Tutorial for Wave Propagation and Full Waveform Inversion

Christian Wieners, Johannes Ernesti, Christian Rheinbay

Getting and compiling the Code

The code can be obtained from

```
http://www.math.kit.edu/ianm3/seite/mfoseminar/en
```

Following `README` the code is installed by

```
cd FWI_Tutorial
mkdir build
cd build
cmake ..
make -j
```

and started in the directory `FWI_Tutorial/build` with

```
mpirun -n <procs> M++
```

using the default configuration file in `../conf/` specified in `../conf/m++conf`.
In the following, we use `procs` $= 4$.
The logfiles are in `log/`, the plotting data in `data/vtk/` and can be viewed with `paraview`. The seismogram data in `data/FWI/` are displayed with

```
python3 ../tools/seismogram_lib.py data/FWI/<sname1> data/FWI/<sname2> ...
```

A) Time stepping methods for DG approximations of the acoustic wave equation

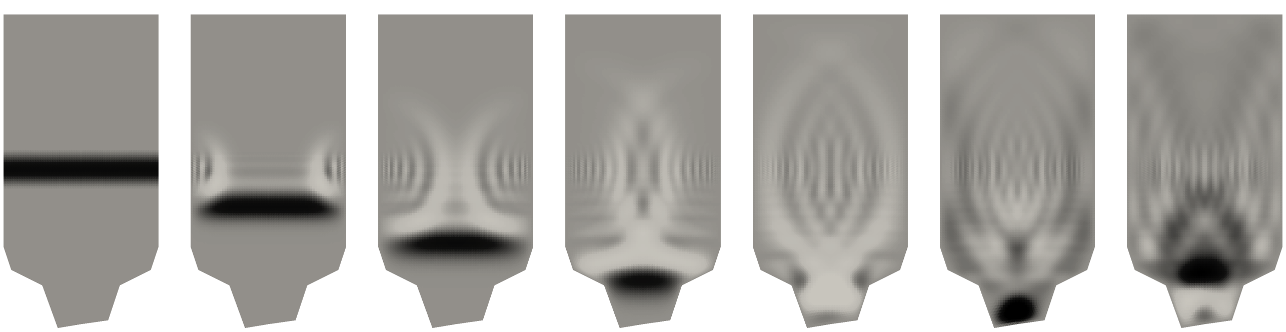
We consider an acoustic wave $\partial_t \mathbf{v} = \nabla p$ and $\partial_t p = \operatorname{div} \mathbf{v}$ in $\Omega \subset (0,4) \times (-2.1, 6) \subset \mathbb{R}^2$, cf. Fig. 1. At $t = 0$ we start with the initial values

$$\begin{pmatrix} p(x_1,x_2) \\ \mathbf{v}(x_1,x_2) \end{pmatrix} = \begin{cases} \exp(-4(y_{\text{mid}} - x_2)^2)\,(1 - 4(y_{\text{mid}} - x_2)^2) \begin{pmatrix} 100 \\ -100 \\ 0 \end{pmatrix} & \text{for } |y_{\text{mid}} - x_2| < \frac{1}{2}\,, \\ \begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix} & \text{else.} \end{cases}$$

for all $\mathbf{x} = (x_1, x_2)^\top \in \Omega$. The location in x_2-direction of the plane wave is controlled by the variable $y_{\text{mid}} \in \mathbb{R}$.
The final time is $T = 6$.
Starting with

```
mpirun -n 4 M++ acoustic
paraview &
```

yields the results in Fig. 1 by loading the data `P` in `data/vtk/` in `paraview`.



$t = 0$ $t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$ $t = 6$

FIGURE 1. Solution for the pressure distribution of the acoustic wave propagation.

We consider the linear goal functional in the region of interest $S = \{T\} \times (1,3) \times (-1,0)$

$$E(p,\mathbf{v}) = \frac{1}{2}\int_S p(T,\mathbf{x})\,\mathrm{d}\mathbf{x}$$

(a) Compare the results for the linear goal functional $E(p,\mathbf{v})$ for different mesh size $\triangle x$, time step size $\triangle t$, and polynomial degree q of the DG method. Therefore, modify in `../conf/acoustic.conf` the parameters `level`, `dt`, and `deg`.

(b) Convergence in space is tested with `plevel` < `level` and `Scales`= 1 with `ScaleFactor`= 1; convergence in time is tested with `plevel` = `level` and `Scales`> 1 with `ScaleFactor`= 0.5. Select the parameters $(\triangle x, \triangle t)$ so that the discretization error in space and time is equilibrated.

(c) By comparing the results, estimate the asymptotic value for the goal functional. Then, find out which parameters $(\triangle x, \triangle t, q)$ are sufficient to obtain $E(p,\mathbf{v})$ up to an accuracy of approx. 1%.

(d) Compare the results for different time integrators (implicit midpoint rule, classical explicit Runge-Kutta method, polynomial Krylov method, and Arnoldi method with shift). Therefore, modify in `../conf/acoustic.conf` the parameter `rkorder` and find out for $q = 2$, which parameters $(\triangle x, \triangle t)$ yield for $E(p,\mathbf{v})$ up an accuracy of approx. 1%.

Challenge. Which configuration $(\triangle x, \triangle t, q)$ and time integrator is optimal with respect to the computing time.

B) Adaptive space-time approximation of the acoustic wave equation

Now start for the example in A) the space-time method

```
mpirun -n 4 M++ spacetime_acoustic
paraview &
```

and compare space-time results in `data/vtk/` with the time-stepping results. The number of adaptive cycles can be set in `../conf/spacetime_acoustic.conf` with `refinement_steps`, and the marking of cells is controlled by the parameter `theta`. Time slices of the space-time solution `P.vtk` can be visualized with the `clip` method, and the refinement plots show the distribution of the polynomial degree.

Challenge. Compute the value $E(p,\mathbf{v})$ up to an accuracy of approx. 1% with the adaptive space-time method. Find out by numerical experiments how many degrees of freedom in space and time are required and can be saved in comparison with a computation on a uniform discretizations in time and space.

C) Solving the forward problem in seismic imaging

In this simple example, we consider the 2D acoustic wave equation $\rho_0\partial_t\mathbf{v} = \nabla p$ and $\partial_t p = \kappa\,\mathrm{div}\,\mathbf{v} + b$ with constant density ρ_0 and variable permeability κ. The first-order system is discretized with a discontinuous Galerkin method of degree `deg` $= q$ in space and by the implicit midpoint rule in time.

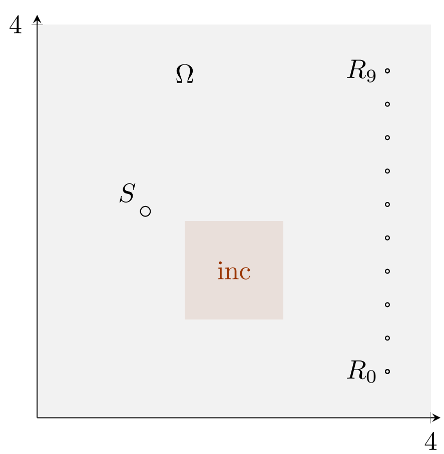


FIGURE 2. The wave that is exited at $S \in \Omega$ propagates through the domain and is recorded at the receiver positions $R_0,\ldots,R_9 \in \Omega$.

We start with an approximate point source b located at

```
source_x = 1.1;
source_y = 2.1;
```

Parameters of the configuration can be changed in `../conf/forward.conf`.

(a) Run the forward simulation with `procs` $= 4$ processes, e.g.,

```
mpirun -n 4 M++ forward
```

This creates snapshot plots of the wave fields in the directory `data/vtk` and seismograms in `data/FWI` (with `../tools/rm_data`, all previous results are removed).

Use the program `paraview` to visualize the pressure `P` and the material distribution `Kappa` and `Rho` in `data/vtk`.

(b) Now test different source locations and compare the seismograms. Therefore, set different `sname` for different source locations in the configuration file and compare

```
python3 ../tools/seismogram_lib.py data/FWI/<sname1> data/FWI/<sname2> ...
```

By adding `-n` to the call of `seismogram_lib.py` each trace gets scaled by its individual amplitude (making differences more visible).

The parameter `-t r` when calling `seismogram_lib.py` calculates the relative seismogram differences. For further options, see `python3 ../tools/seismogram_lib.py -h`.

How sensitive is this measurement? What is the maximal distance of the sources, so that the relative difference of the seismograms is smaller than 0.8?

(c) In our test, we fix the density ρ and the background permeability κ_{bg}, and in a small region (cf. Fig. 2) the permeability is changed to

```
kappa_inc = 1.5;
```

How sensitive is the measurement? Which difference of κ_{bg} and κ_{inc} is required, so that the relative difference of the seismograms for the source position (1.1/2.1) is smaller than 0.5?

D) Convergence of the forward problem

Solving partial differential equations (PDEs) numerically yields an approximation of the true solution of the PDE. Here, we consider how the approximation behaves under mesh refinement in space and time.
We use a uniform mesh of mesh width $\triangle x$, fixed time increments $\triangle t$, and a fixed polynomial degree in the ansatz spaces. In `../conf/forward.conf`, the discretization parameters are defined by

```
level = 4;
deg = 1;
dt = 0.01;
```

Use `sname = automatic` in the configuration file to let the program pick a name according to the discretization. Further, reset the source position to `source_x` $= 1.1$, `source_y` $= 2.1$, and the inclusion to `kappa_inc = 1.5`. Make a few tests with different discretization parameters and compare the seismograms.

(a) Run simulations for `deg = 1, 2` and `dt = 0.01` with `level = 3, 4` and compare the resulting seismograms to the reference solution provided by `FWI/s_ref_1.1_2.1__kappa_inc_1.5`.
Which convergence order do you observe by refining the `level`?

(b) Determine the convergence order in time by computing the solution for `deg = 1, 2` on `level = 4` using different time step sizes.
Compare the relative differences of the seismograms for the time step sizes `dt`$= 0.08, 0.04, 0.02, 0.01$.

(c) Find a configuration which is accurate up to a relative difference in the seismograms smaller than 0.5 compared to the reference solution.

E) The inverse problem

We use the forward simulation to create synthetic data for the inverse problems. In this example, we use the same discretization for both, generation of data and inversion.
The data for three shots are computed with

```
mpirun -n 4 M++ generate
```

and, starting with $\kappa \equiv \kappa_0$, we recover the material distribution approximately by

```
mpirun -n 4 M++ invert
```

using CG-REGINN combining Newton's method and conjugate gradients for the normal equation in every Newton step.
Here, $m = (\rho, \kappa) \in \mathcal{P} \subset \mathrm{L}_\infty(\Omega; \mathbb{R}^2)$ are the material parameters,

$$\mathbf{y} = (p, \mathbf{v}) \in Y \subset \mathrm{C}^1\big(0, T; \mathrm{H}^1(\Omega) \times \mathrm{H}(\mathrm{div}, \Omega)\big) \subset \mathrm{L}_2(Q; \mathbb{R} \times \mathbb{R}^2)$$

are the state variables in the space-time cylinder $Q = (0, T) \times \Omega$, and

$$L_m \mathbf{y} = (\rho \partial_t \mathbf{v} - \nabla p, \partial_t p - \kappa \operatorname{div} \mathbf{v})$$

is the first-order system corresponding to the acoustic wave equation for the parameter distribution $m = (\rho, \kappa)$. This defines the parameter-to-wave-field map

$$\mathcal{F} \colon \mathcal{P} \longrightarrow Y\,, \qquad m \longmapsto \mathbf{y} \quad \text{with} \quad L_m \mathbf{y} = (\mathbf{0}, b)\,.$$

Given a set of space-time receivers $\mathcal{R} \subset Q$, every wave solution defines a seismogram in $\mathcal{S} = \mathbb{R}^{\mathcal{R}}$. The corresponding measurement operator

$$\Psi \colon \mathrm{L}_2(Q; \mathbb{R} \times \mathbb{R}^2) \longrightarrow \mathcal{S}\,, \qquad \mathbf{y} \longmapsto \sum\nolimits_{r \in \mathcal{R}} (\boldsymbol{\varphi}_r, \mathbf{y})_Q$$

is defined by measurement kernels $\boldsymbol{\varphi}_r$ approximating point evaluations at $r \in \mathcal{R}$.
Together, this defines the parameter-to-seismogram map $\Phi = \Psi \circ \mathcal{F} \colon \mathcal{P} \longrightarrow \mathcal{S}$.
In this setting, the problem of Full-Waveform-Inversion reads as follows:

$$\text{Given } s_{\text{obs}} \in \mathcal{S}, \text{ find } m \in \mathcal{P} \text{ with} \quad \Phi(m) = s^{\delta}_{\text{obs}}\,.$$

This problem is solved approximately by the Newton method.
Therefore, let $\Phi'(m) \colon \mathcal{P} \longrightarrow \mathcal{S}$ by the linearized parameter-to-seismogram map.

Algorithm 1 Newton's method with approximate updates

Choose $m^0 \in \mathcal{P}$, $k \leftarrow 0$
while not converged **do**
    $r^k \leftarrow s_{\text{obs}} - \Phi(m^k) \in \mathcal{S}$
    Select $\vartheta_k > 0$.
    Find $\triangle m^k \in \mathcal{P}$ with $\|\Phi'(m^k)[\triangle m^k] - r^k\| \le \vartheta_k \|r^k\|$.
    $m^{k+1} \leftarrow m^k + \triangle m^k$
    $k \leftarrow k + 1$

Since the linearization in this application in general is ill-conditioned, the approximate update $\triangle m^k \in \mathcal{P}$ is computed by the CG iteration for the corresponding normal equation.

Algorithm 2 Conjugate gradient algorithm for $\Phi'(m)^*\Phi'(m)[\triangle m] = \Phi'(m)^* r$

$j \leftarrow 0$, $\beta \leftarrow 0$, $r^0 \leftarrow r \in \mathcal{S}$
$p^0, \triangle m^0 \leftarrow 0 \in \mathcal{P}$
while not converged **do**
    $j \leftarrow j + 1$
    $d \leftarrow \Phi'(m)^*[r^{j-1}] \in \mathcal{P}$
    $p^j \leftarrow d + \beta\, \|d\|^2\, p^{j-1}$
    $q \leftarrow \Phi'(m)[p^j] \in \mathcal{S}$
    $\alpha \leftarrow \|d\|^2 / \|q\|^2$
    $\triangle m^j \leftarrow \triangle m^{j-1} + \alpha\, p^j$
    $r^j \leftarrow r^{j-1} - \alpha\, q$
    $\beta \leftarrow 1/\|d\|^2$

In our example we fix the density $\rho \equiv \rho_0$ and the a small number source functions b_n $(n = 0, 1, ..., N_{\text{shots}} - 1)$ so that the inverse algorithm only tries to recover κ. In the Newton step k, we use the source function b_n and seismogram $s_{n,\text{obs}}$ alternating with $n = k \mod N_{\text{shots}}$.

(a) Compare the seismograms in `data/FWI/`. How large is the relative difference in the observed seismogram and the initial seismogram of the inversion?
(b) How many steps in the inverse method are required to obtain a relative difference in the seismograms smaller than 0.1?
(c) Which accuracy can be obtain with the method? Why is the accuracy limited and what is required to obtain better results?