  $K = 0, 1$  and  $2$  (the parasite modes have been magnified by a factor 100 for visibility).

the behaviour of the numerical wavenumber is again similar to the one of other schemes, showing a slower propagation than the physical wave. The compact finite difference schemes of high order show a very good behaviour in comparison with the high order finite element schemes H3 and H5.

Table 1.3 gives the values of  $\eta_{div}$  for which the relative error between numerical and theoretical wavenumbers is equal to 1%. It is clear that the standard finite differences schemes Hc2 and Hc4 present a lower  $\eta_{div}$  than the finite element schemes. Even the linear element P1 offers a higher  $\eta_{div}$  than the fourth order Hc4 scheme. The best results come from the two Hermite elements H3 and H5 and from the compact finite difference schemes of high order. Lumping the mass matrix of the finite element scheme P3 allows to remove the singularity but leads to an even smaller  $\eta_{div}$  than with the original mass matrix. However, the relative error of the lumped P3 scheme remains within an acceptable range ( $\eta_{div}/\eta_{max} < 15\%$ ) up to a larger wavenumber.



**Figure 1.7:** (top left) Modified numerical wavenumber  $\eta_{num}$  as a function of  $\eta$ , (top right) the relative error between the two wavenumbers and (bottom) the phase velocity.

	$\eta_{max}$	$\frac{\eta_{div}}{\eta_{max}}$
FD Hc2	$\pi$	0.0781
FD Hc4	$\pi$	0.2396
FE P1	$\pi$	0.3554
FE P3	$3\pi$	0.1968
FE P3 lumped	$3\pi$	0.0591
FE H3	$2\pi$	0.676
FE H5	$3\pi$	0.7878
FD compact optimal	$\pi$	0.836
FD compact order 10	$\pi$	0.6817
FD compact order 8	$\pi$	0.6187
FD compact order 6	$\pi$	0.502

**Table 1.3:** Values of  $\eta_{div}$  for which the relative error between numerical and theoretical wavenumbers is equal to 1%.

## Chapter 2

# Numerical results

### 2.1 Shock formation from a sinus wave

#### 2.1.1 Physical behaviour

This section is concerned with the formation of a shock starting from the initial condition  $u(x, 0) = \sin(2\pi x/L)$ . For this purpose, the forcing term is removed for the following results. The formation of the shock from a sine wave is linked with the nonlinear advective term. The maximum and minimum of the sine wave will tend to travel towards each other and join at the middle of the domain, where  $u = 0$ . Then, the viscosity effect will take over and damp the maxima of the shock. This standard test case is used as a validation for the different discretizations discussed in the previous section.

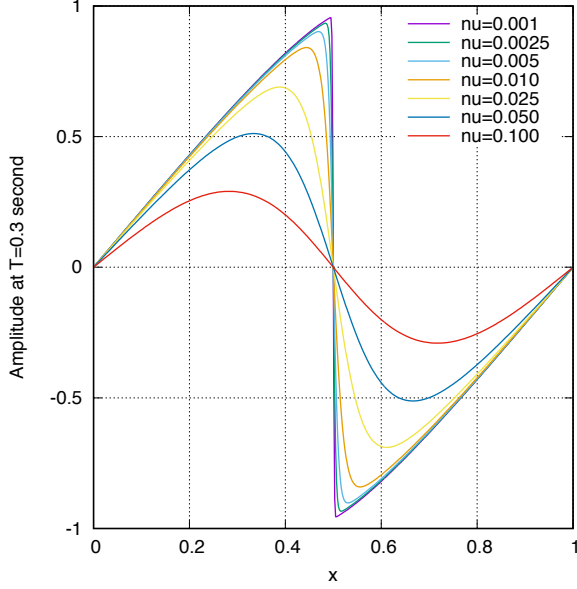
Figure 2.1a shows the amplitude of the solution for several values of the viscosity (from  $\nu = 0.001$  to  $0.1m^2/s$ ) at the same time  $T = 0.3s$ . It is clear that the shock forms only at the lower values of viscosity. For the higher values of  $\nu$ , the viscosity takes over the nonlinear effect and rapidly damps the sine wave.

In the case where  $N$  is too low, oscillations appear around the shock. This phenomenon is shown on Fig. 2.1b for the finite element Lagrange P1 and P3 and for the Hermite H3 and H5. The computations were performed at an identical problem size until  $T = 0.3s$ . The solution is reinterpolated inside the elements for the high order polynomials. The linear Lagrange P1 interpolation has the strongest oscillations but they are rapidly damped away from the shock. The same observation is made for the cubic Lagrange element. For this element, we clearly see that the derivatives of the solution are not conserved between the elements. The Hermite H3 and H5 elements ensure respectively the continuity up to the first and second derivatives of the solution across the elements. However, we see that the parasite oscillations propagate farther from the shock than the two other interpolations. For the Hermite H5 element, the damping of these oscillations is very weak so that they tend to propagate farther away of the shock than any other discretization.

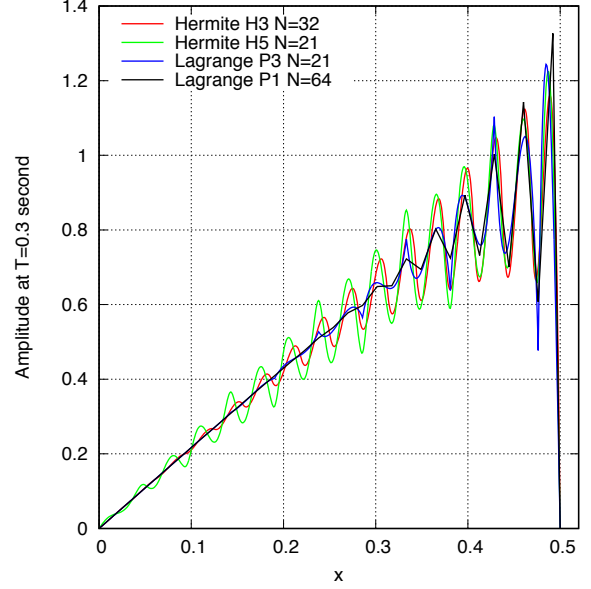
Figure 2.1c shows the evolution of the difference between the two maxima as a function of time for different values of viscosity. In addition analytical expressions are also plotted. It is observed that the amplitude of the solution behaves like  $\exp(-4t/10\nu)$  before the formation of the shock. After the formation of the shock, the amplitude of the solution behaves like  $0.9/(t + 0.1)$ . For  $\nu = 0.1m^2/s$ , the solution is governed by the initial evolution  $\exp(-4t/10\nu)$  because no shock is created.

It can be proved that there exists a relation between the viscosity  $\nu$  and the distance  $\epsilon$  between the two maxima. This relation takes the form  $\log(\epsilon) = \alpha \log(\nu) + \beta$  where  $\alpha$  and  $\beta$  are constants. Figure 2.1d shows the distance between the two maxima as a function of the viscosity

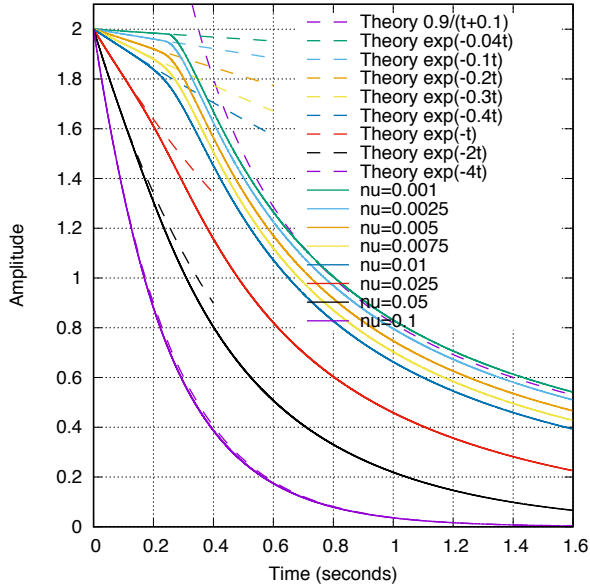
at different times in logarithmic axes. We see that the theoretical relation is indeed respected for the lowest values of  $\nu$  but not for the highest ones.



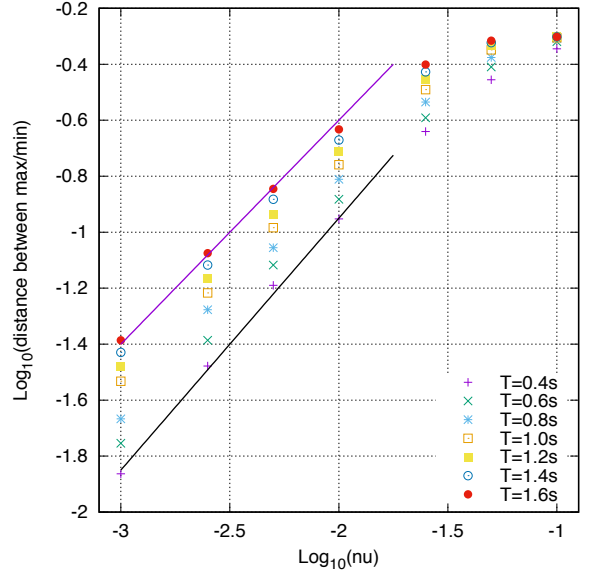
(a) Sinus wave for Lagrange P1 at  $T=0.3s$  and 512 elements



(b) Sinus wave for  $\nu = 0.001$  at  $T=0.3s$  and reinterpolation for Lagrange P3 and Hermite

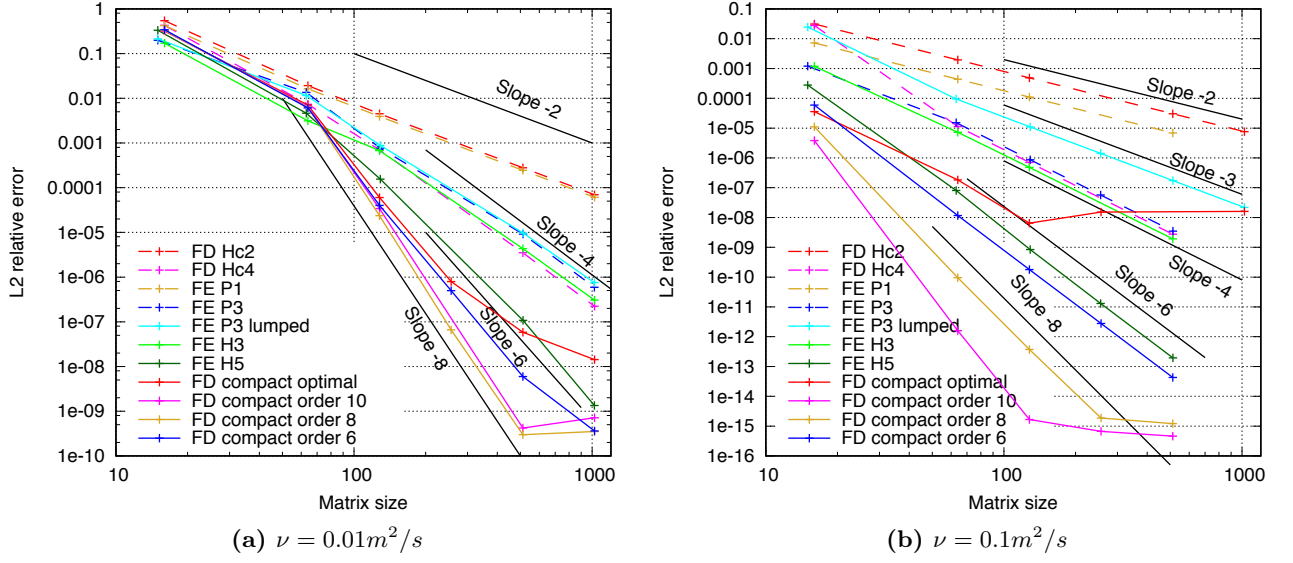


(c) Sinus wave, amplitude vs. time for Lagrange P1 and 512 elements



(d) Sinus wave, gap vs. viscosity for Lagrange P1 and 512 elements

**Figure 2.1:** Formation of a shock from a sinusoidal initial condition



**Figure 2.2:** Formation of a shock from a sinusoidal initial condition - Convergence curves for three values of the viscosity ( $\nu = 0.01, 0.1$  and  $1m^2/s$ ).

### 2.1.2 Convergence study

The formation of the shock possesses an analytical solution in the case of an initial solution  $u(x, 0) = \sin(\pi x)$  where  $0 \leq x \leq 1$ . This solution is defined by the following expression:

$$u_{th}(x, t) = \frac{4\pi\nu \sum_{n=1}^{\infty} n I_n \left( \frac{1}{2\pi\nu} \right) \sin(n\pi x) \exp(-n^2\nu\pi^2 t)}{I_0 \left( \frac{1}{2\pi\nu} \right) + 2 \sum_{n=1}^{\infty} I_n \left( \frac{1}{2\pi\nu} \right) \cos(n\pi x) \exp(-n^2\nu\pi^2 t)} \quad (2.1)$$

where  $I_n$  denotes the modified Bessel function of the first kind and of order  $n$ . Because of the exponential growth of  $I_n$ , it is difficult to evaluate the analytical solution for a viscosity  $\nu < 0.01m^2/s$ . Numerical computations were performed with the four finite element discretizations with an initial solution  $u(x, 0) = \sin(2\pi x)$  and  $0 \leq x \leq 2$ . The numerical solution is then compared with the analytical one in the domain  $0 \leq x \leq 1$ . The comparison is based on the L2 norm of the relative error

$$\epsilon_{L2}^{rel}(t) = \frac{\sqrt{\sum_i [u_{num}(x_i, t) - u_{th}(x_i, t)]^2}}{\sqrt{\sum_i u_{th}^2(x_i, t)}} \quad (2.2)$$

This relative error is plotted in Fig. 2.2 for 2 values of the viscosity ( $\nu = 0.01$  and  $0.1$ ). The computations were all performed at identical matrix sizes until the physical time  $T = 1s$  is reached. At the highest matrix sizes, the slopes of the curves are equal to

- -2 for the finite element Lagrange P1 and finite difference Hc2 schemes
- -3 for the lumped version of the finite element Lagrange P3 scheme
- -4 for the finite element Lagrange P3 and Hermite H3 and the finite difference Hc4 schemes



- -6 for the finite element Hermite H5 scheme and the compact finite difference scheme of order 6
- -8 for the compact finite difference scheme of order 8
- -10 for the compact finite difference scheme of order 10, at least for  $\nu = 0.1$
- unknown for the compact optimal finite difference scheme with spectral-like resolution

It is also clear that all the high order schemes lose their advantage at low values of the viscosity and low numbers of elements. This is because of the presence of parasite oscillations around the shock that appears in the middle of the domain. It is also with this combination of low values for the viscosity and numbers of elements that the relative errors are the highest. At low values of viscosity, the high order schemes retrieve their rate of convergence at high number of elements, when the parasite oscillations disappear.

Depending on the value of the viscosity, the finite element P1 and the finite difference scheme Hc2 show nearly identical errors or an order of magnitude different. The same observation is made for the finite elements P3 and H3 and the finite difference scheme Hc4.

### 2.1.3 A note on the maximum time step

Another interesting property of all the spatial discretization is their different limit on the increment in time which can be used before the explicit time integration scheme becomes unstable. However all the schemes present the same pattern: the limit on the increment in time is

$$\Delta t_{\max} = \text{constant} \frac{\Delta x^2}{\nu} \quad (2.3)$$

The maximum increments in time for each spatial discretization are given in Table 2.1. The results were based on the computations performed in the previous section. It must be underlined that the computations were performed at identical matrix sizes (the H5/P3 schemes need three times less nodes than the P1 scheme to have the same matrix size. The H3 scheme needs two times less nodes than the P1 scheme). The quadratic dependency of  $\Delta t_{\max}$  on the mesh size  $\Delta x$  and the inverse dependency on the viscosity  $\nu$  are clearly visible. The finite difference scheme Hc2 offers the largest increment in time whereas the finite element Lagrange P3 shows the most constraining increment in time. Lumping the mass matrix allows to multiply by a factor 2.4 this limit for the Lagrange P3 element. As a reminder, the lumped version of the P1 scheme is strictly equal to the Hc2 scheme and is thus not used here.

If the computations are performed at identical mesh sizes  $\Delta x$ , then the constant in Eq. (2.3) can be evaluated for the different schemes and are given in Eq. (2.4). It is clear that, at identical mesh sizes  $\Delta x$ , the high order finite element schemes show a very constraining limit on the increment in time in comparison with the P1 element and the finite difference schemes. Lumping the mass matrix helps to increase this limit by multiplying it by a factor 2.4 for the P3 scheme.

$$\begin{aligned} \Delta t_{\max}^{\text{Hc2}} &\simeq 0.687 \frac{\Delta x^2}{\nu} & \Delta t_{\max}^{\text{Hc4}} &\simeq 0.517 \frac{\Delta x^2}{\nu} & \Delta t_{\max}^{\text{P1}} &\simeq 0.231 \frac{\Delta x^2}{\nu} \\ \Delta t_{\max}^{\text{P3}} &\simeq 0.016 \frac{\Delta x^2}{\nu} & \Delta t_{\max}^{\text{P3 lumped}} &\simeq 0.039 \frac{\Delta x^2}{\nu} & & \\ \Delta t_{\max}^{\text{H3}} &\simeq 0.065 \frac{\Delta x^2}{\nu} & \Delta t_{\max}^{\text{H5}} &\simeq 0.03 \frac{\Delta x^2}{\nu} & & \end{aligned} \quad (2.4)$$

Viscosity Matrix size	$\Delta t_{\max}$					
	$\nu=0.01m^2/s$		$\nu=0.1m^2/s$		$\nu=1m^2/s$	
	512	1024	128	512	128	256
Hc2	1.05e-3	2.63e-4	1.67e-3	1.05e-4	1.67e-4	4.17e-5
Hc4	7.94e-4	1.96e-4	1.25e-3	7.94e-5	1.26e-4	3.17e-5
P1	3.57e-4	8.77e-5	5.56e-4	3.45e-5	5.56e-5	1.41e-5
P3	2.22e-4	5.62e-5	3.45e-4	2.22e-5	3.45e-5	8.93e-6
Lumped P3	5.26e-4	1.37e-4	8.33e-4	5.26e-5	8.33e-5	2.17e-5
H3	4e-4	1e-4	6.45e-4	4e-5	6.45e-5	1.61e-5
H5	4e-4	1.02e-4	6.45e-4	4.08e-5	6.45e-5	1.63e-5

**Table 2.1:** Maximum value for the increment in time for the spatial discretization with an explicit four steps Runge-Kutta time integrator at identical matrix sizes.

## 2.2 Turbulent flow

The forced Burgers equation (1.1) is solved through the numerical discretization given in the previous section. A random initial solution evolves towards the unsteady turbulent regime. This turbulent regime is obtained after 10 physical seconds according to the evolution of the kinetic energy. At this moment, the energy spectrum is computed and averaged in time over 200 physical seconds. These energy spectra are then compared with those coming from the spectral method, which is used as a reference. The first section will present the results for computations performed with an insufficient number of nodes to reach the dissipation range. In the second section, resolved DNS are performed. Concerning the cubic Lagrange and Hermite methods, the physical solution is reinterpolated within each element before computing the energy spectra. This ensures that the cubic representation of the solution is taken into account.

### 2.2.1 Under-resolved DNS

In the context of LES, a cutoff wavenumber is chosen in the inertial range at which the numerical result is filtered and the higher wavenumber are modelled. In this context, the best numerical scheme is the one which follow the inertial range the longer possible. In the figures 2.3, the energy spectra are plotted as a function of the wavenumber for three values of the number of nodes ( $N = 64, 128$  and  $256$ ) and two values for the viscosity ( $\nu = 0.0035$  and  $0.0075m^2/s$ ). The Kolmogorov wavenumber is given as an indication ( $k_\nu = 2\pi/L_\nu$  with the Kolmogorov length  $L_\nu = LRe^{-3/4}$  and  $Re = \text{mean}(u)L/\nu$ ). Note that the computations are performed at identical problem sizes. The Lagrange P3/Hermite H5 and Hermite H3 methods require respectively three and two times less nodes than the linear Lagrange polynomial P1.

With  $N = 64$ , the second-order the divergent (Hd) and energy conservative (Hc2) finite difference schemes and the DFD scheme show poor results in the inertial range. The energy spectra move off the reference curves rapidly to reach a maximum. The excitation of the high frequencies is linked with the birth of high-frequency oscillations around the shocks in the physical solution. The fourth order energy conservative scheme Hc4 has comparable results as with the linear Lagrange finite element P1. In overall, the other spatial discretizations of higher order show better results with  $N = 64$ .

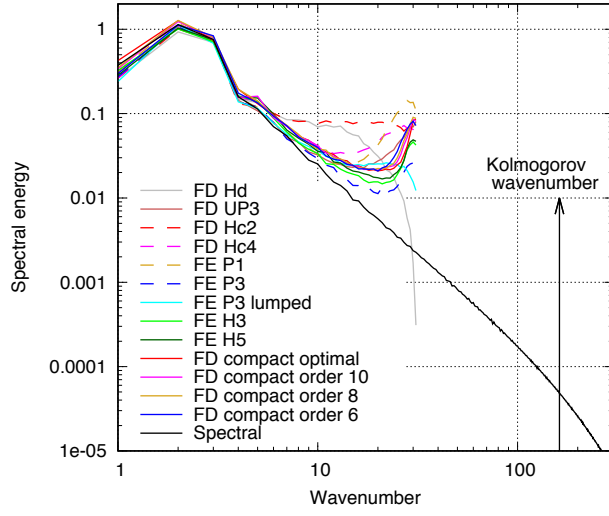
For  $N = 128$  and  $256$ , the inertial range is better captured by the three schemes. Note the progressive damping of the maxima at high wavenumbers due to the disappearance of the oscillations around the shocks. This maximum even disappear for the divergent scheme Hd with  $\nu = 0.0075m^2/s$  and  $N = 256$ . The cubic Lagrange P3 and Hermite H3 and fifth order Hermite H5 finite elements offer the best performances with the break-off from the inertial range

located at higher wavenumbers. The fifth order Hermite H5 element shows stronger peaks at high wavenumbers than the cubic Hermite H3 element. This is due to the weaker damping of the parasite oscillations for the H5 element, as seen in section 2.1. The cubic Lagrange P3 element fits the inertial range up to the same wavenumbers as the Hermite elements but shows weaker excitations of the high wavenumbers than the Hermite polynomials. In the configuration of subfigure 2.3f, the peak disappears and is replaced by an over-damping in comparison with the spectral method.

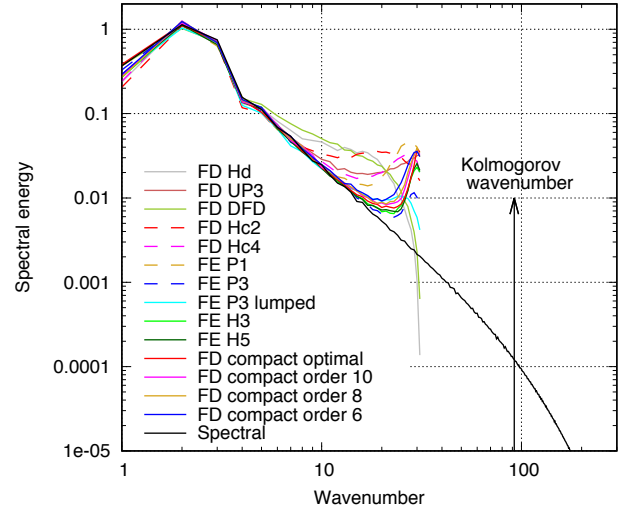
### 2.2.2 Resolved DNS

Figure 2.5 shows the results of the different numerical schemes for resolved DNS computations. The problem size is equal to 2048. The relative error is computed with the spectral method used as the reference. The inertial range is very well computed by all the schemes. The main differences appear at higher wavenumbers than the Kolmogorov wavenumber:

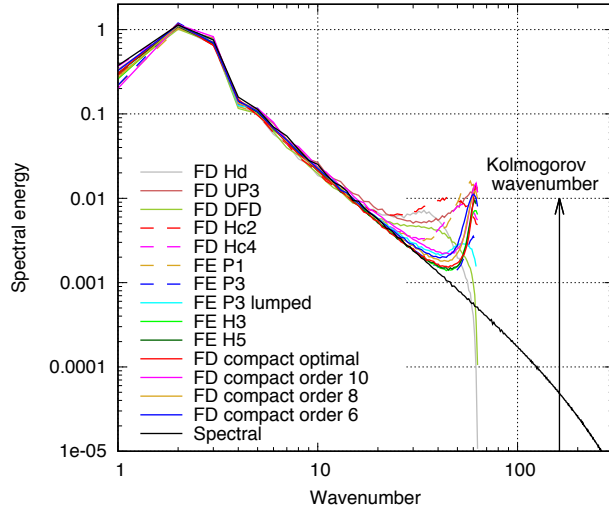
- The second order divergent finite difference scheme Hd has the worst behaviour at large wavenumbers: the diffusive range is not satisfactorily computed with a too high damping.
- The second and fourth order energy conservative scheme Hc2/Hc4, the linear Lagrange finite element P1 and the cubic Lagrange finite element P3 present similar results.
- The linear Lagrange finite element P1 offers the best result for  $\nu = 0.0075m^2/s$  whereas the cubic Hermite finite element H3 is nearly identical to the spectral method for  $\nu = 0.0035m^2/s$ .
- The Hermite H5 element shows disappointing results in overall, in comparison with the third order methods.
- The Hc4, P3, H3 and H5 still have an parasite excitation of the highest wavenumbers, mainly for  $\nu = 0.0075m^2/s$ .



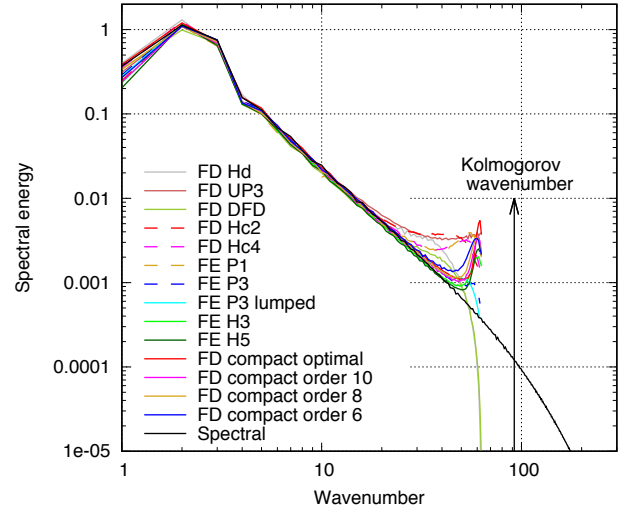
(a) Matrix size=64



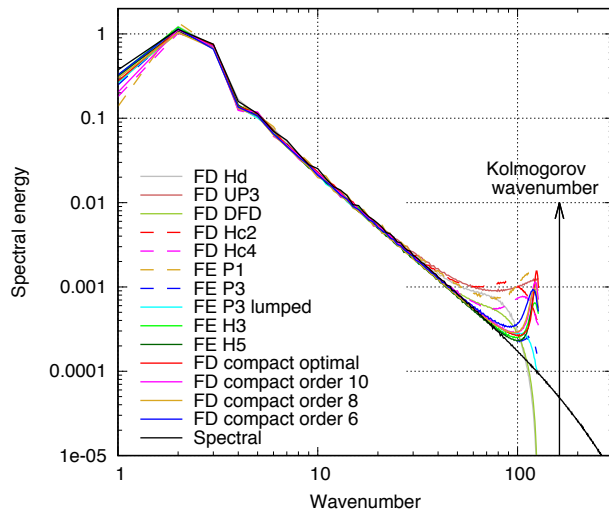
(b) Matrix size=64



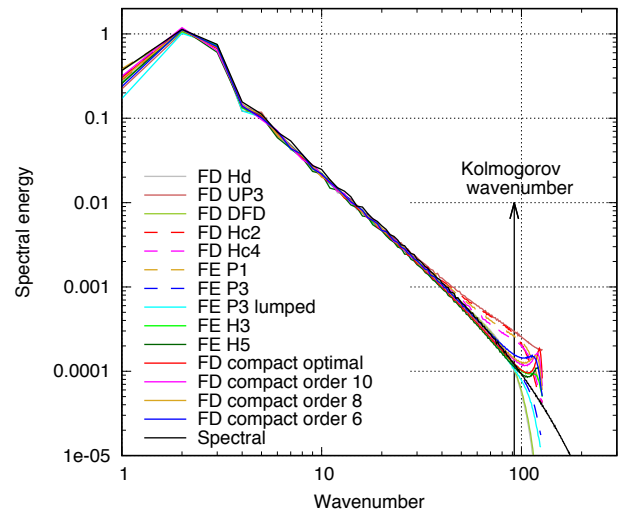
(c) Matrix size=128



(d) Matrix size=128

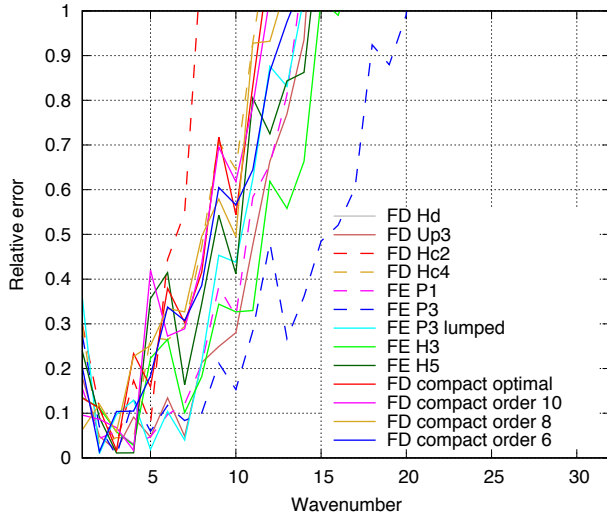


(e) Matrix size=256

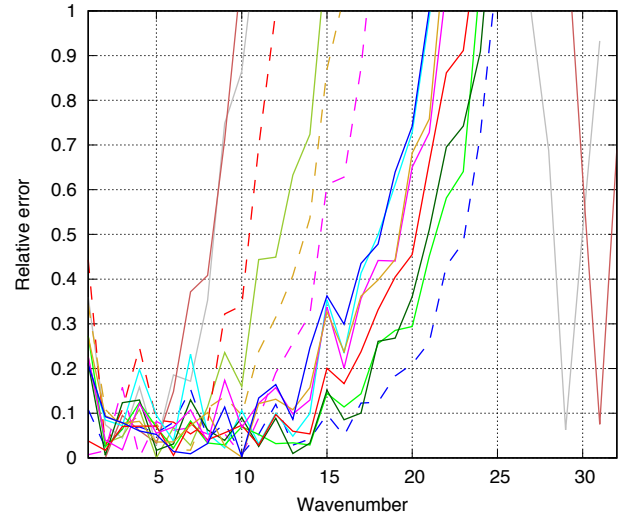


(f) Matrix size=256

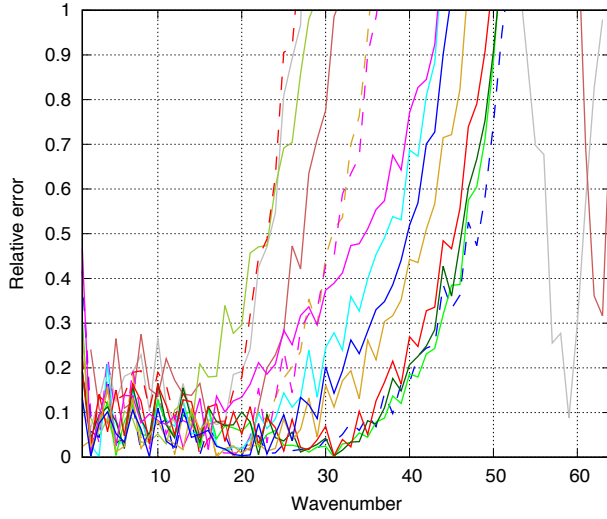
**Figure 2.3:** Under-resolved DNS - Energy spectra for (left)  $\nu = 0.0035m^2/s$  and (right)  $\nu = 0.0075m^2/s$



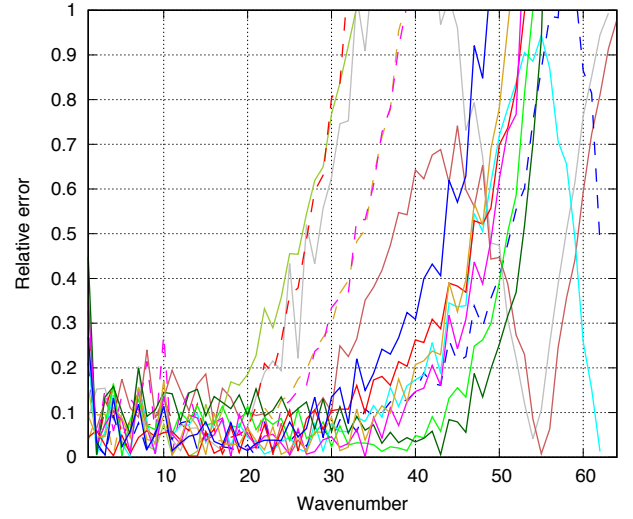
(a) Matrix size=64



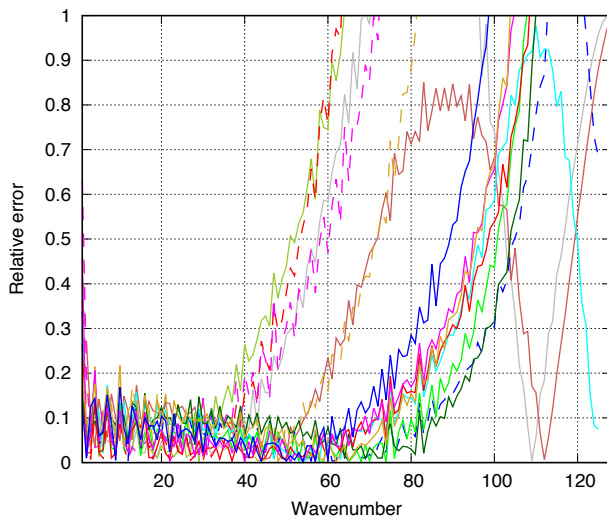
(b) Matrix size=64



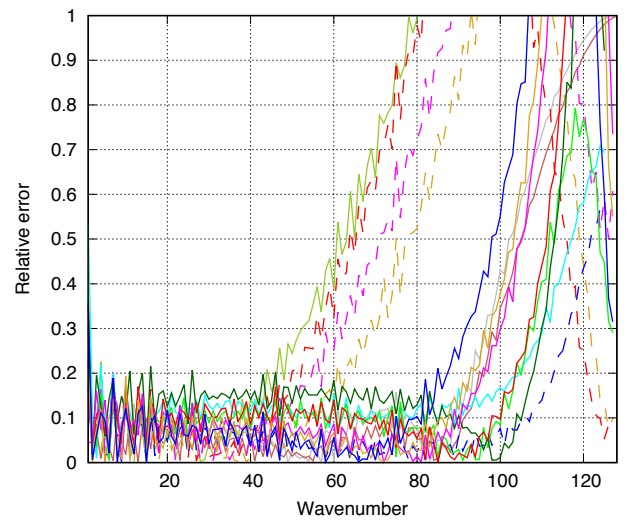
(c) Matrix size=128



(d) Matrix size=128

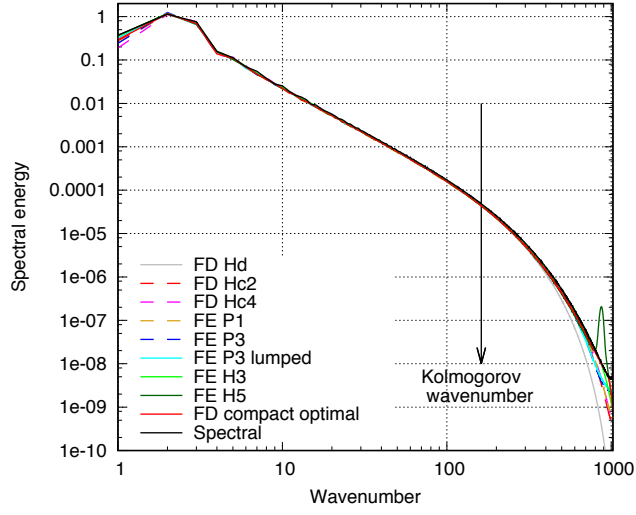


(e) Matrix size=256

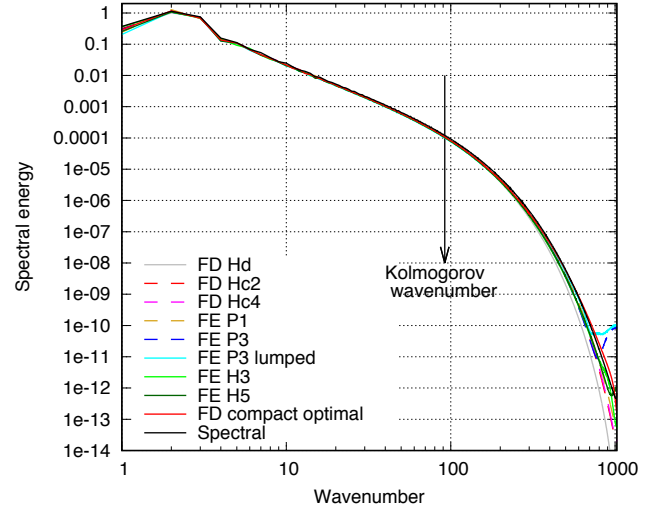


(f) Matrix size=256

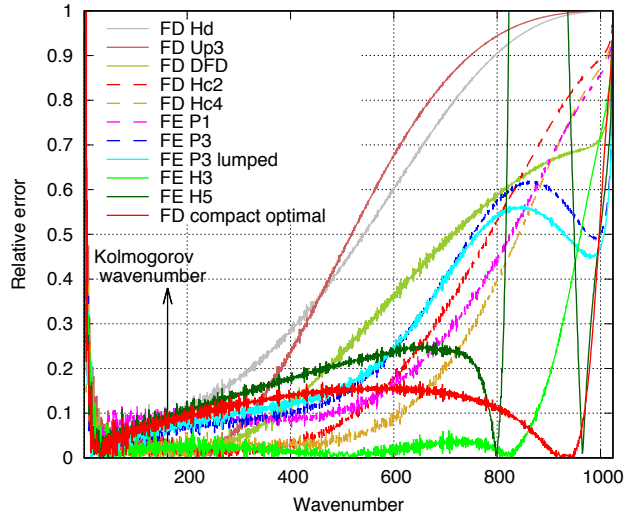
**Figure 2.4:** Under-resolved DNS - Energy spectra for (left)  $\nu = 0.0035m^2/s$  and (right)  $\nu = 0.0075m^2/s$



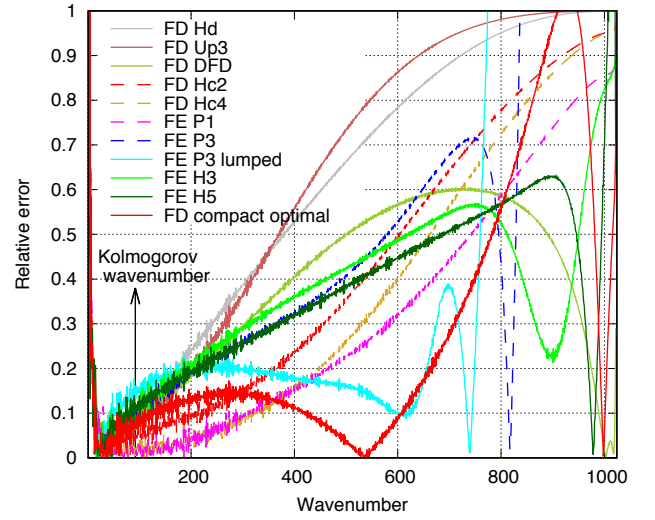
(a) Energy spectra



(b) Energy spectra



(c) Relative error



(d) Relative error

**Figure 2.5:** Resolved DNS - Matrix size=2048 - Energy spectra for (left)  $\nu = 0.0035\text{m}^2/\text{s}$  and (right)  $\nu = 0.0075\text{m}^2/\text{s}$

# Bibliography

- [1] Lele S.K., Compact finite difference schemes with spectral-like resolution, *Journal of Computational Physics* 103 (1992) 16-42.
- [2] Fauconnier D. and Dick E., On the spectral and conservation properties of nonlinear discretization operators, *Journal of Computational Physics* 230 (2011) 4488-4518.
- [3] Wang R. and Spiteri R.j., Linear instability of the fifth-order WENO method, *SIAM J. Numer. Anal.* 45(5) (2007) 1871-1901.
- [4] Ladeinde F. *et al.*, Turbulence spectra characteristics of high order schemes for direct and large eddy simulation, *Applied Numerical Mathematics* 36 (2001) 447-474.