Dispersion Analysis of Numerical Wave Propagation and its Computational Consequences

Alain Sei and William Symes*

June 9, 1994

Abstract

We present in this paper a comparison of the dispersion properties for several finite-difference approximations of the acoustic wave equation. We investigate the compact and staggered schemes of fourth order accuracy in space and of second order or fourth order accuracy in time. We derive the computational cost of the simulation implied by a precision criterion on the numerical simulation (maximum allowed error in phase or group velocity). We conclude that for moderate accuracy the staggered scheme of second order in time is more efficient, whereas for very precise simulation the compact scheme of fourth order in time is a better choice. The comparison increasingly favors the lower order staggered scheme as the dimension increases. In three dimensional simulation, the cost of extremely precise simulation with any of the schemes is very large, whereas for simulation of moderate precision the staggered scheme is the least expensive.

1 Introduction

Asymptotic expansions and formulas like geometric optics or Kirchhoff integral do not contain all reflected, transmitted and refracted waves. The acoustic wave equation on the contrary takes all these phenomena into account. So the numerical methods used to integrate this equation are of great interest for understanding the complex interactions of the wave phenomena. To be of any use such methods need to be accurate, and in particular must overcome or at least control very well the numerical artifacts they introduce.

A well known and major artifact is the numerical dispersion. This phenomenon makes the waves'velocity frequency dependent. After a certain time it is therefore impossible to recognize the shape of a traveling pulse, and to interpret on a seismogram for instance, the presence of this "damaged" signal.

To overcome this essential drawback, the current practice consists of plotting the dispersion curves of the given numerical method and deriving from those curves a "rule of thumb" for the number

*Department of Computational & Applied Mathematics, Rice University, Houston, Texas, 77251

of points per wavelength and the number of points per period one needs to take to have a small error (cf [1]).
A natural idea is to use very accurate methods so that dispersion errors will intrinsicly be very small. For instance if one uses discrete Fourier transforms then there is no numerical dispersion (cf [2], [3], [4], [5]). but spectral methods suffer from other drawbacks which we will not consider here.
The finite difference method is the most widely used numerical method for wave propagation problems in seismology. The higher the order for the finite difference schemes, the less dispersion is experienced by the wave but the more floating point operations are required.
So the question we want to address in this paper is the following: "is higher order better ?". Our measure of quality is the computational cost of the numerical scheme, for a prescribed dispersion level.
Our approach is similar to the one developed in [6], [7] and [8] for staggered schemes, introduced for elastic waves in [10].
A good way to control the numerical dispersion is to use the phase and group velocities to measure the precision of the simulation. Given an *a priori* precision on the phase or group velocity, we show how to choose the number of points per wavelength and the number of points per period, to minimize the number of floating point operations required for that *a priori* precision.
The two most widely used "families" of finite difference schemes are the compact schemes (cf [11] [12] [13] and [14]), and the staggered schemes (cf [6] and [8]). So we compare for the same precision, the cost of the different schemes. We consider among these two families of schemes, those which have either second or fourth order accuracy in time and fourth order accuracy in space. This limitation on the order of the spatial derivatives is a consequence of our previous work (cf [8]). We showed that for staggered schemes, higher order were *not* better.

2 Precision Criteria and Computational Cost

To compare the different numerical schemes, we demand that they all fulfill the same precision requirements. We consider phase and group velocity errors. To simplify the computations of the different precision criteria, we assume that the domain where the wave propagates is a square ($A \times A$) or a cube ($A \times A \times A$), and that the time of propagation T_{max} is equal to the time of a round trip to the bottom of the model. So we have :

$$T_{max} = \frac{2.A}{c} \tag{2.1}$$

where c is the velocity of the wave in this homogeneous medium.
An non-dimensional way to measure the time of propagation is to use the shortest period as unit. We use a band limited wavelet of maximal frequency F_{max}; so the shortest period is $1/F_{max}$. For instance for a Ricker wavelet of central frequency F_0, $F_{max} \simeq 3.F_0$. So let J be the number of shortest period and λ_{min} the shortest wavelength in the medium, we can write :

$$T_{max} = J.\frac{\lambda_{min}}{c} \tag{2.2}$$

Therefore we have $A = J/2.\lambda_{min}$. So J appears as a measure of the size of the modeling problem we are interested in. It measures the number of shortest wavelength propagated in the medium.

2.1 Phase Velocity Criterion

We demand that the numerical solution be such that at the end of the simulation, the phase shift between the exact wave (with phase velocity c) and the numerical wave (with phase velocity $c_\varphi(k)$) be less than a fraction of the shortest wavelength. Put in equation we impose that :

$$|c_\varphi(k) - c|.T_{max} \leq \frac{\lambda_{min}}{n} \tag{2.3}$$

where n is an integer. Using (2.2) we can write (2.3) as follows :

$$\begin{aligned} & |c_\varphi(k) - c|.\frac{J.\lambda_{min}}{c} \leq \frac{\lambda_{min}}{n} \\ \Longleftrightarrow \quad & |E_\varphi| = \left|\frac{c_\varphi(k) - c}{c}\right| \leq \frac{1}{n.J} \end{aligned}$$

where E_φ is the normalised phase error.
We will use three instances for the precision on the phase velocity. The quarter of a wavelength is the threshold at which geophysicists are able to separate two arrivals on a seismogramm. We consider a first criterion corresponding to $n = 2$, that is a shift of a half wavelength. Then we consider the threshold in question corresponding to $n = 4$, that is a shift of a quarter of wavelength. And finaly we consider a shift of one sixth of a wavelength for $n = 6$. To relate those criteria to a travel time arrival on a seismogramm let us take the example of a Ricker wavelet with central frequency 30 Hz. Then the shortest period is 11.1 ms. So for $n = 2$ the shift in time between the two arrivals is 5.55 ms, for $n = 4$ it is 2.77 ms and for $n = 6$, 1.85 ms.
We will also use a model of one hundred wavelength ($J = 100$), which is a typical size in geophysical applications.

2.2 Group Velocity Criterion

It is well known that the group velocity governs the propagation of wave packets, that is the propagation of "energy" (cf [15] and [16]). It is therefore important to control the numerical effects on the group velocity.
To do so we impose, like for the phase velocity, that the group shift between the exact and numerical wave be less than a portion of the shortest wavelength. So we impose

$$\begin{aligned} & |c_{gr}(k) - c|.T_{max} \leq \frac{\lambda_{min}}{n} \\ \Longleftrightarrow \quad & |c_{gr}(k) - c|.\frac{J.\lambda_{min}}{c} \leq \frac{\lambda_{min}}{n} \\ \Longleftrightarrow \quad & |E_{gr}| = \left|\frac{c_{gr}(k) - c}{c}\right| \leq \frac{1}{n.J} \end{aligned}$$

where E_{gr} is the normalised group error. Again in this case we will set $J = 100$ and $n = 2, 4, 6$.

2.3 The Computational Cost

The number of floating points operations necessary to complete the simulation defines our computational cost. It is possible to take as a measure of efficency the CPU time, but this aspect of the problem involves dealing with different architectures of machines, considering vectorization, parallelisation problems etc (cf [17] and [18]). Our goal in this paper is limited to compare the numerical methods, and not their machine dependent implementations.
So the computational cost of a simulation will be given by the product of the number of points in space (N_x^d) (because the domain considered is a square $d = 2$ or a cube $d = 3$ and we choose for simplicity $\Delta x = \Delta y = \Delta z$) by the number of time steps (N_t) by the number of operations per grid point and time step (N_{op}) for the given numerical method. We can then write :

$$Cost \;=\; N_{op}.(N_x)^d.N_t \tag{2.4}$$

Introducing the two classical non-dimensional parameters, $N_\lambda = \lambda_{min}/\Delta x$ number of points per shortest wavelength and $N_p = \lambda_{min}/(c.\Delta t)$ number of points per shortest period we can write :

$$N_x \;=\; \frac{A}{\Delta x} \;=\; A.\frac{N_\lambda}{\lambda_{min}} \;=\; J.\lambda_{min}/2.\frac{N_\lambda}{\lambda_{min}} \;=\; \frac{J}{2}.N_\lambda$$

$$N_t \;=\; T_{max}\Delta t \;=\; \frac{J.\lambda_{min}}{c}.\frac{N_p.c}{\lambda_{min}} \;=\; J.N_p$$

so finally we can write the cost in dimension d as follows :

$$Cost \;=\; \frac{J^{d+1}}{2^d}.N_{op}.N_p.N_\lambda^d \tag{2.5}$$

With this definition of the cost we can state more clearly the question we want to answer.
Given *precision* requirements on the phase and group velocities, *what scheme is the most economical ?*. A related question is *Does it make sense, in that context, to use high order schemes ?*.

3 Schemes of Second Order Accuracy in Time

As mentioned in the introduction, we only consider schemes of fourth order accuracy in space. We concentrate on the two most widely used (2-4) schemes, the (2-4) compact scheme (cf [13]) and the (2-4) staggered scheme (cf [8]). Those schemes are used to integrate the acoustic wave equation which is a model for the propagation of small disturbances in an acoustic medium. This kind of medium is caracterized by its density ρ and its bulk modulus K, which are functions of the space coordinates x and y. The acoustic wave equation links the pressure field u to the density and the bulk modulus in the medium as follows:

$$\frac{1}{K}\frac{\partial^2 u}{\partial t^2} - \nabla.\left(\frac{1}{\rho}\nabla u\right) \;=\; 0 \tag{3.1}$$

All the difficulty for the numerical treatment of this equation lies in the treatment of the spatial operator. One has to consider the fact that the equation contains derivatives of the possibly

non differentiable function $\frac{1}{\rho}\frac{\partial u}{\partial x}$ for instance. The two kinds of schemes we treat, deal with that potential problem in two different ways. If one considers the constant density case, then the spatial operator is the Laplacian and numerically the problem is much simpler and much less interesting. Stability issues disappear and with them the subtleties of the numerical treatment of the spatial operator $\nabla.(\frac{1}{\rho}\nabla u)$.
Physically speaking the constant density case is of poor interest as well since, for instance it does not take into account amplitude versus offset (AVO) effects.

3.1 The Compact Schemes

For the homogeneous case when the parameters K and ρ are constant it is possible to find an approximation of the spatial operator (the Laplacian) on nine points in two dimensions (on thirteen points in three dimensions). The compact scheme technique consists in generalizing that idea to the heterogeneous case, that is to find a nine (or thirteen in 3D) points approximation of the operator $\nabla.(\frac{1}{\rho}\nabla u)$.
We will introduce the scheme in the context of a two dimensional propagation problem. The three dimensional case can easily be deduce from it, and we will just give the results for the three dimensional case.
For the numerical scheme, the functions K and ρ are discretized by the following formulae

$$\begin{aligned}
\frac{1}{K_{i+1/2,j+1/2}} &= \frac{1}{\Delta x^2}\int_{x_{i,j}}^{x_{i+1,j}}\int_{x_{i,j}}^{x_{i,j+1}} \frac{1}{K(x)}\,dx\,dy \\
\frac{1}{\rho_{i+1/2,j+1/2}} &= \frac{1}{\Delta x^2}\int_{x_{i,j}}^{x_{i+1,j}}\int_{x_{i,j}}^{x_{i,j+1}} \frac{1}{\rho(x)}\,dx\,dy
\end{aligned}$$

To be able to stay on nine points, one needs to use extra coefficients computed from the given coefficients above. Those extra coefficients are “averages” of the given functions over neighbouring cells caracterised by two parameters β and λ. Setting

$$\begin{aligned}
\frac{1}{\tilde{K}_{i,j}} &= \beta(\frac{1}{K_{i+1,j}}+\frac{1}{K_{i-1,j}}+\frac{1}{K_{i,j+1}}+\frac{1}{K_{i,j-1}})+(1-4\beta)\frac{1}{K_{i,j}} \\
\frac{1}{\tilde{\rho}_{i+1/2,j}} &= (1-2\lambda)\frac{1}{\rho_{i+1/2,j}}+\lambda(\frac{1}{\rho_{i+3/2,j}}+\frac{1}{\rho_{i-1/2,j}}) \\
\frac{1}{\tilde{\rho}_{i,j+1/2}} &= (1-2\lambda)\frac{1}{\rho_{i,j+1/2}}+\lambda(\frac{1}{\rho_{i,j+3/2}}+\frac{1}{\rho_{i,j-1/2}})
\end{aligned}$$

with

$$\begin{aligned}
\frac{1}{\rho_{i+1/2,j}} &= \frac{1}{2}(\frac{1}{\rho_{i+1/2,j+1/2}}+\frac{1}{\rho_{i+1/2,j-1/2}}) \\
\frac{1}{\rho_{i,j+1/2}} &= \frac{1}{2}(\frac{1}{\rho_{i+1/2,j+1/2}}+\frac{1}{\rho_{i-1/2,j+1/2}})
\end{aligned}$$

$$
\begin{aligned}
\frac{1}{\rho_{i,j}} &= \frac{1}{4}\left(\frac{1}{\rho_{i+1/2,j+1/2}}+\frac{1}{\rho_{i+1/2,j-1/2}}+\frac{1}{\rho_{i-1/2,j+1/2}}+\frac{1}{\rho_{i-1/2,j-1/2}}\right)\\
\frac{1}{K_{i,j}} &= \frac{1}{4}\left(\frac{1}{K_{i+1/2,j+1/2}}+\frac{1}{K_{i+1/2,j-1/2}}+\frac{1}{K_{i-1/2,j+1/2}}+\frac{1}{K_{i-1/2,j-1/2}}\right)
\end{aligned}
$$

an iteration of the (2-4) compact scheme can be written as follows:

$$
\begin{aligned}
u_{i,j}^{n+1} &= 2u_{i,j}^{n}+u_{i,j}^{n-1}+\frac{\tilde{K}_{i,j}\Delta t^2}{\Delta x^2}\left[\frac{4}{3}\left\{\frac{1}{\tilde{\rho}_{i+1/2,j}}\left(u_{i+1,j}^{n}-u_{i,j}^{n}\right)-\frac{1}{\tilde{\rho}_{i-1/2,j}}\left(u_{i,j}^{n}-u_{i-1,j}^{n}\right)\right.\right.\\
&+ \left.\frac{1}{\tilde{\rho}_{i,j+1/2}}\left(u_{i,j+1}^{n}-u_{i,j}^{n}\right)-\frac{1}{\tilde{\rho}_{i,j-1/2}}\left(u_{i,j}^{n}-u_{i,j-1}^{n}\right)\right\}\\
&- \frac{1}{6}\left\{\frac{1}{\rho_{i+1,j}}\left(u_{i+2,j}^{n}-u_{i,j}^{n}\right)-\frac{1}{\rho_{i-1,j}}\left(u_{i,j}^{n}-u_{i-2,j}^{n}\right)\right.\\
&- \left.\left.\frac{1}{\rho_{i,j+1}}\left(u_{i,j+2}^{n}-u_{i,j}^{n}\right)-\frac{1}{\rho_{i,j-1}}\left(u_{i,j}^{n}-u_{i,j-2}^{n}\right)\right\}\right]
\end{aligned}
$$

The organisation of the computation can be done in two different ways according to the memory space we have at our disposal. If one computes beforehand and store the extra coefficients necessary to the iteration of the scheme, in two dimensional arrays (we need four of them, $\tilde{K}_{i,j}$, $\tilde{\rho}_{i+1/2,j}$, $\tilde{\rho}_{i,j+1/2}$, $\rho_{i,j}$), the (2-4) compact scheme takes 30 operations to iterate. Therefore in our cost analysis we have $N_{op} = 30$, in 2D. For the 3D case we have to add 12 more operations, so $N_{op} = 42$, in 3D. The stability condition states that the (2-4) compact scheme will be stable if:

$$
\left\{
\begin{aligned}
\frac{c_{max}.\Delta t}{\Delta x} &\leq \frac{\sqrt{3}}{2.\sqrt{2}} \simeq 0.61 \quad in\ 2D\\
\frac{c_{max}.\Delta t}{\Delta x} &\leq 0.5 \quad in\ 3D
\end{aligned}
\right.
\tag{3.2}
$$

where c_{max} is the maximum of the velocities in the medium ($c = \sqrt{K/\rho}$) and λ is one of the parameters used in the computation of the extra coefficients (cf [19]).

3.2 The Staggered Schemes

The approach for the staggered schemes is totally different. It was driven by stability and energy conservation considerations. The idea is to treat the spatial differential operator as the composition of two first order operators. For the homogeneous case (when the spatial operator is the Laplacian) we have a 13 points approximation in 2D and a 19 points approximation in 3D. For the heterogeneous case the extension is straightforward since we have built this scheme to respect the variational form of the continuous spatial operator.
More precisely, let us consider the 2D case. We introduce the fourth order accurate finite-difference

approximation A_x and A_y of the first order derivatives $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$, defined by

$$A_x u_{i+1/2,j} = \frac{9}{8}\left(\frac{u_{i+1,j}-u_{i,j}}{\Delta x}\right) - \frac{1}{24}\left(\frac{u_{i+2,j}-u_{i-1,j}}{\Delta x}\right)$$

$$A_y u_{i,j+1/2} = \frac{9}{8}\left(\frac{u_{i,j+1}-u_{i,j}}{\Delta x}\right) - \frac{1}{24}\left(\frac{u_{i,j+2}-u_{i,j-1}}{\Delta x}\right)$$

To approximate the spatial operator $\nabla.(\frac{1}{\rho}\nabla u) = \frac{\partial}{\partial x}(\frac{1}{\rho}\frac{\partial u}{\partial x}) + \frac{\partial}{\partial y}(\frac{1}{\rho}\frac{\partial u}{\partial y})$ we use the finite difference operator A_x, A_y and their transposes as follows :

$$\nabla.(\frac{1}{\rho}\nabla u) \simeq -\left({}^tA_x(\frac{1}{\rho}A_x) + {}^tA_y(\frac{1}{\rho}A_y)\right) u$$

We can now write the iteration step of that scheme :

$$u_{i,j}^{n+1} = 2u_{i,j}^n + u_{i,j}^{n-1} + K_{i,j}\Delta t^2 \left({}^tA_x(\frac{1}{\rho}A_x) + {}^tA_y(\frac{1}{\rho}A_y)\right) u_{i,j}^n$$

In this case we do not need extra coefficients which would translate into extra arrays in the computation. Factorizing by Δx we see that computing $A_x u_{i+1/2,j}$ requires 5 operations. Therefore an iteration of the scheme takes 28 operations, so $N_{op} = 28$.
This shows that the number of operations per node N_{op} is *not* proportional to the number of points used by the scheme. Indeed on 13 points we have 28 operations per node, and the compact scheme on 9 points, 30 operations per node (if we store the coefficients).
The generalisation of this scheme to three dimensional problems is straightforward. We need to add ${}^tA_z(\frac{1}{\rho}A_z)$ which takes 12 operations. So in 3D, we have $N_{op} = 40$. Again this does not require extra coefficients (arrays to be stored), whereas the compact scheme technique does. This can be a serious drawback of this last appraoch since in three dimensions, the issue of storage and also the issue of fast access to the memory, have to be considered.
The stability condition for the (2-4) staggered scheme is given by

$$(3.3) \qquad \begin{cases} \dfrac{c_{max}.\Delta t}{\Delta x} \leq \sqrt{\dfrac{\rho_{min}}{\rho_{max}}}.\dfrac{6}{7.\sqrt{2}} \simeq 0.606.\sqrt{\dfrac{\rho_{min}}{\rho_{max}}} & in\ 2D \\[2ex] \dfrac{c_{max}.\Delta t}{\Delta x} \leq \sqrt{\dfrac{\rho_{min}}{\rho_{max}}}.\dfrac{6}{7.\sqrt{3}} \simeq 0.495.\sqrt{\dfrac{\rho_{min}}{\rho_{max}}} & in\ 3D \end{cases}$$

where c_{max} is the maximum of the wave velocities, ρ_{min} the minimum of the the density and ρ_{max} its maximum. When the medium has a sharp density contrast, the stability condition will be accordingly less than its maximum 0.606 or 0.495. However as we will see in the sequel, the stability condition has only an indicative role. We will have to choose $p = c.\Delta t/\Delta x$ much less than the maximum allowed by the stability condition to fulfill the precision criteria we have imposed.

3.3 Dispersion Analysis and Computational Cost

Many papers analysed the numerical errors for wave propagation problems in homogeneous media, where precisely dispersion occurs. This is due to the fact that a lot less is known about the

numerical errors in heterogeneous media in general.
It is possible to analyze the errors on reflection and transmission coefficients for two layers media (cf [21], [13]), and also to quantify the error for very rapidly varying medium (cf [9]). Even in heterogeneous media, where other phenomena than dispersion occur, numerical dispersion is a major error contributor. Therefore the control of dispersion is crucial to the numerical simulation. With the precision criteria we have introduced on the phase and group velocities, it is possible to deduce the number of points per wavelength and the number of points per period one needs to choose to fulfill those precision requirements. We will explain the method in 2D and give the results for the 3D case.
The dispersion relation links the time pulsation ω to the wave vector $\vec{k}$. In two dimensions we can use polar coordinates for $\vec{k}$ and caracterize it by its norm $k = ||\vec{k}||$ and its direction with the x-axis θ. For the (2-4) compact scheme the dispersion relation is given by :

$$\left\{ \begin{array}{rcl} \omega(k,\theta) & = & \frac{2}{\Delta t} \arcsin\left(\frac{c.\Delta t}{\Delta x} \sqrt{f_x^2 + f_y^2 + \frac{1}{3}\left(f_x^4 + f_y^4\right)} \right) \\ f_x & = & \sin(\frac{k.\Delta x}{2}\cos(\theta)) \qquad f_y \;=\; \sin(\frac{k.\Delta x}{2}\sin(\theta)) \end{array} \right. \tag{3.4}$$

For the staggered schemes the dispersion relation is given by :

$$\left\{ \begin{array}{rcl} \omega(k,\theta) & = & \frac{2}{\Delta t} \arcsin\left(\frac{c\Delta t}{\Delta x} \sqrt{g^2(\frac{k.\Delta x}{2}\cos(\theta)) + g^2(\frac{k.\Delta x}{2}\cos(\theta)))} \right) \\ g(v) & = & \frac{9}{8}\sin(v) - \frac{1}{24}\sin(3.v) \end{array} \right. \tag{3.5}$$

Error on the phase velocity
Using the classical non-dimensional variables $p = c.\Delta t/\Delta x$ (Courant parameter), $H = \Delta x/\lambda$ (inverse of number of points per wavelength) and the definition of the phase velocity we can write the normalised phase error E_φ as follows for the compact scheme :

$$\left\{ \begin{array}{rcl} E_\varphi & = & \frac{1}{\pi.p.H} \arcsin\left(p.\sqrt{f_x^2 + f_y^2 + \frac{1}{3}\left(f_x^4 + f_y^4\right)} \right) - 1 \\ f_x & = & \sin(\pi.H.\cos(\theta)) \qquad f_y \;=\; \sin(\pi.H.\sin(\theta)) \end{array} \right. \tag{3.6}$$

and for the staggered scheme:

$$\left\{ \begin{array}{rcl} E_\varphi & = & \frac{1}{\pi.p.H} \arcsin\left(p.\sqrt{g_x^2 + g_y^2} \right) - 1 \\ g_x & = & \frac{9}{8}\sin(\pi.H.\cos(\theta)) - \frac{1}{24}\sin(3.\pi.H.\cos(\theta)) \\ g_y & = & \frac{9}{8}\sin(\pi.H.\sin(\theta)) - \frac{1}{24}\sin(3.\pi.H.\sin(\theta)) \end{array} \right. \tag{3.7}$$

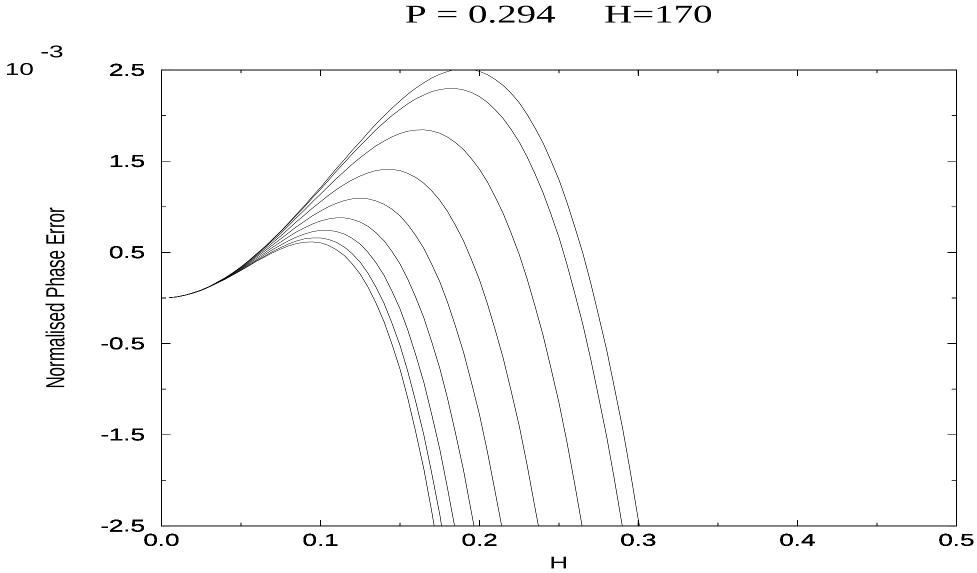
The precision on the phase error is proportional to the inverse of the number of wavelength propagated in the medium (J). For a typical geophysical application, J is of the order of 100. This compels E_φ to be in a neighborhood of 0. Since $E_\varphi(H = 0) = 0$ we can use a Taylor expansion about $H = 0$ to gain more insight in the behaviour of E_φ in that interesting region. We have for the (2-4) compact scheme:

$$E_\varphi = \frac{\pi^2 p^2 H^2}{6} - \frac{61}{90}(\cos^6(\theta) + \sin^6(\theta))\pi^4 H^4 + O(H^6)$$

and for the (2-4) staggered scheme:

$$E_\varphi = \frac{\pi^2 p^2 H^2}{6} - \frac{3}{40}(\cos^6(\theta) + \sin^6(\theta))\pi^4 H^4 + O(H^6)$$

This shows that both schemes are second order in time and fourth order in space (since $p.H$ is proportional to Δt and H to Δx). It is clear with this calculation that the time error has a positive contribution whereas the space error has a negative contribution to the phase error.
We set the error criterion to a quarter of a wavelength ($n = 4$) for one hundred wavelength ($J = 100$). So we must have for all direction of propagation θ, $|E_\varphi| \leq 1/400 = 2.5.10^{-3}$.
So according to the preceeding calculation, we have to choose p small enough so that $E_\varphi \leq 2.5\ 10^{-3}$.
Of course we will choose the biggest p in this category. This will give the biggest time step allowed by the precision criterion. Then we use the dispersion curve to find the maximum value of H admissible, for all directions θ to fulfill $E_\varphi \geq -2.5\ 10^{-3}$.



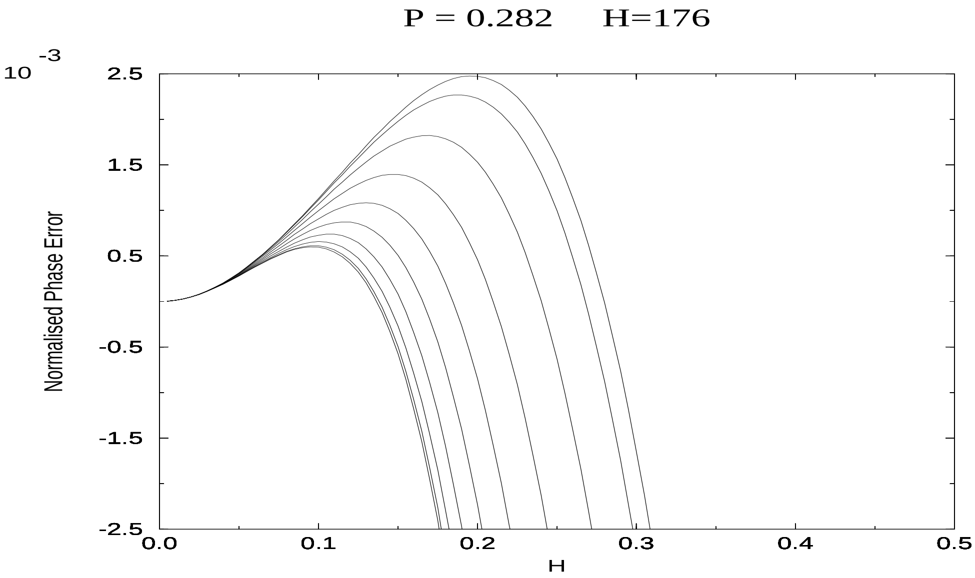


FIG 1 : The normalised phase error for different values of a propagation angle θ with the biggest courant parameter $p = C.\Delta t/h$ respecting the constraint $E_\varphi \leq 2.5\ 10^{-3}$ for the compact scheme (up) and the staggered scheme (down).

The results in 2D are the following

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.345$ $H = 0.203$	$N_p = 14.28$ $N_\lambda = 4.93$	$p = 0.334$ $H = 0.210$	$N_p = 14.25$ $N_\lambda = 4.76$
n = 4	$p = 0.294$ $H = 0.170$	$N_p = 20$ $N_\lambda = 5.88$	$p = 0.282$ $H = 0.176$	$N_p = 20.15$ $N_\lambda = 5.68$
n = 6	$p = 0.266$ $H = 0.152$	$N_p = 24.62$ $N_\lambda = 6.54$	$p = 0.256$ $H = 0.158$	$N_p = 24.64$ $N_\lambda = 6.30$

This means that for a precision of a quarter of wavelength for instance, we must take at least 20 points per period and 5.88 points per wavelength for the compact scheme, and 20.15 points per period and 5.68 points per wavelength for the staggered scheme.
Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 30$	Staggered Scheme $N_{op} = 28$	Staggered/Compact
n=2	$2.60.10^9$	$2.26.10^9$	86.82 %
n=4	$5.18.10^9$	$4.55.10^9$	87.73 %
n=6	$7.89.10^{10}$	$6.86.10^{10}$	86.94 %

So for a *fixed precision* of the normalised phase error the (2-4) staggered scheme will be roughly 13% cheaper than the (2-4) compact scheme, 2D.
For the 3D case, we have

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.285$ $H = 0.187$	$N_p = 18.69$ $N_\lambda = 5.32$	$p = 0.334$ $H = 0.194$	$N_p = 18.90$ $N_\lambda = 5.14$
n = 4	$p = 0.240$ $H = 0.156$	$N_p = 26.59$ $N_\lambda = 6.38$	$p = 0.231$ $H = 0.162$	$N_p = 26.54$ $N_\lambda = 6.14$
n = 6	$p = 0.217$ $H = 0.141$	$N_p = 32.63$ $N_\lambda = 7.09$	$p = 0.209$ $H = 0.146$	$N_p = 32.659$ $N_\lambda = 6.82$

The comparison of the cost in 3D gives :

Cost	Compact Scheme $N_{op} = 42$	Staggered Scheme $N_{op} = 40$	Staggered/Compact
n=2	$1.48.10^{12}$	$1.29.10^{12}$	86.97 %
n=4	$3.62.10^{12}$	$3.08.10^{12}$	84.96 %
n=6	$6.11.10^{12}$	$5.19.10^{12}$	84.93 %

Error on the group velocity

From the dispersion relations (3.4) and (3.5), we can compute the group velocities by taking the derivative of ω with respect to k. We have for the compact scheme :

$$(3.8)\quad \left\{ \begin{array}{lll} E_{gr} & = & \dfrac{\tilde{f}_x f_x.(1+\frac{2}{3}f_x^2) + \tilde{f}_y f_y.(1+\frac{2}{3}f_y^2)}{\sqrt{1-p^2\left(f_x^2+f_y^2+\dfrac{1}{3}(f_x^4+f_y^4)\right)}.\left(f_x^2+f_y^2+\dfrac{1}{3}(f_x^4+f_y^4)\right)} - 1 \\ f_x & = & \sin(\pi.H.\cos(\theta)) \qquad f_y \;=\; \sin(\pi.H.\sin(\theta)) \\ \tilde{f}_x & = & \cos(\theta).\cos(\pi.H.\cos(\theta)) \qquad \tilde{f}_y \;=\; \sin(\theta).\cos(\pi.H.\sin(\theta)) \end{array} \right.$$

and for the staggered scheme as follows :

$$(3.9)\qquad \left\{\begin{array}{rcl} E_{gr} & = & \dfrac{\tilde{g}_x g_x + \tilde{g}_y g_y}{\sqrt{1-p^2\left(g_x^2+g_y^2\right)}.\sqrt{g_x^2+g_y^2)}} \\ g_x & = & \dfrac{9}{8}\sin(\pi.H.\cos(\theta)) - \dfrac{1}{24}\sin(3.\pi.H.\cos(\theta)) \\ g_y & = & \dfrac{9}{8}\sin(\pi.H.\sin(\theta)) - \dfrac{1}{24}\sin(3.\pi.H.\sin(\theta)) \\ \tilde{g}_x & = & \cos(\theta)\left(\dfrac{9}{8}\sin(\pi.H.\cos(\theta)) - \dfrac{1}{8}\sin(3.\pi.H.\cos(\theta))\right) \\ \tilde{g}_y & = & \sin(\theta)\left(\dfrac{9}{8}\sin(\pi.H.\sin(\theta)) - \dfrac{1}{8}\sin(3.\pi.H.\sin(\theta))\right) \end{array}\right.$$

The results in 2D are the following.

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.302$	$N_p = 24.54$	$p = 0.290$	$N_p = 24.68$
	$H = 0.135$	$N_\lambda = 7.41$	$H = 0.139$	$N_\lambda = 7.158$
n = 4	$p = 0.255$	$N_p = 34.86$	$p = 0.245$	$N_p = 34.88$
	$H = 0.112$	$N_\lambda = 8.89$	$H = 0.117$	$N_\lambda = 8.55$
n = 6	$p = 0.231$	$N_p = 42.69$	$p = 0.222$	$N_p = 42.74$
	$H = 0.101$	$N_\lambda = 9.86$	$H = 0.105$	$N_\lambda = 9.48$

Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 30$	Staggered Scheme $N_{op} = 28$	Staggered/Compact
n=2	$10.12.10^9$	$8.85.10^9$	87.49 %
n=4	$20.66.10^9$	$17.85.10^9$	86.40 %
n=6	$31.14.10^9$	$26.82.10^9$	86.13 %

For a *fixed precision* on the normalised group error, the (2-4) staggered scheme in 2D will be roughly 13% cheaper than the (2-4) compact scheme.
The results in 3D are the following.

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.246$	$N_p = 32.72$	$p = 0.237$	$N_p = 32.70$
	$H = 0.124$	$N_\lambda = 8.05$	$H = 0.129$	$N_\lambda = 7.75$
n = 4	$p = 0.209$	$N_p = 46.11$	$p = 0.200$	$N_p = 46.25$
	$H = 0.104$	$N_\lambda = 9.61$	$H = 0.108$	$N_\lambda = 9.25$
n = 6	$p = 0.188$	$N_p = 56.58$	$p = 0.181$	$N_p = 56.72$
	$H = 0.093$	$N_\lambda = 10.67$	$H = 0.097$	$N_\lambda = 10.26$

Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 42$	Staggered Scheme $N_{op} = 40$	Staggered/Compact
n=2	$8.96.10^{12}$	$7.61.10^{12}$	84.93 %
n=4	$21.52.10^{12}$	$18.30.10^{12}$	85.05 %
n=6	$36.10.10^{12}$	$30.68.10^{12}$	86.04 %

Remarks

- The preceeding analysis shows that the (2-4) compact scheme approximate better the time derivatives, since in the two cases (phase and group errors) it required less points per period than the (2-4) staggered scheme. Conversely, the (2-4) staggered scheme is better for the spatial approximation, since for a fixed error level it required less points per wavelength.
- The 3D case implies an increase of points per period and points per wavelength, compared to the 2D case. This is not surprising since the third dimension implies that in the direction of the diagonal of the cube (the worst direction in 3D), one has fewer points per wavelength than in the worst 2D direction (diagonal of the square).

4 Schemes of Fourth Order Accuracy in Time

The fourth order in time schemes are obtained by a Lax-Wendroff correction (cf [23]) of the wave equation, which is often called the modified equation approach (cf [14] and [22]). It consists in decreasing the truncation error of the finite difference in time, by replacing the time derivatives of order higher than two, by their spatial equivalent using the wave equation. The idea can be simply explained in homogeneous media. The truncation error of the finite-difference in time is :

$$\frac{\partial^2 u}{\partial t^2} = \frac{u_{i,j}^{n+1} - 2u_{i,j}^n + u_{i,j}^{n-1}}{\Delta t^2} + \frac{\Delta t^2}{12}\frac{\partial^4 u}{\partial t^4} + O(\Delta t^6)$$

Using the wave equation which in this case of homogeneous medium is

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

we can replace the fourth order derivative in time by

$$\frac{\partial^4 u}{\partial t^4} = c^4 \frac{\partial^4 u}{\partial x^4}$$

With this correction we therefore approximate at the discrete level the modified equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - \frac{c^2.\Delta t^2}{12}\frac{\partial^4 u}{\partial x^4}$$

An important remark for the numerical method comes from the fact that the correction term needs only be approximated at the second order in space, since it is multiplied by Δt^2 and $\Delta t^2.O(\Delta x^2) = O(\Delta t^4) = O(\Delta x^4)$ by the stability condition.

4.1 The Compact Schemes

For the fourth order in time case, an iteration of the scheme can be written as follows

$$\begin{aligned}
u_{i,j}^{n+1} &= 2u_{i,j}^n + u_{i,j}^{n-1} + \frac{\tilde{K}_{i,j}\Delta t^2}{h^2}\left[\frac{4}{3}\left\{\frac{1}{\tilde{\rho}_{i+1/2,j}}\left(u_{i+1,j}^n - u_{i,j}^n\right) - \frac{1}{\tilde{\rho}_{i-1/2,j}}\left(u_{i,j}^n - u_{i-1,j}^n\right)\right.\right.\\
&+ \left.\frac{1}{\tilde{\rho}_{i,j+1/2}}\left(u_{i,j+1}^n - u_{i,j}^n\right) - \frac{1}{\tilde{\rho}_{i,j-1/2}}\left(u_{i,j}^n - u_{i,j-1}^n\right)\right\}\\
&- \frac{1}{6}\left\{\frac{1}{\rho_{i+1,j}}\left(u_{i+2,j}^n - u_{i,j}^n\right) - \frac{1}{\rho_{i-1,j}}\left(u_{i,j}^n - u_{i-2,j}^n\right)\right.\\
&- \left.\left.\frac{1}{\rho_{i,j+1}}\left(u_{i,j+2}^n - u_{i,j}^n\right) - \frac{1}{\rho_{i,j-1}}\left(u_{i,j}^n - u_{i,j-2}^n\right)\right\}\right] - \frac{\Delta t^2}{12}B_h u_{i,j}^n
\end{aligned}$$

with B_h is a second order approximation of the operator B defined by

$$Bu = K\nabla.\left(\frac{1}{\rho}\left(K\nabla.\left(\frac{1}{\rho}\nabla u\right)\right)\right)$$

The best way to compute B_h is to keep it under its factorized form. We will have a clear understanding of the process if we explain it in the "continuous" operator B. We can write :

$$\begin{aligned}
f &= K\nabla.\left(\frac{1}{\rho}\nabla u\right)\\
Bu &= K\nabla.\left(\frac{1}{\rho}\nabla f\right)
\end{aligned}$$

To compute B_h we use the same approach and the following second order approximation.

$$\begin{aligned}
f^n_{i,j} &= \frac{K_{i,j}}{\Delta x^2}.\left[\left(\frac{1}{\rho_{i+1/2,j}}.(u^n_{i+1,j}-u^n_{i,j}) - \frac{1}{\rho_{i-1/2,j}}.(u^n_{i,j}-u^n_{i-1,j})\right)\right. \\
&+ \left.\left(\frac{1}{\rho_{i,j+1/2}}.(u^n_{i,j+1}-u^n_{i,j}) - \frac{1}{\rho_{i,j-1/2}}.(u^n_{i,j-1}-u^n_{i,j})\right)\right] \\
B_h u^n_{i,j} &= \frac{K_{i,j}}{\Delta x^2}.\left[\left(\frac{1}{\rho_{i+1/2,j}}.(f^n_{i+1,j}-f^n_{i,j}) - \frac{1}{\rho_{i-1/2,j}}.(f^n_{i,j}-f^n_{i-1,j})\right)\right. \\
&+ \left.\left(\frac{1}{\rho_{i,j+1/2}}.(f^n_{i,j+1}-f^n_{i,j}) - \frac{1}{\rho_{ij-1/2}}.(f^n_{i,j-1}-f^n_{i,j})\right)\right]
\end{aligned}$$

In this calculation we have implicitly assumed that we can store the temporary variable f. Then the correction part B_h takes 24 operations per node so that the (4-4) compact scheme actually takes $N_{op} = 30 + 24 = 54$ operations per node.
In 3D we have to add, 12 operations for the computations of B_h, and the number of operations for the (2-4) part of the scheme is 42, as we have seen in the previous section. So in 3D, we have $N_{op} = 42 + 24 + 12 = 78$. The stability conditions are given by :

$$\left\{\begin{aligned}
\frac{c_{max}.\Delta t}{\Delta x} &\leq \frac{\sqrt{2}}{2} \simeq 0.707 \ \ in\ 2D \\
\frac{c_{max}.\Delta t}{\Delta x} &\leq \frac{\sqrt{3}}{3} \simeq 0.577 \ \ in\ 3D
\end{aligned}\right. \tag{4.1}$$

4.2 The Staggered Schemes

For the staggered scheme the correction term is straightforward to integrate. We define

$$\nabla_h.(\frac{1}{\rho}\nabla_h u) = -\left({}^tA_x(\frac{1}{\rho}A_x) + {}^tA_y(\frac{1}{\rho}A_y)\right) u$$

Then the modified scheme is given by :

$$u^{n+1}_{i,j} = 2u^n_{i,j} + u^{n-1}_{i,j} + \nabla_h.(\frac{1}{\rho}\nabla_h u)^n_{i,j} + K.\frac{\Delta t^2}{12}\nabla_h.(\frac{1}{\rho}\nabla_h K \nabla_h.(\frac{1}{\rho}\nabla_h u^n_{i,j}))$$

In this case we approximate the correction to fourth order in space, which is more than necessary. The reason we do this is that we have been able to prove the stability of the scheme, in that case. We can use a second order approximation based on the principle of approximating the first order derivative. But then we do not have a stability result, even though numerical experiment shows that the scheme is "stable" under the stability condition of the (2-4) scheme.
The number of operations required by the correction term when we use the fourth order approximation is 25 in 2D, so the total number of operation for the scheme is $N_{op} = 28 + 26 = 54$. We

assume in this calculation that we have stored the value of $B_x = {}^tA_x(\frac{1}{\rho}A_x)$ and $B_y = {}^tA_y(\frac{1}{\rho}A_y)$ so that we can reuse them in the computation of the correction term. With this notation we can write the (4-4) staggered scheme as follows:

$$u_{i,j}^{n+1} = 2u_{i,j}^n + u_{i,j}^{n-1} + (B_x + B_y)u_{i,j}^n + {}^t(B_x + B_y)K(B_x + B_y)u_{i,j}^n$$

So we compute in a first step B_x and B_y, then compute (and store of course) their sum in an array V, then compute ${}^t(B_x + B_y)KV = {}^tB_xKV + {}^tB_yKV$. Each computation of one of the B operators takes 11 operations, so we add 25 operations for the correction term. In 3D, the same procedure gives 38 operations for the correction term. So in 3D, $N_{op} = 40 + 39 = 79$ The stability conditions for that scheme are the same as the (2-4) staggered scheme that is :

$$\begin{cases} \dfrac{c_{max}.\Delta t}{\Delta x} \leq \sqrt{\dfrac{\rho_{min}}{\rho_{max}}}.\dfrac{6}{7.\sqrt{2}} \simeq 0.606.\sqrt{\dfrac{\rho_{min}}{\rho_{max}}} \quad in\ 2D \\[2ex] \dfrac{c_{max}.\Delta t}{\Delta x} \leq \sqrt{\dfrac{\rho_{min}}{\rho_{max}}}.\dfrac{6}{7.\sqrt{3}} \simeq 0.495.\sqrt{\dfrac{\rho_{min}}{\rho_{max}}} \quad in\ 3D \end{cases}$$

4.3 Dispersion Analysis and Computational Cost

The dispersion relations for the (4-4) schemes are given by a modification of the dispersion relations of the (2-4) schemes (as one would expect). For the compact scheme we have :

$$(4.2) \qquad \begin{cases} \omega(k,\theta) = \dfrac{2}{\Delta t}\arcsin\left(\dfrac{c.\Delta t}{\Delta x}\sqrt{(f_x^2 + f_y^2).(1 - \dfrac{1}{3}(\dfrac{c.\Delta t}{\Delta x})^2) + \dfrac{1}{3}\left(f_x^4 + f_y^4\right)}\right) \\[2ex] f_x = \sin(\dfrac{k.\Delta x}{2}\cos(\theta)) \qquad f_y = \sin(\dfrac{k.\Delta x}{2}\sin(\theta)) \end{cases}$$

For the staggered scheme the dispersion relation is given by :

$$(4.3) \qquad \begin{cases} \omega(k,\theta) = \dfrac{2}{\Delta t}\arcsin\left(\dfrac{c\Delta t}{\Delta x}\sqrt{g_x^2 + g_y^2 - \dfrac{1}{3}(\dfrac{c.\Delta t}{\Delta x})^2(g_x^2 + g_y^2)^2}\right) \\[2ex] g_x = \dfrac{9}{8}\sin(\dfrac{3k\Delta x}{2}\cos(\theta)) - \dfrac{1}{24}\sin(\dfrac{3k\Delta x}{2}\cos(\theta)) \\[2ex] g_y = \dfrac{9}{8}\sin(\dfrac{3k.\Delta x}{2}\sin(\theta)) - \dfrac{1}{24}\sin(\dfrac{3k\Delta x}{2}\sin(\theta)) \end{cases}$$

Error on the phase velocity
For the 2D case the results are the following.

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.707$ $H = 0.191$	$N_p = 7.40$ $N_\lambda = 5.23$	$p = 0.606$ $H = 0.162$	$N_p = 10.16$ $N_\lambda = 6.15$
n = 4	$p = 0.707$ $H = 0.160$	$N_p = 8.80$ $N_\lambda = 6.25$	$p = 0.606$ $H = 0.136$	$N_p = 12.13$ $N_\lambda = 7.35$
n = 6	$p = 0.707$ $H = 0.145$	$N_p = 9.75$ $N_\lambda = 6.89$	$p = 0.606$ $H = 0.122$	$N_p = 13.47$ $N_\lambda = 8.16$

Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 54$	Staggered Scheme $N_{op} = 54$	Staggered/Compact
n=2	$2.73.10^9$	$5.10.10^9$	186.33 %
n=4	$4.64.10^9$	$8.68.10^9$	187.00 %
n=6	$6.26.10^9$	$11.88.10^9$	189.91 %

In 3D we have

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.577$ $H = 0.177$	$N_p = 9.78$ $N_\lambda = 5.64$	$p = 0.495$ $H = 0.163$	$N_p = 12.28$ $N_\lambda = 6.11$
n = 4	$p = 0.577$ $H = 0.148$	$N_p = 11.67$ $N_\lambda = 6.73$	$p = 0.495$ $H = 0.136$	$N_p = 14.67$ $N_\lambda = 7.30$
n = 6	$p = 0.577$ $H = 0.133$	$N_p = 12.94$ $N_\lambda = 7.46$	$p = 0.495$ $H = 0.123$	$N_p = 16.29$ $N_\lambda = 8.11$

Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 78$	Staggered Scheme $N_{op} = 79$	Staggered/Compact
n=2	$1.71.10^{12}$	$2.73.10^{12}$	159.87 %
n=4	$3.48.10^{12}$	$5.58.10^{12}$	160.42 %
n=6	$5.25.10^{12}$	$8.49.10^{12}$	161.62 %

Error on the group velocity

The results in 2D are the following.

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.707$ $H = 0.127$	$N_p = 11.13$ $N_\lambda = 7.87$	$p = 0.606$ $H = 0.108$	$N_p = 15.28$ $N_\lambda = 9.26$
n = 4	$p = 0.707$ $H = 0.107$	$N_p = 13.21$ $N_\lambda = 9.34$	$p = 0.606$ $H = 0.090$	$N_p = 18.33$ $N_\lambda = 11.11$
n = 6	$p = 0.707$ $H = 0.096$	$N_p = 14.73$ $N_\lambda = 10.42$	$p = 0.606$ $H = 0.081$	$N_p = 20.37$ $N_\lambda = 12.34$

Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 54$	Staggered Scheme $N_{op} = 54$	Staggered/Compact
n=2	$9.31.10^9$	$26.49.10^9$	284.53 %
n=4	$15.56.10^9$	$45.82.10^9$	294.47 %
n=6	$21.60.10^9$	$62.81.10^9$	290.78 %

In 3D we have

Precision	Compact Scheme		Staggered Scheme	
n = 2	$p = 0.577$ $H = 0.118$	$N_p = 14.68$ $N_\lambda = 8.47$	$p = 0.495$ $H = 0.108$	$N_p = 18.49$ $N_\lambda = 9.20$
n = 4	$p = 0.577$ $H = 0.099$	$N_p = 17.48.$ $N_\lambda = 10.09$	$p = 0.495$ $H = 0.091$	$N_p = 22.01$ $N_\lambda = 10.96$
n = 6	$p = 0.577$ $H = 0.089$	$N_p = 19.45$ $N_\lambda = 11.22$	$p = 0.495$ $H = 0.082$	$N_p = 24.45$ $N_\lambda = 12.18$

Using (2.5) we can compute the cost of the two methods.

Cost	Compact Scheme $N_{op} = 78$	Staggered Scheme $N_{op} = 79$	Staggered/Compact
n=2	$8.71.10^{12}$	$14.03.10^{12}$	161.10 %
n=4	$17.51.10^{12}$	$28.30.10^{12}$	161.57 %
n=6	$26.80.10^{12}$	$43.08.10^{12}$	160.73 %

Remarks

- The analysis in for the (4-4) schemes shows that the compact scheme requires less points per wavelength and less points per period than the staggered (4-4) scheme. So it is not surprising that the compact scheme achieves the same accuracy with a smaller cost.

5 Comparison between the (2-4) and (4-4) cases

First we compare the change in number of points per period when using fourth order accuracy in time. For the Phase error we give below the ratio of number of points per period for the (2-4) schemes / number of points per period for the (4-4) schemes. For the 3D case, we have

Precision	Compact Scheme	Staggered Scheme
n = 2	1.91	1.54
n = 4	2.27	1.81
n = 6	2.52	2.00

For the group error we have

Precision	Compact Scheme	Staggered Scheme
n = 2	2.22	1.76
n = 4	2.63	2.10
n = 6	2.90	2.31

and we have analogous results in 2D. This clearly means that the change to the fourth order in time benefits the compact scheme much more than the staggered scheme.
Let us see now what effect the change to fourth order accuracy in time has on the number of points per wavelength. For the phase error we have in 3D :

Precision	Compact Scheme	Staggered Scheme
n = 2	0.94	0.84
n = 4	0.94	0.84
n = 6	0.94	0.84

and for the group error:

Precision	Compact Scheme	Staggered Scheme
n = 2	0.95	0.84
n = 4	0.95	0.84
n = 6	0.95	0.84

and again analogous results for the 2D case. This means that both scheme require more points per wavelength for a given precision, when the fourth order correction term is used. Again, the compact scheme loses “less” points per wavelength than the staggered scheme.

The goal of the numerical schemes is after all, to give an accurate numerical solution of the wave equation at the lowest possible cost. So, we need to compare the (2-4) schemes and the (4-4) schemes and clearly only the two best of each category. The best of the (2-4) scheme is the staggered (2-4) scheme and the best of the (4-4) scheme is the (4-4) compact scheme. Comparing the cost of each scheme we have the following results.
In 2D for the phase error

Cost	(4-4) Compact Scheme	(2-4) Staggered Scheme	Staggered/Compact
n=2	$2.73.10^9$	$2.26.10^9$	82.78 %
n=4	$4.64.10^9$	$4.55.10^9$	98.06 %
n=6	$6.26.10^9$	$6.86.10^9$	109.58 %

In 2D for the the group error

Cost	(4-4) Compact Scheme	(2-4) Staggered Scheme	Staggered/Compact
n=2	$9.31.10^9$	$8.85.10^9$	95.05 %
n=4	$15.86.10^9$	$17.85.10^9$	112.54 %
n=6	$21.60.10^9$	$26.82.10^9$	124.35 %

In 3D for the phase error

Cost	(4-4) Compact Scheme	(2-4) Staggered Scheme	Staggered/Compact
n=2	$1.7134.10^{12}$	$1.2902.10^{12}$	75.30 %
n=4	$3.4831.10^{12}$	$3.0817.10^{12}$	88.47 %
n=6	$5.2559.10^{12}$	$5.1913.10^{12}$	98.77 %

and for the the group error

Cost	(4-4) Compact Scheme	(2-4) Staggered Scheme	Staggered/Compact
n=2	$8.71.10^{12}$	$7.615.10^{12}$	87.74 %
n=4	$17.51.10^{12}$	$18.30.10^{12}$	104.99 %
n=6	$26.80.10^{12}$	$30.68.10^{12}$	114.46 %

6 Conclusions

In this paper we endeavor to clarify the choice of a numerical scheme for wave propagation problems in seismology. We set a framework in which to compare the cost of the most widely used finite difference schemes to date. These are the compact and staggered schemes, of order two or four in time and of order four in space.
Our main objective is to control dispersion artifacts. So we require that the numerical simulations done with the numerical schemes in question satisfy an a priori precision. The criteria we chose involved the phase and group velocities.
The preceeding cost analysis showed that the choice of the numerical scheme is related to the degree of accuracy one wishes for the solution. When we increase the precision imposed on the phase and group velocities, then the (4-4) compact scheme becomes cheaper.
However the precision levels we choose are very high, since for instance they require for a group shift of a quarter of wavelength, between 20 and 25 points per period and between 11 and 13 points per wavelength in 3D, for the (4-4) scheme. That is a stiff requirement for this scheme, especially in 3D.
For geophysical applications, we can conclude that the (2-4) staggered scheme requires the smallest number of operations for a reasonable precision level. It is the most economical scheme among the family of (2-4) and (4-4) compact or staggered schemes.
We did not address the machine dependent problems of parallel architecture, vectorization or fast access to memory. Those questions need to be address in practice and are very important for the global performance of the finite difference scheme.
We had in mind in this paper to compare the numerical "methods" and not the numerical "algorithm" resulting from them.

Acknowlegement

This work was partially supported by the Texas Geophysical Parallel Computation Project, the Schlumberger Foundation, and The Rice Inversion Project. TRIP Sponsors for 1994 are Advance Geophysical, Amoco Production Co., Conoco Inc., Cray Research Inc., Exxon Production Research Co., Interactive Network Technologies, Mobil Research and Development Corp., and Texaco Inc.

References

[1] R.M Alford, K.R Kelly, D.M Boore. *Accuracy of finite difference modeling of the acoustic wave equation.* Geophysics, Vol 39, 1974.

[2] J. Gazdag *Modeling of the acoustic wave equation with transform methods.* Geophysics, Vol 46, No 6, 1981.

[3] D. D. Kosloff, E. Baysal *Forward modeling by Fourier method.* Geophysics, Vol 47, No 10, 1982.

[4] B. Fornberg. *The pseudospectral method: comparison with the finite differences for elastic wave equation.* Geophysics, Vol 52, No 4, 1985.

[5] M. Reshef, D. Kosloff, M. Edwards, C. Hsiung *Three dimensional acoustic modeling by Fourier method.* Geophysics, Vol 53, No 9, 1988.

[6] A. Sei. *Etude de schemas numeriques pour des modeles de propagation d'ondes en milieu heterogene.* Ph.D Thesis, Universite Paris IX-Dauphine, 1991.

[7] A. Sei, F.Collino *Discretization of the wave equation in Heterogeneous Media.* Proceedings of the First International Conference on the Mathematical and Numerical Aspects of Wave Propagation Phenomena, SIAM, 1991.

[8] A. Sei *Computational cost of finite-difference elastic waves modeling.* Proceedings of the 63rd SEG Annual Meeting, Washington D.C, 1993.

[9] A. Sei *Error analysis of numerical schemes for the wave equation in heterogeneous media.* Submitted to Mathematics of Computation, 1993.

[10] Y. Luo and G. Schuster *A parsimonous staggered grid differencing scheme.* Geophysical Research Letters, Vol 17, N0 2, 1990.

[11] G. Cohen, P. Joly *Fourth order schemes for the 2-D wave equation in homogeneous medium.* Proceedings of the 56^{th} SEG Annual International Meeting and Exposition, Houston, 1986.

[12] G. Cohen, P. Joly *Description and analysis of schemes, fourth order in time and space, for 1-D and 2-D acoustic equations.* Proceedings of the 57^{th} SEG Annual International Meeting and Exposition, New Orleans, 1987.

[13] G. Cohen, P. Joly *Fourth order schemes for the heterogeneous acoustic equation.* Computer Methods in Applied Mechanics and Engineering, 1990.

[14] M. Dablain *High order differencing for the scalar wave equation.* Geophysics, Vol 51, 1986.

[15] G.B Whitham *Linear and Non linear waves.* John Wiley & Sons, 1974.

[16] L. Trefethen *Group velocity in finite difference schemes.* SIAM Reviews, Vol 24, 1982.

[17] M. Kern, W.W Symes *Loop level parallelisation of a seismic inversion code.* Proceedings of the 63rd SEG Annual Meeting, Washington D.C, 1993.

[18] E. Piault, P. Duclos, A. Sei *Gigaflops performance for 2D/3D heterogeneous acoustic modeling on a massively parallel computer.* Proceedings of the First International Conference on the Mathematical and Numerical Aspects of Wave Propagation Phenomena, SIAM, 1991.

[19] I. Mufti *Large-scale three dimensional seismic models and their interpretative significance.* Geophysics, Vol 55, No 9, 1990.

[20] R. Vichnevetsky, J.B Bowles. *Fourier analysis of numerical approximations of hyperbolic equations.* SIAM Studies in Applied Mathematics, SIAM, 1982.

[21] D. Brown *A Note on the numerical solution of wave equation with piecewise smooth coefficients.* Mathematics of Computation, Vol 42 , No 166, April 1984.

[22] G.R Shubin, J.B Bell *A Modified equation approach to constructing fourth order methods for acoustic wave propagation.* SIAM Journal on Scientific and Statistical Computing, Vol 8, March 1987.

[23] P.D Lax, B Wendroff *Difference schemes for hyperbolic equations with high order of accuracy.* Communications on Pure and Applied Mathematics, Vol XVII, 1964.