Dispersion and Cost Analysis of Some Finite Difference Schemes in One-Parameter Acoustic Wave Modelling

William W. Symes* and Quang Huy Tran†

June 10, 1994

Abstract

A systematic comparison is carried out between some standard finite difference schemes, regarding their costs and dispersion properties. To be more specific, given a precision threshold to be imposed on the velocity error and a finite difference scheme, it is possible to determine a time-step and a grid-spacing in an optimal manner, i.e. so as to minimize the computational cost. Using this optimal cost as a criterion, it becomes easy to single out the most economical scheme for the purpose of a synthetic seismic campaign.

This survey represents the preliminary part of the larger project Marmousi 3-D, the purpose of which is to create a synthetic database corresponding to one-parameter media. In this paper, one's attention is focused on the 2-2m family of schemes. Since both homogeneous and heteregenous cases are investigated, the study is expected to provide realistic figures for future simulations.

1 Introduction

The use of numerical methods to produce synthetic seismograms dates back to the seventies with the paper by Kelly et al. [6]. Undeniably, its advantage over other traditional approaches like asymptotic expansion or Kirchhoff integral is that by directly dealing with the acoustic wave equation, it takes into account all physical phenomena. However, for such methods to be valid, it is capital to be able to overcome —or at least to control— the artifacts they are likely to bring about.

One of the major artifacts is dispersion. Since the numerical scheme is only an approximation to the wave equation, a plane-wave calculation [15] reveals that it always makes the phase and group velocities depend on the frequency. This discrepancy in velocity is all the more noticeable as the propagation time is long. After some critical amount of time, the signal is so much distorted [1, 14] that it becomes hopelessly impossible to recognize the shape of the traveling pulse.

For a given propagation time, it is always possible to lessen the effects of dispersion by choosing smaller values for the sampling intervals. Nevertheless, there is a trade-off between the numerical

*Dept. of Comput. and Appl. Math., Rice University, P.O. Box 1892, Houston TX 77251, USA.
†Institut Français du Pétrole, DIMA-DER, B.P. 311, 92506 Rueil-Malmaison Cedex, France.

accuracy thus obtained and the computational cost. The latter issue is of paramount importance, to the extent that we are having a 3-D project in mind. We will therefore indicate an optimal strategy to select the time- and space-steps within the requirement of accuracy. For simplicity, the numerical schemes we will be concerned with are Taylor finite difference schemes whose order in time is 2 and whose order in space is $2m$. Other types of scheme, such as spectral methods [4, 8, 11] or Holberg coefficients [5] are in reality not devoid of drawbacks. These will not be considered in this paper.

It may come as a little surprise to realize that although this approach is fairly elementary, no systematic comparison has been previously carried out for the family of schemes considered. On one hand, the work initiated by Sei [13, 14] applies to stencils of the form A^tA which operate on two-parameter media. Transposing Sei's ideas to the present case is relatively straightforward. On the other hand, all studies regarding the choice of sampling intervals and computer costs [7, 10] have addressed only homogeneous media. This is why we will need to find out a suitable way to proceed in heterogeneous media.

This paper is divided into three parts. First, we begin by recalling some basic notions on finite difference schemes and their dispersive behavior. Then, we will establish some useful theoretical properties which would make it easier to derive an optimal strategy. Finally, a thorough investigation on the cost of the schemes will be taken up in both homogeneous and heteregenous media.

2 Mathematical background

2.1 Approximation of the Laplacian

First, examine the 1-D case. Let v be a function of the variable $x \in \mathbf{R}$. Consider a regularly spaced sequence of nodes $x_i = i\Delta x$ where $i \in \mathbf{Z}$ and $\Delta x > 0$. If $v_i = v(x_i)$, then the ratio

$$(\Delta_x v)_i = \frac{v_{i+1} - 2v_i + v_{i-1}}{\Delta x^2} \tag{2.1}$$

is known to be a second order approximation of $v_{xx}(x_i)$, in the sense that the leading term in the difference between the latter and the former is proportional to Δx^2. More generally, for any integer $p \geq 1$, introduce the discrete Laplacian associated with a larger step $p\Delta x$

$$(\Delta_x^p v)_i = \frac{v_{i+p} - 2v_i + v_{i-p}}{p^2 \Delta x^2}. \tag{2.2}$$

This is also a second order approximation of $v_{xx}(x_i)$, although the actual error between $(\Delta_x^p v)_i$ and $v_{xx}(x_i)$ is p^2 times greater than that between $(\Delta_x v)_i$ and $v_{xx}(x_i)$. It is natural to combine the $(\Delta_x^p v)_i$'s with different p's in order to get a higher order approximation of $v_{xx}(x_i)$. Put another way, $m \geq 1$ being an integer, we look for a linear combination of m elementary discrete Laplacians

$$(D_x^m u)_i = \sum_{p=1}^{m} \alpha_p^m \, (\Delta_x^p v)_i \tag{2.3}$$

which approximates $v_{xx}(x_i)$ at the $2m$-order.

Lemma 2.1 *A necessary and sufficient condition for $(D_x^m u)_i$, defined by (2.3), to be a $2m$-th order approximation of $v_{xx}(x_i)$ is that the coefficients α_p^m are solution to the Van der Monde system*

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 \\ 1^2 & 2^2 & 3^2 & \cdots & \cdots & m^2 \\ 1^4 & 2^4 & 3^4 & \cdots & \cdots & m^4 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1^{2m'} & 2^{2m'} & 3^{2m'} & \cdots & \cdots & m^{2m'} \end{pmatrix} \begin{pmatrix} \alpha_1^m \\ \alpha_2^m \\ \alpha_3^m \\ \cdots \\ \cdots \\ \alpha_m^m \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ \cdots \\ 0 \end{pmatrix} \tag{2.4}$$

where $m' = m - 1$.

PROOF It suffices to carry out the Taylor expansions, and to proceed to termwise identification. Cancelling out the errors in Δx^{2p} for $1 \leq p \leq m$ leads to the $m-1$ last equations. $\lhd$

Let us give some frequently used examples. For $m = 2$, we have

$$\alpha_1^2 = \frac{4}{3} \quad \text{and} \quad \alpha_1^2 = -\frac{1}{3},$$

which gives rise to the fourth order stencil. For $m = 3$, the coefficients are

$$\alpha_1^3 = \frac{3}{2}, \quad \alpha_2^3 = -\frac{3}{5} \quad \text{and} \quad \alpha_3^3 = \frac{1}{10}.$$

For $m = 4$, we end up with

$$\alpha_1^4 = \frac{8}{5}, \quad \alpha_2^4 = -\frac{4}{5}, \quad \alpha_3^4 = \frac{8}{35} \quad \text{and} \quad \alpha_4^4 = -\frac{1}{35}.$$

2.2 Discretization of the wave equation

We wish to solve the wave equation

$$\frac{1}{c^2} u_{tt} - [\, u_{xx} + u_{yy} + u_{zz} \,] = \delta \otimes R \tag{2.5}$$

where δ is the Dirac function and R the time dependency of the source. Take a time step Δt and a space step $\Delta x = \Delta y = \Delta z = h$. Consider a regular grid of points $M_{i,j,k} = (ih, jh, kh)$. Denote by $u_{i,j,k}^n$ the approximate value for u at time $n\Delta t$ and at point $M_{i,j,k}$. Replace the wave equation (2.5) by its discrete version

$$\frac{1}{c_{i,j,k}^2} (\Delta_t u_{i,j,k})^n - \left[\left(D_x^m u_{j,k}^n\right)_i + \left(D_y^m u_{k,i}^n\right)_j + \left(D_z^m u_{i,j}^n\right)_k \right] = \delta_{i,j,k} \otimes R(n\Delta t) \tag{2.6}$$

where $m \geq 1$ is an integer and D^m the operator introduced by (2.3). From the standpoint of accuracy, the numerical scheme (2.6) is a 2-$2m$ scheme, i.e. second order in time and $2m$-th order in space. For the sake of notation conveniency, we will henceforth write (2.6) as

$$\frac{1}{c_{i,j,k}^2} (\Delta_t u_{i,j,k})^n - \sum_{\flat \in \{x,y,z\}} (D_\flat^m u^n)_{i,j,k} = \delta_{i,j,k} \otimes R(n\Delta t) \tag{2.7}$$

This is the family of finite difference schemes we will be concerned with. We will try to precise how to choose suitable values for the steps h and Δt, in connection with dispersion and cost issues.

2.3 Relation of dispersion

Assume for the time being that the medium is homogeneous, i.e. $c_{i,j,k} = c$ for all $M_{i,j,k}$, and that the excitation source R in the right-hand side is zero. To the discrete problem (2.6) we seek harmonic plane wave solutions, i.e. of the form

$$u^n_{i,j,k} \,=\, \exp \imath \left[\omega n \Delta t - (\xi_x i + \xi_y j + \xi_z k)\, h \right] \tag{2.8}$$

where $\omega > 0$ is the pulsation and $\boldsymbol{\xi} = (\xi_x\,,\,\xi_y\,,\,\xi_z\,)$ denotes the wave vector.

Proposition 2.1 *The discrete wave equation (2.6) in a homogeneous medium without excitation source admits the harmonic plane wave (2.8) as a solution if and only if the following condition, called relation of dispersion, is satisfied*

$$\sin^2\left(\frac{\omega\Delta t}{2}\right) \,=\, \left(\frac{c\Delta t}{h}\right)^2 \sum_{p=1}^{m} \frac{\alpha_p^m}{p^2} \sum_{\flat\in\{x,y,z\}} \sin^2\left(\frac{p\xi_\flat h}{2}\right) .$$

PROOF Plug (2.8) into the numerical scheme (2.6), try to factor out $u^n_{i,j,k}$, and make some trigonometrical transformations. ◁

Obviously, the quantities $\dfrac{\omega}{\|\boldsymbol{\xi}\|}$ and $\left\| \dfrac{d\omega}{d\boldsymbol{\xi}} \right\|$, respectively known as phase and group velocities, are not equal to c but depend on the wave vector $\boldsymbol{\xi}$. Thus, each harmonic component travels at its own speed. This phenomenon, which may turn very annoying as far as numerical solutions are concerned, is commonly referred to as dispersion.

We will study more thoroughly the behavior of the velocity errors in the next section. Let us beforehand mention a quite interesting trigonometrical property of which we will make extensive use later.

Lemma 2.2 *There exists a sequence of positive numbers β_p , $p \geq 1$, the values of which are independent of the m, such that*

$$\forall\, m \geq 1\,, \quad \forall\, \theta \in \mathbf{R}\,, \qquad \sum_{p=1}^{m} \frac{\alpha_p^m}{p^2} \sin^2(p\theta) \,=\, \sum_{p=1}^{m} \beta_p \sin^{2p}(\theta) \tag{2.9}$$

where α_p^m , $p \in \{1, \cdots, m\}$, are the coefficients determined by the system (2.4).

PROOF Just carry out the calculations. We end up with

$$\beta_1 = 1\;, \quad \beta_2 = \frac{1}{3}\;, \quad \beta_3 = \frac{8}{45}\;, \quad \beta_4 = \frac{4}{35}\; \cdots$$

These first four terms will be sufficient for our purpose. ◁

This nice transformation provides us with a more usable relation of dispersion

$$\sin^2\left(\frac{\omega\Delta t}{2}\right) \,=\, \left(\frac{c\Delta t}{h}\right)^2 \sum_{p=1}^{m} \sum_{\flat\in\{x,y,z\}} \beta_p \sin^{2p}\left(\frac{\xi_\flat h}{2}\right) . \tag{2.10}$$

2.4 Phase and group velocities

Let ω and $\boldsymbol{\xi}$ be connected by the relation of dispersion (2.10). The phase velocity, defined as

$$c_\varphi = \frac{\omega}{\|\boldsymbol{\xi}\|},$$

is the speed at which propagates the plane $\omega t - \boldsymbol{\xi}.\boldsymbol{r} = Cte$. The group velocity, defined as

$$c_g = \left\| \frac{d\omega}{d\boldsymbol{\xi}} \right\| = \left[\left(\frac{\partial\omega}{\partial\xi_x}\right)^2 + \left(\frac{\partial\omega}{\partial\xi_y}\right)^2 + \left(\frac{\partial\omega}{\partial\xi_z}\right)^2 \right]^{1/2},$$

can be shown [3, 15] to represent the speed at which propagates the energy packet corresponding to the wave vector $\boldsymbol{\xi}$.

3 Choice of discretization parameters

3.1 Accuracy criterion

Let T be lapse of time over which we would like to carry out numerical simulations. The absolute error, measured by the difference in the distance covered after time T, is equal to $T\,(c_g(\boldsymbol{\xi}) - c)$ for the wave component $\boldsymbol{\xi}$. It is desirable to impose an appropriate upper-bound on this error, so as to ensure accuracy to the numerical solution. The upper-bound can, for instance, be a fraction of some characteristic wavelength. In geophysical modelling, it is typical [10, 13, 14] to set this wavelength to

$$\lambda_{\min} = \frac{c}{f_{\max}}, \tag{3.1}$$

where $f_{\max}$ denotes the cutoff frequency of the source wavelet R. With this shortest wavelength is associated the largest wave number

$$\xi_{\max} = \frac{2\pi}{\lambda_{\min}}.$$

By means of these quantities, our criterion can now be expressed as

$$\forall\boldsymbol{\xi},\quad \|\boldsymbol{\xi}\| \le \xi_{\max} \quad\Longrightarrow\quad |\,T\,(c_g(\boldsymbol{\xi}) - c)\,| \le \frac{\lambda_{\min}}{n_\lambda}.$$

where $1/n_\lambda$, the fraction of wavelength, is at our disposal. In order to get an non-dimensional criterion, divide both sides of the above inequality by cT. This yields

$$\forall\boldsymbol{\xi},\quad \|\boldsymbol{\xi}\| \le \xi_{\max} \quad\Longrightarrow\quad \left|\frac{c_g(\boldsymbol{\xi}) - c}{c}\right| \le \frac{1}{n_\lambda}\frac{\lambda_{\min}}{cT} = \frac{1}{n_\lambda}\frac{1}{n_T}, \tag{3.2}$$

where $n_T = cT/\lambda_{\min}$, the number of shortest wavelengths covered after time T, is also at our disposal. Condition (3.2) will be the most important constraint on the parameters h et Δt. Its left-hand side

$$e_g = \frac{c_g - c}{c} \tag{3.3}$$

appears to be the relative group velocity error. Similarly, the relative phase velocity error

$$e_\varphi = \frac{c_\varphi - c}{c} \tag{3.4}$$

can be introduced. In view of (2.10), e_g and e_φ depend a priori on m, c, $\boldsymbol{\xi}$, h and Δt.

3.2 Computational cost

In addition to the accuracy criterion, the parameters h and Δt must be chosen so as to minimize the cost of simulations. Let us elaborate a little more on the concept of cost.

Suppose on one hand that the sizes of the model are $l_x \times l_y \times l_z$. The number of gridpoints needed to sample this model is $l_x/h \times l_y/h \times l_z/h$, so is inversely proportional to h^3. On the other hand, if the simulation time is T, then the number of time-steps necessary to cover this interval is $T/\Delta t$, therefore is inversely proportional to Δt.

Let N_m be the number of elementary floating point operations involved in the 2-2m scheme. The actual value of N_m may depend on the programming tricks used in the implementation of the scheme. In our machine-independent optimized version of the code, we have $N_m = 4 + 7m$. When the 2-2m scheme is employed, the cost of a simulation is proportional to

$$J_m = J_m(h, \Delta t) \propto \frac{N_m}{h^3 \Delta t} \tag{3.5}$$

which is to be minimized under various constraints.

Thus far, we have regarded the number of floating point operations as a measure of the cost. Undoubtedly, it would have been more relevant to look at the CPU time. However, for this aspect of the problem, we would have had to take into account too many architecture-dependent factors such as load/store operations, memory hierarchies or parallelization issues [7, 10]. We believe that the cost J_m defined above is sufficient to compare different numerical schemes.

3.3 First formulation

We are now in a position to state the problem of choosing h and Δt in homogeneous media as a constrained minimization problem. The formulation (**F1**) given below is that which would come right away to our minds, but is not very convenient to work with. It will be made simpler and more exploitable in the next section.

(**F1**) GIVEN

c	velocity of the homogeneous medium
$f_{\max}$	cutoff frequency of the source wavelet
m	half-order in space of the scheme 2-$2m$
ϵ	accuracy threshold $(1/n_\lambda n_T)$

FIND

h and Δt so as to <u>minimize</u> $J_m(h, \Delta t)$ under the following constraints

- Nyquist's frequency

$$h \leq \frac{1}{2}\,\lambda_{\min} = \frac{1}{2}\,\frac{c}{f_{\max}} \tag{3.6}$$

- stability condition

$$\frac{c\Delta t}{h} < \gamma_m^{\max} = \left(3\sum_{p=1}^{m}\beta_p\right)^{-1/2} \tag{3.7}$$

- accuracy criterion

$$\forall \boldsymbol{\xi}, \quad \|\boldsymbol{\xi}\| \leq \xi_{\max} = \frac{2\pi}{\lambda_{\min}} \quad \Longrightarrow \quad |\, e_g(m, c, h, \Delta t, \boldsymbol{\xi})\,| \leq \epsilon \tag{3.8}$$

REMARK 2.1 The stability condition (3.7) can be readily derived [9, 13] from the relation of dispersion (2.10). □

REMARK 2.2 The accuracy threshold ϵ is to be set by the user according to what is desired for n_λ and n_T. □

4 Optimal strategy for the homogeneous case

4.1 Properties of velocity errors

In order to find a handier version for (**F1**) and to ultimately solve it, let us go through some useful properties of the velocity errors e_g and e_φ. To begin with, it is advisable to perform a few transformations so as to express these in terms of the following non-dimensional variables

γ	=	$c\Delta t/h$	stability ratio
$\boldsymbol{\nu}$	=	$\boldsymbol{\xi}/\|\boldsymbol{\xi}\|$	unit vector in the direction of $\boldsymbol{\xi}$
H	=	$h/\lambda_{\min}$	inverse of the number of points per wavelength
s	=	$\|\boldsymbol{\xi}\|\lambda_{\min}/2\pi$	ratio of $\|\boldsymbol{\xi}\|$ to $\xi_{\max}$

Lemma 4.1 *For every integer $m \geq 1$, the relative velocity errors e_g and e_φ are functions of m and the four variables γ, $\boldsymbol{\nu}$, H, s alone. More specifically,*

$$e_\varphi = \frac{1}{\pi\gamma s H} \arcsin\left[\gamma\left(\sum_{p=1}^{m}\sum_{\flat\in\{x,y,z\}} \beta_p \sin^{2p}(\pi\nu_\flat s H)\right)^{1/2}\right] - 1$$

$$e_g = \frac{\gamma\left[\displaystyle\sum_{\flat\in\{x,y,z\}}\left(\sum_{p=1}^{m} 2p\beta_p \sin^{2p-1}(\pi\nu_\flat s H)\cos(\pi\nu_\flat s H)\right)^2\right]^{1/2}}{\sin\left\{2\arcsin\left[\gamma\left(\displaystyle\sum_{p=1}^{m}\sum_{\flat\in\{x,y,z\}}\beta_p \sin^{2p}(\pi\nu_\flat s H)\right)^{1/2}\right]\right\}} - 1$$

PROOF Take the square root of both sides of (2.10) and pull out ω. Divide by $c\|\boldsymbol{\xi}\|$ and subtract one to get the first formula.

As far as the second formula is concerned, take the derivative of both sides of (2.10) with respect to $\xi_\flat$ for $\flat \in \{x, y, z\}$ while keeping in mind that ω depends on $\xi_\flat$. This allows us to deduce $\dfrac{\partial\omega}{\partial\xi_\flat}$ and to proceed further. $\triangleleft$

In the above equations, the ratio γ must be less than the stability threshold $\gamma_m^{\max}$ defined in (3.7). It should also be strictly positive. By passage to the limit when $\gamma\downarrow 0$, we obtain the formulae corresponding to the semi-discrete case. The general behavior of e_φ and e_g with respect to γ is given by

Proposition 4.1 *For any integer $m \geq 1$ and for fixed $\boldsymbol{\nu}$, s, H, the errors e_φ and e_g are strictly increasing functions of $\gamma \in]\,0, \gamma_m^{\max}\,[$.*

PROOF C_1, C_2 and C_3 denoting various positive constants, we see from Lemma 4.1 that

$$e_\varphi = \frac{\arcsin(C_1\gamma)}{C_2\gamma} - 1 = \frac{C_1}{C_2}\,\frac{\arcsin(C_1\gamma)}{C_1\gamma} - 1$$

where $C_1\gamma \in]\,0, 1\,[$ in view of $\gamma_m^{\max}$'s value. Now, the function $\dfrac{\arcsin x}{x}$ is strictly increasing over $]\,0, 1\,[$. So, the same holds true for the phase error e_φ. As for the group error, we have

$$e_g = \frac{C_3\gamma}{\sin[2\arcsin(C_1\gamma)]} = \frac{C_3}{2C_1}\,\frac{1}{\sqrt{1 - C_1^2\gamma^2}}$$

with again $C_1\gamma \in]\,0, 1\,[$. The function $\dfrac{1}{\sqrt{1-x^2}}$ is also strictly increasing over $]\,0, 1\,[$, which completes the proof of this Proposition. $\triangleleft$

REMARK 3.1 In the polynomial expansion of the expressions previously given for e_φ and e_g, γ always appears as a square, which testifies to the second order accuracy in time. $\square$

The next Proposition is concerned with the extrema of e_φ with respect to the directions $\boldsymbol{\nu}$. A good knowledge of these extrema would allow us to eliminate the angle parameters in the accuracy criterion.

Proposition 4.2 *For $m \geq 1$ and γ, s, H fixed in such a way that $sH \leq 1/2$, the phase velocity relative error e_φ attains*

- *its minimum algebraic value for the coordinate directions*

$$\boldsymbol{\nu}^- \;=\; (\pm 1,\, 0,\, 0) \quad or \quad (0,\, \pm 1,\, 0) \quad or \quad (0,\, 0,\, \pm 1)$$

- *its maximum algebraic value for the principal diagonal directions*

$$\boldsymbol{\nu}^+ \;=\; (\pm 1/\sqrt{3},\, \pm 1/\sqrt{3},\, \pm 1/\sqrt{3})$$

- *its saddle-point value for the secondary diagonal directions*

$$\boldsymbol{\nu}^0 \;=\; (\pm 1/\sqrt{2},\, \pm 1/\sqrt{2},\, 0) \quad or \quad (\pm 1/\sqrt{2},\, 0,\, \pm 1/\sqrt{2}) \quad or \quad (0,\, \pm 1/\sqrt{2},\, \pm 1/\sqrt{2})$$

In addition, there is no other direction for which e_φ reaches a local extremum.

PROOF Let $\chi_\flat \;=\; \pi \nu_\flat s H$ for $\flat \in \{x, y, z\}$. Introduce, for $\chi \in [0,\, \pi/2]$, the function

$$\varphi_m(\chi) \;=\; \sum_{p=1}^{m} \beta_p \, \sin^{2p}(\chi) \tag{4.1}$$

so that the search for extremal values of e_φ is equivalent to that for extremal values of

$$\mathcal{F}_m(\chi_x,\, \chi_y,\, \chi_z) \;=\; \varphi_m(\chi_x) + \varphi_m(\chi_y) + \varphi_m(\chi_z)\,. \tag{4.2}$$

The main reason justifying this reduction is that $x \mapsto \arcsin x$ is an increasing function, so we only have look for the extremal values of its argument, or even the square of its argument, which coincides with $\mathcal{F}_m$ within a multiplicative factor γ^2.

Since the $\chi_\flat$'s are subject to the constraint

$$\chi_x^2 + \chi_y^2 + \chi_z^2 \;=\; \pi^2 s^2 H^2\,,$$

we will apply Lagrange's multiplier rule to solve this optimization problem. It follows that for $\boldsymbol{\chi} = (\chi_x,\, \chi_y,\, \chi_z)$ to be a direction for which $\mathcal{F}_m$ attains an extremum, it is necessary that

$$\exists\, \lambda \in \mathbf{R} \;\mid\; \forall \flat \in \{x, y, z\}\,, \quad \frac{\partial \mathcal{F}_m}{\partial \chi_\flat}(\boldsymbol{\chi}) \;=\; 2\lambda\, \chi_\flat$$

or, because of (4.2),

$$\exists\, \lambda \in \mathbf{R} \;\mid\; \forall \flat \in \{x, y, z\}\,, \quad \varphi_m'(\chi_\flat) \;=\; 2\lambda\, \chi_\flat\,.$$

It is clear from (4.1) that φ_m is an even function and therefore $\varphi'_m(0) = 0$. The optimality condition for $\boldsymbol{\chi}$ can then be expressed without the multiplier λ as

$$\frac{\varphi'_m(\chi_x)}{\chi_x} = \frac{\varphi'_m(\chi_y)}{\chi_y} = \frac{\varphi'_m(\chi_z)}{\chi_z} \qquad \text{and} \qquad \chi_x^2 + \chi_y^2 + \chi_z^2 = \pi^2 s^2 H^2 , \tag{4.3}$$

provided that for $\chi_\flat = 0$, the ratio $\dfrac{\varphi'_m(\chi_\flat)}{\chi_\flat}$ is allowed to take any real value. Now, a careful study of $\chi \mapsto \dfrac{\varphi'_m(\chi)}{\chi}$ reveals that this function is strictly decreasing over $\left]0, \dfrac{\pi}{2}\right[$. As a result, the situation described in (4.3) can occur under only three circumstances:

1. Two of the $\chi_\flat$'s vanish and the remaining one is equal to $\pm\pi sH$. This corresponds to the family of coordinate directions the first part of $\boldsymbol{\nu}^-$.
2. One of the $\chi_\flat$'s vanishes and the remaining two are both equal to $\pm\pi sH/\sqrt{2}$. This corresponds to the family of secondary diagonal directions $\boldsymbol{\nu}^0$.
3. None of the $\chi_\flat$'s is zero and the three of them are all equal to $\pm\pi sH/\sqrt{3}$. This corresponds to the family of principal diagonal directions $\boldsymbol{\nu}^+$.

This implies that there is no other direction for which $\mathcal{F}_m$ is extremal. Examining the quadratic form induced by the Hessian matrix (which is diagonal in the present context) over the tangent plane of the constraint sphere at these points, we can further classify them into maximum, minimum or saddle points as indicated in the Proposition. $\triangleleft$

This property enables us to get rid of the anisotropy of the errors by directly dealing with their maximal and minimal values. Define

$$\begin{aligned} e_\varphi^-(m,\gamma,H') &= \frac{1}{\pi\gamma H'} \arcsin\left[\gamma\left(\sum_{p=1}^{m} \beta_p \sin^{2p}(\pi H')\right)^{1/2}\right] - 1 \\ e_\varphi^+(m,\gamma,H') &= \frac{1}{\pi\gamma H'} \arcsin\left[\gamma\left(3\sum_{p=1}^{m} \beta_p \sin^{2p}\left(\frac{\pi H'}{\sqrt{3}}\right)\right)^{1/2}\right] - 1 \end{aligned} \tag{4.4}$$

as functions of $m \geq 1$, $\gamma \in \left]0, \gamma_m^{\max}\right[$ and $H' \in [0, 1/2]$. Physically speaking, e_φ^- is the phase velocity relative error in the *slowest* direction, while e_φ^+ is the phase velocity relative error in the *fastest* direction. It would be highly interesting to show similar results for the group velocity relative error e_g. Unfortunately, transposing the proof of Proposition 4.2 to e_g is far from being easy, especially because e_g does not have the special form (4.2). We must therefore content ourselves with extensive numerical computations, which nonetheless lead us to suggest that

Hypothesis 4.1 *The group velocity relative error e_g attains its extrema for the same directions in the space of wave vectors as does its phase counterpart e_φ. Furthermore, the minimum algebraic value is reached for $\boldsymbol{\nu}^-$, and the maximum algebraic value is reached for $\boldsymbol{\nu}^+$.*

We wish to insist on the fact that although no analytic proof has been found, Hypothesis 4.1 was experimentally confirmed for a wide range of m, γ and sH. Once this has been taken for granted, it becomes natural to introduce the extremal values

$$
\begin{aligned}
e_g^-\,(m,\gamma,H') &= \frac{\gamma \displaystyle\sum_{p=1}^{m} \beta_p \, \sin^{2p-1}(\pi\nu_\flat H')\cos(\pi H')}{\sin\left\{ 2\arcsin\left[\gamma\left(\displaystyle\sum_{p=1}^{m} \beta_p \, \sin^{2p}\,(\pi H')\right)^{1/2}\right]\right\}} \quad - \quad 1 \\
e_g^+\,(m,\gamma,H') &= \frac{\gamma\sqrt{3} \displaystyle\sum_{p=1}^{m} 2p\beta_p \, \sin^{2p-1}\left(\frac{\pi H'}{\sqrt{3}}\right)\cos\left(\frac{\pi H'}{\sqrt{3}}\right)}{\sin\left\{ 2\arcsin\left[\gamma\sqrt{3}\left(\displaystyle\sum_{p=1}^{m} \beta_p \, \sin^{2p}\left(\frac{\pi H'}{\sqrt{3}}\right)\right)^{1/2}\right]\right\}} \quad - \quad 1
\end{aligned}
\tag{4.5}
$$

for $m \geq 1$, $\gamma \in \,]\,0,\, \gamma_m^{\max}\,[$ and $H' \in [0,\, 1/2]$.

4.2 Design of an algorithm

Taking advantage of the properties enumerated above, we can reformulate problem (**F1**) as

(**F2**) GIVEN

c	velocity of the homogeneous medium
m	half-order in space of the scheme 2-2m
ϵ	accuracy threshold $(1/n_\lambda n_T)$

FIND

$$\gamma \text{ and } H \text{ so as to minimize } \bar{J}_m = \frac{N_m}{\gamma H^4} \text{ under the following constraints}$$

- Nyquist's frequency

$$0 < H \leq \frac{1}{2}$$

- stability condition

$$0 < \gamma < \gamma_m^{\max}$$

- accuracy criterion

$$\forall\, H' \in [0,\, H], \quad e_g^+(m,\gamma,H') \leq \epsilon \quad \text{and} \quad e_g^-(m,\gamma,H') \geq -\epsilon \tag{4.6}$$

REMARK 3.2 Once γ and H have been determined, the parameters h and Δt are deduced via c and $f_{\max}$. The latter is not explicitly involved in the statement of (**F2**). □

In the denominator of $\bar{J}_m$, the objective function, H is raised at the fourth power while Δt stays as it is. Since H seems to exert a much stronger influence on $\bar{J}_m$ than Δt does, we are tempted to make a bold step by replacing $\bar{J}_m$ by $\tilde{J} = \dfrac{1}{H}$. In other words, we will try to maximize H under the prescribed constraints. A rigorous justification of this step would require a sensitivity study of e_g^+ and e_g^- with regard to γ and H, which we have not taken up. However, this simplification can be intuitively understood as follows. The schemes considered are second order in time and $2m$-th order in space. Therefore, Δt must be very small so as to produce an error Δt^2 of the same order of magnitude as h^{2m}, the error produced by h. As a result, it is a pointless effort to take Δt into account in the objective function.

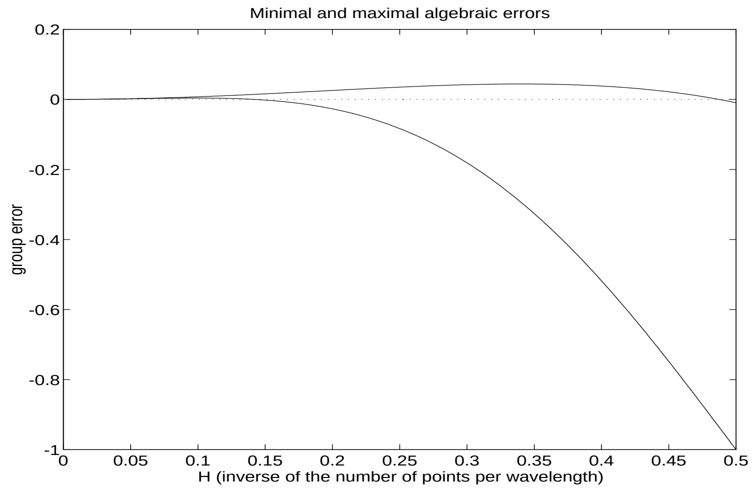

Figure 4.1: Behavior of the extremal errors $e_g^+(m,\gamma,H)$ and $e_g^-(m,\gamma,H)$ with respect to H for the 2-4 scheme ($m=2$) and a stability ratio $\gamma=0.4$.

Prior to maximizing H, it is helpful to have an idea about the behavior of $e_g^+(m,\gamma,H)$ and $e_g^-(m,\gamma,H)$ with respect to H for fixed m and γ. Figure 1 illustrates the typical shapes of e_g^+ and e_g^- as functions of H for given m and γ. Starting from positive values for H close to 0, both e_g^+ and e_g^- climb to their maximal values at some point before decreasing steadily and perhaps taking negative values. Given m and γ, for $\epsilon>0$ small enough, it is possible to consider

$$\begin{aligned} H_\epsilon^+(m,\gamma) &\equiv \text{ smallest value of } H \text{ for which } e_g^+(m,\gamma,H)=\epsilon \\ H_\epsilon^-(m,\gamma) &\equiv \text{ smallest value of } H \text{ for which } e_g^-(m,\gamma,H)=-\epsilon \end{aligned} \tag{4.7}$$

Lemma 4.2 *For fixed $m\geq 1$ and $\epsilon>0$ small enough for H_ϵ^+ and H_ϵ^- to be well-defined. These are continuous functions of γ. Furthermore, $H_\epsilon^+(m,\gamma)$ is a decreasing function of γ, whereas $H_\epsilon^-(m,\gamma)$ is an increasing function of γ.*

PROOF The existence of $H_\epsilon^\pm(m,\gamma)$ for ϵ small enough stems from the behavior of e_g^+ and e_g^- with respect to H. Their continuity is a consequence of the implicit functions theorem. Their mononoty with respect to γ comes from Proposition 4.1. $\triangleleft$

We are at last ready for the optimal strategy. On the grounds of the above properties, we intuitively feel that γ and H must be determined in some simultaneous way so as to maximize H under the accuracy criterion, which appears to be the most important constraint.

Theorem 4.1 *Consider problem* (**F2**) *in which the objective function $\bar{J}_m$ has been replaced by $1/H$. Assume that the threshold ϵ is small enough. Then, the couple (γ, H) is solution to* (**F2**) *if and only if $H_\epsilon^+(m,\gamma) = H_\epsilon^-(m,\gamma) = H$.*

PROOF Let (γ, H) be the optimal parameters. If $H_\epsilon^+(m,\gamma) < H_\epsilon^-(m,\gamma)$, then for the accuracy criterion to be met, we must have $H = H_\epsilon^+(m,\gamma)$. However, by virtue of the properties known for e_g^+ and e_g^-, it is possible to change γ to $\gamma' > \gamma$, so that

$$H_\epsilon^+(m,\gamma) < H_\epsilon^+(m,\gamma') < H_\epsilon^-(m,\gamma') < H_\epsilon^-(m,\gamma')$$

and thus (γ', H') with $H' = H_\epsilon^+(m,\gamma')$ would be a better candidate. Likewise, assuming $H_\epsilon^+(m,\gamma) > H_\epsilon^-(m,\gamma)$ also leads to a contradiction. This implies the equality $H_\epsilon^+(m,\gamma) = H_\epsilon^-(m,\gamma)$. There is no difficulty in showing the converse. ◁

Let us give some insights into the practical procedure of finding γ and H. Given $m \geq 1$ and $\epsilon > 0$, we have to grope a little while before getting the correct value for γ. Figure 2*a* represents what happens when γ is smaller than the optimal value. In such an instance, $H_\epsilon^+(m,\gamma) > H_\epsilon^-(m,\gamma)$ and so γ must be bigger. Figure 2*b* depicts the situation when γ is larger than the optimal value. In this case, $H_\epsilon^+(m,\gamma) < H_\epsilon^-(m,\gamma)$ and so γ must be smaller. We keep on adjusting γ until $H_\epsilon^+(m,\gamma) = H_\epsilon^-(m,\gamma)$ is satisfied, as in panel *c*.

4.3 Results and comments

One of the preliminary questions of the Marmousi 3-D campaign is to know whether or not the 2-8 scheme is more economical than the 2-4 one. Put another way, our task is to compare $m = 4$ with $m = 2$. Even in homogeneous media, the answer unavoidably depends on the accuracy threshold ϵ we want to impose to the group velocity relative error. Typically, it is useful to consider ϵ ranging from 0.001 to 0.01. The rationale of such a range is that for a simulation to be useful, in general $n_T \geq 100$ and $n_\lambda \geq 4$. A precision $\epsilon = 0.001$ means, for instance, that after the wave front has propagated over $n_T = 200$ shortest wavelengths, the relative discrepancy between the numerical solution and the exact solution is less than $1/n_\lambda = 1/5$.

Given m and ϵ, the optimal strategy sketched out in Theorem 4.1 is applied to determine the optimal paramters γ and H. Then, the minimal cost $J_m^\star = \dfrac{N_m}{\gamma H^4}$ is computed. Figure 3 plots $J_m^\star$ versus ϵ for $m = 2$ and From the standpoint of elementary floating point operations, $N_4 = 32$ is much larger than $N_2 = 18$. However, this is sufficiently compensated for by the gain in H obtained via the 2-8 scheme so that the latter turns out to be mostly better than the 2-4 one. Table 1 supplies us with more details about this comparison. Note that $1/H$ is exactly the number of points per shortest wavelength. We wish to emphasize that the stability ratio γ is associated with the sampling ratio $1/H$ in such a way that the couple (γ, H) is optimal. In other words, any change in γ at fixed H will result in a violation of the accuracy criterion (3.8).

As is clearly seen from Figure 3 and Table 1, the more demanding we are on the accuracy, the more expensive the simulations, and the more the 2-8 scheme ascends over the 2-4 one. Up to $\epsilon = 0.01$, which is not such a tremendous precision for a large scale model, it takes the 2-4 scheme twice as many operations than the 2-8 one to run simulations. Note that this conclusion is exactly opposite to that drawn by Sei [14] for schemes of the form $A^t A$. The key feature which accounts

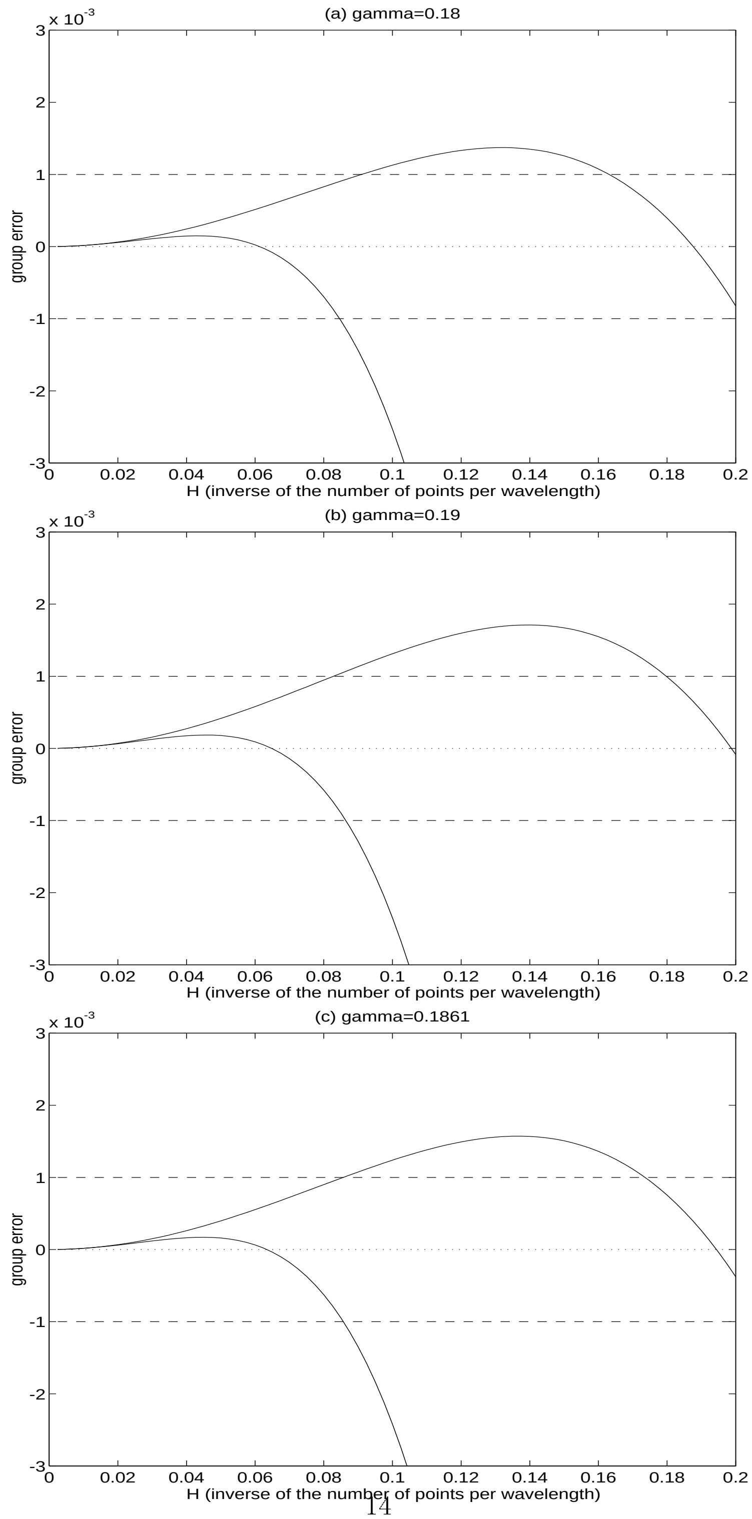

Figure 4.2: Determination of γ and H for $m = 2$ and $\epsilon = 10^{-3}$. In (a), γ is a little smaller than the optimal value. In (b), it is a little larger. In (c), it is equal to the optimal value.

Precision	Stability ratio		Sampling ratio		Cost ratio
ϵ	γ_2	γ_4	$1/H_2$	$1/H_4$	$J_2^\star/J_4^\star$
0.001	0.186	0.083	11.7	5.7	4.2
0.002	0.221	0.107	9.8	5.4	3.3
0.003	0.244	0.124	8.8	5.1	2.9
0.004	0.261	0.138	8.2	4.8	2.6
0.005	0.275	0.149	7.7	4.6	2.4
0.008	0.308	0.176	6.8	4.3	2.0
0.010	0.324	0.190	6.5	4.2	1.9

Table 4.1: Detailed comparison of the 2-4 scheme ($m = 2$) and the 2-8 sheme ($m = 4$) in homogeneous media.

for this situation lies in the number of elementary operations of the schemes, which implies that the greatest attention should be paid to the family of schemes under study.

Although this comparison speaks up strongly in favor of the 2-8 scheme, we should be aware of the fact that real-life media are not homogeneous. This is the reason why we are now going to propose a generalization of the optimal strategy to the heterogeneous case.

5 Generalization to the heterogeneous case

5.1 New formulation

Let $c_{\min}$—$c_{\max}$ be the velocity range of the heteregeneous medium at hand. The basic idea consists in assimilating this medium as merely a series of several homogeneous media c between $c_{\min}$ and $c_{\max}$. The discretization parameters h and Δt are required to satisfy the constraints of problem (**F1**) for all $c \in [c_{\min}, c_{\max}]$. Concretely speaking, we have to solve

(**G1**) GIVEN

$c_{\min}$—$c_{\max}$	velocity range of the heterogeneous medium
$f_{\max}$	cutoff frequency of the source wavelet
m	half-order in space of the scheme 2-2m
ϵ	accuracy threshold ($= 1/n_\lambda n_T$)

FIND

h and Δt so as to <u>minimize</u> $J_m(h, \Delta t)$ under the following constraints

- Nyquist's frequency

$$h \le \frac{1}{2}\lambda_{\min} = \frac{1}{2}\frac{c_{\min}}{f_{\max}} \tag{5.1}$$

- stability condition

$$\frac{c_{\max}\Delta t}{h} < \gamma_m^{\max} \tag{5.2}$$

- accuracy criterion

$$\forall \boldsymbol{\xi}\,,\ \ \|\boldsymbol{\xi}\| \le \xi_{\max} \implies \forall\, c \in [c_{\min}, c_{\max}]\,,\ \ |\,e_g(m, c, h, \Delta t, \boldsymbol{\xi})\,| \le \epsilon \tag{5.3}$$

In order to somehow make use of the optimal strategy for the homogeneous case, we first have to reformulate (**G1**) in terms of non-dimensional variables γ and H. In this process, a slight difficulty may appear. Since we are faced with a whole range of velocity instead of one single velocity, we need to know the velocity value with respect to which γ and H are to be defined. For conveniency purposes, we can rule out that this is always the smallest velocity $c_{\min}$. So, the discretization parameters and the non-dimensional variables are connected by

$$\gamma = \frac{c_{\min}\Delta t}{h} \qquad \text{and} \qquad H = \frac{h}{\lambda_{\min}} = \frac{h f_{\max}}{c_{\min}}\,. \tag{5.4}$$

We will also need to introduce the velocity contrast

$$\sigma = \frac{c_{\max}}{c_{\min}}\,. \tag{5.5}$$

Now, consider

(G2) GIVEN

$c_{\min}$	minimum velocity of the medium
σ	velocity contrast
m	half-order in space of the scheme 2-$2m$
ϵ	accuracy threshold $(1/n_\lambda n_T)$

FIND

$$\gamma \text{ and } H \text{ so as to minimize } \bar{J}_m = \frac{N_m}{\gamma H^4} \text{ under the following constraints}$$

- Nyquist's frequency

$$0 < H \leq \frac{1}{2}$$

- stability condition

$$0 < \gamma < \frac{1}{\sigma}\gamma_m^{\max} \tag{5.6}$$

- accuracy criterion

$$\begin{array}{lll} \displaystyle\sup_{\sigma' \in [1,\sigma]} \sup_{H' \in [0,H]} e_g^+\left(m,\, \sigma'\gamma,\, \frac{H'}{\sigma'}\right) & \leq & \epsilon \\ \displaystyle\inf_{\sigma' \in [1,\sigma]} \inf_{H' \in [0,H]} e_g^-\left(m,\, \sigma'\gamma,\, \frac{H'}{\sigma'}\right) & \geq & -\epsilon \end{array} \tag{5.7}$$

REMARK 4.1 In (5.4), the stability ratio γ is associated with $c_{\min}$ and not $c_{\max}$ as is traditionally the case. This accounts for the division by σ in the right-hand side of (5.6). □

REMARK 4.2 It is easily seen that condition (5.7) is equivalent to criterion (5.3). Indeed, for any $c' = \sigma' c_{\min}$ belonging to the velocity range, $\sigma'\gamma$ and $\dfrac{H}{\sigma'}$ are the non-dimensional variables corresponding to h and Δt via c' and $f_{\max}$. □

5.2 Search procedure

Analogously to the homogeneous case, the idea is first and foremost to replace the objective function $\bar{J}_m(\gamma, H)$ by $1/H$, which is tantamount to maximizing H under the three constraints. Secondly, a method for determining optimal γ and H will be proposed, based on the following property.

Lemma 5.1 *For fixed $m \geq 1$, $\sigma \geq 1$, $\gamma \in \left]0, \frac{1}{\sigma}\gamma_m^{\max}\right[$ and $H \in \left]0, \frac{1}{2}\right[$, the mappings*

$$\sigma' \mapsto e_g^+\left(m,\, \sigma'\gamma,\, \frac{H}{\sigma'}\right) \qquad \text{and} \qquad \sigma' \mapsto e_g^-\left(m,\, \sigma'\gamma,\, \frac{H}{\sigma'}\right)$$

are increasing functions of the variable $\sigma' \in [1, \sigma]$.

PROOF By the chain rule, we have

$$\frac{de_g^{\pm}}{d\sigma'} = \gamma \, \frac{\partial e_g^{\pm}}{\partial \gamma} - \frac{H}{\sigma'^2} \, \frac{\partial e_g^{\pm}}{\partial H'} \tag{5.8}$$

where the derivatives are taken at point $(m,\, \sigma'\gamma,\, H/\sigma')$. According to Proposition 4.1, $e_g^{\pm}$ is an increasing function of γ. Hence, $\dfrac{\partial e_g^{\pm}}{\partial \gamma} \geq 0$. If the other derivative $\dfrac{\partial e_g^{\pm}}{\partial H'}$ is also negative, evidently $\dfrac{de_g^{\pm}}{d\sigma'} \geq 0$ and so $e_g^{\pm}$ is an increasing function of σ'.

In the case $\dfrac{\partial e_g^{\pm}}{\partial H'} > 0$, we still want to prove that $\dfrac{de_g^{\pm}}{d\sigma'} \geq 0$. Thanks to (5.8), the desired inequality is equivalent to

$$\frac{\partial e_g^{\pm}}{\partial \gamma}(m, \sigma'\gamma, H/\sigma') \left[\frac{\partial e_g^{\pm}}{\partial H'}\right]^{-1}(m, \sigma'\gamma, H/\sigma') \;\geq\; \frac{H}{\sigma'}\,(\sigma'\gamma)^{-1} \,.$$

By the change of variable $\gamma' = \sigma'\gamma$ and $H' = H/\sigma'$, what we have to show is

$$\frac{\partial e_g^{\pm}}{\partial \gamma}(m, \gamma', H') \left[\frac{\partial e_g^{\pm}}{\partial H'}\right]^{-1}(m, \gamma', H') \;\geq\; \frac{H'}{\gamma'} \tag{5.9}$$

for all $\gamma' \in \,]\,0, \gamma_m^{\max}\,[$ and $H' \in \,]\,0, 1/2\sigma\,[$ such that $\dfrac{\partial e_g^{\pm}}{\partial H'}(m, \gamma', H') > 0$. But, by the implicit functions theorem, the left-hand side of (5.9) represents the opposite of the derivative with respect to γ' of the function $\gamma' \mapsto H_m^{\star}(\gamma')$ defined implicitly by $e_g^{\pm}(m, \gamma', H_m^{\star}(\gamma')) = \eta$ where η is some positive number, small enough such that $H' = H_m^{\star}(\gamma')$. Inequality (5.9) can now be rewritten as

$$-\frac{dH_m^{\star}}{d\gamma'}(\gamma') \;\geq\; \frac{H_m^{\star}(\gamma')}{\gamma'}$$

which, after some algebra, amounts to saying that the mapping $\gamma' \mapsto \gamma' H_m^{\star}(\gamma')$ has to be a decreasing function of $\gamma' \in \,]\,0, \gamma_m^{\max}\,[$ for $\eta > 0$ sufficiently small.

On one hand, for the same reasons as those invoked for $H_{\epsilon}^{\pm}(m,\gamma)$, which were defined earlier by (4.7), $H_m^{\star}(\gamma')$ is a decreasing function of γ'. On the other hand, it follows from the asymptotic expansions of the formulae (4.5) that

$$e_g^{\pm}(m, \gamma', H') = A^{\pm}\,(\gamma' H')^2 + B_m^{\pm}\,(H')^{2m} + \text{higher powers of } H'$$

where $A^{\pm} > 0$ and $B_m^{\pm} < 0$ are various constants. To get an idea about different orders of magnitude, replace the primary definition of $H_m^{\star}$ by the approximate characterization

$$A^{\pm}\,[\gamma' H_m^{\star}(\gamma')]^2 + B_m^{\pm}\,[H_m^{\star}(\gamma')]^{2m} \approx \eta \tag{5.10}$$

Arguing that $A^{\pm} > 0$, $B_m^{\pm} < 0$ and that $H_m^{\star}(\gamma')$ decreases with γ', we can infer from (5.10) that $\gamma' H_m^{\star}(\gamma')$ is also a decreasing function of γ', which finally completes the proof. $\triangleleft$

The geometrical interpretation of Lemma 5.1 is the following. For fixed $\gamma \in \,]\,0,\, \gamma_m^{\max}/\sigma\,[$ and any $\sigma' \in [1, \sigma]$, let $\mathcal{E}^{\pm}(\sigma')$ be the dispersion curves representing $e_g^{\pm}(m, \sigma'\gamma, H/\sigma')$ as functions of $H \in [\,0, 1/2]$. The result of Lemma 5.1 ensures that if $\sigma'' > \sigma'$, then $\mathcal{E}^{+}(\sigma'')$ always lies above $\mathcal{E}^{+}(\sigma')$, and $\mathcal{E}^{-}(\sigma'')$ always lies above $\mathcal{E}^{-}(\sigma')$. This geometrical property is illustrated in Figure 4.

Lemma 5.1 enables us to simplify the heterogeneous accuracy criterion (5.7) as

$$\forall\, H' \in [0, H]\,, \quad e_g^{+}(m,\, \sigma\gamma,\, H'/\sigma) \leq \epsilon \quad \text{and} \quad e_g^{-}(m,\, \gamma,\, H') \geq -\epsilon \tag{5.11}$$

The search procedure is readily inspired from the homogeneous case by defining

$$\begin{array}{rcl} H_\epsilon^{+}(m, \sigma', \gamma) & \equiv & \text{smallest value of } H \text{ for which } \; e_g^{+}(m, \sigma'\gamma, H/\sigma') = \epsilon \\ H_\epsilon^{-}(m, \sigma', \gamma) & \equiv & \text{smallest value of } H \text{ for which } \; e_g^{-}(m, \sigma'\gamma, H/\sigma') = -\epsilon \end{array}$$

As before, it can be shown that for fixed $m \geq 1$, $\epsilon > 0$ and $\sigma' \in [\,1, \sigma]$, the mapping $\gamma \mapsto H_\epsilon^{+}(m, \sigma', \gamma)$ is a decreasing function of $\gamma \in \,]\,0,\, \sigma_m^{\max}/\sigma\,[$, while $\gamma \mapsto H_\epsilon^{-}(m, \sigma', \gamma)$ is an increasing function. The optimal parameters are given by

Theorem 5.1 *Consider problem* (**G2**) *in which the objective function $\bar{J}_m$ has been replaced by $1/H$. Assume that the threshold ϵ is small enough. Then, the couple (γ, H) is solution to* (**G2**) *if and only if (i) $H_\epsilon^{+}(m, \sigma, \gamma) = H_\epsilon^{-}(m, 1, \gamma) = H$.*

PROOF Similar to the proof of Theorem 4.1 $\triangleleft$

5.3 Results and comments

We have applied the above procedure to heterogeneous media with $\sigma = 2$. The experimental approach is much the same as in the homogeneous case: for ϵ ranging from 0.001 to 0.01, we compare the optimal costs corresponding to the 2-4 and 2-8 schemes.

Figure 5 summarizes the results obtained, while Table 2 gives more details on the comparison. It is pleasant to notice that the optimal discretization parameters are just a trifle smaller than those of the homogeneous case. This can be physically understood as follows. If the medium were homogeneous with $c = c_{\min}$, then γ and H would be given by Table 1, which yields some values for h and Δt. When another homogeneous medium $c' > c_{\min}$ is superimposed to the initial medium, the stability ratio, computed as $c'\Delta t/h$ would be worsened. Meanwhile, the sampling ratio $c'/f_{\max}h$ would be improved. Thus, there is no need to select a really smaller space-step h for numerical simulations in this medium. The only issue that matters then is to lower the time-step Δt so as to respect the stability condition.

The method proposed in Theorem 5.1 suffers from a serious drawback for large ϵ, for which the stability threshold is saturated before one could figure out an appropriate γ for the equality $H_\epsilon^{+}(m, \sigma, \gamma) = H_\epsilon^{-}(m, 1, \gamma)$ to be satisfied. In such cases, the exact solution to (**G2**) is $\gamma = \gamma_m^{\max}$ and $H = H_\epsilon^{-}(m, 1, \gamma_m^{\max})$. Table 2 shows that this situation occurs to the 2-4 scheme for $\epsilon \geq \epsilon_\sigma = 0.005$. Anyhow, the cost ratio still remains definitely in favor of the 2-8 scheme. The stronger the contrast σ, the smaller the level ϵ_σ and the sooner the stability condition is saturated.

Precision	Stability ratio		Sampling ratio		Cost ratio
ϵ	γ_2	γ_4	$1/H_2$	$1/H_4$	$J_2^\star/J_4^\star$
0.001	0.172	0.082	12.0	5.8	5.0
0.002	0.204	0.106	10.1	5.5	3.9
0.003	0.225	0.122	9.1	5.2	3.4
0.004	0.241	0.136	8.4	4.9	3.1
0.005	$0.250^\dagger$	0.147	8.0	4.6	2.9
0.008	$0.250^\dagger$	0.173	7.3	4.3	3.2
0.010	$0.250^\dagger$	0.187	7.0	4.2	3.3

Table 5.1: Detailed comparison of the 2-4 scheme ($m = 2$) and the 2-8 sheme ($m = 4$) in heterogeneous media with $c_{\max} = 2c_{\min}$. The symbol $\dagger$ indicates that the stability ratio has been saturated.

6 Conclusion

Throughout this paper, we have set up what we believe is a good framework for comparing the cost of numerical schemes of the type 2-$2m$ under the same accuracy constraint. In the process of presenting the optimal strategy, we have even derived some useful theoretical results, namely the behavior of the group or phase velocity relative error with respect to various parameters, such as the stability ratio or the angle in the wave vector space. As is a widely common practice in geophysics, the accuracy criterion has been based on the group velocity error, since this represents the velocity at which energy propagates.

The numerical results, obtained for the 2-4 and 2-8 schemes in both homogeneous and heterogeneous media, lead us to recommend the 2-8 scheme as the more economical method for the purposes of the 3-D modelling campaign envisaged. In addition, the optimal discretization parameters are explicitly given as functions of the accuracy level desired. The existence of an optimal strategy to determine the discretization parameters in heterogeneous media is of particular interest.

However, it must be reminded that this is a comparison merely between numerical methods, and not between actual CPU times. Implementing the 2-8 scheme turns out to be much a harder task than coding the 2-4 one, insofar as the boundaries of the computational domain require more special treatments. Notwithstanding, this study provides an excellent overview on how the simulation cost is about to change with the accuracy level.

Acknowledgments

The authors are grateful to Alain Sei, who initiated this kind of comparison study in his thesis, for many fruitful discussions. We also wish to thank all of those who patiently reviewed the preliminary versions of this manuscript.

References

[1] R. M. ALFORD, K. R. KELLY and D. M. BOORE, Accuracy of Finite-Difference Modeling of the Acousic Wave Equation, *Geophysics* **39**, 834-842 (1974).

[2] A. BAMBERGER, G. CHAVENT and P. LAILLY, *Étude de Schémas Numériques pour les Équations de l'Élastodynamique Linéaire*, Rapport de Recherche **41**, INRIA, Rocquencourt, 1980.

[3] L. BRILLOUIN, *Wave Propagation and Group Velocity*, Academic Press, New-York, 1960.

[4] M. A. DABLAIN, The Application of High-Order Differencing to the Scalar Wave Equation, *Geophysics* **51**, 54-66 (1986).

[5] O. HOLBERG, Computational Aspects of the Choice of Operator and Sampling Interval for Numerical Differentiation in Large-Scale Simulation of Wave Phenomena, *Geophys. Prosp.* **35**, 629-655 (1987).

[6] K. R. KELLY, R. W. WARD, S. TREITEL and R. M. ALFORD, Synthetic Seismograms: a Finite-Difference Approach, *Geophysics* **41**, 2-27 (1976).

[7] M. KERN and W. W. SYMES, Loop Level Parallelisation of a Seismic Inversion Code, *Proceedings of the 63rd SEG Annual Meeting*, Washington D. C., 1993.

[8] D. KOSLOFF and E. BAYSAL, Forward Modeling by a Fourier Method, *Geophysics* **47**, 1402-1412 (1982).

[9] I. MUFTI, Large-Scale Three-Dimensional Seismic Models and Their Interpretative Significance, *Geophysics* **55**, 1166-1182 (1990).

[10] E. PIAULT, P. DUCLOS and A. SEI, Gigaflops Performance for 2-D/3-D Heterogeneous Acoustic Modelling on a Massively Parallel Computer, *Proceedings of the 1st International Conference on Mathematical and Numerical Aspects of Wave Propagation Phenomena*, SIAM, 1991.

[11] M. RESHEF, D. KOSLOFF, M. EDWARDS and C. HSIUNG, Three-Dimensional Acoustic Modeling by Fourier Method, *Geophysics* **53**, 1175-1183 (1988).

[12] R. RICHTMYER and K. MORTON, *Difference Methods for Initial-Value Problems*, Wiley-Interscience, New-York, 1967.

[13] A. SEI, *Étude de Schémas Numériques pour les Modèles de Propagation d'Ondes en Milieux Hétérogènes*, Thèse de Doctorat, Université de Paris IX-Dauphine, 1991.

[14] A. SEI, Computational Cost of Finite Difference Elastic Waves Modelling, *Proceedings of the 63rd SEG Annual Meeting*, Washington D. C., 1993.

[15] L. TREFETHEN, Group Velocity in Finite Difference Schemes, *SIAM Review* **24**, 113-136 (1982).

[16] R. VICHNEVETSKY and J. B. BOWLES, *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*, SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1982.

[17] G. B. WHITHAM, *Linear and Nonlinear Waves*, John Wiley & Sons, New-York, 1974.

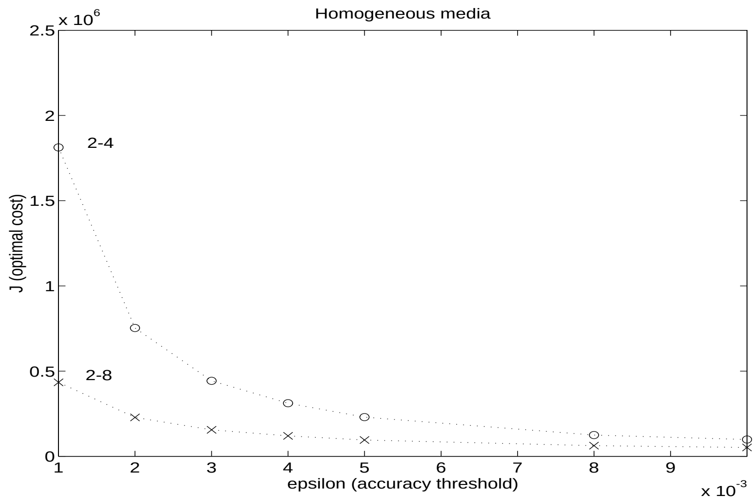

Figure 4.3: Optimal cost $J_m^\star$ in normalized unit versus accuracy threshold ϵ for (o) the 2-4 scheme and (x) the 2-8 scheme in homogeneous media.

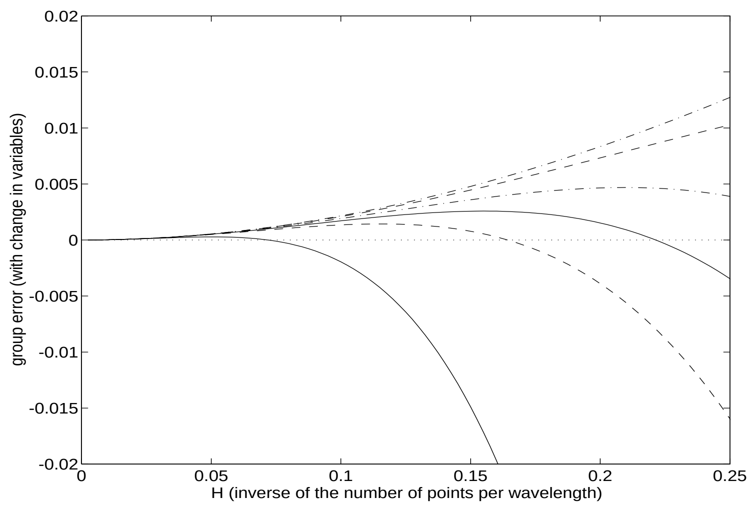

Figure 5.1: Illustration of Lemma 4.1. Group errors $e_g^{\pm}(m, \sigma'\gamma, H/\sigma')$ as functions of H for $m = 2$, $\gamma = 0.1$ and $\sigma' = 1$ (solid line), $\sigma' = 1.5$ (dash line) and $\sigma' = 2$ (dash-dotted line).

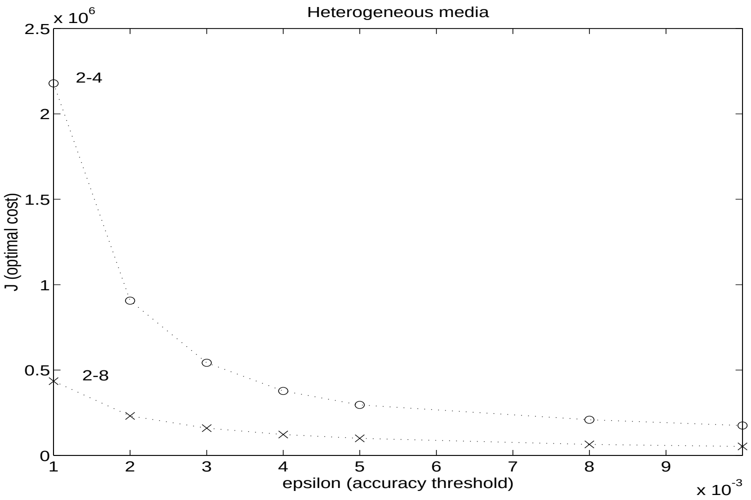

Figure 5.2: Optimal cost $J_m^\star$ in normalized unit versus accuracy threshold ϵ for (o) the 2-4 scheme and (x) the 2-8 scheme in heterogeneous media with $\sigma = 2$.