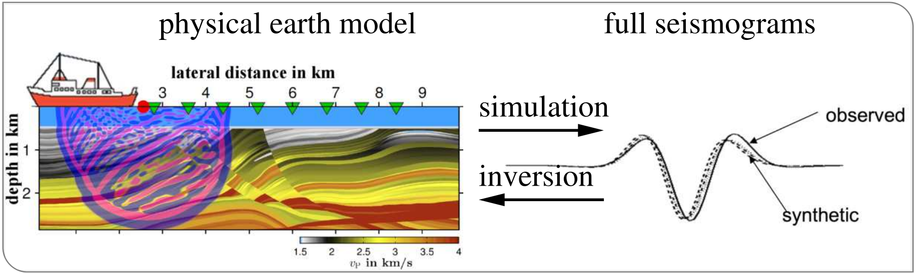
Full Waveform Inversion

Introduction and Born approximation

Thomas Bohlen



Agenda

Karlsruhe Institute of Technology

Agenda

Agenda

Goals of FWI

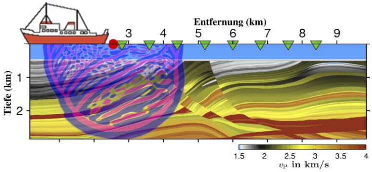
Find all earth models that predict all signals by full wave modelling !

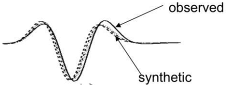
State of the art: Find one numerical model that predicts selected signals at low frequencies by full wave modelling.

Goals of FWI

Find all earth models that predict all signals by full wave modelling !

State of the art: Find one numerical model that predicts selected signals at low frequencies by full wave modelling.





P–waves

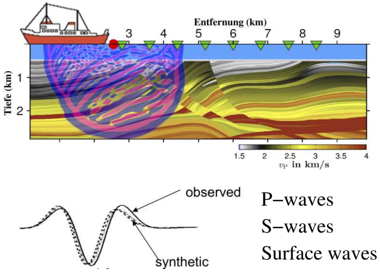
S–waves

Surface waves

Goals of FWI

Find all earth models that predict all signals by full wave modelling !

State of the art: Find one numerical model that predicts selected signals at low frequencies by full wave modelling.

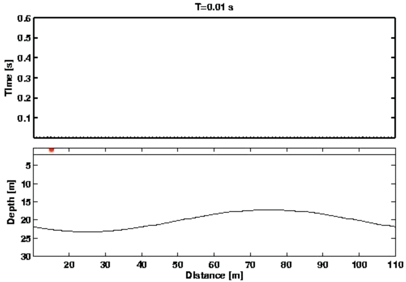


Benefits

1. **Improved resolution:** $\approx \frac{\lambda}{2}$ ☺
2. **Multi-parameter reconstruction:**
   - a. P-wave velocity ☺
   - b. S-wave velocity 😐?
   - c. Attenuation 😐?
   - d. Anisotropy 😐?
   - e. Density ☹
3. **Improved petrophysical characterization**

Elastic wave propagation is complex

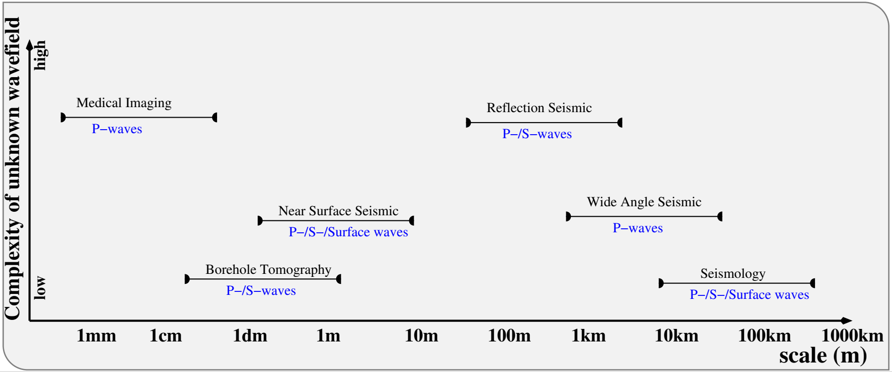
Observed seismograms contain signals of P-waves, S-waves, surface waves, mode conversions,...



Click on frame to play movie

Applications of FWI

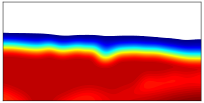
In recent 20 years FWI has been applied sucessfully to different wave types and a broad range of spatial scales covering 9 orders of magnitude



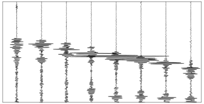
Agenda

FWI: iterative data fitting procedure

Initial model m_0

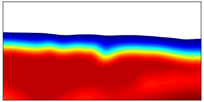


Field data d

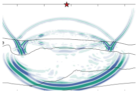


FWI: iterative data fitting procedure

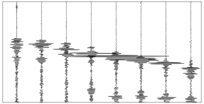
Initial model m_0



Forward modelling

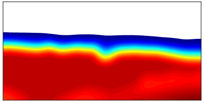


Field data d

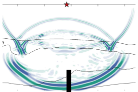


FWI: iterative data fitting procedure

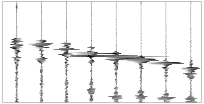
Initial model m_0



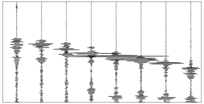
Forward modelling



Field data d

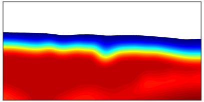


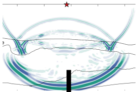
Synthetic data d'



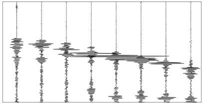
FWI: iterative data fitting procedure

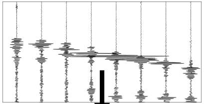
Initial model $\mathbf{m}_0$ → **Forward modelling**





Field data d → Synthetic data d'

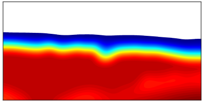




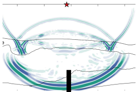
Misfit $E(\mathbf{m})$

FWI: iterative data fitting procedure

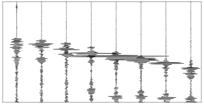
Initial model $\mathbf{m}_0$



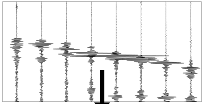
Forward modelling



Field data d



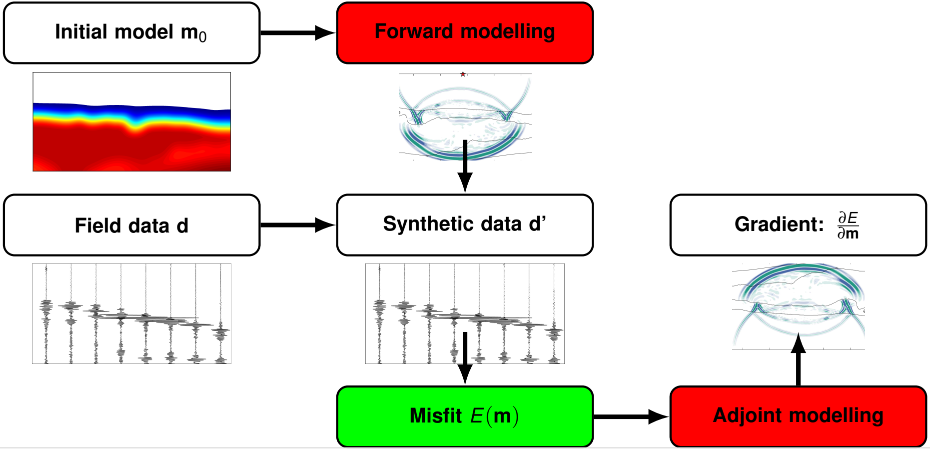
Synthetic data d'



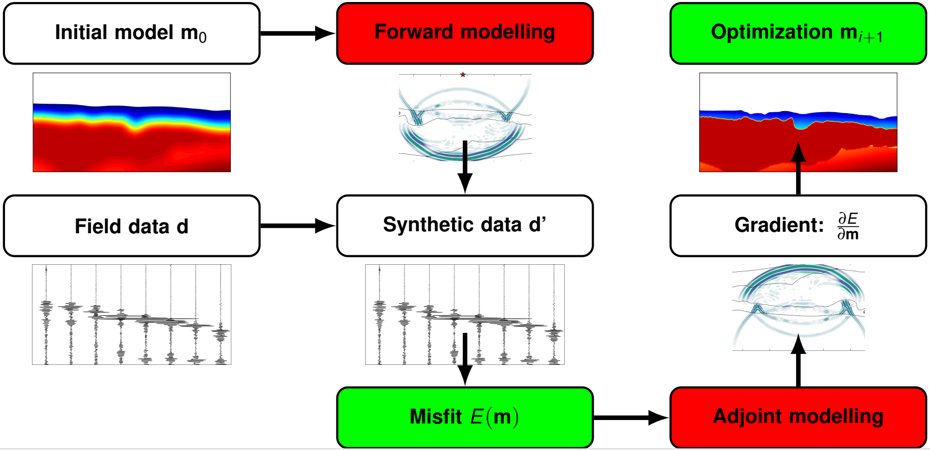
Misfit $E(\mathbf{m})$

Adjoint modelling

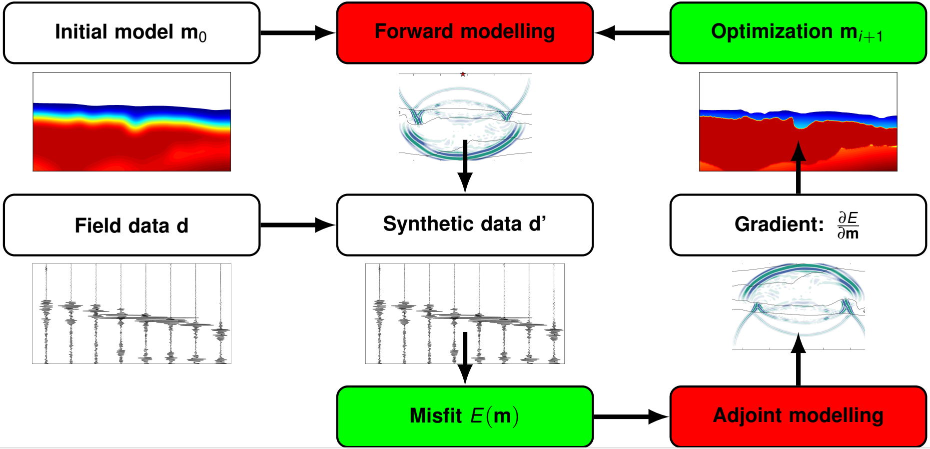
FWI: iterative data fitting procedure



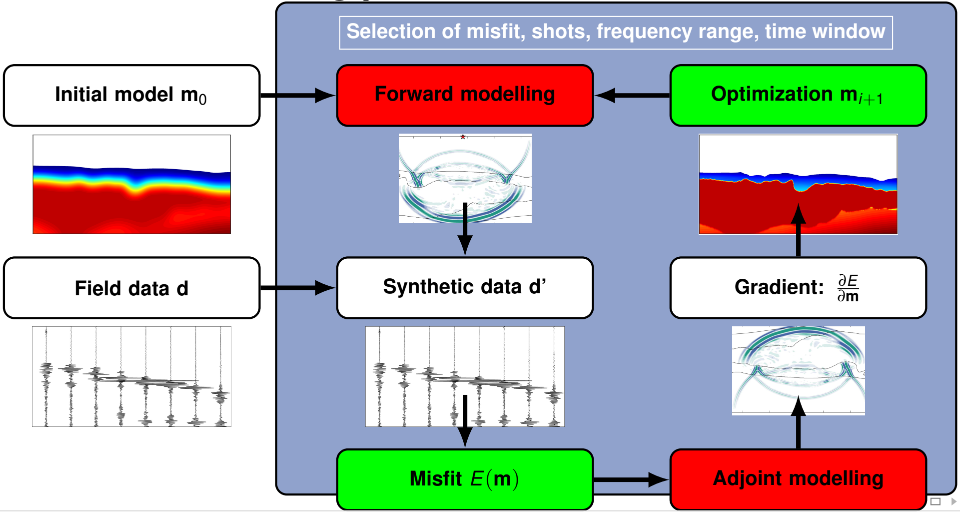
FWI: iterative data fitting procedure



FWI: iterative data fitting procedure



FWI: iterative data fitting procedure

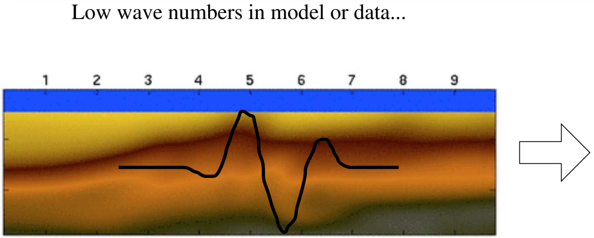


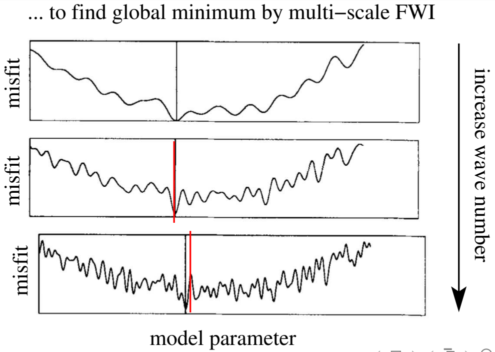
Agenda

Challenges of FWI (1/6)

Mitigate non-linearities by multi-scale approach

- we need sufficient low wave numbers in the initial model or the observed data

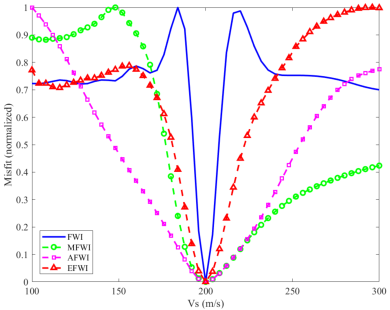




Challenges of FWI (2/6)

Suitable misfit definition

- to measure the misfit of the relevant signals
- **Normalized L2**, envelope, optimal transport,...
- defines the adjoint sources
- tradeoff between robustness (against noise, cycle skipping) and resolution



Challenges of FWI (3/6)

Appropriate physics for wave propagation

- to model the relevant signals
- multi-parameter reconstruction
- consider forward and adjoint equations

Acoustic

$$\frac{\partial^2 p}{\partial t^2} = c^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right)$$

P–waves

Elastic

$$p_{ij} = \lambda \theta \delta_{ij} + 2\mu\epsilon_{ij}$$

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$$

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial p_{ij}}{\partial x_j} + f_i$$

P–waves, S–waves
Surface waves

Visco–elastic

$$\begin{aligned} \dot{\sigma}_{ij} &= \frac{\partial v_k}{\partial x_k} \{ M(1 + L\tau^p) - 2\mu(1 + L\tau^s) \} + 2\frac{\partial v_i}{\partial x_j} \mu (1 + L\tau^s) + \sum_{l=1}^{L} r_{ijl} \quad \text{if } i = j \\ \dot{\sigma}_{ij} &= \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \mu (1 + L\tau^s) + \sum_{l=1}^{L} r_{ijl} \quad \text{if } i \neq j \ , \end{aligned}$$

$$\begin{aligned} \dot{r}_{ijl} &= -\frac{1}{\tau_{\sigma l}} \left\{ (M\tau^p - 2\mu\tau^s) \frac{\partial v_k}{\partial x_k} + 2\frac{\partial v_i}{\partial x_j} \mu\tau^s + r_{ijl} \right\} \quad \text{if } i = j \ , \\ \dot{r}_{ijl} &= -\frac{1}{\tau_{\sigma l}} \left\{ \mu\tau^s \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + r_{ijl} \right\} \quad \text{if } i \neq j \ . \end{aligned}$$

$$\varrho \frac{\partial v_i}{\partial t} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \ ,$$

P–waves, S–waves
Surface waves
Attenuation

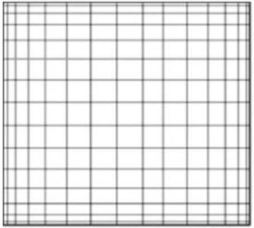
$\longrightarrow$ Computational requirements

Challenges of FWI (4/6)

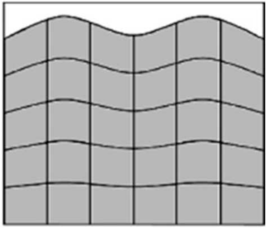
Numerical solution and space discretization

- **Finite-Differences**, Spectral elements
- Boundary condition (free surface topography is challenge with FD)

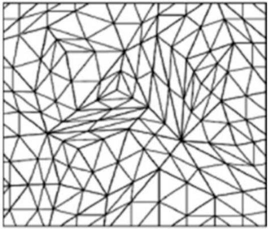
FD: Cartesian grid



FD: Stretched grid



Specfem: Triangular



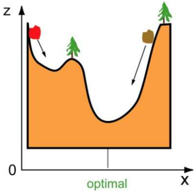
(?)

Challenges of FWI (5/6)

Optimization method

- efficient calculation of gradients by the adjoint method
- available methods: steepest-descent, conjugate gradient, L-BFGS, Gauß-Newton, Truncated Newton etc.
- consider global strategy if number of parameters is small (uncertainty estimation)

Gradient–based (local)

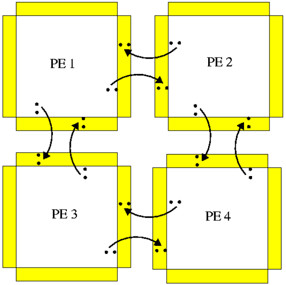


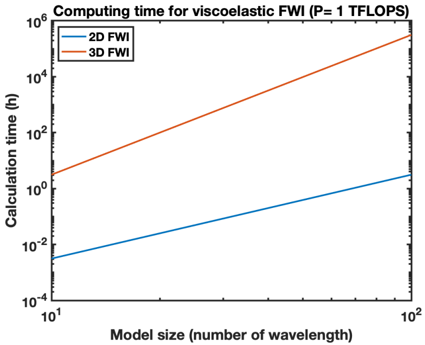
Global

Challenges of FWI (6/6)

High Performance Computing

- Efficient forward and adjoint simulation on heterogeneous architectures (CPU/GPU)





Agenda

Agenda

Karlsruhe Institute of Technology

Scattering Series and Born approximation

We use a linear relation between perturbations in the model and data space. This is called the Born approximation. It is the first linear term of a scattering series. We will derive the Born approximation for the acoustic wave equation. We define the model parameter

$$m(x) := \frac{1}{c^2(x)}, \tag{1}$$

and split it up into

$$m(x) = m_0(x) + \varepsilon m_1(x), \tag{2}$$

where $m_0(x)$ is the background model and $\varepsilon m_1(x)$ a small perturbation of this model. We do the same for the data $u(x)$, splitting it up into

$$u(x) = u_0(x) + u_{sc}(x), \tag{3}$$

where $u_0(x)$ is the background wavefield (propagating in model $m_0(x)$) and $u_{sc}(x)$ is the scattered wavefield produced by $m_1(x)$.

Wave equations

We now have two wave equations:

$$m_0 \frac{\partial^2 u_0}{\partial t^2} - \Delta u_0 = f(x,t) \qquad \text{background wavefield} \tag{4}$$

$$m \frac{\partial^2 u}{\partial t^2} - \Delta u = f(x,t) \qquad \text{total wavefield} \tag{5}$$

Subtracting equation (4) from (5) yields

$$\begin{aligned}
& m \frac{\partial^2 u}{\partial t^2} - m_0 \frac{\partial^2 u_0}{\partial t^2} - \Delta u + \Delta u_0 &= 0 \\
\Leftrightarrow \quad & (m_0 + \varepsilon m_1) \frac{\partial^2 u}{\partial t^2} - m_0 \frac{\partial^2}{\partial t^2}(u - u_{sc}) - \Delta(u - u_0) &= 0 \\
\Leftrightarrow \quad & m_0 \frac{\partial^2 u_{sc}}{\partial t^2} - \Delta u_{sc} &= -\varepsilon m_1 \frac{\partial^2 u}{\partial t^2} \qquad (6)
\end{aligned}$$

Solutions in terms of Greens functions

In the following we will describe solutions of the wave equations using so-called Greens functions. We first briefly describe the concept and properties of these functions.

Let us consider the scalar acoustic wave equation of the form

$$\left[\frac{\partial^2}{\partial t^2} - c^2(\mathrm{x})\Delta\right] u(x,t) = f(x,t) \tag{7}$$

with $u(x,0) = \frac{\partial u}{\partial t}(x,0) = 0$. u is the pressure, c the sound velocity and f the source function.

Note that x denotes a space vector $x \in \mathbb{R}^3$ and $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$.

Solutions in terms of Greens functions

Karlsruhe Institute of Technology

The Green's function $G(x, y, t)$ of the wave equation is its impulse response, which means it is the solution of the equation with a delta-function as source function:

$$\left[\frac{\partial^2}{\partial t^2} - c^2(x)\Delta_x\right] G(x, y, t) = \delta(x - y)\delta t \tag{8}$$

The source function describes a delta-impulse at the location $x = y$ and at the time $t = 0$. The Green's function depends on the receiver position x, the source position y and the time t.

If we know the Green's function, we can calculate the pressure field for any arbitrary source function by convolving the Green's function with the source function and integrating over the whole model space:

$$u(x, t) = \int_0^t \int_{\mathbb{R}^3} G(x, y, t - s) f(y, s)\, dy\, ds \tag{9}$$

Solutions in terms of Greens functions

The solution of equation (6)

$$m_0 \frac{\partial^2 u_{sc}}{\partial t^2} - \Delta u_{sc} = -\varepsilon m_1 \frac{\partial^2 u}{\partial t^2}$$

can thus be found by convolving the Green's function with the source term

$$u_{sc}(x,t) = \int_0^t \int_{\mathbb{R}^3} G_0(x,y,t-s)\left(-\varepsilon m_1 \frac{\partial^2 u}{\partial t^2}(y,s)\right) dy\, ds \tag{10}$$

It has to be considered, that the Green's function depends on the respective model. In this case, the Green's function G_0 for the background model m_0 is used.
In the following, we will use a shorter notation for the space-time integral by introducing an operator $\hat{G}_0$ that replaces the convolution and the space integral. Thus, equation (10) is written as

$$u_{sc} = -\varepsilon \hat{G}_0 m_1 \frac{\partial^2 u}{\partial t^2} = u - u_0 \tag{11}$$

Scattering operators

We obtain the solution for the total wavefield

$$u = u_0 - \varepsilon \hat{G}_0 m_1 \frac{\partial^2 u}{\partial t^2} \tag{12}$$

This equation gives an implicit relation for the wavefield u and is called the *Lippmann-Schwinger equation*. We rearrange equation 12 and use operators

$$\begin{aligned} \left[\hat{I} + \varepsilon \hat{G}_0 m_1 \frac{\partial^2}{\partial t^2} \right] u &= u_0 \\ \Leftrightarrow \quad u &= \underbrace{\left[\hat{I} + \varepsilon \hat{G}_0 m_1 \frac{\partial^2}{\partial t^2} \right]^{-1}}_{\text{scattering operator}} u_0 \end{aligned} \tag{13}$$

with the identity operator $\hat{I}$. The scattering operator describes the relation between the background field u_0 and the total field u.

Scattering series

For an operator $\hat{A}$ we can develop the expression $[\hat{I} + \hat{A}]^{-1}$ in a Neumann series (Bronštejn 2015):

$$[\hat{I} + \hat{A}]^{-1} = \hat{I} - \hat{A} + \hat{A}^2 - \hat{A}^3 + \dots \tag{14}$$

With $\hat{A} = \varepsilon \hat{G}_0 m_1 \frac{\partial^2}{\partial t^2}$, equation (13) can be written as a series, which is called the *Born series*:

$$\begin{aligned} u &= u_0 - \varepsilon(\hat{G}_0 m_1 \frac{\partial^2}{\partial t^2})u_0 + \varepsilon^2(\hat{G}_0 m_1 \frac{\partial^2}{\partial t^2})(\hat{G}_0 m_1 \frac{\partial^2}{\partial t^2})u_0 + \dots \\ &= u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots \end{aligned} \tag{15}$$

The first order term $u_1 = -\hat{G}_0 m_1 \frac{\partial^2 u_0}{\partial t^2}$, corresponds to single scattering, while the second order term $u_2 = (\hat{G}_0 m_1 \frac{\partial^2}{\partial t^2})(\hat{G}_0 m_1 \frac{\partial^2}{\partial t^2})u_0$ corresponds to double scattering, meaning that one wave is scattered at two different points on its travel path.

Born approximation

The first order approximation $u = u_0 + \varepsilon u_1$, that is considering only single scattering, is called the *Born approximation* and will be used in the following. With this approximation, the scattered wavefield is given by $u_{sc} = \varepsilon u_1$. Inserting this in equation (6) results in

$$m_0 \frac{\partial^2 \varepsilon u_1}{\partial t^2} - \Delta \varepsilon u_1 = -\varepsilon m_1 \frac{\partial^2 (u_0 + \varepsilon u_1)}{\partial t^2}$$

$$\Leftrightarrow \quad \boxed{m_0 \frac{\partial^2 u_1}{\partial t^2} - \Delta u_1 = -m_1 \frac{\partial^2 u_0}{\partial t^2}} \tag{16}$$

where we neglected the term of the order ε^2. This is a wave equation for the single scattered field. In the source term the background wavefield u_0 appears, so the scattered wavefield is induced by the background wavefield scattered at the model perturbations m_1. This secondary scattered wavefield is then propagating in the background model m_0. The background wavefield u_0 can be calculated with equation (4).

Born approximation

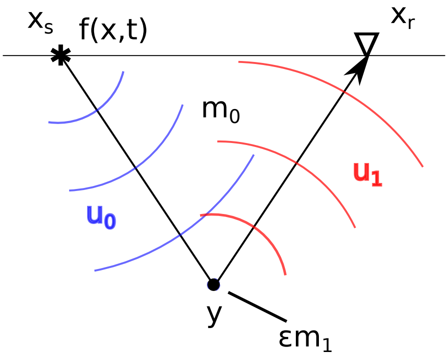


Figure: Principle of the scattering theory: The background wavefield u_0 induced by source f is scattered at the model perturbations εm_1 in point y, producing the scattered wavefield u_1.

Born approximation

Figure 1 shows the principle of the scattering theory: The source $f(x,t)$ is initiating the background wavefield u_0. At model perturbations εm_1, the background field is scattered, which leads to the additional wavefield u_1. Because we have different positions of sources and receivers, the Green's function differs for both parts of the wavefield. For the background field, we use the notation $\hat{G}_0 = G(y, x_s, t)$, and for the scattered field with the secondary source $\hat{G}'_0 = G(x_r, y, t)$. Thus, the wavefields are calculated with the Green's functions as follows:

$$u_0 = \hat{G}_0 f \tag{17}$$

$$u_1 = -\hat{G}'_0 m_1 \frac{\partial^2 u_0}{\partial t^2} = -m_1 \hat{G}'_0 \frac{\partial^2}{\partial t^2}(\hat{G}_0 f) \tag{18}$$

Frechet-derivative

With equation (18), we finally have a linear relation between the scattered wavefield u_1 and the model perturbations m_1. For this linear function, it is easy to calculate the derivative of the wavefield u_1 with respect to the model parameters m_1, which is called the *Frechet-derivative*:

$$\frac{\partial u_1}{\partial m_1} = -\hat{G}_0' \frac{\partial^2}{\partial t^2} \left(\hat{G}_0 f \right) \tag{19}$$

If we transform the Frechet-derivative into the frequency domain, the second partial derivative $\frac{\partial^2}{\partial t^2}$ will be replaced by $-\omega^2$, so that we get

$$\frac{\partial u_1}{\partial m_1} = \omega^2 \hat{\mathcal{G}}_0' \hat{\mathcal{G}}_0 f, \tag{20}$$

where $\hat{\mathcal{G}}_0$ denotes the Fourier transform of the operator $\hat{G}_0$. This derivative will be used for the inversion as described in the following section.

Summary of first lecture

- FWI in its simpelst form is a least-squares iterative data fitting procedure that can be applied on a broad range of spatial scales
- The aim of FWI to exploit the full information content of seismic data to invert multi-parameter models is a challenging task, e.g.
  - local minima of the misfit function
  - high computational efforts
- FWI is based on linear perturbation theory, i.e. the Born approximation
  - only single scattering is assumed
  - allows efficient calculation of the gradient (next lecture)

Karlsruhe Institute of Technology

Thank you for your attention

Thomas.Bohlen@kit.edu

http://www.gpi.kit.edu/

References

Bronštejn, I. N. (2015), *Handbook of Mathematics*, 6th edn, Springer, Berlin, Heidelberg.