Time-stepping beyond CFL: a locally one-dimensional scheme for acoustic wave propagation

Leonardo Zepeda-Núñez[1](*) , Russell J. Hewett[1], Minghua Michel Rao[2], and Laurent Demanet[1]
[1]Dept. of Mathematics and Earth Resources Lab, Massachusetts Institute of Technology, [2]Ecole Polytechnique

SUMMARY

In this abstract, we present a case study in the application of a time-stepping method, unconstrained by the CFL condition, for computational acoustic wave propagation in the context of full waveform inversion. The numerical scheme is a locally one-dimensional (LOD) variant of alternating dimension implicit (ADI) method. The LOD method has a maximum time step that is restricted only by the Nyquist sampling rate. The advantage over traditional explicit time-stepping methods occurs in the presence of high contrast media, low frequencies, and steep, narrow perfectly matched layers (PML). The main technical point of the note, from a numerical analysis perspective, is that the LOD scheme is adapted to the presence of a PML. A complexity study is presented and an application to full waveform inversion is shown.

INTRODUCTION

In seismic imaging, full waveform inversion (FWI) is the recovery of Earth's physical parameters by fitting numerical solution of wave equations, e.g., in the time or in the frequency domain, to data. Time domain solvers for the wave equation can be implemented with minimal storage cost, but they involve a time-stepping scheme that is computationally expensive when the time step is small. While the equivalent frequency-domain problem, the Helmholtz equation, can be solved with sophisticated, memory-intensive (particularly in the 3D case) preconditioners, it can alternatively be solved by computing the Fourier transform of a solution to the time-domain equation.

Let $c_{\max}$ be the maximum wave velocity in the domain and let $\omega_{\max}$ be the maximum frequency in the source signature. For traditional explicit schemes, the size of the time step Δt is limited by two factors:

1. the *Courant-Friedrichs-Lewy (CFL) condition*, which requires $\Delta t < \alpha \Delta x / c_{\max}$ for numerical *stability*,

2. the *Nyquist condition*, which requires $\Delta t < \beta 2\pi / \omega_{\max}$ for *accuracy*,

where α and β are implementation-specific scaling factors. Implicit time-stepping schemes are unconditionally stable, i.e., the CFL condition does not apply, but they require the solution of a large linear system of equations at each time step and in practice the added cost hinders their practicality.

Alternating-dimension implicit (ADI) methods offer a practical middle ground between explicit and implicit methods (Peaceman and Rachford, 1955). ADI methods can be unconditionally stable while requiring the solution of more tractable linear systems at each time step. While not all ADI methods

are adequate for computational wave propagation (Lees, 1962; Fairweather and Mitchell, 1965; Qin, 2009), so-called *locally-one dimensional* (LOD) variants exist for which time-stepping method is unconditionally stable for the acoustic wave equation, in 2D and 3D, independently of the model velocity.

The LOD method was used for solving parabolic equations in Karaa (2005). Kim and Lim (2007) developed a LOD method for hyperbolic equations and proved unconditional stability in a simple setting (see also, Geiser (2008)). In this note we extend the LOD method so that it works in the presence of a perfectly matched layer (PML), report on some computational benchmarks, and discuss its relevance in the context of full waveform inversion.

CFL CONDITION vs NYQUIST CONDITION

The LOD method's main advantage is that it is not constrained by the CFL condition, which can be much more restrictive than the Nyquist condition. In order to understand why that is the case, consider that $\omega_{\max}$ is also the maximum frequency contained in the acoustic wavefield, i.e., the frequency above which the temporal Fourier transform becomes negligible. The corresponding maximum wave number occurs in regions where the wave speed is minimal and obeys $k_{\max} = \omega_{\max}/c_{\min}$, with corresponding wavelength $\lambda_{\min} = 2\pi/k_{\max}$. If q points per wavelength are sufficient to resolve $\lambda_{\min}$[1], then $\Delta x \leq \lambda_{\min}/q$. However, the grid must also resolve the fine-grained heterogeneity of the medium L, which gives $\Delta x \leq L$. Together with the CFL condition, these restrictions on grid spacing imply

$$\Delta t \leq \min\left[\alpha\frac{1}{q}\frac{c_{\min}}{c_{\max}}\frac{2\pi}{\omega_{\max}}, \alpha\frac{L}{c_{\max}}\right].$$

Thus, the maximum time step due to the CFL condition may be may be substantially smaller than the maximum time step due the Nyquist rate, $\beta 2\pi/\omega_{\max}$. This situation happens when, for example:

- the velocity profile has a high contrast $c_{\max}/c_{\min}$,

- the fine-grained details in the velocity profile have length scale smaller than the representative wavelength $\lambda_{\min}$,

- q is large, to ensure accuracy of the spatial discretization[2],

- a high-order spatial discretization, for example a *p*-order discontinuous Galerkin with, e.g., $p > 1$, which causes α to be small,

---

[1]In practice, $q = 6$ is a good choice.

[2]Finite differences of order p require $O((\Delta x)^{-1/p})$ grid points per wavelength, per dimension, to reach a fixed accuracy level.

- a grid with high spatial resolution is used to solve a low frequency problem, for which the Nyquist rate is now $\beta 2\pi/\omega$ with $\omega < \omega_{\max}$,

- a PML is used to approximate absorbing boundary conditions, causing an additional constraint on Δt linked to the strength of the absorption in the layer.

As a consequence, using the LOD scheme for time-stepping allows problems with high velocity contrasts, sharp medium heterogeneities, accurate discretizations, or steep PMLs, to be solved without further penalty on the time step.

LOD FORMULATION

In this paper we restrict ourselves to two spatial dimensions. Let $u(\mathbf{x},t)$ be the solution of the acoustic wave equation, with constant density, given by

$$m(\mathbf{x})\partial_{tt}u(\mathbf{x},t) = \triangle u(\mathbf{x},t) + f(\mathbf{x},t), \qquad (1)$$

$$u(\mathbf{x},0) = 0, \qquad (2)$$

$$\partial_t u(\mathbf{x},0) = 0, \qquad (3)$$

where $\mathbf{x} = (x,z)$, $m(\mathbf{x}) = 1/c^2(\mathbf{x})$ is the model with units of squared slowness, and $f(\mathbf{x},t) = g(\mathbf{x})w(t)$ is a separable source with point impulse g and essentially bandlimited wavelet w. For now, we disregard spatial boundary conditions. Without loss of generality, we assume that $\Delta x = \Delta z$.

While there are many different approaches for discretizing (1)–(3), in this paper we use finite differences to approximate the derivatives both in time and in space. Let $t^n = n\Delta t$ and $h = \Delta x$, and let $\mathbf{u}^n$ and $\mathbf{m}$ be the discrete approximations of $u(\mathbf{x},t^n)$ and $m(\mathbf{x})$, respectively. Let δ_x^h, δ_z^h, δ_{xx}^h, and δ_{zz}^h be spatial finite difference operators, of order p, for the first (x and z) and second derivatives (xx and zz). Then, the simplest explicit time-stepping scheme for approximating solutions to (1) is

$$\mathbf{mu}^{n+1} = \mathbf{m}\left(2\mathbf{u}^{n+1} - \mathbf{u}^{n-1}\right) + \Delta t^2 \left((\delta_{xx}^h + \delta_{zz}^h)\mathbf{u}^n + \mathbf{f}^n\right), \qquad (4)$$

which is second order accurate in time when u is smooth enough. Each time step in this scheme is computationally inexpensive, requiring only $\mathcal{O}(N)$ floating point operations and $\mathcal{O}(N)$ storage, where $N = N_x \times N_z$ is the number of unknowns. However, this scheme is limited by the CFL condition.

Implicit methods operate on the principle that the terms involving spatial differences can also be evaluated at time t^{n+1}, which results in a linear system of equations for $\mathbf{u}^{n+1}$. ADI methods split the time step into two or more sub-steps and treat a single dimension implicitly in each sub-step. As a result, the linear systems due to the implicit part are simple, tightly banded systems.

The LOD method we describe in this article has similar structure, but the sub-steps are conceptually different. Disregarding

boundary conditions, the LOD scheme is

$$\left(\mathbf{m} - \eta\Delta t^2 \delta_{xx}^h\right)\widetilde{\mathbf{u}} = \mathbf{m}\left(2\mathbf{u}^n - \mathbf{u}^{n-1}\right) + \Delta t^2 \left(\delta_{xx}^h \mathbf{u}_\eta^n + \mathbf{f}^n\right), \qquad (5)$$

$$\left(\mathbf{m} - \eta\Delta t^2 \delta_{zz}^h\right)\mathbf{u}^{n+1} = \mathbf{m}\widetilde{\mathbf{u}} + \Delta t^2 \left(\delta_{zz}^h \mathbf{u}_\eta^n\right), \qquad (6)$$

where η is a constant, $\mathbf{u}_\eta^n = ((1-2\eta)\mathbf{u}^n + \eta\mathbf{u}^{n-1})$. This LOD scheme is unconditionally stable if $\eta > 0.25$ (Geiser, 2008).

In contrast to traditional ADI methods (Lees, 1962; Fairweather and Mitchell, 1965), the sub-step solution $\widetilde{\mathbf{u}}$ is not a solution at $t^{n+1/2}$, rather, it is a full step approximation of the solution at t^{n+1}, which is then corrected by (6). Moreover, the LOD method use information from a single derivative term at each step, whereas ADI methods usually involve information from both derivative terms at each half time-step.

The LOD scheme is easily extended to 3D, by adding a third equation to (5) and (6), to account for the correction in y. The stability condition on η remains unchanged in the 3D case.

LOD formulation with PML

Open boundary conditions for wave equations are approximated by an absorbing condition on the boundary of the computational domain. A naive truncation of the domain introduces unphysical reflections that pollute the solution and cause severe imaging artifacts. One approach for mitigating unphysical reflections is to pad the domain with an absorption layer. An efficient realization of such a scheme is the perfectly matched layer (PML), introduced by Bérenger (1994). The PML approach involves complex coordinate stretching, which results in the introduction of auxiliary wavefields to absorb the energy of the outgoing wave. Because it requires fewer auxiliary wavefields, we consider the PML method for the scalar wave equation introduced by Sim (2010),

$$m\left(\partial_{tt}u + \sigma_{x+z}\partial_t u + \sigma_{xz}u\right) = \triangle u + \partial_x\phi_x + \partial_z\phi_z + f, \qquad (7)$$

$$\partial_t\phi_x = -\sigma_x\phi_x + \sigma_{z-x}\partial_x u, \qquad (8)$$

$$\partial_t\phi_z = -\sigma_z\phi_z + \sigma_{x-z}\partial_z u, \qquad (9)$$

$$\phi_x(\mathbf{x},0) = 0, \quad \phi_z(\mathbf{x},0) = 0, \quad u(\mathbf{x},0) = 0, \quad \partial_t u(\mathbf{x},0) = 0, \quad (10)$$

where σ_x and σ_z are the PML profile functions, $\sigma_{xz} = \sigma_x\sigma_z$, $\sigma_{x+z} = \sigma_x + \sigma_z$, etc., and ϕ_x and ϕ_z are the auxiliary wavefields.

LOD solver for the wave equation with PML
We extend the LOD solver in (5) and (6) to the PML-augmented scalar wave equation in (7)–(10). Discretizing (7)–(10) as before leads to a system of equations for $\mathbf{u}$ and the the discrete auxiliary wavefields, $\boldsymbol{\phi}_x$ and $\boldsymbol{\phi}_z$. In the PML formulation, the evolution equations for $\boldsymbol{\phi}_x$ and $\boldsymbol{\phi}_z$ involve spatial operators only in x and in z, respectively. Thus, with the auxiliary fields already decoupled and the property that the LOD method treats dimensions independently, we formulate a decoupled system defined by the operators

$$L^x = \begin{bmatrix} \mathbf{m} - \eta\Delta t^2\delta_{xx}^h & -\Delta t^2\delta_x^h \\ -\Delta t\eta\sigma_{z-x}\delta_x^h & I + \gamma_x \end{bmatrix}, \qquad (11)$$

$$L^z = \begin{bmatrix} \mathbf{m}(I + \gamma_{x+z}) - \eta\Delta t^2(\delta_{zz}^h - m\sigma_{xz}) & -\Delta t^2\delta_z^h \\ -\Delta t\eta\sigma_{x-z}\delta_z^h & I + \gamma_z \end{bmatrix}, \qquad (12)$$

in which L^x and L^z contain only spatial differences in x and in z respectively, and where $\gamma_x = \Delta t \sigma_x / 2$, $\gamma_z = \Delta t \sigma_z / 2$, and $\gamma_{x+z} = \Delta t \sigma_{x+z} / 2$.

The complete LOD scheme for the PML formulation is

$$L^x \begin{bmatrix} \widetilde{\mathbf{u}} \\ \boldsymbol{\phi}_x^{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{m}(2\mathbf{u}^n - \mathbf{u}^{n-1}) + \Delta t^2 \left(\delta_{xx}^h \mathbf{u}_\eta^n + \mathbf{f}^n \right) \\ (1 - \gamma_x) \boldsymbol{\phi}_x^n + \Delta t \sigma_{z-x} \delta_x^h \mathbf{u}_\eta^n \end{bmatrix} \tag{13}$$

$$L^z \begin{bmatrix} \mathbf{u}^{n+1} \\ \boldsymbol{\phi}_z^{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{m}\widetilde{\mathbf{u}} + \Delta t^2 \left(\delta_{zz}^h - \mathbf{m}\sigma_{xz} \right) \mathbf{u}_\eta^n + \gamma_{x-z} \mathbf{u}^{n-1} \\ (1 - \gamma_z) \boldsymbol{\phi}_z^n + \Delta t \sigma_{x-z} \delta_z^h \mathbf{u}_\eta^n \end{bmatrix}, \tag{14}$$

where $\gamma_{x-z} = \Delta t \sigma_{x-z} / 2$.

For $\eta > 0.25$, we claim that this scheme is stable, provided that $\Delta t < 1 / \max \left(\|\sigma_x\|_{\ell_\infty}, \|\sigma_z\|_{\ell_\infty} \right)$. The first restriction is from the boundary-free LOD scheme and the second condition is due to the stiffness of the evolution equations ((8) and (9)) for ϕ_x and ϕ_z. As a result, the LOD method for the wave equation with PML is no longer unconditionally stable, but the stability requirement due to the PML does not involve the grid spacing h, hence the restriction is very mild when compared to the original CFL condition.

Efficient inversion

The matrices defined in (11) and in (12) are expensive to invert as is. Precomputing the inverse requires a large amount of memory because of fill-in. To overcome this difficulty, we precompute the inverse of the Schur complement as a sparse banded matrix. For

$$L = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \tag{15}$$

we have

$$L_{\text{Schur}} = A - BD^{-1}C. \tag{16}$$

Using (16), a closed form expression for the inverse of an operator in the form of (15) (e.g., (11) and in (12)) is

$$L^{-1} = \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} \begin{bmatrix} L_{\text{Schur}}^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix}. \tag{17}$$

The block involving L_{Schur}^{-1} is not necessarily sparse and may be difficult to evaluate, but when D is diagonal and $A, B,$ and C are sparse and tightly banded, as in (11) and in (12), the Schur complement is also sparse and tightly banded. Thus, for our formulation, the LU factorization of (16) is sparse for both L^x and L^z. In practice, solving the Schur complement is done by two back-substitutions using the precomputed LU factors, which are stored in $\mathcal{O}(N)$ memory.

As a result, the complexity of solving (13) and (14) in each time step is $\mathcal{O}(N)$ and the new method has the same asymptotic complexity as the explicit scheme. In the next section, we examine complexity in more detail.

RESULTS

Complexity

The explicit time-stepping scheme (with PML) requires three sparse matrix-vector products, of size at most $N + M \times N + M$, where M is the number of new degrees of freedom added due to the PML and auxiliary wavefields, and a negligible number of vector arithmetic operations. M is usually a percentage of N, so we consider the cost of the explicit method to be $\mathcal{O}(N)$, per time step[3].

The dominant cost for the LOD method is in the $2N_x + 2N_z$ back substitutions, each of which has a respective cost of $\mathcal{O}(N_z)$ and $\mathcal{O}(N_x)$. When combined with a non-negligible number of vector arithmetic operations and a few smaller sparse matrix-vector products, the overall cost is still $\mathcal{O}(N)$, but the constant is larger.

The explicit time-stepping scheme (with PML) requires three wavefields per time index ($\mathbf{u}, \phi_x$ and ϕ_z), and three time snapshots for second order accuracy in time. Additionally, the three sparse update matrices must be stored, at $\mathcal{O}(N)$ storage cost. The LOD method requires one additional wavefield ($\widetilde{\mathbf{u}}$) and 16 sparse matrices, with a maximum number of nonzeros per row that depends on the spatial accuracy of the finite difference approximation (e.g., 5 for fourth-order accuracy).

To understand the range of grid spacing Δx and frequency $\omega_{\max}$ at which the LOD method would provide an advantage over traditional explicit time-stepping, we apply both methods to the left 22km of the 2004 BP Velocity-Analysis Benchmark (Billette and Brandsberg-Dahl, 2005) with $\Delta x = \Delta z = 12.5, 25,$ 50, and 100 m and compare the runtime per time step. The data shown are for Δt determined by the CFL condition with $\alpha = 1/6$. The same Δt was used for both methods so the number of time steps remained constant and both methods are fourth-order accurate in space. All results in this paper are obtained using the authors' Python Seismic Imaging Toolbox (PySIT). The results in Figure 1 are for final time $T = 5$s and averaged over 25 runs per grid scale. Empirically, our implementation of the LOD is approximately $r = 3.25$ times slower than the explicit method *per time step*.

When to use LOD

The LOD method is favored over an explicit method when the smaller overall number of time steps compensates for the additional stepwise complexity. This situation is illustrated in Figure 2, where for $c_{\min} = 1500$ m/s, $c_{\max} = 4500$ m/s, $\alpha = 1$, $\beta = 1/6$, $q = 6$, and $r = 3.25$ we consider the valid ranges for Δt and Δx for the explicit and LOD methods at $\omega_{\max} = 15\pi$ (resp. $\omega'_{\max} = 6\pi$). In region A (resp. A'), the explicit method is stable and in region B (resp. B') the LOD method is valid, but the explicit method is more efficient, even at smaller Δt. In region C (resp. C'), the LOD method is faster than the explicit method. When

$$r \leq \beta q \frac{1}{\alpha} \frac{c_{\max}}{c_{\min}}, \tag{18}$$

region C extends below the diagonal and we have that the LOD method is always favored over an explicit method.

---

[3]Sparse matrix-vector products have complexity $\mathcal{O}(N_r + nnz)$, for matrices with N_r rows and nnz non-zeros. For this problem, both are on the order of N.

LOD Methods

Figure 3 shows the solution to the Helmholtz equation, computed by the Fourier transform of the wavefields determined by the explicit and LOD methods. We used a single 2.5 Hz Ricker wavelet source in the left 22 km of the BP model, with PML boundaries on all sides and a free surface boundary on top. The LOD experiment had equivalent computational cost but required over a third fewer time steps. At higher resolution, the cost reduction is magnified. At low frequency, the two solutions are virtually indistinguishable. At higher frequency, there are differences that we attribute to differing levels of numerical dispersion in the two methods.

Application to Imaging

For this numerical experiment, we use our LOD solver to recover a 370×170 patch of the Marmousi2 p-wave velocity model (G. Martin and Marfurt, 2006). We use 24 sources and 360 receivers, equispaced at the top of the domain with a Ricker wavelet centered at 7.5 Hz. The time-domain least-squares misfit objective function was reduced by approximately two orders of magnitude in 50 L-BFGS iterations from a linearly increasing gradient initial model. The simulated measured data are computed via the explicit method with a smaller time step to prevent the so-called inverse crime. The true model, initial model, and final reconstruction are shown in Figure 4.

CONCLUSION

We present a locally one-dimensional (LOD) solver for the scalar time-dependent wave equation with PML boundary conditions. The method has very favorable stability properties not restricted by the CFL condition. It can be an efficient alternative to explicit time-stepping, e.g., in settings of low frequencies or high velocity contrasts. We demonstrate the potential of the LOD method by contrasting it to an explicit method for the numerical solution of the Helmholtz equation. Full waveform inversion tests also validate the solver.
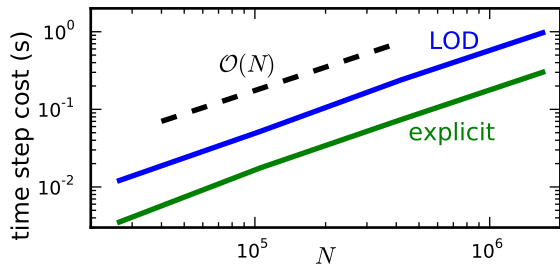
ACKNOWLEDGEMENTS

Figure 1: Run time per time step for explicit (green) and LOD (blue) as function of number of unknowns.
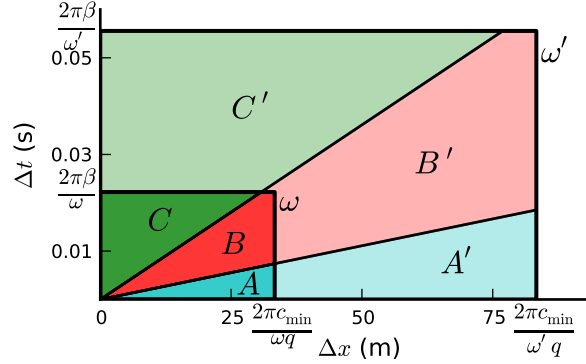

Figure 2: Boxes for $\omega = 15\pi$ (resp. $\omega' = 6\pi$) correspond to sampling restrictions in time and space. Region A (resp. A') satisfies the CFL stability condition $\Delta t \leq \alpha \Delta x / c_{\max}$, region B (resp. B') has LOD method valid, but slower than explicit method with smaller Δt, region C (resp. C') has LOD method more efficient than explicit method.
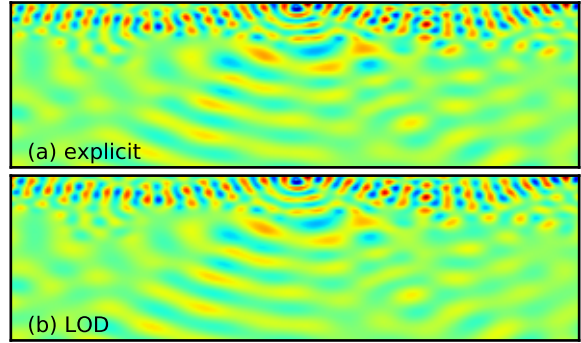

Figure 3: Real part of solution to the Helmholtz equation at 2 Hz, computed by discrete Fourier transform of temporal wavefields from (a) explicit method and (b) LOD method. LOD method is computed with 3 times fewer time steps.
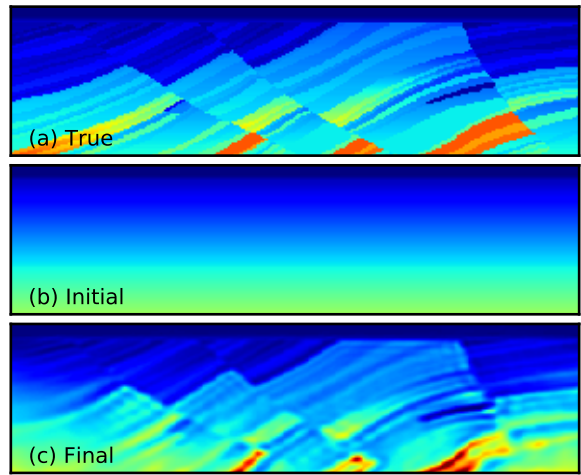

Figure 4: (a) True, (b) initial, and (c) recovered model from time-domain full waveform inversion, using LOD solver, after 50 L-BFGS iterations.

REFERENCES

Bérenger, J.-P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: J. C. P., **114**.

Billette, F., and S. Brandsberg-Dahl, 2005, The 2004 BP velocity benchmark.: EAGE.

Fairweather, G., and A. Mitchell, 1965, A high accuracy alternating direction method for the wave equation: IMA Journal of Applied Mathematics, **1(4)**.

G. Martin, R. W., and K. Marfurt, 2006, An elastic upgrade for Marmousi: The Leading Edge, Society for Exploration Geophysics, **25**.

Geiser, J., 2008, Fourth-order splitting methods for time-dependent differential equations: Numer. Math. Theor. Meth. Appl., **1(3)**.

Karaa, S., 2005, An accurate LOD scheme for two-dimensional parabolic problems: Appl. Math. Comput., **170**.

Kim, S., and H. Lim, 2007, High-order schemes for acoustic waveform simulation: Appl. Numer. Math., **57**.

Lees, M., 1962, Alternating direction methods for hyperbolic differential equations: J. SIAM, **10(4)**.

Peaceman, D., and H. Rachford, 1955, The numerical solution of parabolic and elliptic differential equations: J. SIAM, **3(1)**.

Qin, J., 2009, The new alternating direction implicit difference methods for the wave equations: J. Comput. Appl. Math., **230(1)**.

Sim, I., 2010, Nonreflecting boundary conditions for the time-dependent wave equation: PhD thesis, University of Basel.