Parallel 3-D viscoelastic finite difference seismic modelling

Thomas Bohlen*

Kiel University, Institute of Geosciences, Geophysics, Otto–Hahn–Platz 1, D–24118 Kiel, Germany

Received 5 June 2001; received in revised form 23 November 2001

Abstract

Computational power has advanced to a state where we can begin to perform wavefield simulations for realistic (complex) 3-D earth models at frequencies of interest to both seismologists and engineers. On serial platforms, however, 3-D calculations are still limited to small grid sizes and short seismic wave traveltimes. To make use of the efficiency of network computers a parallel 3-D viscoelastic finite difference (FD) code is implemented which allows to distribute the work on several PCs or workstations connected via standard ethernet in an in-house network. By using the portable message passing interface standard (MPI) for the communication between processors, running times can be reduced and grid sizes can be increased significantly. Furthermore, the code shows good performance on massive parallel supercomputers which makes the computation of very large grids feasible. This implementation greatly expands the applicability of the 3-D elastic/viscoelastic finite-difference modelling technique by providing an efficient, portable and practical C-program. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Seismic wave attenuation; Seismic wave dispersion; Seismic wave scattering; Parallel computing; Message passing interface (MPI)

1. Introduction

In order to extract information about the 3-D structure and composition of the crust from seismic observations, it is necessary to be able to predict how seismic wavefields are affected by complex structures. Since exact analytical solutions to the wave equations do not exist for most subsurface configurations, the solutions can be obtained only by numerical methods. Synthetic seismograms are helpful in predicting and understanding the kinematic and dynamic properties of seismic waves propagating through models of the crust. With the increased amount of detailed information required from seismic data, seismic modelling has become an essential tool for the evaluation of seismic measurements. It helps in every stage of a seismic investigation. It can help determine optimal recording parameters in data acquisition. Synthetic datasets can be computed to test processing procedures. The comparison of synthetic and field seismograms leads to a better understanding of seismic measurements and thus, finer details can be extracted from seismic field recordings. In seismic inversion procedures, modelling is the kernel of the inversion process.

Various techniques for seismic wave modelling in realistic (complex) media have been developed. Such methods include wavenumber integration, e.g. the Reflectivity method (Müller, 1985), Ray-tracing (Červený et al., 1977), finite elements (Chen, 1984), Fourier or pseudospectral methods (Kosloff and Baysal, 1982), hybrid methods (Emmerich, 1992), and finite differences (FD) (Alterman and Karal, 1968; Alford et al., 1974; Kelly et al., 1976).

Explicit finite difference methods have been widely used to model seismic wave propagation in 2-D elastic media, because of their ability to accurately model seismic waves in arbitrary heterogeneous media. Kelly et al. (1976) and Kelly (1983) used a displacement formulation developed from the second-order elastic
equations, and Madariaga (1976), Virieux (1984, 1986) and Levander (1988) formulated a staggered-grid, finite difference scheme based on a system of first-order coupled elastic equations where the variables are stresses and velocities, rather than displacements. These elastic finite difference simulators mentioned so far calculate synthetic seismograms for 2-D elastic models of the earth crust, but fail to model the earth’s anelastic behaviour, i.e. attenuation and dispersion of seismic waves.

Day and Minster (1984) made the first attempt to incorporate anelasticity into a 2-D time-domain modelling methods by applying a Padé approximant method. Emmerich and Korn (1987), however, found this method being of poor quality and computationally inefficient. They suggested a new approach based on the rheological model called “generalized standard linear solid” (GSLS), and developed a 2-D finite difference algorithm for scalar wave propagation. Robertsson et al. (1994b) described a staggered grid, velocity-stress finite difference technique, which is also based on the GSLS, to model the propagation of P-SV waves in 2-D viscoelastic media. Their algorithm was also extended to the 3-D situation (Robertsson et al., 1994a).

The main drawback of the FD method is that modelling of realistic 3-D models consumes vast quantities of computational resources. Such computational requirements are generally beyond the resources for sequential platforms (single PC or workstation) and even supercomputers with shared-memory configurations. In recent years it has become feasible to use clusters of workstations or PCs for scientific computing. This paper shows how FD modelling can benefit from this technique by describing a message passing implementation of a 3-D viscoelastic FD algorithm. Using the free and portable message passing interface (MPI) the calculations are distributed on PCs or workstations which are connected by an in-house network. By clustering a set of processors, for example PCs running Linux, wall-clock times can be decreased and possible grid sizes can be increased significantly. Furthermore, the code shows excellent performance on a massive parallel supercomputer (CRAY T3E). On these platforms the computation of large-scale 3-D grids (500³ gridpoints) becomes now possible in acceptable running times.

The paper is organized as follows: Section 2 provides a short review of the underlying theory of seismic wave propagation. The basic methodology of the FD technique is explained thereafter. The parallelization using MPI, the role of communication between processors, and the performance on different parallel platforms are discussed. In the last part an application to simulate seismic wave transmission through a 3-D heterogeneous elastic and viscoelastic medium is described.

2. Theory

2.1. Attenuation model

In order to include viscoelastic effects in a modelling algorithm, it is necessary to define a model of the absorption mechanism. Liu et al. (1976) showed that a linear viscoelastic rheology based on a GSLS gives a realistic framework which can explain experimental observations of wave propagation through earth materials. A GSLS can be used to model any frequency dependence of the quality factor \( Q \).

The schematic diagram (Fig. 1) shows that the GSLS is composed of \( L \) Maxwell bodies (spring \( k_i \) and dashpot \( \eta_i \) in series; \( i = 1, \ldots, L \)) connected in parallel with a spring \( k_0 \). \( \eta_i \) (\( i = 0, \ldots, L \)) and \( \eta_i \) (\( i = 1, \ldots, L \)) represent elastic moduli and Newtonian viscosities, respectively. The complex modulus \( M \) of a GSLS can be expressed in the frequency-domain as

\[
M(\omega) = k_0 \left\{ 1 - \sum_{i=1}^{L} \frac{1 + i\omega\tau_{ot}}{1 + i\omega\tau_{ot}} \right\},
\]

where \( \omega \) denotes angular frequency. The stress relaxation times \( \tau_{ot} \) and strain retardation times \( \tau_{ot} \) for the \( L \)th Maxwell body of the GSLS are connected with the

![Fig. 1. Schematic diagram of generalized standard linear solid (GSLS) composed of \( L \) so-called relaxation mechanisms or Maxwell bodies. \( k_l \) and \( \eta_l \) (\( l = 1, \ldots, L \)) represent elastic moduli and Newtonian viscosities, respectively. Stress relaxation times \( \tau_{ot} \) and strain retardation times \( \tau_{ot} \) for the \( L \)th relaxation mechanism are connected with the constituents \( (k_l, \eta_l) \) via Eqs. (2) and (3).](image-url)
constituents \((k_i, \eta_i)\) of the GSLS by (Zener, 1948)

\[
\tau_{al} = \frac{\eta_l}{k_l}
\]

and

\[
\tau_{al} = \frac{\eta_l}{k_0} + \frac{\eta_l}{k_l}.
\]

The attenuation properties of rocks are described in terms of the so-called seismic quality factor \(Q\), defined as (O’Connell and Budiansky, 1978)

\[
Q = \frac{M_R}{M_I},
\]

where \(M_R\) and \(M_I\) are the real and imaginary parts, respectively, of the complex modulus related to the elastic wave type under consideration. With Eqs. (1) and (4) the seismic quality factor \(Q\) for a GSLS reads

\[
Q(w, \tau_{al}, \tau) = \frac{1 + \sum_{l=1}^{L} \frac{w^2 \tau_{al}^2}{1 + w^2 \tau_{al}^2} \tau}{\sum_{l=1}^{L} \frac{w^2 \tau_{al}^2}{1 + w^2 \tau_{al}^2} \tau},
\]

where the variable

\[
\tau = \frac{\tau_{al}}{\tau_{al}^2} - 1,
\]

introduced by Blanch et al. (1995), is used to save memory and reduce calculations in FD modelling. Eq. (5) is the key for finding the \(L + 1\) parameters \(\tau_{al}, \tau\) which describe a constant \(Q\)-spectrum within a limited frequency range by a limited number of Maxwell bodies. These optimization variables \(\tau_{al}, \tau\) are determined by a least-squares inversion, i.e. the following function is minimized numerically in a least-squares sense:

\[
J(\tau_{al}, \tau) := \int_{\omega_0}^{\omega_1} \left[ (Q^{-1}(\omega, \tau_{al}, \tau) - \tilde{Q}^{-1})^2 \right] d\omega,
\]

where \(\tilde{Q}^{-1}\) is the desired constant \(Q\). The advantage of this procedure compared with the so-called “\(\tau\)-method” suggested by Blanch et al. (1995) is that (1) it works also for strong absorption \((Q < 10)\), and (2) the \(L\) retardation times \(\tau_{al}\) are also optimized. This optimization has to be performed for the desired \(Q\)’s for both P- and S-waves which yields \(\tau\)-values for P- and S-waves which are denoted by \(\tau^P\) and \(\tau^S\) in the following. The same relaxation times \(\tau_{al}\) can be used for both wave types.

In the situations of a GSLS consisting of only one Maxwell body \((L = 1)\) connected in parallel with the spring \((k_0)\), a good estimate for \(\tau\) is

\[
\tau = 2/Q.
\]

A GSLS with \(L = 1\) is also called standard linear solid or single relaxation mechanism. The \(Q\)-spectrum has the form of a Deby-peak with a centre frequency at \(2\pi/\tau_{al}\) Hz.

### 2.2. 3-D Viscoelastic wave equations

In this section the velocity-stress formulation of the system of differential equations which were the basis for the FD implementation is described. A derivation of these equations can be found for example in Robertsson et al. (1994b) and Carcione et al. (1988). Following Blanch et al. (1995), I use the variable \(\tau\) (see Eq. (6)) in the wave equation formulation.

The stress–strain relation for a generalized standard linear solid reads

\[
\sigma_{ij} = \frac{\partial v_k}{\partial x_k} \left\{ M(1 + \tau^P) - 2 \mu(1 + \tau^S) \right\} + \frac{2 \partial v_i}{\partial x_j} \mu(1 + \tau^P)
\]

\[
+ \sum_{l=1}^{L} r_{ijl} \quad \text{if } i = j,
\]

\[
\sigma_{ij} = \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \mu(1 + \tau^P) + \sum_{l=1}^{L} r_{ijl} \quad \text{if } i \neq j,
\]

with the so-called memory equations:

\[
r_{ijl} = -\frac{1}{\tau_{al}} \left\{ (M \tau^P - 2 \tau^S) \frac{\partial v_k}{\partial x_k} + \frac{2 \partial v_i}{\partial x_j} \mu \tau^P + r_{ijl} \right\} \quad \text{if } i = j,
\]

\[
r_{ijl} = -\frac{1}{\tau_{al}} \left\{ \mu \tau^P \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + r_{ijl} \right\} \quad \text{if } i \neq j.
\]

The equation of momentum conservation:

\[
\frac{\partial v_i}{\partial t} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i,
\]

completes the system of first-order coupled partial differential equations which describe seismic wave propagation in a 3-D viscoelastic medium. The meaning of the symbols is as follows:

- \(\sigma_{ij}\) denotes the \(ij\)th component of the stress tensor \((i, j = 1, 2, 3)\),
- \(v_i\) denote the components of the particle velocities,
- \(x_i\) indicate the three spatial directions \((x, y, z)\),
- \(r_{ijl}\) are the \(L\) memory variables \((l = 1, \ldots, L)\) which correspond to the stress tensor \(\sigma_{ij}\),
- \(f_i\) denotes the components of external body force,
- \(\tau_{al}\) are the \(L\) stress relaxation times for both P- and S-waves,
- \(\tau^P, \tau^S\) define the level of attenuation for P- and S-waves, respectively,
- \(\rho\) is the density.

The dot over symbols indicates partial differentiation with respect to time. The moduli \(M\) and \(\mu\) are used to define the phase velocity models \(v_{po}\) and \(v_{so}\) at the reference frequency \(\omega_0\) for P- and S-waves, respectively. \(\omega_0\) should equal the centre frequency of the source so that the main frequencies travel with \(v_{po}\) and \(v_{so}\). In order to achieve this the moduli \(M\) and \(\mu\) can be
calculated by

\[
M = \mathbf{v}_{\rho_0}^2 \mathcal{R}^2 \left( \frac{1}{1 + \frac{L}{\sum_{i=1}^{L} 1 + i \omega_{t} \tau_{al}} - \tau_{rr}} \right),
\]

\[
\mu = \mathbf{v}_{\sigma_0}^2 \mathcal{R}^2 \left( \frac{1}{1 + \frac{L}{\sum_{i=1}^{L} 1 + i \omega_{t} \tau_{al}} - \tau_{rr}} \right),
\]

where the symbol \( \mathcal{R}() \) denotes the real part of the argument. The parameters \( \tau_{al}, \tau_{rr} \) and \( \tau_{rr} \) which are optimized for the desired \( Q \)-spectra in a least-squares sense (see Eq. (7)) have to be inserted into Eqs. (12).

3. Finite difference algorithm

3.1. Discretization

The coupled system of continuous differential Eqs. (9)–(11) were transformed into discretized equivalents using a staggered-grid approach. Following Levander (1988) and Robertsson et al. (1994b), a second-order centred difference scheme was applied to approximate the time derivatives, and a fourth-order staggered scheme with centred differences to approximate the spatial derivatives (\( O(2,4) \)). The explicit finite difference equations used for updating the wavefield are given in Appendix A.

Fig. 2 shows the staggered spatial locations of the \( 12 + 6L \) dynamic (time-dependent) variables (six stress components \( \sigma_{ij} \), \( 6L \) memory variables \( r_{ij} \), three components of particle velocity \( v_{ij} \), three body force components \( f_{ij} \)) and the five material parameters (moduli \( M \) and \( \mu \), level of P- and S-wave attenuation \( \tau \), \( \tau \), and density \( \rho \)) on a cubic finite difference cell. The \( 12 + 6L \) dynamic variables describe wave propagation, and the five constants define the structure and material properties of the viscoelastic medium. \( L \) is the number of Maxwell bodies which are considered in the computations. Note that the \( L \) stress relaxation times \( \tau_{al} \) need not be defined within each finite difference cell assuming that the frequency dependency of \( Q \) does not vary in the model, which is the case for example when \( Q = \text{const} \). The level of attenuation, i.e. \( Q \), however, can be heterogeneous since it is quantified by \( \tau_{rr} \) and \( \tau_{rr} \) via Eq. (7). The indices \( i, j, k \) represent the node number in x-, y- and z-direction, respectively, \( n \) is the time level. The grid is staggered in both space and time, i.e. the variables and constants are not known at the same discretized point. The spatial grid size of the square grid is \( h \) and the temporal discretization interval is \( \Delta t \).

3.2. Stability and numerical dispersion

Finite difference modelling is often regarded as a realistic modelling method for arbitrary complex models. Inaccuracies, i.e. numerical grid dispersion and grid anisotropy, may occur if the spatial and temporal sampling is not fine enough. For the \( O(2,4) \) scheme the spatial grid spacing \( h \) must be \( < \frac{\tau_{min}}{6} \) where \( \tau_{min} \) denotes the minimum shear wavelength at maximum frequency within the model. This guarantees that the error due to numerical dispersion is smaller than \( 5\% \) (Robertsson et al., 1994b). The 3-D FD algorithm is stable provided that the time step \( \Delta t \) fulfills the criterion

\[
\Delta t \leq \frac{6h}{7 \sqrt{3 v_{p}^{max}}},
\]

where \( v_{p}^{max} \) is the maximum P-wave velocity at maximum frequency travelling within the model Blanch, 1995).

3.3. Boundary conditions

The geologic model that is used as input to the FD equation is restricted to a finite number of grid points. A free surface is imposed at the top of the model by applying the imaging method suggested by Levander (1988) (see also Robertsson, 1996). The discretized model likewise is bounded laterally and at the bottom, but these edges do not represent real boundaries. To reduce reflections from the edges of the numerical grid various techniques have been proposed. Generally a dissipative frame around the grid is used in which particle velocities and stress values are decreased smoothly by multiplying with a factor smaller than \( l \) (Cerjan et al., 1985). Another possibility would be to introduce a frame with low \( Q \)-values around the grid Robertsson et al., 1994b; Liao and McMechan, 1996). However, this procedure is less efficient than the method of Cerjan et al. (1985).

4. Parallelization

4.1. The message passing interface (MPI) standard

MPI is a library specification for message passing, proposed as a standard by a broadly based committee of vendors, implementors, and users\(^1\). MPI was designed for high performance on both massively parallel machines and on workstation clusters. MPI is a public domain software which is available in FORTRAN and C. MPI provides source-code portability of message-

passing programs across a variety of architectures. More information is available on the web\(^2\). The FD code was developed on a PC Linux cluster running local area multicomputer (LAM), a MPI implementation for Linux based networks\(^3\). The same MPI code is running on a massive parallel supercomputer (CRAY T3E) without modifications.

### 4.2. Grid decomposition

The decomposition of the global 3-D model into subvolumes is illustrated in Fig. 3. Each processing element (PE) is updating the wavefield using the Eqs. (21)–(34) within its portion of the grid. The processors lying at the top of the global grid generally apply a free surface boundary conditions while the processors lying at the other edges apply absorbing boundary conditions or periodic boundary conditions. At the internal edges the processors must exchange the wavefield information, i.e. the stress tensor \( \tau \) and particle velocities \( v \). Memory variables do not have to be exchanged because no spatial derivates of these variables are required during wavefield update. For the communication at the internal edges a two-point-thick

---


padding layer has to be introduced. The thickness of this padding layer is generally half of the length of the spatial finite difference operator which is two for a fourth-order FD operator. Data exchange at the internal edges has to be performed after each timestep. After communication the padding layer always contains the most recently updated wavefield received from the neighbouring processor. This information is required to calculate the spatial derivatives of stress and particle velocity at the first two gridpoints within the internal grid.

4.3. Communication

Communication plays an important role in any application which is network dependent. The critical factor is the amount of data (network load) which has to be exchanged at each timestep between processors. If the ratio of communication time versus computation time is high, parallelization may be counterproductive. Since explicit MPI calls were used for the communication between PEs the amount of communication could be quantified. The total amount of data, denoted by \( D \), which has to be transmitted after each timestep in a 3-D viscoelastic FD run of a global cubic grid with \( N \) gridpoints divided into \( p \) cubic subgrids is

\[
D = 240(N^2p)^{1/3} \text{ bytes.} \tag{14}
\]

\( D \) does not depend on the number of Maxwell bodies \( (L) \) used in the simulation. Thus, communication times for elastic and viscoelastic simulations are equal. Plots of Eq. (14) as a function of the number of PEs \( (p) \) are shown in Fig. 4 for different grid sizes \( (N = 200^3, 500^3, 750^3) \) as solid lines. The amount of communication increases significantly with grid size \( (N) \).

For example, in a simulation of \( N = 500^3 \) gridpoints with more than 200 PEs the amount of data which has to be exchanged between PEs exceeds 300 Mbytes per timestep. Thus, a fast network is required to achieve good parallelization.

4.4. Memory requirements

3-D FD simulations generally require a large amount of memory which is distributed on different PEs in a parallel simulation. The local memory requirements (on each PE) can be estimated by

\[
\text{memory/PE} \approx \frac{4(17 + 6L)N}{p} \text{ bytes,} \tag{15}
\]

where \( L \) is the number of Maxwell bodies used in the simulation. Eq. (15) is useful to determine the minimum number of PEs \( (p) \) required to store a grid with \( N \) gridpoints.

Plots of Eq. (15) as a function of the number of PEs \( (p) \) are shown in Fig. 4 for different grid sizes \( (N = 200^3, 500^3, 750^3) \) as dashed lines. The dashed curves show that the computation of large grids \( (N > 500^3) \) becomes possible only when using a large number of PEs \( (p > 50) \), which are available on massive parallel supercomputers only. Intermediate grid sizes \( (N \approx 200^3) \), however, can be computed with a comparatively small number of PEs \( (10–20) \), for example on a cluster of PCs and workstations.

4.5. Performance results

Amdahl’s law (Amdahl, 1967) provides an estimate of how much faster an algorithm will run when executed in parallel. It states that if only a fraction \( f \) of the operations in a programme can be carried out in parallel, the maximum speedup, i.e. serial execution time divided by parallel execution time, on \( p \) processors is

\[
S = \frac{p}{f + p(1 - f)} \tag{16}
\]

This is bounded by \( 1/(1 - f) \), regardless of the number of processors.

4.5.1. CRAY T3E

The speedup of the parallel viscoelastic FD code \( (L = 1) \) on the massive parallel supercomputer CRAY T3E LC 384 at ZIB Berlin was investigated. A maximum number of 384 DEC Alpha EV5.6 PEs were available, the slowest PE ran with 450 Mhz. The transfer-rate of the network was 480 MB/s. The influence of the number of PEs on wall-clock times for wavefield update and communication were measured. The results for grids sizes \( N = 200^3 \) and \( N = 500^3 \) are plotted in Fig. 5. The wall-clock time required for one timestep decreases
significantly with increasing number of PEs. A large grid with \( N = 500^3 \) gridpoints (total memory required: 11 GBytes) needs only 10 s on 343 PEs for the computation of one timestep. Note that a single PE would need approximately 57 min for one timestep. Parallelization allows modelling of large-scale models in acceptable running times. Note that the communication time can always be neglected.

The observed speedups for the CRAY T3E are shown in Fig. 6. Surprisingly, the solid speedup curve for \( N = 200^3 \) gridpoints lies above the linear speedup line (dashed). This means we observe superlinear speedup, i.e. a run with 343 PEs is 370 times faster than a serial execution, resulting in a parallelizable fraction \( f \) of 1.0002 (Eq. (16)). The parallelizable fraction for \( N = 500^3 \) gridpoints (dashed curve in Fig. 6) is \( f = 0.9999 \).

### 4.5.2. Linux cluster

The same performance analysis was carried out on an open in-house network of 20 Linux PCs connected by a 100 Mb/s Ethernet switch. The slowest PC ran with 300 Mhz. Due to the large amount of memory only the \( N = 200^3 \) grid could be calculated. Measured wall-clock times and speedups are plotted in Fig. 7 and Fig. 8, respectively. As expected, wall-clock times for the wavefield update (dashed line) decrease with increasing number of PEs. However, communication time is increasing for \( p > 12 \) significantly. This leads to stationary total computation times (solid line). Consequently, the speedup curve shown in Fig. 8 shows no improvement for \( p > 12 \). The reason of this poor performance is the strong increase of communicated data (Fig. 4) leading to an increase of communication times within the (slow) ethernet. Even for high communication times (high network load) and high number of PCs our network remains stable.

### 5. Examples

In this section it is described how the parallel viscoelastic FD programme has been applied to simulate
seismic wave propagation through a 3-D random medium. Effects of scattering and intrinsic attenuation are studied using two acquisition geometries: a simple plane wave transmission geometry, and a vertical seismic profile (VSP) geometry. These examples demonstrate the capability of the program to efficiently model seismic wave propagation in arbitrary heterogeneous viscoelastic media. The parameter files which were used in the examples are included in the provided program package. The user should thus be able to reproduce the results described below.

5.1. 3-D random media model

The 3-D random medium used in the wave propagation simulations (Fig. 9) contains $240 \times 240 \times 600$ grid-points in $x$, $y$, $z$-direction, respectively. The grid spacing is 2 m. The P-wave velocity ($v_p$) is Gaussian distributed about a mean of 3000 m/s. The standard deviation of the P-velocity perturbation is 5%.

Shearwave velocity ($v_s$) and density values ($\rho$) are derived from P-velocities by applying the following empirical relations which were derived for sandstones:

$$v_s = 314.59 + 0.61v_p$$

and

$$\rho = 1498.0 + 0.22v_p$$

(Han, 1986). The medium parameters are of the order of reservoir rocks which are targets in hydrocarbon exploration.

The isotropic autocovariance function of the medium fluctuations is of the form

$$A(r) = \frac{\sigma^2}{2^{n-1}I(n)}(\frac{r}{a})^n K_0(\frac{r}{a}),$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the lag, $n$ the so-called Hurst coefficient, $a$ the correlation length, $\sigma$ the standard deviation of the fluctuations, $I$ the gamma function, and $K_0$ the modified Bessel function of the second kind of order 0. Eq. (17) is a so-called von Karman autocovariance function which characterizes stochastic processes which are self-affine, or fractal, at scales smaller than $a$. The fractal dimension $D$ of a 3-D medium is $D = 4 - n$. For $n = 1$ one obtains smooth fluctuations (fractal dimension $D = 3$), whereas for $n = 0.1$ the fluctuations are rough ($D = 3.9$). A comparative study of upper-crystal sonic logs revealed that P-wave velocity fluctuations are self-affine with Hurst numbers ($n$) lying between 0.1 and 0.2 (Holliger, 1987). The 3-D random medium used in the numerical experiment was generated using a Hurst number of $n = 0.15$ and a correlation length of 45 m (Fig. 9). The random medium is generated by 3-D Fourier transforming the autocovariance function 17, multiplying the square root of the amplitude spectrum with a random phase between $-\pi$ and $\pi$, and inverse 3-D Fourier transformation.
In a last step the Gaussian medium fluctuations are scaled to the desired standard deviation $\sigma$.

5.2. Transmission experiment

A simple transmission experiment is used to study wave propagation through the random medium. The parallel viscoelastic FD program is used to simulate a plane compressional wave with a dominant frequency of approximately $f_c = 70$ Hz starting at the top of the random medium (Fig. 9) and propagating downwards. To avoid artificial damping of the plane wave with traveltime, periodic boundary conditions at all edges except at the top and bottom were applied. At the top and bottom absorbing boundaries were installed. 3-D modelling was performed on the massive parallel supercomputer CRAY T3E LC 384 at ZIB Berlin. The total memory requirement was approximately 4 Gbytes. Computing time was approximately 5 h for 2000 time-steps on 128 CPUs.

The transmitted wavefield is recorded at geophones lying within a plane perpendicular to the propagation direction (vertical direction) (dashed line in Fig. 9). The travel distance is $L = 980$ m corresponding to 22 times the correlation length. Elastic and viscoelastic simulations were performed. The quality factor $Q$ as a function of frequency, shown in Fig. 11, was applied everywhere in the viscoelastic model for both P- and S-waves.

In Fig. 10 synthetic seismograms (vertical component of particle velocity) for the elastic and viscoelastic case are compared. The direct wave arrives at approximately 0.33 s at the receivers. Due to scattering effects the primary wave shows significant lateral fluctuations of amplitude and traveltime. In the elastic and viscoelastic case amplitudes of the direct pulse decrease with travelpath due to scattering attenuation. In the viscoelastic case seismic wave amplitudes are additionally attenuated by intrinsic attenuation with an intrinsic $Q$ of approximately 50 in the investigated frequency range (Fig. 11). Intrinsic attenuation leads to a significant loss of high frequencies with travel distance (low-pass filter effect). Since scattering effects depend strongly on frequency content of the incident wave, intrinsic attenuation leads to a different overall wavefield (compare Fig. 10A and B).

5.3. VSP experiment

In the second example seismic wavefield is generated by an explosive point source (black dot) located at the top of the model (Fig. 9), and recorded along a vertical seismic profile (VSP) indicated by the solid vertical line. A free surface boundary condition is applied at the top of the model, while absorbing boundary conditions (Cerjan et al., 1985) are applied at the other edges. 2-D and 3-D finite-difference modelling results are compared for the elastic and viscoelastic case in Fig. 12. The 2-D simulations were performed using a 2-D implementation

Fig. 10. Scattered wavefield (vertical component) of plane wave which has travelled $L = 980$ m through 3-D random medium shown in Fig. 9: (A) Elastic case, (B) viscoelastic case. Amplitudes in (B) are scaled by factor of 4.9. $Q$ as function of frequency simulated in (B) is shown in Fig. 11.

Fig. 11. Quality factor $Q$ as function of frequency (solid line) used in viscoelastic modelling ($L = 1$, $\tau = 0.04$, and relaxation frequency $f_l = 2\pi/\tau_\ell = 70$ Hz in Eq. (5)). Dashed line represents amplitude spectrum of source wavelet.
of the 3-D program. The 2-D random medium used in the 2-D modelling is simply a vertical slice containing the receivers and the shot position through the 3-D model. In the viscoelastic simulations the frequency dependence of the quality factor \( Q \) as shown in Fig. 11 was applied at all gridpoints.

Three main events can be clearly identified in the seismic sections shown in Fig. 12: (1) the downgoing (direct) P-wave denoted by P, (2) the downgoing (direct) S-wave denoted by S, and (3) the Rayleigh-wave generated by the free surface denoted as R. Scattered waves which follow the main events are more pronounced in the elastic case than in the viscoelastic case due to intrinsic attenuation. Different waveforms of the main events in the elastic and viscoelastic modelling can be observed, especially for the direct S-wave. Differences in waveform for elastic and viscoelastic modelling were also observed in the transmission experiment described in the previous example (Fig. 10). Thus, both examples lead to the conclusion that intrinsic attenuation should be considered for full wavefield interpretations in complex media.

The difference between 2-D and 3-D modelled wavefields is moderate. In 3-D model the seismic coda which is generated by multiple scattering is more pronounced than in 2-D. In the viscoelastic case this is not that severe since multiple scattered waves are stronger attenuated in the viscoelastic medium. The deviation between the 2-D and 3-D modelled direct wavefield is increasing with traveltime.

6. Conclusions

The examples demonstrate the capability of the viscoelastic finite-difference method to efficiently model seismic wave propagation in 3-D heterogeneous viscoelastic media. For full wavefield interpretation in 3-D complex media viscoelastic effects (intrinsic attenuation) should be considered in the modelling.

In this paper a parallel implementation of a 3-D viscoelastic FD code is described. Parallelization is based on domain decomposition. Communication is performed by using the message passing interface (MPI) standard. It is shown that by parallel FD modelling computing times can be reduced and possible grid sizes can be increased significantly. The use of parallel computer technology opens new avenues to the study of 3-D seismic wave propagation in complex media.

A software package containing the source code, various utilities, description of programme usage, and the parameter-files used to generate the numerical results presented above, is provided\(^4\). The program can be run on a cluster of conventional PCs connected via standard Ethernet or on massive parallel supercomputers. On massive parallel supercomputers the performance is excellent even for large grids \((N > 500^3)\). On a PC cluster, however, a fast network is required to achieve good performance for large grids.

Acknowledgements

I am grateful to Bernd Milkereit for discussion and comments on the original version of the manuscript. Modelling was performed on the massive parallel supercomputer CRAY T3E LC 384 at ZIB Berlin, and on an in-house network of 20 Linux PCs.

\(^4\)Source Code of FDMPI plus Users Guide. [http://www.geophysik.uni-kiel.de/~tbohlen/fdmpi](http://www.geophysik.uni-kiel.de/~tbohlen/fdmpi)
Appendix A. 3-D viscoelastic finite difference equations

For the approximation of the spatial partial derivatives in the wave Eqs. (9)–(11) a fourth-order staggered forward operator \( D_x^+ \) and a backward operator \( D_x^- \) are applied (Levander, 1988):

\[
\frac{\partial f(x)}{\partial x} \bigg|_{(i+1/2)h} \approx D_x^+[f(i)] = \frac{1}{24h} \left(-f(i) + 27f(i+1) - f(i+2)ight),
\]

\[
\frac{\partial f(x)}{\partial x} \bigg|_{(i-1/2)h} \approx D_x^-[f(i)] = \frac{1}{24h} \left(-f(i) + 27f(i-1) - f(i-2)\right),
\]

(\text{A.1})

The operators \( D_x^+ \) and \( D_x^- \) approximate the partial derivative of a continuous function \( f(x) \) at \( i^+ = (i + 1/2)h \) and \( i^- = (i - 1/2)h \), respectively. The distance between two gridpoints is denoted by \( h \) so that \( i = x/h \).

The temporal partial derivatives \( (\partial / \partial t) \) are approximated by a Crank–Nicholson scheme:

\[
\frac{\partial f^\tau(x,y,z,t)}{\partial t} \bigg|_{(i,j,k)} \approx \frac{f^\tau(i,j,k) - f^\tau(i,j,k)}{\Delta t},
\]

(\text{A.2})

where \( i,j,k \) are the indices for the three spatial directions \( (x,y,z) \) and time \( t \), respectively. \( \Delta t \) denotes the size of a timestep.

By applying these operators to the differential Eqs. (9)–(11) one obtains the following explicit FD scheme:

\[\sigma_{xy}^\tau(i^+,j^-,k) = \sigma_{xy}^\tau(i^+,j^-,k) + \mu(i^+,j^-) + M(i^+,j^-), \]

\[\Delta t \{ 1 + L^\tau(i^+,j^-,k) \} \{ D_x^+ [v_x^\tau(i^+,j^-,k)] \}
+ D_x^- [v_x^\tau(i,j,k)] + \frac{\Delta t}{2} \left( r_{xxy}^\tau(i^+,j^-,k) \right), \]

\[\sigma_{zz}^\tau(i,j^+,k) = \sigma_{zz}^\tau(i,j^+,k) + \mu(i,j^+) + M(i,j^+), \]

\[\Delta t \{ 1 + L^\tau(i,j^+,k) \} \{ D_x^+ [v_x^\tau(i,j^+,k)] \}
+ D_x^- [v_x^\tau(i,j^-,k)] + \frac{\Delta t}{2} \left( r_{xz}^\tau(i,j^+,k) \right), \]

\[\sigma_{zz}^\tau(i,j^+,k) = \sigma_{zz}^\tau(i,j^+,k) + \mu(i,j^+) + M(i,j^+), \]

\[\Delta t \{ 1 + L^\tau(i,j^+,k) \} \{ D_x^+ [v_x^\tau(i,j^+,k)] \}
+ D_x^- [v_x^\tau(i,j^-,k)] + \frac{\Delta t}{2} \left( r_{xz}^\tau(i,j^+,k) \right), \]

(\text{A.3})

(\text{A.4})

with the 6L memory variables \( (l = 1, \ldots, L) \):

\[r_{xxy}^\tau(i,j^+,k) = \left( 1 + \frac{\Delta t}{2 \tau_0} \right)^{-1} \left\{ \left( 1 - \frac{\Delta t}{2 \tau_0} \right) r_{xxy}^\tau(i,j^+,k) \right\}, \]

(\text{A.5})
\[ r_{x(i,j),k}^{(m)} = \left(1 + \frac{\Delta t}{2\tau_{ol}}\right)^{-1} \left\{ \left(1 - \frac{\Delta t}{2\tau_{ol}}\right) r_{x(i,j),k}^{(m)} \right. \right. \\
\left. \left. - \frac{M(i,j,k)\Delta t}{\tau_{ol}} e^{(i,j),k} \right. \right. \\
\left. \left. + D_x[v_x^{(i,j),k}] + D_y[v_y^{(i,j),k}] \right. \right. \\
\left. \left. + D_z[v_z^{(i,j),k}]\right) \right\}, \quad (A.12) \]

\[ r_{y(i,j),k}^{(m)} = \left(1 + \frac{\Delta t}{2\tau_{ol}}\right)^{-1} \left\{ \left(1 - \frac{\Delta t}{2\tau_{ol}}\right) r_{y(i,j),k}^{(m)} \right. \right. \\
\left. \left. - \frac{M(i,j,k)\Delta t}{\tau_{ol}} e^{(i,j),k} \right. \right. \\
\left. \left. + D_x[v_x^{(i,j),k}] + D_y[v_y^{(i,j),k}] \right. \right. \\
\left. \left. + D_z[v_z^{(i,j),k}]\right) \right\}, \quad (A.13) \]

\[ r_{z(i,j),k}^{(m)} = \left(1 + \frac{\Delta t}{2\tau_{ol}}\right)^{-1} \left\{ \left(1 - \frac{\Delta t}{2\tau_{ol}}\right) r_{z(i,j),k}^{(m)} \right. \right. \\
\left. \left. - \frac{M(i,j,k)\Delta t}{\tau_{ol}} e^{(i,j),k} \right. \right. \\
\left. \left. + D_x[v_x^{(i,j),k}] + D_y[v_y^{(i,j),k}] \right. \right. \\
\left. \left. + D_z[v_z^{(i,j),k}]\right) \right\}, \quad (A.14) \]

\[ v_x^{(i,j),k} = v_x^{(i,j),k-1} + \frac{\Delta t}{\tau_{ol}(i,j,k)} \left( D_x[v_x^{(i,j),k}] + \frac{M(i,j,k)\Delta t}{\tau_{ol}} e^{(i,j),k} \right) \]

\[ v_y^{(i,j),k} = v_y^{(i,j),k-1} + \frac{\Delta t}{\tau_{ol}(i,j,k)} \left( D_y[v_y^{(i,j),k}] \right) \]

\[ v_z^{(i,j),k} = v_z^{(i,j),k-1} + \frac{\Delta t}{\tau_{ol}(i,j,k)} \left( D_z[v_z^{(i,j),k}] \right) \]

\[ \left( D_x[v_x^{(i,j),k}] + D_y[v_y^{(i,j),k}] + D_z[v_z^{(i,j),k}] \right) \]

This scheme requires 9 + 6L dynamic (time-dependent) variables (six stress components plus 6L memory variables plus three components of particle velocity) and five material parameters (\(\mu, M, g, \tau^x, \tau^y\)) to be stored in every cell (see Fig. 2). A directive force can be implemented by assigning the body force components \( f_i \) at the source point with the source wavelet.

**References**


